

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO (UFRJ)
INSTITUTO DE QUÍMICA (IQ)
DEPARTAMENTO DE QUÍMICA ANALÍTICA (DQA)
TRABALHO DE CONCLUSÃO DE CURSO

**PROPOSTA DA APLICAÇÃO DA ANÁLISE POR COMPONENTES PRINCIPAIS
NO ROTEIRO DE UMA AULA DE ANÁLISE INSTRUMENTAL DO CURSO DE
GRADUAÇÃO EM QUÍMICA**

Aluna: Laryssa dos Santos

DRE: 113046430

Orientadores: M.Sc. Humberto B. de Novaes, D. Sc. Vinicius Kartnaller e D. Sc. Paula F. de
Aguiar

RIO DE JANEIRO

Março 2020

LARYSSA DOS SANTOS

**PROPOSTA DA APLICAÇÃO DA ANÁLISE POR COMPONENTES PRINCIPAIS
NO ROTEIRO DE UMA AULA DE ANÁLISE INSTRUMENTAL DO CURSO DE
GRADUAÇÃO EM QUÍMICA**

Trabalho de conclusão de curso apresentado ao Instituto de Química da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do grau de Químico com Atribuições Tecnológicas.

Orientadores: M Sc. Humberto Brevilato Novaes (Técnico em Química DQA/UFRJ)
D. Sc. Vinicius Kartnaller Montalvão (GQA/UFF)
D. Sc. Paula Fernandes de Aguiar (DQA/UFRJ)

Rio de Janeiro

Março 2020

AGRADECIMENTOS

À minha mãe por ser um exemplo de mulher forte, o qual venho tentando seguir, e por todo suporte dado no caminho das minhas conquistas e sonhos, independente dos sacrifícios exigidos.

À minha família, irmãos, primas, madrinha, padrinho, por toda a ajuda nos momentos difíceis, aos risos nos momentos felizes e por terem acreditado em mim em momentos cruciais.

À minha irmã por ser outro exemplo de mulher forte e por ter me ajudado a priorizar a minha saúde mental diante de obstáculos.

Aos meus amigos que são responsáveis por muitos momentos importantes de alegria e descontração, mesmo com a minha resistência em relaxar um pouco.

À professora e orientadora Paula Fernandes de Aguiar por ter me guiado não só ao longo desse trabalho, mas ao longo de quase toda a graduação com “brincadeiras” e elogios quando merecidos.

Ao orientador Humberto Brevilato Novaes por ter disponibilizado todas as condições operacionais para a realização desse e de outros trabalhos e por compartilhar um pouco do seu vasto conhecimento em técnicas instrumentais.

Ao professor e orientador Vinicius Kartnaller pela paciência e pelas dicas valiosas tanto para ferramenta quimiométrica selecionada quanto para a organização e parte escrita do trabalho.

Ao Laboratório LAM (Laboratório de Análises Ambientais e Minerais) por me fornecer o suporte para a realização dos experimentos.

À banca, professor Carlos Alberto da Silva Riehl e professor Rodolfo Santos Barboza, por ter aceitado o convite.

Ao Instituto de Química (IQ) e à Universidade Federal do Rio de Janeiro (UFRJ) por terem me acolhido e possibilitado o aprendizado de conhecimentos que vão além dos conceitos ensinados no curso de Química.

RESUMO

PROJETO DE CURSO

TÍTULO: PROPOSTA DA APLICAÇÃO DA ANÁLISE POR COMPONENTES PRINCIPAIS NO ROTEIRO DE UMA AULA DE ANÁLISE INSTRUMENTAL DO CURSO DE GRADUAÇÃO EM QUÍMICA

ALUNA: Laryssa dos Santos

ORIENTADORES: M. Sc. Humberto Brevilato Novaes (Técnico em Química DQA/UFRJ), D. Sc. Vinicius Tadeu Kartnaller Montalvão (GQA/UFF) e D. Sc. Paula Fernandes de Aguiar (DQA/UFRJ)

DATA DA DEFESA: 03/03/2020

PALAVRAS-CHAVE: Análise por Componente Principais, Graduação em Química Águas minerais e Cromatografia de Íons.

No final do século passado, o avanço computacional e da instrumentação científica promoveu o surgimento de uma grande quantidade de dados em diversos níveis de complexidade. Neste contexto, e com o intuito de extrair, mediante a uma abordagem multivariada, uma maior quantidade de informações relevantes dos dados gerados no laboratório, auxiliando na sua interpretação e propondo estratégias pré-experimentais, originou-se uma área nomeada Quimiometria.

Uma das frentes de pesquisa da Quimiometria e um dos focos deste trabalho é o reconhecimento de padrões, o qual tem um dos principais métodos a Análise por Componentes Principais (ACP ou PCA, do inglês *Principal Component Analysis*). Esta ferramenta, permite a visualização gráfica de todo o conjunto de dados, principalmente, quando o número de amostras e variáveis é elevado. Contudo, esta técnica faz uso, principalmente, de conceitos de álgebra e estatística, áreas que têm foco relativamente restrito nos cursos de Química em geral.

Entretanto, o uso de programas estatísticos bem construídos permite substituir a complexidade matemática por uma simplicidade didática. Dessa forma, este trabalho visou propor um roteiro de aula, no qual introduz os conceitos

principais da ferramenta estatística selecionada (ACP), usando uma rotina computacional que pode ser aplicada à uma plataforma de uso gratuito (GNU Octave).

A disciplina selecionada para a utilização do roteiro foi a Análise Instrumental I (código IQA366) ministrada para os alunos do curso de Química com Atribuições Tecnológicas da Universidade Federal do Rio de Janeiro (UFRJ), na prática experimental de análise da composição de águas minerais por meio da técnica de cromatografia de íons.

As variáveis selecionadas foram as concentrações de 7 ânions (F^- , Cl^- , NO_2^- , Br^- , NO_3^- , PO_4^{3-} e SO_4^{2-}) em 16 águas minerais de marcas brasileiras e estrangeiras. Ao quantificar os teores dos íons nas águas minerais, os íons Br^- e PO_4^{3-} se apresentaram abaixo dos seus respectivos limites de quantificação e não foram utilizados no estudo da ACP.

A ferramenta quimiométrica (ACP) foi aplicada através de uma rotina preparada na plataforma GNU Octave 5.1.0, o qual possibilitou o cálculo e construção dos gráficos para os *scores* e *loadings* das componentes principais PC1 e PC2, as quais explicam aproximadamente 80% da estrutura dos dados originais. A interpretação desses dados viabilizou, inicialmente, a redução da dimensão do conjunto dos dados para o estudo em um plano PC1 versus PC2 e a distinção da água mineral da marca S.Pellegrino devido ao seu alto teor de sulfato em relação ao restante. Além disso, permitiu a distinção das marcas São Lourenço e Evian, por apresentarem as maiores concentrações de nitrito e a Perrier, a maior concentração de nitrato.

Ao longo da realização dos experimentos, tratamento de dados e uso da plataforma foram feitas anotações, as quais auxiliaram, junto com o procedimento atual descrito na apostila da disciplina, na elaboração do roteiro proposto para a aula. Este roteiro, sugere a aplicação da ACP com uma rotina do GNU Octave para amostras águas minerais brasileira e estrangeira como uma forma de iniciar a implementação destas ferramentas no currículo do curso de Química.

Autorizamos a divulgação no Repositório Institucional da UFRJ-Pantheon.

Aluna: _____

Orientadora: _____

Assinatura

Assinatura

ABSTRACT

At the end of the last century, computational advancement and scientific instrumentation promoted the emergence of a large amount of data at various levels of complexity. In this context, an area named Chemometrics emerged to extract, through a multivariate approach, a greater amount of relevant information from the data generated in the laboratory, assisting in its interpretation and proposing pre-experimental strategies.

One of the research frontiers of Chemometrics and one of the focuses of this work is the recognition of patterns, which has one of the main methods as the Principal Component Analysis (PCA). This tool allows graphical visualization of the entire dataset, especially when the number of samples and variables is high. However, this technique makes use, mainly, of algebra and statistical concepts, areas that have a relatively restricted focus on chemistry courses in general.

Nevertheless, the use of well-built statistical programs allows to replacement of mathematical complexity with didactic simplicity. Thus, this work aimed to propose a lesson script, in which it introduces the main concepts of the selected statistical tool (PCA), using a computational routine that can be applied to the free use platform (GNU Octave). The discipline selected for the use of the script was the Instrumental Analysis I (code IQA366), taught to chemistry students for the Federal University of Rio de Janeiro (UFRJ), in the experimental practice of analysis of the composition of mineral waters by means of ion chromatography technique.

The variables selected were the concentrations of 7 anions (F^- , Cl^- , NO_2^- , Br^- , NO_3^- , PO_4^{3-} and SO_4^{2-}) in 16 mineral waters of Brazilian and foreign brands. By quantifying the ion contents in mineral waters, Br^- and PO_4^{3-} ions were below their respective quantification limits and were not used in the ACP study.

The chemometric tool (PCA) was applied through the free GNU Octave 5.1.0 software, which enabled the calculation and construction of the graphs for the scores and loadings of the main PC1 and PC2 components, which explained approximately 80% of the structure of the original data. The interpretation of these data initially made it possible to reduce the dimension of the data set for the study in a PC1 versus PC2 plan and, then, the distinction of mineral water of the S.Pellegrino brand due to its high sulfate content in relation to the rest. Evenfurther, it allowed the

distinction of the brands São Lourenço and Evian, which had higher concentrations of nitrite and the brand Perrier, with the highest concentration of nitrate.

During the performance of the experiments, the data processing and the use of the software, notes were made, which assisted, together with the current procedure described in the handout of the discipline, in the elaboration of the script proposed for the class. This script suggests the application of the PCA with a GNU Octave routine, for samples of Brazilian mineral water and foreign to start the implementation of these tools in the Chemistry course curriculum.

LISTA DE SIGLAS

\hat{y}_i	Previsão da reta
\bar{x}	Média de todos os valores de concentração x_i
\bar{y}	Média de todos os valores obtidos para a resposta y_i
A	Matriz de autovetores
ACP/PCA	Análise por Componente Principais (<i>Principal Component Analysis</i>)
ANOVA	<i>Analysis of Variance</i> (Análise de Variância)
ANVISA	Agência Nacional de Vigilância Sanitária
AOAC	<i>Association of Official Agricultural Chemists</i>
b_0	Ponto de interseção da reta ao eixo y
b_1	Inclinação da reta
CEDAE	Companhia Estadual de Águas e Esgotos do Rio de Janeiro
CI	Cromatografia de Íons
c_k	contribuição do k-ésimo componente principal
CMD	Concentração Média Determinada
CP/PC	Componente Principal (<i>Principal Component</i>)
CV	Coeficiente de Variação
D_A	Coeficiente de distribuição da espécie iônica A
DMEA	Dimetiletanolamina
DPR_{lin}	Desvio padrão relativo do coeficiente angular (linearidade)
DPR_{rep}	Desvio padrão relativo (repetibilidade)
DVS/SVD	Decomposição de Valor Singular (<i>Single Value Decomposition</i>)
e_i	Resíduo entre o valor medido e previsão da reta
F	Valor de F no teste de análises de variâncias
F_{tab}	Valor de tabelado para os graus de liberdades das variâncias analisadas
g_i	Autovetores associados à um autovalor da matriz de covariância
H_0	Hipótese inicial a ser testada
H_1	Hipótese secundária em contrário à H_0
INMETRO	Instituto Nacional de Metrologia
IQ	Instituto de Química

LAM	Laboratório de Análises A mbientais e M inerais
LD	Limite de D etecção
LQ	Limite de Q uantificação
m_j	Média da coluna j
MMA	M etilm et acrilato
MQ_{regr}	M édia quadrática da Soma dos Q uadrados da R egressão
MQ_{res}	M édia quadrática da Soma dos Q uadrados dos R esíduos
n	Número de experimentos
p	Número de variáveis
P	Matriz dos <i>loadings</i>
PS-DVB	P olystyrene- d ivinyl b enzene (Poliestireno-divinilbenzeno)
QAT	Q uímica com A tribuições T ecnológicas
R^2	Coeficiente de Determinação
R_{soma}	Soma dos quadrados dos resíduos
S	Matriz de <i>scores</i>
s	Desvio padrão
S_b	Desvio padrão do coeficiente angular
S_{branco}	Desvio padrão do sinal do branco
S_{ip}	<i>Scores</i>
SQ_{regr}	S oma dos Q uadrados da R egressão
SQ_{res}	S oma dos Q uadrados dos R esíduos
SQT	S oma dos Q uadrados T otais
$S_{y/x}$	Erros aleatórios na direção y
TMA	T rimetilamina
U	Matriz relacionada aos <i>scores</i>
UFRJ	U niversidade F ederal do R io de J aneiro
V	Matriz relacionada aos <i>loadings</i>
VIM	V ocabulário I nternacional de M etrologia
V_{ip}	<i>Loading</i> da variável
W	Matriz de valores singulares
X	Matriz de dados
y_i	Resposta experimentais
Z	Forma vetorial das p variáveis em cada n amostras após o pré-

	tratamento
Z_{ij}	Valor obtido de cada amostra após o tratamento de dados
$\bar{\delta}_j$	Desvio padrão da coluna
Λ	Matriz de autovalores
λ_i	Autovalores associados a matriz de covariância
Σ	Matriz de covariância

LISTA DE FIGURAS

Figura 3.1: Diagrama do processo de troca iônica. Esquerda: troca catiônica. Direita: troca aniônica. Adaptado de: EITH <i>et al.</i> , 2007.....	25
Figura 3.2: Estrutura dos materiais de empacotamento mais usados. Esquerda: poliestireno-divinilbenzeno (PS-DVB). Direita: polimetilmetacrilato (PMMA). Adaptado de: NAZARIO & LANÇAS, 2013.....	27
Figura 3.3: Grupos funcionais mais importantes para cromatografia de ânions. Esquerda: trimetilamina (TMA). Direita: dimetiletanolamina (DMEA).....	27
Figura 3.4: Representação esquemática de um sistema de cromatografia de íons com supressão de condutividade. Adaptado de: GUEDES, 2010.....	29
Figura 3.5: Representação da reta obtida por regressão linear. Fonte: MASSART, 1997.....	31
Figura 3.6: Exemplos de gráficos de resíduos. Fonte: MASSART, 1997.....	33
Figura 3.7: Projeções de duas para uma dimensão. (a) há perda de informação na projeção; (b) PC1 - a informação é preservada. Fonte: MASSART, 1997.....	42
Figura 3.8: (a) – PC2 é ortogonal à PC1; (b) – Gráfico construído com os componentes principais PC1 e PC2; (c) – PC1. Fonte: MASSART, 1997.....	43
Figura 3.9: Matriz de dados n x m. Adaptado de: MASSART, 1997.....	44
Figura 3.10: Gráfico de scores de 3 tipos de vinhos italianos (Δ = Barolo, \bullet = Barbera, x = Grignolino). Fonte: MASSART, 1997.....	51
Figura 3.11: Gráfico de loading de algumas variáveis para o exemplo de poluição do ar. Fonte: MASSART, 1997.....	52
Figura 4.1: CI à esquerda, computador usado para o tratamento de dados e, à direita, o módulo do equipamento.....	56
Figura 5.1: Cromatograma obtido na análise da solução com $0,20 \text{ mg.L}^{-1}$ com os sete íons estudados.....	62
Figura 5.2: Curva analítica para o íon fluoreto.....	64
Figura 5.3: Gráfico de resíduos construído para a curva analítica do íon fluoreto.....	65
Figura 5.4: Curvas analíticas para o íon cloreto. A – Faixa baixa. B – Faixa alta.....	68
Figura 5.5: Curva analítica para o íon nitrito.....	68
Figura 5.6: Curva analítica para o íon brometo.....	68

Figura 5.7: Curvas analíticas para o íon nitrato. A – Faixa baixa. B – Faixa alta.....	69
Figura 5.8: Curva analítica para o íon fosfato.....	69
Figura 5.9: Curva analítica para o íon sulfato.....	70
Figura 5.10: Gráfico de resíduos construído para as curvas analíticas do íon cloreto. A – faixa baixa. B – faixa alta.....	70
Figura 5.11: Gráfico de resíduos construído para as curvas analíticas do íon nitrito.....	71
Figura 5.12: Gráfico de resíduos construído para as curvas analíticas do íon brometo.	71
Figura 5.13: Gráfico de resíduos construído para as curvas analíticas do íon nitrato. A – Faixa baixa. B – Faixa alta.....	71
Figura 5.14: Gráfico de resíduos construído para as curvas analíticas do íon fosfato.	72
Figura 5.15: Gráfico de resíduos construído para as curvas analíticas do íon sulfato.	72
Figura 5. 16: Teores de sulfato, nitrato, cloreto e fluoreto apresentados na análise e no rótulo para três águas minerais brasileiras.....	80
Figura 5.17: Teores de sulfato, nitrato e cloreto apresentados na análise e no rótulo para três águas minerais estrangeiras.....	81
Figura 5.18: Percentagem cumulativa da contribuição dos componentes principais.....	84
Figura 5.19: Gráfico de <i>scores</i> (PC1 x PC2)	85
Figura 5.20: Gráfico dos <i>loadings</i> no plano PC1 x PC2.....	87
Figura 5.21: Tela do programa com as principais funções destacadas.....	102
Figura 5.22: Tela do programa com as funções para carregar os dados destacadas.....	107

LISTA DE QUADROS

Quadro 3.1: Classes de águas minerais quanto à composição química (Fonte de: JEBER & PROFETA, 2018)	22
Quadro 3.2: Classificação das fontes de água mineral (Fonte de: JEBER & PROFETA,2018).....	23
Quadro 3.3: Principais resinas de troca de íons para cromatografia de íons (Fonte: GUEDES, 2010).....	28
Quadro 3.4: Quadro de análise de variância para o ajuste do modelo quadrático*	35
Quadro 3.5: Parâmetros de validação conforme o tipo de ensaio. Fonte: INMETRO, 2018.....	36
Quadro 3.6: Critérios de aceitação para a repetibilidade. (Fonte: <i>Association of Official Agricultural Chemists,2016</i>).....	40
Quadro 3.7: Critérios de aceitação para recuperação. Fonte: AOAC, 2016.....	41
Quadro 4.1: Marcas em que foram realizadas diluições, alíquotas e volumes finais.....	55
Quadro 4.2: Características e condições operacionais do CI.....	57
Quadro 4.3: Faixas de concentração selecionadas para cada íon para a construção da curva analítica.....	58
Quadro 4.4 – Métodos utilizados para determinar LD e LQ.....	59
Quadro 5.1: Áreas dos picos referentes ao íon fluoreto nas soluções padrões da curva analítica.....	63
Quadro 5.2: Valores de G_{calc} obtidos para os dados da curva analítica do fluoreto.....	63
Quadro 5.3: Valores de variância obtidos para os dados da curva analítica do fluoreto.....	64
Quadro 5.4: ANOVA para a curva analítica do íon fluoreto.....	65
Quadro 5.5: Valores de G_{calc} obtidos para os dados das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato.....	66
Quadro 5.6: Valores de variância obtidos para os dados das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato.....	66

Quadro 5.7: Valores calculados para a maior variância, o somatório das variâncias e o C_{calc} para as curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato.....	67
Quadro 5.8: Valores calculados para F_{calc} para as curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato.....	67
Quadro 5.9: Resumo das equações das retas para os íons fluoreto, cloreto, nitrito, brometo, nitrato, fosfato e sulfato e seus respectivos R^2	73
Quadro 5.10: Desvio padrão relativo do coeficiente angular para cada íon.....	74
Quadro 5.11: Valores obtidos para as áreas dos picos para a análise de três replicatas do branco.....	74
Quadro 5.12: Limites de quantificação e limites de detecção obtidos pela relação sinal/ruído.....	75
Quadro 5.13: Limites de detecção e quantificação decorrentes do método do desvio padrão do branco e inclinação da reta.....	75
Quadro 5.14: Desvios padrões relativos calculados para as soluções padrões com 0,10 mg.L ⁻¹ , 0,30 mg.L ⁻¹ e 0,50 mg.L ⁻¹ dos íons em estudo.....	76
Quadro 5.15: Recuperação analítica calculada para as soluções padrão com 0,10 mg.L ⁻¹ , 0,30 mg.L ⁻¹ e 0,50 mg.L ⁻¹ dos íons em estudo.....	77
Quadro 5.16: Concentrações calculadas para os íons fluoreto, cloreto, nitrito, nitrato e sulfato para todas as 16 águas minerais.....	78
Quadro 5.17: Autovetores relacionados com a matriz de covariância dos dados pré-tratados.....	82
Quadro 5.18: Autovalores relacionados com a matriz de covariância dos dados pré-tratados.....	83
Quadro 5.19: Contribuição de cada componente principal para a explicação da variância total.....	83
Quadro 5.20: Percentagens cumulativas ao adicionar cada componente principal na explicação da variação dos dados originais.....	84
Quadro 5.21: <i>Loadings</i> nas duas componentes principais (PC1 e PC2).....	86

Sumário

1. Introdução	18
2. Objetivo.....	20
3. Fundamentação teórica.....	21
3.1. Água mineral	21
3.2. Princípios da técnica de cromatografia de íons	25
3.3. Métodos matemáticos e testes estatísticos.....	30
3.3.1. Teste de Grubbs	30
3.3.2. Teste de Cochran	30
3.3.3. Regressão linear.....	31
3.3.4. Gráfico de resíduos	32
3.3.5. Análise de variância (ANOVA)	33
3.4. Validação parcial do método.....	36
3.4.1. Linearidade.....	37
3.4.2. Limites de detecção (LD) e limites de quantificação (LQ).....	38
3.4.3. Repetibilidade.....	39
3.4.4. Recuperação analítica.....	40
3.5. Quimiometria e suas ferramentas.....	41
3.5.1. Análise por componentes principais (ACP)	41
3.6. Plataforma	52
3.6.1. GNU octave	53
4. Material e métodos.....	54
4.1. Material.....	54
4.2. Seleção da amostra a ser estudada	54
4.2.1. Determinação da quantidade e das marcas das águas minerais.....	55
4.2.2. Procedimento analítico das análises das águas minerais	55
4.3. Seleção da técnica analítica e condições operacionais da cromatografia de íons	56
4.4. Construção das curvas analíticas e aplicação dos testes estatísticos.....	57
4.5. Validação parcial	58
4.5.1. Linearidade.....	59

4.5.2. Métodos de determinação de LD e LQ e procedimento de aplicação	
59	
4.5.3. Repetibilidade.....	60
4.5.4. Recuperação analítica.....	60
4.6. Seleção da ferramenta quimiométrica.....	60
4.7. Seleção e uso da plataforma selecionada para a aplicação do PCA	60
4.8. Elaboração do roteiro de aula e proposta de implementação no curso	61
5. Resultados e discussão	62
5.1. Construção da curva analítica e aplicação dos métodos matemáticos e testes estatísticos.....	62
5.2. Validação parcial	73
5.2.1. Linearidade.....	73
5.2.2. Limites de detecção (LD) e quantificação (LQ)	74
5.2.3. Repetibilidade.....	76
5.2.4. Recuperação analítica.....	76
5.3. Análise das águas minerais	77
5.4. Aplicação da técnica quimiométrica.....	81
5.5. Elaboração do roteiro de aula e proposta de implementação no curso	87
6. Conclusão	91
7. Referências bibliográficas	92
8. ANEXO.....	95
ANEXO A – Passo a passo da aplicação do teste de Grubb.....	95
ANEXO B - Quadro com os valores críticos de G.....	97
ANEXO C – Passo a passo para a aplicação do teste de Cochran	98
ANEXO D - Quadro com os valores críticos de C.....	99
ANEXO E – Quadro com os valores críticos de F*	100
ANEXO F - Nomes, fontes e concentrações dos íons estudados descritos nos rótulos de cada marca.....	101
9. Apêndice.....	102
Apêndice A - Modo de utilização do software GNU Octave	102
Apêndice B - Áreas dos picos referentes aos íons fluoreto, cloreto, nitrito, brometo, nitrato, fosfato e sulfato nas soluções padrões das curvas analíticas.	104
Apêndice C - Tabelas ANOVA das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato	105

Apêndice D – Desvios padrões (DP) e concentrações médias determinadas (CMD) das soluções padrões 0,10 mg.L ⁻¹ , 0,30 mg.L ⁻¹ e 0,50 mg.L ⁻¹	106
Apêndice E – Comandos utilizados na plataforma do GNU Octave para a aplicação do ACP nos dados gerados na análise das 16 águas minerais	107
Apêndice F – Scores calculados para cada componente principal	113

1. Introdução

No final do século passado, o avanço computacional e da instrumentação científica promoveu o surgimento de uma grande quantidade de dados em diversos níveis de complexidade. Entretanto, ainda que seja possível retirar mais informação de uma amostra em estudo, a interpretação de todos os dados se tornou uma tarefa árdua (VALDERRAMA *et al.*, 2015).

Com o intuito de reduzir à poucas variáveis e facilitar a compreensão dos dados, muitos analistas ainda recorrem à estatística univariada, na qual as variáveis são estudadas isoladamente. Contudo, quando um fenômeno depende de muitos fatores simultaneamente, geralmente este tipo de análise falha, pois as relações existentes entre eles não são notadas e os efeitos antagônicos ou sinérgicos existentes dificultam a interpretação do fenômeno (NETO, 2004).

Dessa forma, originou-se uma área, nomeada Quimiometria, na década de 1970, a qual se dedicava, com uma abordagem multivariada, a extrair uma maior quantidade de informações relevantes dos dados gerados no laboratório, auxiliando na sua interpretação e propondo estratégias pré-experimentais (VALDERRAMA *et al.*, 2015).

Uma das frentes de pesquisa da Quimiometria e um dos focos deste trabalho é o reconhecimento de padrões, o qual tem um dos principais métodos a Análise por Componentes Principais (ACP – sigla adotada neste trabalho - ou PCA, do inglês *Principal Component Analysis*). Esta ferramenta, permite a visualização gráfica de todo o conjunto de dados, principalmente, quando o número de amostras e variáveis é elevado (CORREIA & FERREIRA, 2007). Contudo, esta técnica faz uso, principalmente, de conceitos álgebra e estatística, áreas que têm foco relativamente restrito nos cursos de Química em geral.

Os cursos de Química das universidades brasileiras começaram a incluir de forma discreta ferramentas como esta em seus cursos. Segundo uma pesquisa realizada no início desta década, com os cursos de graduação em Química de 70 Instituições de Ensino Superior no Brasil (PEREIRA *et al.*, 2014), em apenas em 20 delas aparece uma disciplina específica de Quimiometria.

Isto se deve ao fato do enfoque matemático e estatístico necessário para o entendimento e aplicação desse método ser, demandando a presença de pessoas

capacitadas no assunto e implicando, em alguns casos, na diminuição do tempo de aula de outras disciplinas, razões que dificultam a incorporação desse assunto nas grades curriculares atuais (NETO, 2004).

Entretanto, o uso de programas estatísticos bem construídos permite substituir a complexidade matemática por uma simplicidade didática (NETO, J. M. M., 2004). Deste modo, é possível trabalhar mais a parte da inferência estatística, ou seja, o que se pode afirmar a partir dos dados disponíveis.

Assim, este trabalho visa propor um roteiro de aula para o Instituto de Química (IQ) da Universidade Federal do Rio de Janeiro (UFRJ), no qual introduz os conceitos principais da ferramenta estatística selecionada (PCA), usando rotina computacional que pode ser aplicada à plataforma de uso gratuito aplicadas à técnica de cromatografia de íons na análise da composição de águas minerais.

2. Objetivo

Propor um roteiro de aula o qual permita introduzir os conceitos principais da ferramenta estatística escolhida ACP (Análise de Componentes Principais), aplicada à técnica de cromatografia de íons na análise das diferenças de composição química de águas minerais nacionais e oriundas de outros países para a disciplina de Análise Instrumental I, de código IQA366, ministrada para os alunos do curso de Química com Atribuições Tecnológicas (QAT).

3. Fundamentação teórica

3.1. Água mineral

Segundo o Código de Águas Minerais (Decreto-Lei nº 7.841, de 08 de agosto de 1945), águas minerais são aquelas provenientes de fontes naturais ou de fontes artificialmente captadas que possuem composição química ou propriedades físicas ou físico-químicas distintas daquelas águas comuns, com características que lhes conferem uma ação medicamentosa.

De acordo com esse Código, as águas minerais podem ser classificadas com base na sua composição química (Quadro 3.1).

Quadro 3.1: Classificação de águas minerais quanto à composição química (Fonte: JEBER & PROFETA, 2018)

Classes	Tipos	Características
I	Oligominerais	Quando apesar de não atingirem os limites estabelecidos, forem classificadas como minerais pelo disposto nos parágrafos 2º e 3º do art. da presente Lei.
II	Radíferas	Quando contiverem substâncias radioativas dissolvidas que lhes atribuam radioatividade permanente.
II	Alcalino-bicarbonatadas	As que contiverem, por litro, uma quantidade de compostos alcalinos equivalente, no mínimo a 0,2 g de bicarbonato de sódio.
IV	Alcalino-terrosas	As que contiverem, por litro, uma quantidade de alcalino terrosos equivalentes no mínimo a 0,12 g de carbonato de cálcio, distinguindo-se: a) alcalino-terrosas cálcicas, as que contiverem, por litro, no mínimo 0,048 g do cátion Ca^{2+} , sob forma de bicarbonato de cálcio; b) alcalino-terrosas magnesianas, as que contiverem, por litro, no mínimo, 0,030 g do cátion Mg^{2+} , sob a forma de bicarbonato de magnésio.
V	Sulfatadas	As que contiverem, por litro, no mínimo, 0,1 g do ânion SO_4^{2-} , combinado aos cátions Na^+ , K^+ e Mg^{2+} .
VI	Sulfurosas	As que contiverem, por litro, no mínimo, 0,1 g do ânion S^{2-} .
VII	Nitratadas	As que contiverem, por litro, no mínimo, 0,1 g do ânion NO_3^- de origem mineral.
VIII	Cloretadas	As que contiverem, por litro, no mínimo, 0,5 g de NaCl.
IX	Ferruginosas	As que contiverem, por litro, no mínimo, 0,005 g de Fe^{2+}/Fe^{3+} .
X	Radioativas	As que contiverem radônio em solução, nos seguintes limites: a) fracamente radioativa: entre 5 e 10 unidades Mache por litro, a 20 °C e 760 mm Hg de pressão; b) radioativas: entre 10 e 50 unidades Mache por litro, a 20 °C e 760 mm Hg de pressão; c) fortemente radioativas: superior a 50 unidades Mache por litro, a 20 °C e 760 mm Hg de pressão.
XI	Toriativas	As que possuírem um teor em Torônio em solução equivalente em unidades eletrostáticas a 2 unidades Mache por litro, no mínimo.
XII	Carbogasosas	As que contiverem, por litro, 200 ml de gás carbônico livre dissolvido a 20 °C e 760 mm Hg de pressão.

As fontes de água mineral, por sua vez, podem ser classificadas quanto a presença de gases e quanto à temperatura (Quadro 3.2).

Quadro 3.2: Classificação das fontes de água mineral (Fonte: JEBER & PROFETA, 2018)

Critério	Classificação	Características
Presença de Gases	Fontes Radioativas	- Fracamente radioativas: as que apresentarem, no mínimo, uma vazão gasosa de 1 litro por minuto com um teor em radônio compreendido entre 5 a 10 unidades Mache, por litro de gás espontâneo, a 20 °C e 760 mm Hg de pressão; - Radioativas: as que apresentarem, no mínimo, uma vazão gasosa de 1 litro por minuto com um teor em radônio compreendido entre 10 e 50 unidades Mache, por litro de gás espontâneo, a 20 °C e 760 mm Hg de pressão; - Fortemente radioativas: as que apresentarem, no mínimo, uma vazão gasosa de 1 litro por minuto com um teor em radônio superior a 50 unidades Mache, por litro de gás espontâneo, a 20 °C e 760 mm Hg de pressão.
	Fontes Toriativas	As que apresentarem, no mínimo, uma vazão gasosa de 1 litro por minuto, com um teor em torônio na emergência equivalente em unidades eletrostáticas, a 2 unidades Mache por litro.
	Fontes Sulfurosas	As que possuírem na emergência desprendimentos definidos de gás sulfídrico.
Temperatura	Frias	Quando sua temperatura for inferior a 25 °C.
	Hipotermais	Quando sua temperatura estiver compreendida entre 25 e 33 °C.
	Mesotermais	Quando sua temperatura estiver compreendida entre 33 e 36 °C.
	Isotermais	Quando sua temperatura estiver compreendida entre 36 e 38 °C.
	Hipertemais	Quando sua temperatura for superior a 38 °C.

Ao confrontar algumas características das fontes de água mineral com a composição química resultante, pode-se notar alguns fatores que justificam a presença e a quantidade de algumas espécies químicas, por exemplo:

- grande parte da água mineral brasileira envasada provém de sistemas de fluxos de natureza rasa (< 70 m) e de rápido tempo de trânsito, resultando em águas de baixa mineralização e de pH ácido (BERTOLO, 2007);
- as águas de regiões litorâneas, nas quais nota-se a influência do mar pela maior presença do sal cloreto de sódio que é transportado para as áreas continentais juntamente com a umidade marinha;
- as águas da região centro-sul do Brasil, por sua vez, são mais influenciadas pelo tipo de rocha percolada e pela profundidade de captação de água. As águas minerais que são originadas do intemperismo de minerais de rochas

alcalinas, por exemplo, têm a oxidação de pirita como sua fonte de sulfato (BERTOLO, 2007).

À exceção desses fatores naturais que podem ser relacionados com a composição química, há também os fatores antrópicos, no qual o nitrato é a principal espécie que indica a alteração de origem antrópica na qualidade da água. As concentrações elevadas de nitrato na água indicam a ocorrência de contaminação advinda de águas servidas, principalmente em ambiente urbano e por utilizações de fertilizantes inorgânicos e orgânicos, em ambiente rural (MILANEZ *et al.*, 2015).

Nos aspectos legais aplicáveis às águas minerais, podem-se destacar ainda as Resoluções 274 e 275, da Agência Nacional de Vigilância Sanitária (ANVISA). A primeira resolução regula a identidade e, principalmente, os limites máximos de substâncias químicas que representam risco à saúde, e é aplicada às águas envasadas e fabricação de gelo destinados ao consumo humano; a segunda trata das características microbiológicas (como a imposição da ausência das bactérias *Escherichia coli* ou coliformes termotolerantes) das águas Mineral Natural e Natural (JEBER & PROFETA, 2018).

De acordo com Jeber & Profeta (2018), o setor de Água Mineral no Brasil sofreu um aumento significativo a partir da década de 70, se tornando, em 2014, o quinto maior mercado de água engarrafada do mundo. Algumas das principais razões desse incremento são: o crescimento da população, a diminuição da oferta de água potável pelos governos e a busca por um estilo de vida mais saudável por parte dos consumidores (MOURA *et al.*, 2011).

Um exemplo desse incremento tem sido vivenciado no início de 2020 na cidade do Rio de Janeiro, com o aumento da demanda da água mineral vendida, devido a problemas de gerenciamento e tratamento da água fornecida pela Companhia Estadual de Águas e Esgotos do Rio de Janeiro (CEDAE) (G1, 2020). O aumento na demanda resultou também no aumento de vendas de águas de baixa qualidade e sem especificação, o que mostra a necessidade das análises químicas de tais produtos.

À vista disso, analisar a composição química e verificar as distinções entre as marcas disponíveis no Brasil e no exterior se faz necessário no sentido de impulsionar o estudo de um produto com grande demanda no mercado brasileiro e do mundo.

3.2. Princípios da técnica de cromatografia de íons

Uma das técnicas instrumentais mais utilizadas para avaliação de íons em análises de água é a cromatografia de troca iônica ou cromatografia de íons (CI).

O princípio da técnica de CI tem como base uma interação entre os íons de uma solução e uma substância sólida contendo os grupos funcionais (fase estacionária), que podem reter íons devido a atrações eletrostáticas (PEREIRA, 2007).

Esse tipo de reação é chamada de troca iônica, uma vez que, primeiramente, após a injeção da amostra e à medida que esta percola pela coluna, há a troca entre o contra-íon original do grupo funcional da fase estacionária com o íon da amostra, de acordo com a afinidade destes com o grupo funcional (EITH *et al.*, 2007).

A Figura 3.1 apresenta um diagrama do processo de troca iônica para ânions e cátions, nos quais a letra A representa o íon da amostra e a letra E, o eluente.

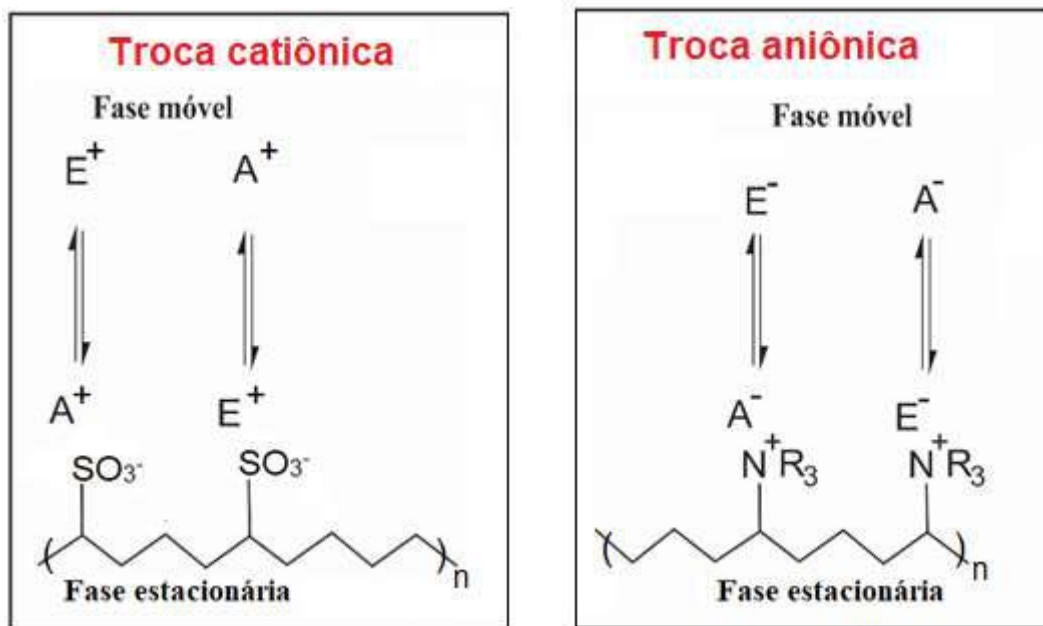


Figura 3.1: Diagrama do processo de troca iônica. Esquerda: troca catiônica. Direita: troca aniônica. Adaptado de: EITH *et al.*, 2007.

O processo de troca iônica, representado na Figura 3.1, ocorre na condição de equilíbrio e a eficiência da separação dos íons participantes depende da

afinidade destes em relação aos grupos funcionais da fase estacionária (EITH *et al.*, 2007).

Assim, quando uma mistura de íons é submetida à separação cromatográfica, um equilíbrio de distribuição é formado entre as fases estacionária e móvel para cada íon. A separação desses íons ocorre apenas quando os coeficientes de distribuição (D_A – Equação 1) diferirem suficientemente uns dos outros, de forma que o maior D_A está associado a íons que são retidos mais fortemente do que os íons associados a valores de D_A menores.

$$D_A = \frac{[A]_S}{[A]_M} \quad (\text{Equação 1})$$

Sendo:

$[A]_S$ = íon A na fase estacionária

$[A]_M$ = íon A na fase móvel

As fases estacionárias, por sua vez, requerem materiais de empacotamento feitos de partículas muito pequenas os quais devem apresentar um rápido processo de troca iônica. Há muitos materiais orgânicos e inorgânicos diferentes que são adequados para essa função (resinas de polímeros orgânicos modificadas, sílica gel modificada, sais inorgânicos – polifosfato, vidros, zeólitas, entre outros).

Atualmente, no entanto, os materiais de empacotamento mais usados são baseados nos copolímeros do poliestireno-divinilbenzeno (PS-DVB do inglês *polystyrene-divinylbenzene*) ou polímeros de metilmetacrilato (MMA) devido à estabilidade química desses compostos em toda escala de pH (Figura 3.2). De acordo com seu tipo de aplicação (cromatografia de ânions ou cátions), as fases estacionárias são modificadas, com a inclusão de sítios ativos à estrutura da resina.

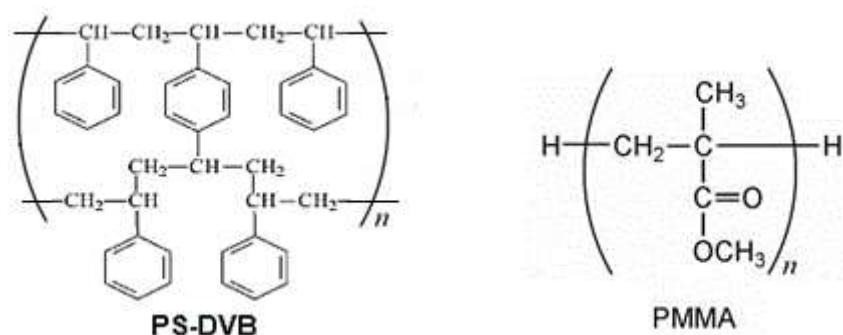


Figura 3.2: Estrutura dos materiais de empacotamento mais usados. Esquerda: poliestireno-divinilbenzeno (PS-DVB). Direita: polimetilmetacrilato (PMMA). Adaptado de: NAZARIO & LANÇAS, 2013.

Na troca de ânions, os grupos funcionais baseados em nitrogênio são praticamente os únicos utilizados, em virtude da sua estabilidade química e um vasto número de substituintes possíveis no átomo de nitrogênio. Dentre eles se destacam derivados de trimetilamina (TMA) e dimetiletanolamina (2-dimetilaminaetanol, DMEA), os quais tem as suas estruturas apresentadas na Figura 3.3 (EITH *et al.*, 2007).

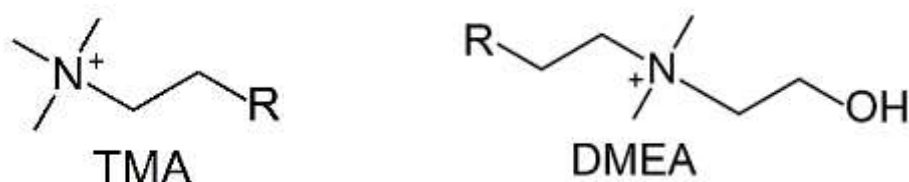


Figura 3.3: Grupos funcionais mais importantes para cromatografia de ânions. Esquerda: trimetilamina (TMA). Direita: dimetiletanolamina (DMEA)

Por sua vez, os trocadores catiônicos mais usados pertencem a duas grandes categorias: resinas trocadoras catiônicas fortes, onde temos as resinas sulfônicas; e as resinas trocadoras catiônicas fracas, onde as mais usadas são as carboxílicas (GUEDES, 2010).

No Quadro 3.3 encontra-se um resumo das principais resinas de troca iônica para cromatografia de íons.

Quadro 3.3: Principais resinas de troca de íons para cromatografia de íons (Fonte: GUEDES, 2010)

Tipo de resina	Constituição química	Grupo Trocador usual disponível
Trocador de cátions fortemente ácido	Ácido sulfônico ligado ao estireno e divinilbenzeno	- SO ₃ ⁻ H ⁺
Trocador de cátions fracamente ácido	Ácido carboxílico ligado ao acrílico e divinilbenzeno	- CH ₂ -COO ⁻ Na ⁺
Trocador de ânions fortemente básico	Amônio quartenário ligado ao estireno e divinilbenzeno	- CH ₂ N ⁺ (CH ₃)Cl ⁻
Trocador de ânions fracamente básico	Polialquilamina	- N ⁺ HR ₂ Cl ⁻

Esta técnica permite o uso de diversos métodos de detecção tais como: por condutividade, amperométrico, potenciométrico, fotométrico, por fluorescência, entre outros. A detecção por condutividade, método usado neste trabalho, monitora todas espécies iônicas por meio da condutância elétrica, visto que esta é uma propriedade comum a todas as espécies iônicas em solução (GUEDES, 2010).

Todavia, a alta condutividade dos íons presentes na fase móvel, principalmente devido ao íon sódio, presente nos eluentes utilizados na determinação de ânions, dificulta a quantificação de outros íons. Nesse contexto, o uso de um supressor da condutividade do eluente, posicionado entre a coluna de separação e o detector, surge como uma solução a este problema, pois os íons do eluente são convertidos em espécies que contribuam para uma condutância baixa ou nula (KAPPES, 2019).

Para o bicarbonato de sódio, eluente aplicado neste trabalho, o processo de supressão química pode ser explicado pelo seguinte equilíbrio:



A supressão é realizada com um trocador de cátion com sítios ativos de troca protonados. De acordo com a Reação 1, o eluente é neutralizado, uma vez que os íons de sódio são substituídos por prótons, diminuindo drasticamente a condutividade relacionada ao eluente (KAPPES, 2019).

A Figura 3.4 mostra um resumo de todas as etapas em cromatografia de íons na forma de um esquema.

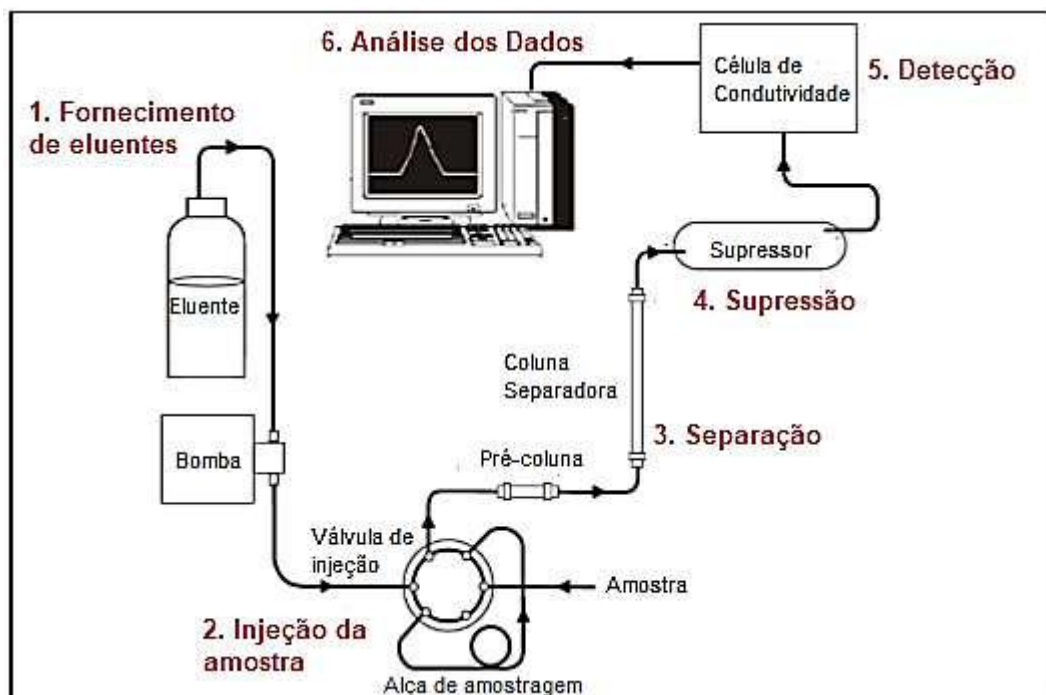


Figura 3.4: Representação esquemática de um sistema de cromatografia de íons com supressão de condutividade. Adaptado de: GUEDES, 2010.

Da mesma maneira que é mostrado na Figura 3.4, o eluente é alimentado através de uma ou mais bombas que formam um sistema que assegura o controle da vazão da fase móvel ao longo dos diversos componentes do sistema cromatográfico.

A amostra é, então, introduzida por meio de uma seringa no “loop” ou alça de amostragem, posicionado na válvula de injeção, na posição de carga. Ao girar o manípulo da válvula para a posição de injeção, dá-se uma alteração no trajeto do eluente no interior da válvula e a amostra é, então, introduzida na corrente do eluente, onde seus íons são separados à medida que se movem através da coluna analítica, que contém a fase estacionária. Dependendo de sua interação com a fase estacionária, o tempo de eluição dos ânions será maior ou menor. Após passar pela coluna, os analitos juntamente com o eluente chegam à coluna supressora, que suprime a condutividade do eluente, melhorando a detecção dos íons da amostra. Ao sair do sistema de supressão, os analitos chegam ao detector, que registra a condutância dos íons transmitindo o sinal para o computador (GUEDES, 2010).

3.3. Métodos matemáticos e testes estatísticos

Uma vez que o objetivo do trabalho era fazer a avaliação da análise águas minerais usando um método de redução de variáveis (ACP), alguns pontos relevantes para o entendimento da discussão dos resultados serão abordados neste tópico. Inicialmente, o foco será na redução de variáveis pela quantificação da concentração dos íons em solução, onde será descrito, de forma sucinta, a construção da regressão linear e a avaliação do modelo proposto.

3.3.1. Teste de Grubbs

O teste de Grubbs tem como função, detectar valores discrepantes (*outliers*) em um conjunto de dados, auxiliando, assim, na construção da curva analítica. Existem dois testes de Grubbs; o aplicado para verificar a presença de um valor discrepante por vez, e outro que permite detectar um par de valores discrepantes de cada vez.

As etapas do procedimento para se aplicar o teste de Grubbs estão descritas no Anexo A (página 94), sendo uma adaptação do procedimento descrito em Massart (1997). Os parâmetros tabelados utilizados no teste encontram-se no Anexo B (página 96).

3.3.2. Teste de Cochran

O método dos mínimos quadrados (Regressão Linear) tem como um dos requisitos para sua aplicação a homocedasticidade do conjunto de dados a ser utilizado, ou seja, que as variâncias, obtidas para cada concentração da curva analítica, sejam homogêneas (MASSART,1997).

Para verificar esta característica, aplica-se o teste de Cochran, o qual compara as variâncias de cada concentração da curva e verifica se são estatisticamente equivalentes ou não. Todos os passos para aplicação desse teste podem ser encontrados no Anexo C (página 97) e são uma adaptação do procedimento descrito em Massart (1997). Os parâmetros tabelados utilizados no teste encontram-se no Anexo D (página 98).

3.3.3. Regressão linear

A regressão seleciona qual a melhor curva a ser traçada que se ajusta a um conjunto de dados experimentais, podendo ser uma reta ou uma curva qualquer, na qual a distância entre os pontos e a reta/curva seja a menor possível (MASSART, 1997; OLIVEIRA, 2013).

Uma importante aplicação da regressão linear é a construção de curvas de calibração para determinar a concentração de um analito em uma amostra. Nesse contexto, a variável dependente, y , representa a resposta e a variável independente, x , representa as concentrações das soluções padrão (MASSART, 1997).

Ao assumir que a relação entre a resposta e a concentração é uma reta (Figura 3.5) e com o intuito de obter um modelo matemático que se ajuste aos dados experimentais, o seguinte modelo é proposto:

$$\hat{y} = b_0 + b_1x \quad (\text{Equação 3})$$

Sendo:

\hat{y} = resposta obtida pela regressão linear

b_0 = ponto de interseção da reta ao eixo y (coeficiente linear)

b_1 = inclinação da reta (coeficiente angular)

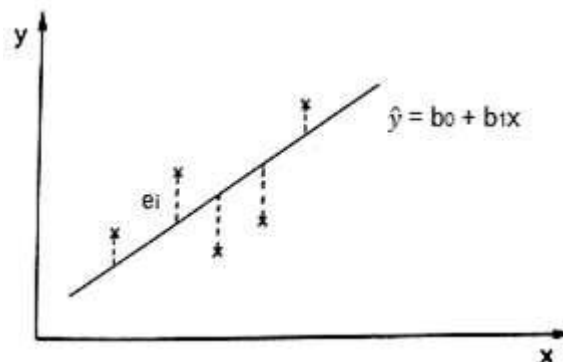


Figura 3.5: Representação da reta obtida por regressão linear. Fonte: MASSART, 1997.

Os parâmetros b_0 e b_1 podem ser calculados através das Equações 4 e 5:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (\text{Equação 4})$$

$$b_0 = (\bar{y} - b_1 \bar{x}) \quad (\text{Equação 5})$$

Sendo:

\bar{y} = média de todos os valores obtidos para a resposta y_i (Equação 6)

\bar{x} = média de todos os valores de concentração x_i (Equação 7)

n = número de experiências realizadas

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (\text{Equação 6})$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (\text{Equação 7})$$

A reta estimada construída a partir dos dados experimentais, também chamada de linha dos mínimos quadrados, advém do método dos mínimos quadrados, o qual minimiza a soma dos quadrados dos resíduos, S (Equação 8).

$$R_{soma} = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \quad (\text{Equação 8})$$

Sendo:

R_{soma} = soma dos quadrados dos resíduos

e_i = resíduo (calculado pela Equação 3 na seção 3.3.3 – Gráfico de resíduos – página 29)

3.3.4. Gráfico de resíduos

Através da análise de resíduos (RODRIGUES, 2016), é possível avaliar se o modelo empregado está correto, e ainda, detectar problemas como, por exemplo:

- Não-linearidade da relação entre x e y ;
- Variância não-constante dos erros (heterocedasticidade);
- Correlação entre os erros;
- Presença de valores discrepantes (*outliers*).

O resíduo (e_i) é a diferença entre o valor medido (y_i) e a previsão da reta (\hat{y}_i), e pode ser descrito pela Equação 2:

$$e_i = y_i - \hat{y}_i \quad (\text{Equação 2})$$

Após estimar os resíduos, é construído um gráfico de valores de x versus valores dos resíduos, no qual é analisada a presença, ou não, de algum dos problemas citados anteriormente (MASSART,1997). Alguns exemplos de gráficos de resíduos são descritos na Figura 3.6.

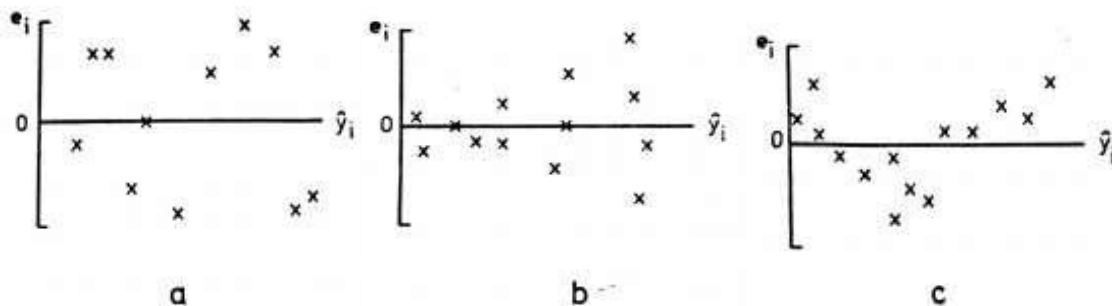


Figura 3.5: Exemplos de gráficos de resíduos. a – distribuição aleatória dos resíduos, b – resíduos que indicam heterocedasticidade dos dados, c – resíduos que indicam que os dados seriam melhor representados por um modelo quadrático.
Fonte: MASSART,1997.

Na Figura 3.6-a, os resíduos estão dispersos aleatoriamente e com um número de resíduos positivos similar ao número de resíduos negativos indicando, assim, que os dados apresentam um comportamento homocedástico e o modelo foi empregado corretamente. Na Figura 3.6-b, no entanto, como os resíduos encontram-se mais dispersos com o aumento de \hat{y} , há o indício de que a condição de homocedasticidade não é obedecida. Na Figura 3.6-c, por fim, o formato em U dos resíduos indica que seria mais adequado que os dados fossem representados por uma curva e não por uma reta indicando uma possível falta de ajuste e não linearidade dos dados.

3.3.5. Análise de variância (ANOVA)

Segundo Teófilo e Ferreira (2006), a maneira mais confiável de se avaliar a qualidade do ajuste de um modelo é empregando a análise de variância.

Análise de variância (frequentemente abreviada para ANOVA do inglês *Analysis of Variance*) é uma técnica estatística que pode ser usada para separar e estimar as diferentes causas de variação (MILLER & MILLER, 2010).

Como consequência da ANOVA, a variabilidade dos dados experimentais é dividida em dois termos: soma dos quadrados da regressão (SQ_{regr}) e soma dos quadrados dos resíduos (SQ_{res}). O primeiro termo mede a quantidade de variabilidade explicada pelo modelo de regressão e, o segundo, a variação residual que não é explicada pelas variáveis independentes (MONTALVÃO, 2014).

A variabilidade total dos dados experimentais é dada pela Soma dos Quadrados Totais (SQT) descrita na Equação 9.

$$SQT = \sum y^2 - \frac{(\sum y_i)^2}{N} = SQ_{regr} + SQ_{res} \quad (\text{Equação 9})$$

Assim sendo, a Equação 9 exhibe a relação da variabilidade dos valores experimentais em relação ao seu valor médio.

No tocante aos graus de liberdade de cada soma, ao seguir o raciocínio de que um sistema, inicialmente, possui (N-1) graus de liberdade e, ao fazer a regressão linear, (p-1) graus são perdidos em decorrência da estimação dos coeficientes desconhecidos (b_1, b_2, \dots, b_p), tem-se, ao final, (N-p) graus de liberdade para a estimativa do erro.

A soma dos quadrados do erro descreve a variabilidade não explicada pela regressão e pode ser definida como:

$$SQ_{res} = \sum (y_i - \hat{y}_i)^2 \quad (\text{Equação 10})$$

A soma dos quadrados da regressão, por sua vez, pode ser definida por:

$$SQ_{regr} = \sum y_i^2 - \frac{(\sum \hat{y}_i)^2}{N} \quad (\text{Equação 11})$$

As médias quadráticas de cada soma são calculadas por meio da divisão dos seus valores pelo seu respectivo grau de liberdade (Equações 12 e 13).

$$MQ_{res} = \frac{SQ_{res}}{n-p-2} \quad (\text{Equação 12})$$

$$MQ_{regr} = \frac{SQ_{regr}}{p-1} \quad (\text{Equação 13})$$

Com a finalidade de analisar a variância na regressão, aplica-se o teste estatístico-F Snedecor, o qual averigua se há uma correlação entre a resposta e algum dos coeficientes de maneira que se pode declarar que a regressão é significativa. Para o uso do teste F, um teste de hipóteses é elaborado, sendo:

$$H_0: b_1 = b_2 = \dots = b_p = 0$$

H_1 : Pelo menos um dos coeficientes b_1, b_2, \dots, b_p não é zero

O valor de F pode ser encontrado pela razão a seguir:

$$F = \frac{MQ_{regr}}{MQ_{res}} \quad (\text{Equação 14})$$

O F_{tab} (Anexo E, página 99) é localizado através dos graus de liberdade no numerador ($p-1$), graus de liberdade no denominador ($n-p$) e o nível de significância (α) com o valor, geralmente, de 5%.

Caso $F > F_{tab}$, a hipótese nula é rejeitada e é confirmada a correlação entre a resposta e algum dos coeficientes.

O Quadro da análise de variância (Quadro 3.4) para o ajuste de um modelo quadrático é composta, então, por 5 colunas com os termos descritos acima e pode ser obtida através da planilha eletrônica Excel® com o uso de pacote de ferramentas “Análise de Dados”.

Quadro 3.4: Quadro de análise de variância para o ajuste do modelo quadrático*

Fonte de variação	Soma Quadrática	Nº de graus de liberdade	Média Quadrática	F
Regressão	SQ_{regr}	$p-1$	MQ_{regr}	MQ_{regr}/MQ_{res}
Resíduos	SQ_{res}	$n-p$	MQ_{res}	
Total	SQT	$n-1$		

* SQ_{regr} = soma dos quadrados dos desvios para a regressão, SQ_{res} = soma dos quadrados dos desvios para os resíduos, SQT = soma dos quadrados dos Totais, p = número de coeficientes, n = número de experiências realizadas, MQ_{regr} = quadrado médio da regressão, MQ_{res} = quadrado médio dos resíduos, F = F calculado para o teste de análises de variâncias.

3.3.6. Coeficiente de determinação (R²)

O valor do coeficiente de determinação (Equação 15) é um parâmetro para avaliar o poder de explicação do modelo de regressão, ou seja, se toda variação em torno da média foi explicada pela regressão (MASSART, 1997).

$$R^2 = 1 - \frac{SQ_{res}}{SQT} \quad (\text{Equação 15})$$

O valor de R² representa a fração da variação que é explicada pela falta de ajuste do modelo. Quanto mais próximo de 1 o valor do coeficiente estiver, melhor estará o ajuste do modelo às respostas observadas (TEÓFILO & FERREIRA, 2006, página 346-347).

3.4. Validação parcial do método

A validação consiste na demonstração de que o método possui as características necessárias para a obtenção de resultado com a qualidade exigida (INMETRO, 2018). Durante esse processo, alguns critérios devem ser avaliados de acordo com o ensaio a ser realizado (Quadro 3.5).

Quadro 3.5: Parâmetros de validação conforme o tipo de ensaio. Fonte: INMETRO, 2018

Parâmetros	Tipo de ensaio	
	Qualitativo	Quantitativo
Seletividade	x	x
Linearidade/faixa de trabalho/Faixa linear de trabalho/Sensibilidade		x
Limite de detecção	x	x
Limite de quantificação		x
Tendência/recuperação		x
Precisão	x	x
Robustez	(*)	(*)

(*) *parâmetro opcional.*

Neste trabalho foram abordados apenas os parâmetros: linearidade, limite de detecção, limite de quantificação, repetibilidade e recuperação analítica.

3.4.1. Linearidade

De acordo com o INMETRO (2018), a linearidade de um procedimento analítico é a habilidade em obter resultados que sejam diretamente proporcionais à concentração do analito na amostra em uma determinada faixa de trabalho.

Tal parâmetro pode ser avaliado através da construção do gráfico dos resultados de resposta medida em função da concentração do analito e, em seguida, com a determinação da equação da reta que relaciona as duas variáveis através do método de regressão linear (Seção 3.3.4 Regressão Linear – página 30).

Deve-se observar que a regressão linear deve ser aplicada após a verificação da ausência de valores discrepantes para cada nível de concentração da curva (Seção 3.3.1. Teste de Grubbs – página 28) e da homocedasticidade (Seção 3.3.2. Teste de Cochran – página 28) dos dados.

Antes de qualquer avaliação estatística para este parâmetro, segundo o Guia para tratamento estatístico da validação analítica da ANVISA (2017), é importante realizar uma análise visual para verificar se os pontos, quando dispostos no gráfico, têm relação linear aparente.

Outra abordagem para a avaliação da linearidade se refere a verificação da dispersão dos dados em torno da linha de tendência dos dados através do desvio padrão relativo (DPR_{lin}) do coeficiente angular (Equação 16) (GONZÁLEZ & HERRADOR, 2007).

$$DPR_{lin} = \frac{s_b}{b_1} \quad (\text{Equação 16})$$

Sendo:

s_b = desvio padrão do coeficiente angular (Equação 17)

b_1 = coeficiente angular

$$s_b = \frac{s_{y/x}}{\sqrt{\sum_i (x_i - \bar{x})^2}} \quad (\text{Equação 17})$$

Sendo: $s_{y/x}$ - erros aleatórios na direção de y (Equação 18)

$$s_{y/x} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}} \quad (\text{Equação 18})$$

O valor do DPR_{lin} é considerado com outro parâmetro de desempenho e, geralmente, o valor máximo adotado é de 5%.

3.4.2. Limites de detecção (LD) e limites de quantificação (LQ)

Segundo o INMETRO (2018), o limite de detecção (LD) de um procedimento analítico individual é a menor quantidade de analito na amostra que pode ser detectada, mas não necessariamente quantificada sob as condições estabelecidas para o ensaio.

Já o limite de quantificação (LQ), consiste na menor quantidade do analito na amostra que pode ser quantitativamente determinada com precisão e exatidão aceitáveis.

Existem diversos modos de se calcular os limites de detecção. Algumas das abordagens mais utilizadas são:

- método da relação sinal/ruído: sucessivas diluições da amostra são feitas até se encontrar a razão sinal/ruído de aproximadamente 3:1;
- método do desvio padrão e inclinação da reta: o limite de detecção é determinado pela Equação 19:

$$LD = (3,3 \cdot s_{branco})/b_1 \quad (\text{Equação 19})$$

Onde:

s_{branco} = desvio padrão do sinal do branco

b_1 = inclinação da curva analítica

Já para o LQ, algumas das abordagens mais comumente utilizadas consistem nos:

- método da relação sinal/ruído: sucessivas diluições da amostra são feitas até se encontrar a razão sinal/ruído de aproximadamente 10:1;
- método do desvio padrão e inclinação da reta: o limite de detecção é determinado pela Equação 20:

$$LD = (10 \cdot s_{branco})/b_1 \quad (\text{Equação 20})$$

Onde:

s_{branco} = desvio padrão do sinal do branco

b_1 = inclinação da curva analítica

Ao colocar b em evidência nas Equações 19 e 20 e igualando, tem-se:

$$\frac{3,3}{LD} = \frac{10}{LQ}$$

$$LQ \approx 3 LD \text{ ou } LD = \frac{LQ}{3} \quad (\text{Equação 21})$$

Tanto o limite de detecção quanto o de quantificação são características do método e, por isso, não possuem critério de aceitação.

3.4.3. Repetibilidade

De acordo com o Vocabulário Internacional de Metrologia (VIM - 2012), as condições de repetibilidade de medição devem incluir o mesmo procedimento de medição, os mesmos operadores, o mesmo sistema de medição, as mesmas condições de operação e o mesmo local, assim como o período entre as medições repetidas deve ser curto.

A repetibilidade pode ser expressa quantitativamente em termos da característica da dispersão dos resultados e pode ser analisado por meio de análise de padrões, material de referência ou adição do analito ao branco da amostra, em concentrações baixa, média e alta da faixa de trabalho (INMETRO, 2018).

Para avaliar a repetibilidade é necessário calcular, através da Equação 22, o desvio padrão relativo (DPR_{rep}), também chamado de coeficiente de variação (CV), obtido sob condições de repetibilidade.

$$CV = DPR_{rep} = \frac{s}{CMD} \times 100 \quad (\text{Equação 22})$$

Sendo:

s = desvio padrão

CMD = concentração média determinada

De forma similar a recuperação analítica, os critérios para aceitação do DPR estão atrelados ao nível de concentração (Quadro 3.6).

Quadro 3.6: Critérios de aceitação para a repetibilidade. (Fonte: *Association of Official Agricultural Chemists, 2016*)

Analito %	Fração Mássica (C)	Unidade	DPR, %
100	1	100%	1,3
10	10 ⁻¹	10%	1,9
1	10 ⁻²	1%	2,7
0,1	10 ⁻³	0,1%	3,7
0,01	10 ⁻⁴	100 ppm (mg.kg ⁻¹)	5,3
0,001	10 ⁻⁵	10 ppm (mg.kg ⁻¹)	7,3
0,0001	10 ⁻⁶	1 ppm (mg.kg ⁻¹)	11
0,00001	10 ⁻⁷	100 ppb (µg.kg ⁻¹)	15
0,000001	10 ⁻⁸	10 ppb (µg.kg ⁻¹)	21
0,0000001	10 ⁻⁹	1 ppb (µg.kg ⁻¹)	30

3.4.4. Recuperação analítica

Matematicamente, recuperação analítica é a razão entre o valor observado e o valor esperado para um analito (Equação 23) (INMETRO, 2018).

$$\text{Recuperação analítica} = \frac{\text{valor observado}}{\text{valor esperado}} \times 100 \quad (\text{Equação 23})$$

Normalmente, os critérios de aceitação para a recuperação estão atrelados ao nível de concentração. Na Quadro 3.7, estão apresentados os critérios sugeridos pela AOAC.

Quadro 3.7: Critérios de aceitação para recuperação. Fonte: AOAC, 2016

Analito %	Fração Mássica (C)	Unidade	Recuperação analítica, %
100	1	100%	98 – 102
10	10 ⁻¹	10%	98 - 102
1	10 ⁻²	1%	97 -103
0,1	10 ⁻³	0,1%	95 - 105
0,01	10 ⁻⁴	100 ppm (mg.kg ⁻¹)	90 – 107
0,001	10 ⁻⁵	10 ppm (mg.kg ⁻¹)	80 - 110
0,0001	10 ⁻⁶	1 ppm (mg.kg ⁻¹)	80 – 110
0,00001	10 ⁻⁷	100 ppb (µg.kg ⁻¹)	80 - 110
0,000001	10 ⁻⁸	10 ppb (µg.kg ⁻¹)	60 – 115
0,0000001	10 ⁻⁹	1 ppb (µg.kg ⁻¹)	40 - 120

3.5. Quimiometria e suas ferramentas

A Quimiometria, a qual consiste na aplicação de métodos matemáticos e técnicas estatísticas para coletar informações significativas de dados químicos, pode ser dividida em diversas frentes de pesquisa: planejamento e otimização de experimentos, reconhecimento de padrões e desenvolvimento de modelos de classificação, resolução de sinais analíticos, calibração multivariada, entre outros (VALDERRAMA *et al.*, 2015).

No contexto da retirada de informações relevantes a partir dos dados experimentais, os métodos de reconhecimento de padrões merecem destaque, dentre eles, principalmente a análise por componentes principais.

3.5.1. Análise por componentes principais (ACP)

Há uma crescente demanda por métodos que diminuam a dimensionalidade dos dados produzidos atualmente, de maneira a facilitar a sua interpretação e, ao mesmo tempo, preservar a maior parte da informação contida nesses dados. Nesse contexto, apesar do desenvolvimento de muitas técnicas para este propósito, a análise de Componentes Principais (ACP ou do inglês *PCA – Principal Components Analysis*) é uma das ferramentas mais antiga e mais usada pela comunidade científica (JOLLIFFE & CADIMA, 2016).

Uns dos primeiros artigos na literatura nos quais essa ferramenta foi utilizada, datam de 1901 com Pearson (*On lines and planes of closest fit to systems of points in space*) e 1933 com Hotteling (*Analysis of a complex of statistical variables into principal componentes*), mas foi apenas depois de décadas, quando os computadores se tornaram mais disponíveis, que ocorreu um aprofundamento maior do assunto (JOLLIFFE & CADIMA, 2016 *apud* PEARSON, 1901; HOTTELING, 1933).

A ACP tem a capacidade de reduzir a dimensionalidade de um conjunto de dados ao realizar combinações lineares das variáveis originais, estabelecendo, assim, novas variáveis ortogonais entre si, denominadas componentes principais (CPs ou PCs – sigla adotada neste trabalho), as quais preservam a maior quantidade de informação (variância) possível (CORREIA & FERREIRA, 2007).

Uma forma de entender, de maneira intuitiva, como funciona a ferramenta ACP, é através do caso mais simples de redução de variáveis: de duas variáveis para uma. Nesse caso, deve-se apenas projetar os pontos do espaço bidimensional (plano) para o espaço unidimensional (linha). Todavia a direção dessa linha é importante, pois pode haver perda de informação ao longo desse processo de projeção (Figura 3.7 – a).

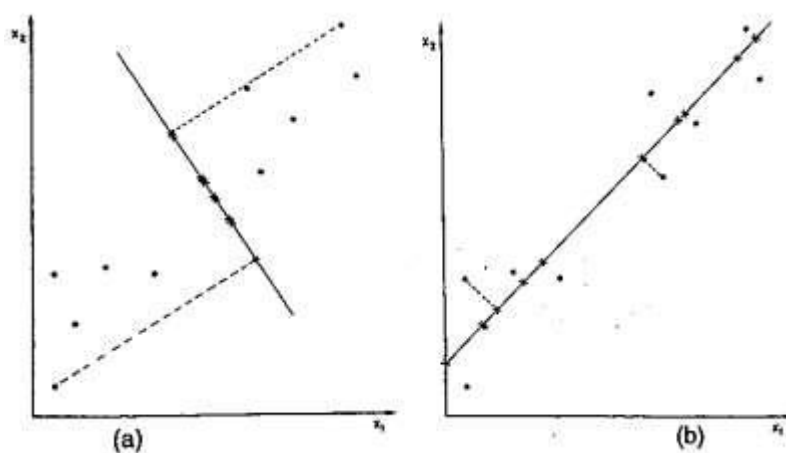


Figura 3.7: Projeções de duas para uma dimensão. (a) há perda de informação na projeção; (b) PC1 - a informação é preservada. Fonte: MASSART, 1997.

Ao comparar a Figura 3.7 – a com b, não é possível verificar, por exemplo, na Figura 3.7 - a a existência de dois grupos de dados devido aos cruzamentos ao longo da linha, já, na Figura 3.7 - b, a distinção entre eles torna-se perceptível.

Desse modo, de maneira a minimizar a perda de informações, deve ser identificada a direção onde há maior variação dos dados para que a linha seja desenhada ao longo deste eixo, chamado de primeiro componente, PC1, o qual explica a maior variação possível dos dados. As projeções dos pontos do espaço original x_1 - x_2 no PC1 são chamadas de *scores* dos objetos. Os resíduos, r_i , apresentam a variação restante inexplicada e podem ser expressos em um segundo eixo, PC2, que, por definição, é ortogonal ao primeiro, no qual também são projetados os dados originais (Figura 3.8 - a).

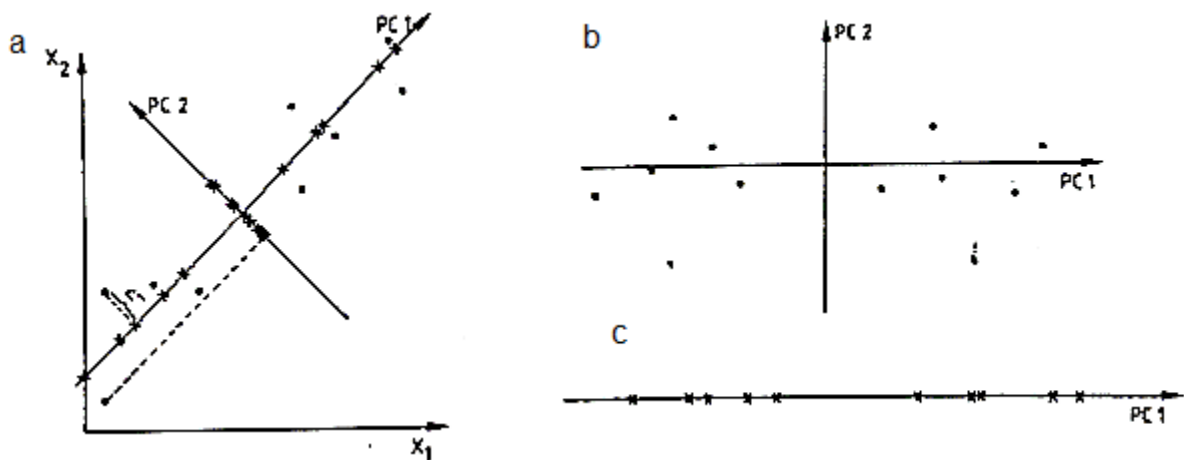


Figura 3.8: (a) – PC2 é ortogonal à PC1; (b) – Gráfico construído com os componentes principais PC1 e PC2; (c) – PC1. Fonte: MASSART, 1997

PC1 e PC2 podem ser considerados os novos eixos no mesmo espaço bidimensional (Figura 3.8 - b). Entretanto, como o intuito da ACP é reduzir o número de variáveis, o eixo PC1 pode ser o único a ser utilizado, pois este representa quase toda a estrutura dos dados originais sem uma perda significativa de informação (Figura 3.8 - c). Além disso, PC1 pode ser chamado de variável latente em comparação às variáveis x_1 e x_2 , as quais são chamadas de variáveis manifestas ou variáveis expostas. Nesse caso específico, PC2 expressa essencialmente ruído.

A elucidação completa do processo matemático da ACP pode ser encontrada em Jolliffe (1986). O presente trabalho vai apresentar tal elucidação de forma superficial, através das seguintes etapas:

- 1) Construção da matriz de dados X ($n \times p$);
- 2) Pré-tratamento de dados;
- 3) Construção da matriz de covariância;

- 4) Determinação dos autovalores e autovetores associados a matriz de covariância;
- 5) Determinação das componentes principais;
- 6) Scores e loadings;
- 7) Cálculo da contribuição de cada componente principal;
- 8) Interpretação gráfica dos scores;
- 9) Interpretação gráfica dos loadings.

3.5.1.1. Construção da matriz de dados X ($n \times p$)

Antes de aplicar a ACP, os dados devem ser organizados em uma tabela de forma que n objetos/amostras constituem as linhas e p variáveis constituem as colunas, construindo uma matriz de dados n versus p (Figura 3.9).

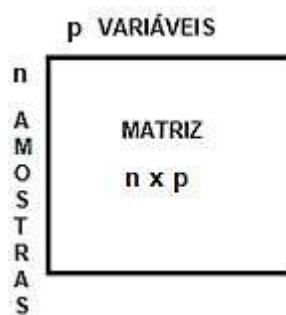


Figura 3.9: Matriz de dados $n \times p$. Adaptado de: MASSART, 1997.

O número de dimensões dessa matriz é equivalente ao número de variáveis medidas para cada amostra, dessa forma, para visualizar todos os dados seria necessário construir um gráfico com p -dimensões. Contudo, para um p maior que 3, a visualização desse gráfico não é possível, tornando necessário uma redução para 3 ou menos variáveis (MASSART, 1997).

3.5.1.2. Pré-tratamento de dados

Algumas propriedades da ACP têm consequências indesejáveis quando as variáveis originais possuem diferentes unidades de medida. O fato de a ACP ser definido por um critério (variância), a qual depende das unidades de medida, faz com que os componentes principais baseados na matriz de covariância sofram

mudanças se a unidade de uma ou mais variáveis sejam diferentes (JOLLIFFE & CADIMA 2016).

Com o intuito de superar esse possível obstáculo, o uso de um pré-tratamento dos dados originais é muito comum, sendo, a centralização pela média, escalamento por variância e o auto escalamento, alguns dos pré-tratamentos mais usados. Estes são descritos como:

- Centralização na média (*mean centering* – Equação 24): a média da coluna correspondente (m_j – Equação 25) é subtraída de cada elemento (x_{ij}) de X.

$$z_{ij} = x_{ij} - m_j \quad (\text{Equação 24})$$

$$m_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (\text{Equação 25})$$

Para:

z = valor obtido de cada amostra após o tratamento de dados

x = valor medido para cada amostra

i = 1, ..., n

j = 1, ..., p

O efeito da centralização da média é a translação paralela dos pontos que representam os dados originais, fazendo com que seu centro coincida com a origem do espaço.

- Escalamento por variância (*variance scaling* – Equação 26): cada elemento é dividido pelo desvio padrão da coluna correspondente (δ_j - Equação 27):

$$z_{ij} = \frac{x_{ij}}{\delta_j} \quad (\text{Equação 26})$$

$$\delta_j^2 = \frac{1}{n} \sum_{i=1}^n (x_{ij} - m_j)^2 \quad (\text{Equação 27})$$

Através do escalamento por variância, as variáveis são colocadas em uma mesma escala de variâncias, tornam os dados com valores invariantes com respeito à unidade utilizada originalmente.

- Auto escalamento (*auto scaling* – Equação 28): cada elemento (x_{ij}) é subtraído da média da coluna (m_j) e dividido pelo desvio padrão da coluna correspondente (δ_j):

$$z_{ij} = \frac{x_{ij} - m_j}{\delta_j} \quad (\text{Equação 28})$$

O auto escalamento combina os efeitos tanto da centralização da média quanto do escalamento por variância, permitindo, assim, que pequenas distinções entre as variáveis fiquem mais evidentes.

O auto escalamento é indicado quando as variáveis têm diferentes unidade ou quando a faixa de variação dos dados é expressivo, pois este pré-tratamento possibilita igualar o impacto de cada uma das variáveis, minimizando o efeito de uma variável dominante nos resultados da análise.

3.5.1.3. Construção da matriz de covariância

Uma das etapas para a determinação dos componentes principais é a construção de uma matriz de covariância (Σ). Em teoria da probabilidade e na estatística, a covariância (Equação 29), é uma medida do grau de interdependência (ou inter-relação) numérica entre duas variáveis aleatórias (X, Y).

$$cov(x, y) = \frac{\sum_{i=1}^n [(x - \bar{x})(y - \bar{y})]}{n} \quad (\text{Equação 29})$$

Sendo:

x = valor medido para a variável x

y = valor medido para a variável y

\bar{x} = valor da média das amostras para a variável x

\bar{y} = valor da média das amostras para a variável y

n = número de amostras

Considerando a situação em que se deseja calcular a cov (x, x), ou seja, a covariância entre uma variável e ela mesma, o resultado obtido equivale a variância de x (Equação 30).

$$var(x) = \frac{\sum_{i=1}^n (x - \bar{x})^2}{n} \quad (\text{Equação 30})$$

Para a construção da matriz, as p variáveis em cada n amostras, após o pré-tratamento, podem ser expressas na forma vetorial por $X = [X_1, X_2, \dots, X_p]'$, assim como as médias dos dados $m = [m_1, m_2, \dots, m_p]'$, sendo o vetor associado as médias dos dados calculado pela Equação 31.

$$m_j = \frac{1}{n} \sum_{i=1}^n z_i \quad (\text{Equação 31})$$

Sendo:

z_i = valor obtido pelo pré-tratamento correspondente ao valor original (x_i) da medida de cada amostra

$j = 1, \dots, p$

A Equação 32, representa, então, a matriz de covariância Σ .

$$\Sigma = \begin{bmatrix} \frac{1}{n}[(X_1 - m_1)(X_1 - m_1)] & \frac{1}{n}[(X_1 - m_1)(X_2 - m_2)] & \dots & \frac{1}{n}[(X_1 - m_1)(X_n - m_n)] \\ \frac{1}{n}[(X_2 - m_2)(X_1 - m_1)] & \frac{1}{n}[(X_2 - m_2)(X_2 - m_2)] & \dots & \frac{1}{n}[(X_2 - m_2)(X_n - m_n)] \\ & \vdots & \ddots & \vdots \\ \frac{1}{n}[(X_n - m_n)(X_1 - m_1)] & \frac{1}{n}[(X_n - m_n)(X_2 - m_2)] & \dots & \frac{1}{n}[(X_n - m_n)(X_n - m_n)] \end{bmatrix} \quad (\text{Equação 32})$$

Dessa forma, cada elemento (i, j) da matriz da covariância consiste na covariância entre o i -ésimo elemento com o j -ésimo elemento, quando $i \neq j$, e a variância do j -ésimo elemento, quando $i = j$.

3.5.1.4. Determinação dos autovalores e autovetores associados a matriz de covariância

A etapa seguinte à construção da matriz de covariância é a determinação dos pares de autovalores (λ_p) e autovetores (g_p): $(\lambda_1, g_1), (\lambda_2, g_2), \dots, (\lambda_p, g_p)$, em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, associados a Σ .

Diz-se que um vetor g_i é um autovetor de uma matriz quadrada se a multiplicação desta matriz pelo vetor resulta em um múltiplo do próprio vetor g_i , isto é, em $\lambda \times g_i$ (multiplicação de um escalar pelo vetor). Nesse caso, λ é chamado de autovalor da matriz quadrada associado ao autovetor g_i .

Os autovalores podem ser obtidos através da Decomposição de Valor Singular (DVS ou em inglês *Single Value Decomposition* – SVD) descrita pela Equação 33, a qual representa a decomposição de uma matriz de dados X:

$$X = U \times W \times V^T \quad (\text{Equação 33})$$

Sendo:

Z = forma vetorial das p variáveis em cada n amostras, após o pré-tratamento

U = matriz relacionada aos scores

V = matriz relacionada aos *loadings*

W = matriz de valores singulares (Equação 34)

$$\Lambda = W^2 \quad (\text{Equação 34})$$

Sendo:

Λ = matriz de autovalores (valores de autovalores encontram-se nas diagonais dessa matriz)

3.5.1.5. Determinação dos componentes principais;

Após a determinação dos autovalores e dos autovetores, pode-se determinar as componentes principais (as direções de maiores variações dos dados), onde o i -ésimo componente principal pode ser definido por (Equação 35):

$$PC_i = g_i'X = g_{i1}X_1 + g_{i2}X_2 + \dots + g_{ip}X_p \quad (\text{Equação 35})$$

Sendo:

X_j = forma vetorial das p variáveis em cada n amostras, após o pré-tratamento

g_i = autovetor associado à um autovalor da matriz de covariância

A variável PC_i é uma variável latente, isto é, não é mensurada a partir do experimento ou levantamento amostral. Deve-se notar que como $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, os primeiros autovetores possuem maiores autovalores, e explicam, necessariamente, maior variabilidade experimental, definindo então a ordem das componentes principais.

3.5.1.6. Scores e loadings

Como foi dito anteriormente, as projeções dos dados nos novos eixos, correspondentes aos componentes principais, são chamados de pontuações, ou *scores* (s_{ip}). No componente principal, os *scores* são uma soma ponderada (Equação 36), ou seja, combinações lineares das variáveis originais (x_{ij}). Os pesos da soma são chamados de *loadings* e contêm informação sobre as variáveis e o quanto cada uma contribui para a nova variável (PC).

$$s_{ip} = \sum_j v_{ip} \times x_{ij} \quad (\text{Equação 36})$$

Sendo:

v_{ip} = peso (*loading*) da variável

x_{ij} = valor do i -ésimo objeto da para a variável original ou do valor correspondente após o pré-tratamento de dados (z_{ij})

Assim sendo, a Equação 36 relaciona os componentes principais às variáveis originais.

Em notação de matriz (Equação 37):

$$S_{n \times p} = X_{n \times p} \times V_{p \times p} \quad (\text{Equação 37})$$

Sendo:

S = matriz dos *scores*

X = matriz das variáveis originais ou matriz correspondente após o pré-tratamento de dados (Z)

V = matriz relacionada aos *loadings*

Para obter os valores dos *loadings* utiliza-se a Equação 38.

$$P = A \times \sqrt{\Lambda} \quad (\text{Equação 38})$$

Sendo:

P = matriz dos *loadings*

A = matriz de autovetores

Λ = matriz de autovalores (valores de autovalores encontram-se nas diagonais dessa matriz)

3.5.1.7. Cálculo da contribuição de cada componente principal.

A contribuição de, por exemplo, o k-ésimo componente principal (PC_i) é dada por (Equação 39):

$$c_k = \frac{\text{var}(CP_i)}{\sum_{l=1}^p \text{var}(CP_l)} \times 100 = \frac{\lambda_i}{\sum_{l=1}^p \lambda_l} \times 100 = \frac{\lambda_i}{\text{tr}(\Sigma)} \times 100 \quad (\text{Equação 39})$$

Pela proporção de explicação da variância total, que o modelo de k componentes principais está associado, é possível determinar o número de componentes que devem ser utilizados. Em muitos casos, adota-se modelos que explicam pelo menos 80% da variação total.

3.5.1.8. Interpretação gráfica dos *scores*

Em uma dada situação com p variáveis, pode-se continuar extraindo componentes principais até que p PC's tenham sido extraídas, entretanto, apenas as

primeiras PC's contêm uma maior variação, ou seja, um valor maior de pontuações, o que implica em uma maior quantidade de informação.

Dessa forma, as projeções dos dados no plano PC1 x PC2 (ou até outros planos que incluam PC's que contribuem significativamente para explicação da variância total) podem ser computadas e exibidas em um gráfico, chamado de gráfico de pontuações, ou *score plot* (exemplo na Figura 3.10). Um exemplo de aplicação é mostrado na Figura 3.10, onde tem-se 100 amostras com 3 tipos de vinhos os quais foram avaliados utilizando-se 8 variáveis com o intuito de diferenciá-los.

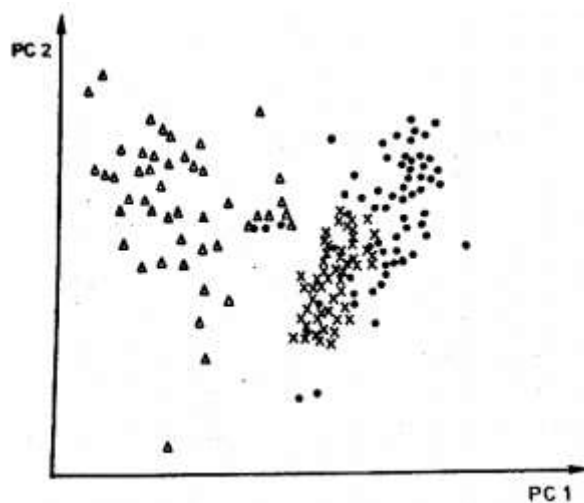


Figura 3.10: Gráfico de *scores* de 3 tipos de vinhos italianos (Δ = Barolo, \bullet = Barbera, \times = Grignolino). Fonte: MASSART, 1997.

Com base na Figura 3.10, pode-se observar que as variáveis foram adequadas para distinguir o vinho Barolo dos outros dois tipos de vinho uma vez que no espaço bidimensional definido pelas componentes PC1 e PC2, eles se encontram visualmente diferenciados dos outros vinhos.

3.5.1.9. Interpretação gráfica dos *loadings*

A informação contida nos *loadings* pode ser apresentada na forma de um gráfico. Um exemplo pode ser visto na Figura 3.11, onde, na avaliação da poluição do ar, foram estudadas concentrações de algumas substâncias em amostras de ar.

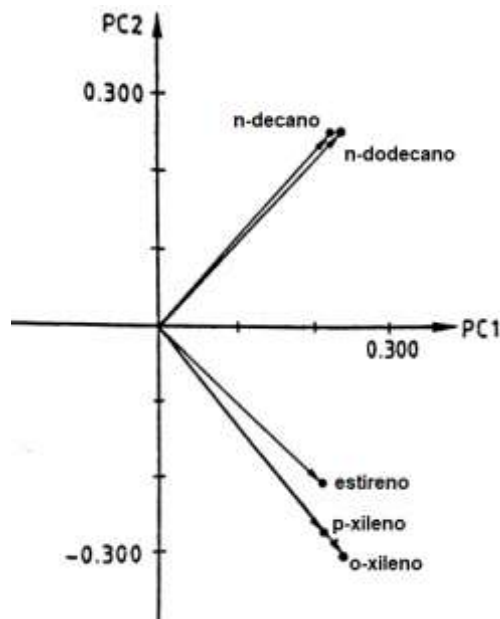


Figura 3.11: Gráfico de *loading* de algumas variáveis para o exemplo de poluição do ar. Fonte: MASSART, 1997.

Assim como demonstrado na Figura 3.11, a interpretação do gráfico é baseada na direção onde as variáveis estão em relação à origem. Ou seja, duas variáveis são correlacionadas significativamente quando há um pequeno ângulo entre as linhas conectando-as à origem. No exemplo apresentado, as variáveis relacionadas as concentrações dos poluentes n-decano e n-dodecano estão correlacionadas fortemente entre si, assim como as concentrações do estireno, orto-xileno e para-xileno.

3.6. Plataforma

Há uma grande variedade de programas voltados para a análise multivariada e a análise de componentes principais está presente na sua maioria. Dentre os programas comerciais pode-se citar: Unscrambler[®], Statistica[®], Pirouette[®] e Matlab[®]. Na categoria de programas livres, destaca-se o GNU Octave[®] (LYRA, et. Al, 2010).

3.6.1. GNU octave

O GNU Octave é uma linguagem de alto nível, destinada principalmente a cálculos numéricos. Ele fornece uma interface de linha de comando conveniente para resolver problemas lineares e não lineares numericamente usando uma linguagem que é principalmente compatível com o Matlab®.

A motivação original para escrever o Octave era fornecer um software para acompanhar um livro de graduação em engenharia química (Análise de Reatores Químicos e Designs Fundamentais, em inglês - *Chemical Reactor Analysis and Design Fundamentals*, publicado em 2002 e escrito por Jim Rawlings e John Ekerdt). No entanto, Octave evoluiu, com o tempo, tornando-se uma ferramenta útil para a área de computação numérica e de construção de gráficos, as quais são usadas para uma grande variedade de tarefas (EATON, 2012).

O Octave, além disso, é uma plataforma livre desenvolvida por uma comunidade de usuários, ou seja, é distribuída sob termos que garantem certas liberdades para seus usuários, como a liberdade de executar, copiar, distribuir, estudar, alterar e melhorar a plataforma (EATON, 2012).

4. Material e métodos

4.1. Material

O equipamento, material e reagentes utilizados ao longo do trabalho estão descritos a seguir:

- ✓ Cromatógrafo de íons Thermo Scientific™ Dionex™;
- ✓ Sal Na_2CO_3 da marca Merck® para o preparo do eluente;
- ✓ H_2SO_4 P.A. da marca Merck® para o preparo do regenerante;
- ✓ Balão volumétrico de 2L para preparo do eluente e do regenerante;
- ✓ Solução padrão misto contendo 1000 mg.L^{-1} dos íons F^- , Cl^- , NO_2^- , Br^- , NO_3^- , PO_4^{3-} e SO_4^{2-} da marca SPECSOL®
- ✓ Pipeta automática de 15 - 300 μL da marca Eppendorf Xplorer®
- ✓ Pipeta automática de 100 - 1000 μL da marca Transferpette®
- ✓ Amostras (16 águas minerais de marcas diferentes);
- ✓ Tubos falcon de 15 mL.

4.2. Seleção da amostra a ser estudada

As amostras selecionadas para o presente estudo foram águas minerais de marcas nacionais e estrangeiras, nas quais foram analisados, em cada marca, os teores dos ânions: F^- , Cl^- , NO_2^- , Br^- , NO_3^- , PO_4^{3-} e SO_4^{2-} .

Os sete íons foram analisados nas águas visto que o objetivo é propor um roteiro para disciplina de Análise Instrumental I, na qual todos estes íons são estudados mesmo que os rótulos das águas minerais indicarem, geralmente, apenas os teores dos ânions F^- , Cl^- , NO_3^- e SO_4^{2-} .

Dentre os diversos produtos que são analisados na disciplina de Análise Instrumental, a água mineral, além de ser um produto barato e de fácil obtenção, também se apresenta como um sistema bastante complexo, por apresentar, por exemplo, diversos sais em sua composição, sendo, dessa forma, uma amostra oportuna para a aplicação da técnica quimiométrica selecionada.

4.2.1. Determinação da quantidade e das marcas das águas minerais

Com o intuito de obter uma diversidade nos resultados das análises, buscou-se adquirir águas minerais de diferentes fontes e de marcas nacionais e estrangeiras. Desse modo, 10 águas de marcas nacionais e 6 de marcas estrangeiras foram selecionadas, nas quais 14 foram compradas em uma loja de supermercados do Rio de Janeiro e 2 vieram de Portugal (Luso e Fastio). No Anexo F (página 100), encontram-se os nomes, fontes e concentrações dos íons estudados descritos nos rótulos de cada marca.

4.2.2. Procedimento analítico das análises das águas minerais

Todas as análises das águas minerais foram realizadas em triplicata. As marcas que foram analisadas sem diluições foram: Bonafont, Pouso Alto (sem gás), Sferriê, Petrópolis e Canção Nova (com gás).

Em algumas marcas, além da análise das soluções originais, foram feitas diluições com base no rótulo, em água ultra purificada (tipo I) Milli Q, de forma que a concentração obtida do íon de interesse ficasse no centro da curva analítica construída para esse íon em questão. As soluções foram preparadas com o auxílio de pipetas automáticas e avolumadas nos tubos falcons.

As marcas em que foram realizadas diluições, alíquotas e volumes finais das soluções preparadas podem ser encontradas no Quadro 4.1.

Quadro 4.1: Marcas em que foram realizadas diluições, alíquotas e volumes finais.

Marca	Alíquota (mL)	Volume final (mL)
São Lourenço	2,00	10,00
Nestlè	2,25	10,00
Minalba Premium	5,00	10,00
Qualitá	4,25	10,00
Pouso alto (com gás)	3,80	10,00
Fastio	2,50	7,00
Luso	2,00	10,00
Evian	1,00	10,00
Acqua Panna	0,70	10,00
S. Pellegrino*	0,06	15,00
	0,30	10,00
	5,00	10,00
Perrier	0,45	10,00

*Foram feitas 3 diluições para a marca S.Pellegrino devido a presença de alguns íons em altas concentrações.

Alíquotas de 15 mL, de cada água mineral com gás foram transferida, separadamente, para cada tudo falcon e colocada por 45 minutos no ultrassom, de maneira a eliminar todo gás carbônico dissolvido.

4.3. Seleção da técnica analítica e condições operacionais da cromatografia de íons (CI)

A técnica de Cromatografia de Íons foi selecionada (Figura 4.1), pois, além de ser de uma das técnicas abordadas nas práticas da disciplina escolhida, é bastante conhecida e utilizada, devido, principalmente, à facilidade com que é feita a separação, identificação e quantificação das espécies químicas.



Figura 4.1: Cromatógrafo de Íons à esquerda, computador usado para o tratamento de dados e, à direita, o módulo do equipamento.

Os experimentos foram realizados parte no laboratório onde as aulas são ministradas (laboratório 511, no quinto andar do Instituto de Química da Universidade Federal do Rio de Janeiro - UFRJ) e parte no laboratório LAM (Laboratório de Análise Ambiental e Mineral – laboratório 513), na mesma localização.

No Quadro 4.2, encontram-se as principais características do equipamento e condições operacionais utilizadas no CI, as quais correspondem às condições que são utilizadas durante a aula de Análise Instrumental I.

Quadro 4.2: Características e condições operacionais do CI

Características e condições operacionais do CI	
Marca	Thermo Scientific™ Dionex™
Modelo	ICS-1600
Coluna*	AS9-HC
Pré-coluna*	AG9-HC
Detector*	DS6
Supressor*	ASRS
Software	Chromeleon 6.80
Corrente aplicada ao supressor	45 mA
Bomba*	Isocrática
Volume de injeção	200 µL
Temperatura aplicada ao forno de coluna	30 °C
Eluente	Solução com Na ₂ CO ₃ 9 mM
Regenerante	Solução com H ₂ SO ₄ 25 mM
Vazão	1 mL/min

* componentes do CI que também pertencem a marca Thermo Scientific™ e Dionex™

4.4. Construção das curvas analíticas e aplicação dos testes estatísticos

As curvas analíticas foram construídas a partir dos valores de concentração, principalmente, para os íons F⁻, Cl⁻, NO₃⁻ e SO₄²⁻, nos rótulos de cada garrafa. Como o padrão disponível para o preparo das soluções consistia em uma mistura dos padrões com uma concentração de 1000 mg.L⁻¹ de todos os íons (F⁻, Cl⁻, NO₂⁻, Br⁻, NO₃⁻, PO₄³⁻ e SO₄²⁻) e os valores de concentração nos rótulos das garrafas variavam significativamente, principalmente para os íons Cl⁻ e NO₃⁻, buscou-se uma ou duas faixas que incorporassem a maioria dos valores de concentração sem a necessidade de muitas diluições das amostras.

Dessa forma, foram selecionadas duas faixas de concentração, de 0,10 à 0,50 mg.L⁻¹ (faixa baixa) e de 0,50 à 2,50 mg.L⁻¹ (faixa alta), para o preparo das soluções padrões, utilizando cada faixa de acordo com as concentrações dos rótulos (Quadro 4.3).

Os níveis de cada faixa de concentração foram:

- Para a faixa de 0,10 à 0,50 mg.L⁻¹: 0,10 – 0,20 – 0,30 – 0,40 – 0,50
- Para a faixa de 0,5 à 2,5 mg.L⁻¹: 0,50 – 1,00 – 1,50 – 2,00 – 2,50

Quadro 4.3: Faixas de concentração selecionadas para cada íon para a construção da curva analítica

Íon	Faixa de concentração selecionada
Fluoreto	0,10 à 0,50 mg.L ⁻¹
Cloreto	Ambas as faixas foram utilizadas
Nitrito	0,10 à 0,50 mg.L ⁻¹
Brometo	0,10 à 0,50 mg.L ⁻¹
Nitrato	Ambas as faixas foram utilizadas
Fosfato	0,10 à 0,50 mg.L ⁻¹
Sulfato	0,50 à 2,50 mg.L ⁻¹

Cada nível foi preparado com o auxílio de pipetas automáticas e tubos falcon e analisado em três replicatas falsas, isto é, três alíquotas da mesma solução foram analisadas.

O software do equipamento além de fornecer o cromatograma resultante da análise, apresenta como resultado diversos parâmetros da análise como tempo de retenção de cada pico, a altura, a área, resolução, entre outros.

Com os dados das áreas dos picos de cada íon, fornecidos pelo software do equipamento, foi realizado, para cada nível de concentração, primeiramente, o teste de Grubbs para um valor suspeito e, depois, com todos os níveis, o teste de Cochran.

Após aplicar todos os testes e confirmar que todos os dados são homocedásticos, foi feita a regressão linear e as curvas analíticas foram construídas. Por fim, foi construído o gráfico de resíduos, a tabela ANOVA, determinado o R² e a equação da reta de cada curva, avaliando, desse modo, o modelo proposto.

4.5. Validação parcial

Neste trabalho, realizou-se uma validação parcial do método com o objetivo de demonstrar, por meio de alguns parâmetros de validação, que o método analítico tem algumas características necessárias para obtenção de resultados com a qualidade exigida.

Os seguintes parâmetros foram determinados: linearidade, limite de detecção, limite de quantificação, repetibilidade e recuperação analítica.

4.5.1. Linearidade

Logo após a realização dos testes estatísticos, Grubbs e Cochran, e os métodos matemáticos, foi feita a avaliação visual das curvas construídas e, em seguida, a avaliação estatística por meio da avaliação do desvio padrão relativo do coeficiente angular verificando se o valor calculado ficou abaixo de 5%, critério sugerido como máximo.

4.5.2. Métodos de determinação de LD e LQ e procedimento de aplicação

Os métodos escolhidos e seus respectivos procedimentos para a determinação dos limites de detecção (LD) e quantificação (LQ) estão no Quadro 4.4.

Quadro 4.4 – Métodos utilizados para determinar LD e LQ

Método	LD	LQ
Relação sinal/ruído	Aos resultados obtidos para LQ através da relação sinal/ruído foi aplicada a Equação: $LD = \frac{LQ}{3}$	Foram realizadas sucessivas diluições até se encontrar a razão sinal/ruído de aproximadamente 10:1
Desvio padrão e inclinação da reta (regressão linear)	A partir da análise em triplicata da água Mili Q e da construção da curva analítica foi aplicada a seguinte equação: $LD = (3,3 \cdot s_{\text{branco}})/b$	A partir da análise em triplicata da água Mili Q e da construção da curva analítica foi aplicada a seguinte equação: $LQ = (10 \cdot s_{\text{branco}})/b$

* s_{branco} = desvio padrão do branco e b = coeficiente angular da curva analítica

Os valores de LD para o método da relação sinal/ruído foram determinados, indiretamente, pelos valores de LQ devido a uma maior confiabilidade nos valores de sinais obtidos pela multiplicação pelo fator 10.

4.5.3. Repetibilidade

A repetibilidade foi avaliada através do cálculo do desvio padrão relativo das concentrações baixa ($0,10 \text{ mg.L}^{-1}$), média ($0,30 \text{ mg.L}^{-1}$) e alta ($0,50 \text{ mg.L}^{-1}$) da faixa de trabalho ($0,10 - 0,50 \text{ mg.L}^{-1}$) para todos os íons e comparando com as faixas permitidas desse parâmetro de acordo com a concentração (analito presente em $0,00001\%$ - $\% \text{DPR} < 15\%$).

4.5.4. Recuperação analítica

A recuperação analítica foi avaliada, de forma similar a repetibilidade, através do cálculo da recuperação analítica das concentrações $0,10 \text{ mg.L}^{-1}$, $0,30 \text{ mg.L}^{-1}$ e $0,50 \text{ mg.L}^{-1}$ para todos os íons e comparando com as faixas sugeridas como permitidas desse parâmetro de acordo com a concentração (analito presente em $0,00001\%$, $80\% - 110\%$).

4.6. Seleção da ferramenta quimiométrica

Apesar do crescente uso da quimiometria para resolução de problemas em diversos campos da pesquisa, poucos são os artigos dedicados a uma abordagem didática sobre a ACP e que demonstram, por meio de uma plataforma, uma aplicação dessa ferramenta a uma matriz X de dados provenientes de variáveis discretas (VALDERRAMA *et al.*, 2015).

Além disso, os conhecimentos em quimiometria funcionam como uma ferramenta que permite uma maior agilidade na interpretação de dados multivariados, os quais são analisados por profissionais na área da Química quase que diariamente.

4.7. Seleção e uso da plataforma selecionada para a aplicação do PCA

A plataforma selecionada foi o GNU Octave 5.1.0 por ser um programa livre e, com isso, ser de fácil obtenção tanto pelos professores quanto para os alunos. Ademais, há uma plataforma virtual que contém orientações, pacotes construídos

por colaboradores com funções pré-construídas, entre outras facilidades que auxiliam o uso do software em questão.

O programa foi obtido na plataforma do GNU através do link:

<https://www.gnu.org/software/octave/download.html>

O pacote de estatística, o qual possui as funções necessárias para a aplicação da ACP, foi obtido na plataforma do GNU através do link:

<https://octave.sourceforge.io/statistics/index.html>

A lista de todas as funções e seus comandos são encontrados no link:

https://octave.sourceforge.io/list_functions.php?sort=alphabetic

Com base nas funções indicadas pela plataforma do GNU Octave foi possível a elaboração de um modo de utilização do GNU (Apêndice A, página 101).

Os resultados alcançados foram comparados com uma plataforma de estatística multivariada comercial.

4.8. Elaboração do roteiro de aula e proposta de implementação no curso

Ao longo da realização dos experimentos, tratamento de dados e uso do software foram feitas anotações, as quais auxiliaram, junto com o procedimento atual descrito na apostila da disciplina, na elaboração do método utilizado e do roteiro proposto para a aula.

5. Resultados e discussão

Inicialmente, discorrer-se-á sobre os testes e métodos matemáticos, da construção da curva analítica através da Regressão linear e da avaliação do modelo proposto.

Em um segundo momento, a avaliação dos parâmetros selecionados para a validação parcial é abordada, definindo se o método utilizado possui as características necessárias para a obtenção de resultado com a qualidade exigida.

A aplicação da ACP para o sistema escolhido é abordada apenas com os resultados obtidos pela plataforma.

5.1. Construção da curva analítica e aplicação dos métodos matemáticos e testes estatísticos

Os cromatogramas obtidos para todas as soluções padrão analisadas exibiram o formato similar com o apresentado a seguir construído para a análise da solução, por exemplo, com $0,20 \text{ mg.L}^{-1}$ com todos os íons (Figura 5.1).

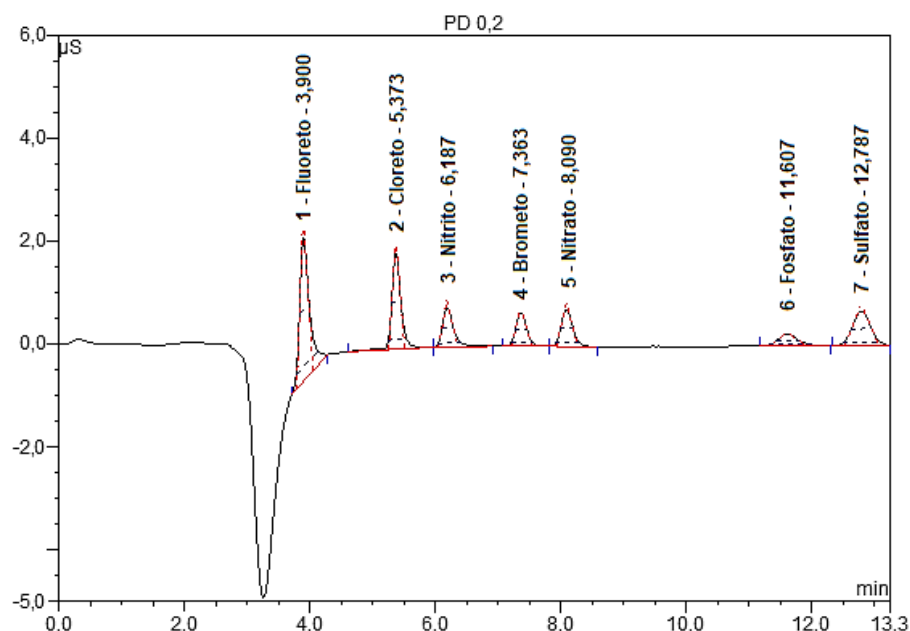


Figura 5.1: Cromatograma obtido na análise da solução padrão com concentração de $0,20 \text{ mg.L}^{-1}$ com os sete íons estudados

A ordem de eluição dos íons apresentada na Figura 5.1 (F⁻, Cl⁻, NO₂⁻, Br⁻, NO₃⁻, PO₄³⁻ e SO₄²⁻) foi a mesma para todas as análises realizadas.

A fim de ilustrar como se sucedeu a aquisição e a discussão dos resultados para a construção de todas as curvas de calibração, serão mostrados, primeiramente, os resultados para o caso do fluoreto.

O Quadro 5.1 apresenta os valores obtidos, para cada replicata, pelo software do equipamento, para as áreas dos picos referentes ao íon fluoreto nas soluções padrões da curva analítica.

Quadro 5.1: Áreas dos picos referentes ao íon fluoreto nas soluções padrões da curva analítica.

Concentração (em mg.L ⁻¹)	Área do pico (em µS*min)		
	n=1	n=2	n=3
0,10	0,279	0,287	0,295
0,20	0,512	0,499	0,521
0,30	0,751	0,696	0,749
0,40	0,975	0,976	1,289
0,50	1,494	1,231	1,292

* n = replicata

Em cada nível de concentração, com os dados organizados em ordem crescente, foi realizado o teste de Grubbs para um valor suspeito e foram obtidos os seguintes valores de G_{calc} (Quadro 5.2):

Quadro 5.2: Valores de G_{calc} obtidos para os dados da curva analítica do fluoreto

Concentração (em mg.L ⁻¹)	G _{calc}
0,10	1,014
0,20	1,044
0,30	1,154
0,40	1,154
0,50	1,125

O valor tabelado (G_{tab}) é igual a 1,155, vide Anexo B (página 96). Comparando os valores de G_{calc}, verifica-se que, em relação a todos os valores, G_{calc} < G_{tab} e, com isso, nenhum valor discrepante foi encontrado nos dados do fluoreto.

Os valores de variância, para cada concentração, podem ser encontrados no Quadro 5.3.

Quadro 5.3: Valores de variância obtidos para os dados da curva analítica do fluoreto

		Concentração (em mg.L ⁻¹)				
		0,10	0,20	0,30	0,40	0,50
Variância		6,48E-05	1,21E-04	9,69E-04	3,26E-02	1,90E-02

Com base no maior valor de variância (3,26E-02), calculado para a concentração de 0,40 mg.L⁻¹, e o somatório de todas as variâncias ($\sum_{i=1}^n s_i^2 = 0,0527$), foi possível utilizar o Teste de Cochran para avaliar a homocedasticidade dos dados. Para isso, calculou-se o C_{calc} (0,618) e comparou-se com o valor tabelado ($C_{tab} = 0,684$), para 5 pontos na curva e 3 replicatas por ponto. Confrontando os valores, verifica-se que $C_{calc} < C_{tab}$, ou seja, pode-se dizer que as variâncias são estatisticamente equivalentes e, dessa forma, os dados têm um comportamento homocedástico.

A curva analítica, a equação da reta e o coeficiente de correlação estão representados na Figura 5.2.

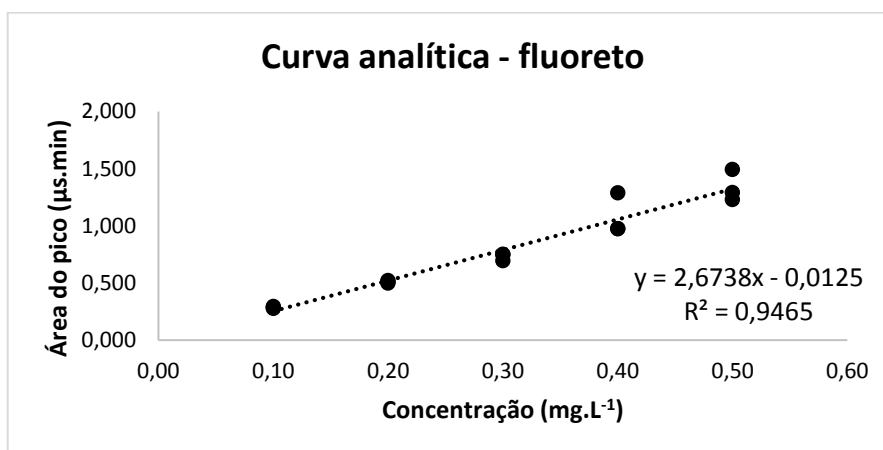


Figura 5.2: Curva analítica para o íon fluoreto

A partir do valor de R^2 (equações 15 – página 36), o qual se apresenta bem próximo à 1, constata-se que há uma correlação forte entre as variáveis estudadas.

O Quadro 5.4 representa a ANOVA para a curva analítica do íon fluoreto.

Quadro 5.4: ANOVA para a curva analítica do íon fluoreto.

Fonte de variação	Soma Quadrática	Nº de graus de liberdade	Média Quadrática	F
Regressão	2,160	4	0,540	51,19
Resíduos	0,105	10	0,010	
Total	2,266	14		

Relembrando, o teste-F no ANOVA permite a comparação da capacidade de explicação da variabilidade experimental pela regressão em relação aos resíduos. Ou seja, compara a média quadrática da regressão (SQ dividido pelo respectivo grau de liberdade) com a média quadrática dos resíduos.

O valor de F tabelado ($F_{(0,05;4,10)}$) é igual a 3,48, vide Anexo E (página 99). Como o valor de $F_{\text{calc}} > F_{\text{tab}}$ ($51,19 > 3,48$), regressão é significativa e o modelo é adequado para descrever a região estudada do sistema.

A Figura 5.3 mostra o gráfico de resíduos construído com base nos dados adquiridos para o íon fluoreto.

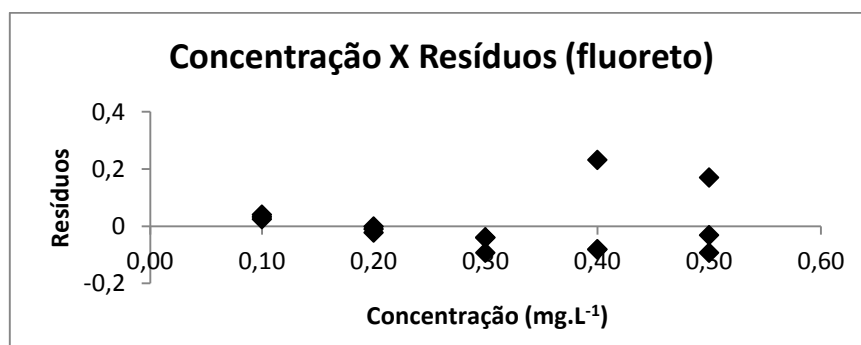


Figura 5.3: Gráfico de resíduos construído para a curva analítica do íon fluoreto

O valor associado a cada resíduo é pequeno, indicando que o valor experimental não se afastou tanto do valor previsto pela reta.

Visto que o procedimento da aplicação da Regressão linear e verificação do modelo proposto foi realizado de um modo similar ao feito para íon fluoreto, os resultados obtidos para os íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato serão apresentados nesse tópico paralelamente.

O Apêndice B (página 103) apresenta os valores obtidos para as áreas dos picos referentes a todos os íons estudados nas soluções padrões das curvas analíticas.

O Quadro 5.5 apresenta os valores de G_{calc} resultantes da aplicação do teste de Grubb's para um valor suspeito em todos os níveis de concentração, com os

dados organizados em ordem crescente, de todas as curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato.

Quadro 5.5: Valores de G_{calc} obtidos para os dados das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato

		Concentração (em mg.L ⁻¹)								
		0,10	0,20	0,30	0,40	0,50	1,00	1,50	2,00	2,50
G_{calc}	Cloreto	1,154	1,136	1,149	1,154	1,149	1,145	1,034	1,069	1,154
	Nitrito	1,121	1,094	1,014	1,064	1,152	-	-	-	-
	Brometo	1,044	1,123	1,153	1,149	1,151	-	-	-	-
	Nitrato	1,074	1,060	1,038	1,098	1,136	1,148	1,015	1,133	1,144
	Fosfato	1,008	1,095	1,109	1,145	1,150	-	-	-	-
	Sulfato	-	-	-	-	1,153	1,154	1,085	1,150	1,082

O valor tabelado (G_{tab}) é igual a 1,155, vide Anexo B (página 96). Comparando os valores de G_{calc} , verifica-se que, em relação a todos os valores, $G_{calc} < G_{tab}$ e, com isso, nenhum valor discrepante foi encontrado.

Como alguns íons necessitaram de uma curva analítica com uma faixa de concentração maior, os resultados das variâncias, para ambas as baixas (0,10 – 0,50 mg.L⁻¹ e 0,50 – 2,50 mg.L⁻¹, podem ser encontrados no Quadro 5.6)

Quadro 5.6: Valores de variância obtidos para os dados das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato

		Concentração (em mg.L ⁻¹)	Variância				
			Cloreto	Nitrito	Brometo	Nitrato	Fosfato
Faixa baixa	0,10	1,41E-04	4,33E-06	3,08E-06	4,25E-05	1,52E-05	-
	0,20	1,16E-05	5,50E-06	2,25E-05	2,79E-05	1,00E-05	-
	0,30	8,74E-04	3,12E-04	1,09E-04	2,20E-04	1,05E-05	-
	0,40	6,68E-05	1,31E-04	2,88E-05	8,49E-06	3,80E-06	-
	0,50	4,95E-04	3,83E-05	1,09E-04	5,21E-05	5,11E-06	-
Faixa alta	0,50	4,95E-04	-	-	5,21E-05	-	5,60E-05
	1,00	2,09E-03	-	-	9,80E-05	-	2,99E-04
	1,50	3,04E-03	-	-	4,91E-04	-	1,14E-03
	2,00	1,10E-03	-	-	3,26E-05	-	8,77E-05
	2,50	2,08E-03	-	-	1,38E-04	-	7,11E-04

No Quadro 5.7 são exibidos os valores, para cada curva analítica construída, da maior variância, do somatório das variâncias e o resultado do Teste de Cochran (C_{calc}).

Quadro 5.7: Valores calculados para a maior variância, o somatório das variâncias e o C_{calc} para as curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato

		Cloreto	Nitrito	Brometo	Nitrato	Fosfato	Sulfato
Faixa baixa	Maior valor de variância	8,74E-04	3,12E-04	1,09E-04	2,20E-04	1,52E-05	-
	$\sum_{i=1}^n s_i^2$	1,59E-03	4,91E-04	2,72E-04	3,51E-04	4,47E-05	-
	C_{calc}	0,550	0,635	0,400	0,627	0,340	-
Faixa alta	Maior valor de variância	3,04E-03	-	-	4,91E-04	-	1,14E-03
	$\sum_{i=1}^n s_i^2$	8,81E-03	-	-	8,11E-04	-	2,30E-03
	C_{calc}	0,345	-	-	0,605	-	0,498

Confrontando todos os valores de C_{calc} e o valor de C_{tab} (0,684 – Anexo D, página 98), verifica-se que $C_{calc} < C_{tab}$, ou seja, pode-se dizer que os dados têm um comportamento homocedástico.

O Quadro 5.8 apresenta todos os valores de F_{calc} obtidos pelos quadros para ANOVA de todas as curvas analíticas, os quais encontram-se reunidos no Apêndice C (página 97).

Quadro 5.8: Valores calculados para F_{calc} para as curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato

	Faixa baixa (0,10 – 0,50 mg.L ⁻¹)					Faixa alta (0,50 – 2,50 mg.L ⁻¹)		
	CLORETO	NITRITO	BROMETO	NITRATO	FOSFATO	CLORETO	NITRATO	SULFATO
F_{calc}	406,93	494,95	598,24	672,73	1272,24	3520,77	9989,31	5800,278

O valor de F tabelado ($F_{(0,05;4,10)}$) é igual a 3,48, vide Anexo E (página 99). Como o valor de $F_{calc} > F_{tab}$, para todas as curvas analíticas, as regressões são significativas e os modelos são adequados para descrever as regiões estudadas.

As curvas analíticas, as equações das retas e os coeficientes de correlação dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato estão representados nas Figuras 5.4, 5.5, 5.6, 5.7, 5.8 e 5.9, respectivamente.

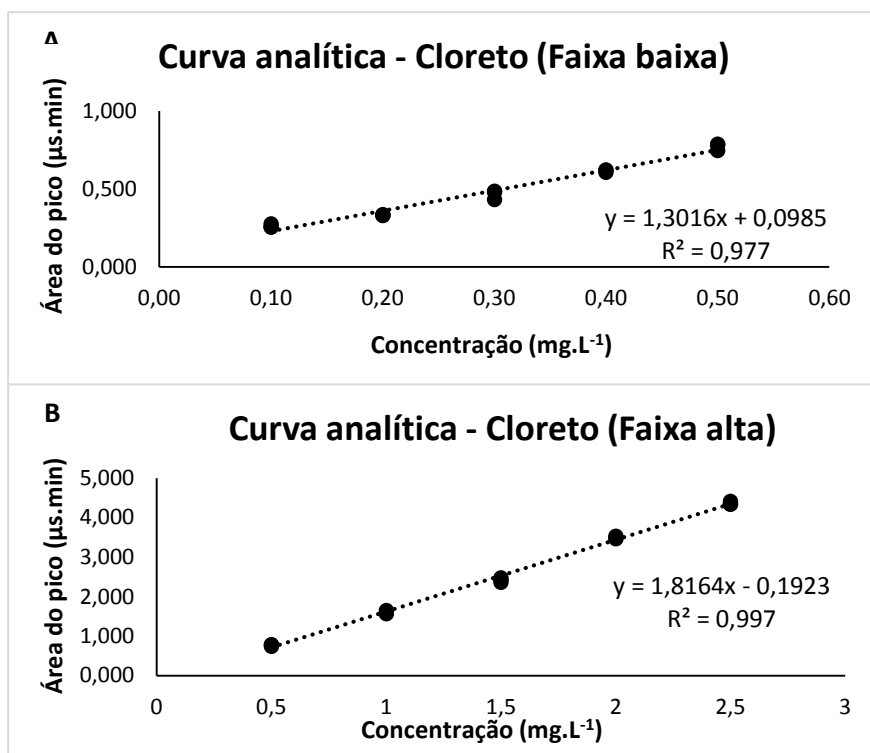


Figura 5.4: Curvas analíticas para o íon cloreto. A – Faixa baixa. B – Faixa alta

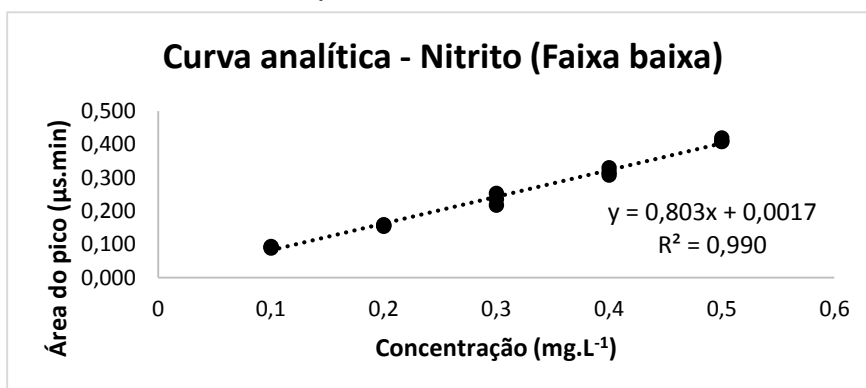


Figura 5.5: Curva analítica para o íon nitrito.

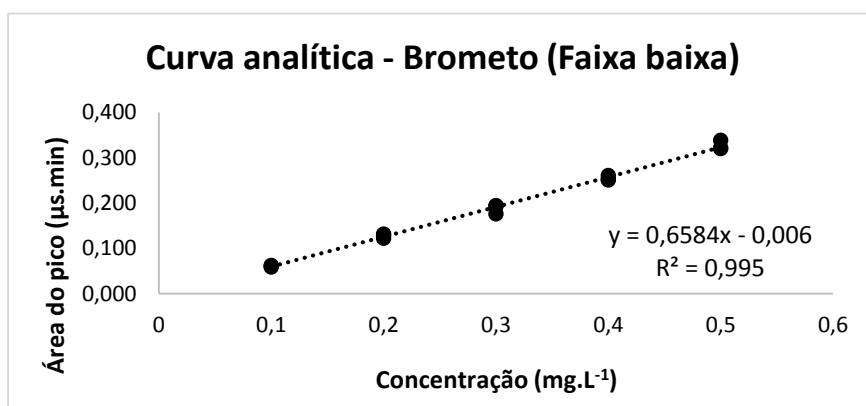


Figura 5.6: Curva analítica para o íon brometo.

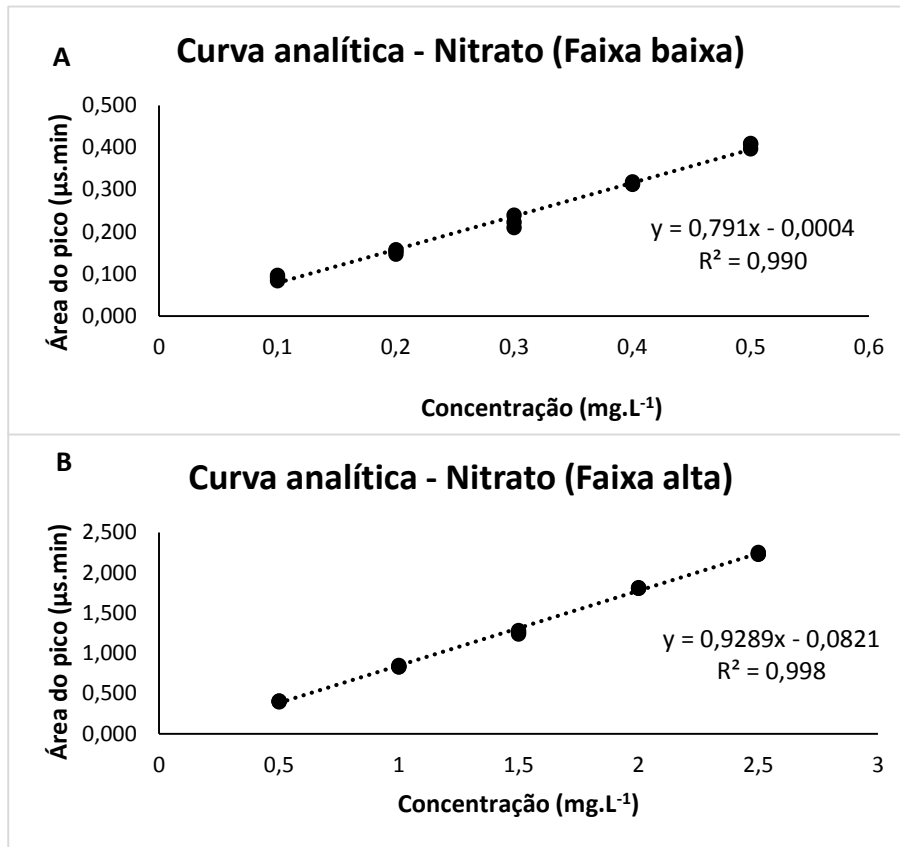


Figura 5.7: Curvas analíticas para o íon nitrato. A – Faixa baixa. B – Faixa alta

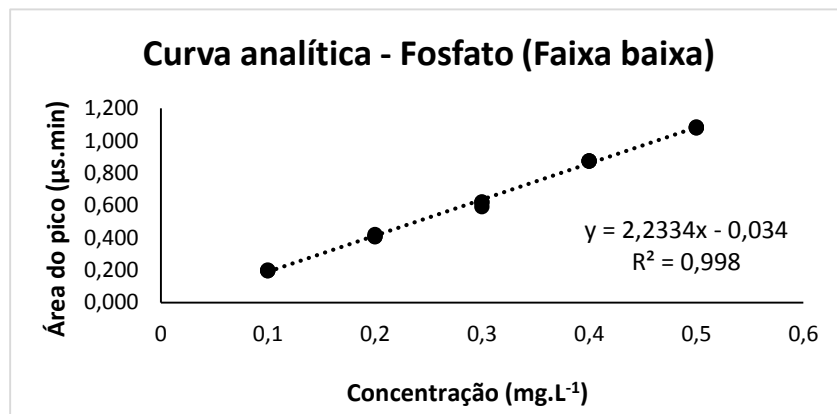


Figura 5.8: Curva analítica para o íon fosfato.

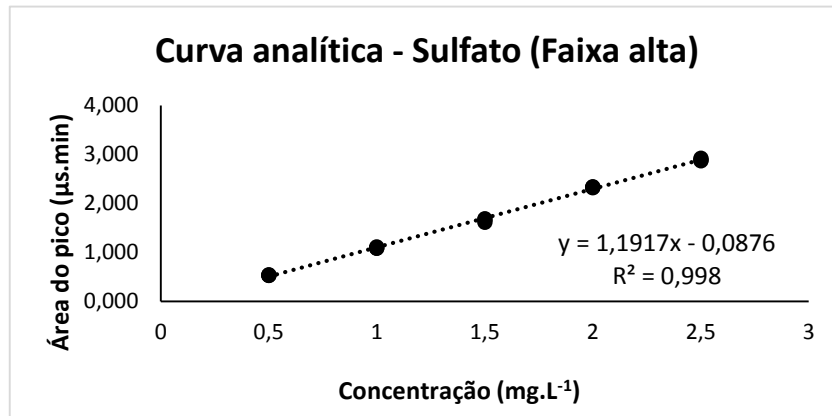


Figura 5.9: Curva analítica para o íon sulfato.

Com base no valor de R^2 , o qual se apresenta acima de 0,977, ou seja, bem próximo a 1, constata-se que há uma correlação forte entre as variáveis estudadas.

As Figuras 5.10, 5.11, 5.12, 5.13, 5.14 e 5.15 mostram os gráficos de resíduos construído com base nos dados adquiridos para os íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato, respectivamente.

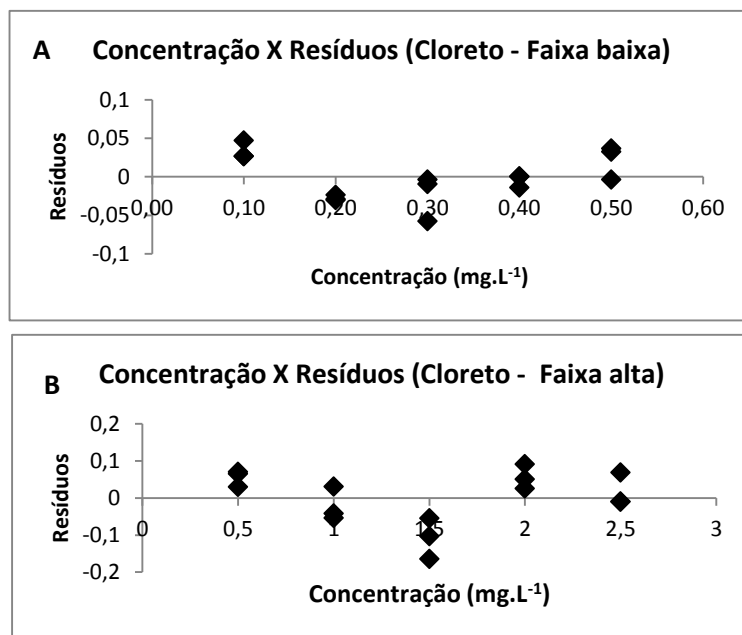


Figura 5.10: Gráfico de resíduos construído para as curvas analíticas do íon cloreto. A – faixa baixa. B – faixa alta.

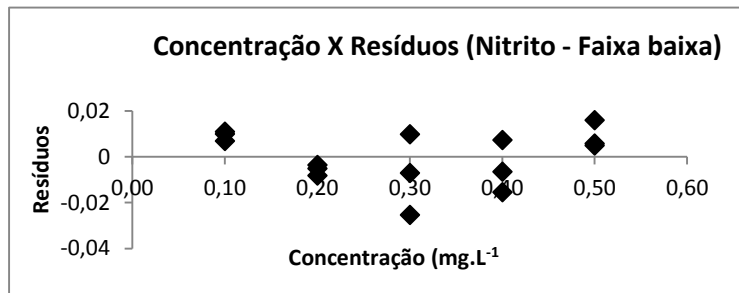


Figura 5.11: Gráfico de resíduos construído para as curvas analíticas do íon nitrito.

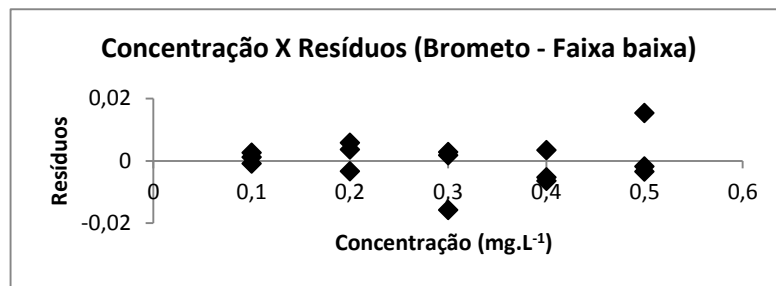


Figura 5.12: Gráfico de resíduos construído para as curvas analíticas do íon brometo.

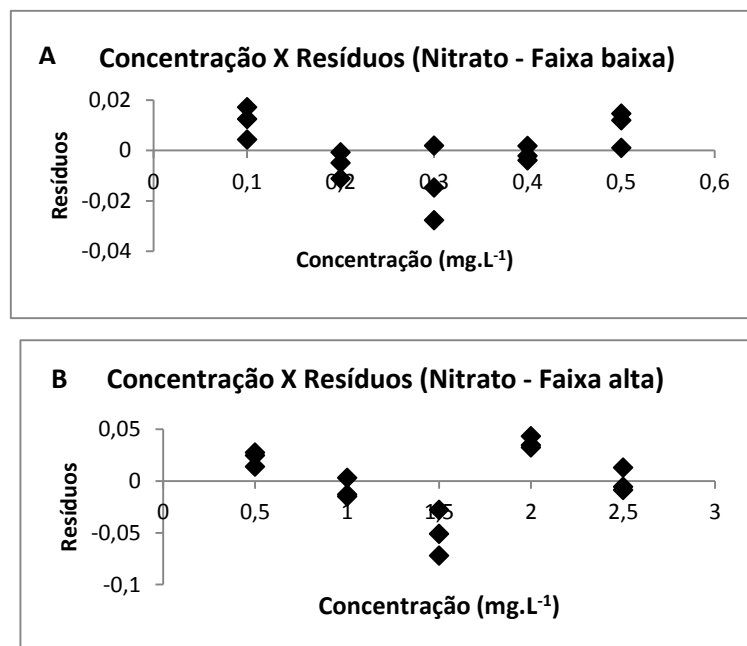


Figura 5.13: Gráfico de resíduos construído para as curvas analíticas do íon nitrato. A – Faixa baixa. B – Faixa alta

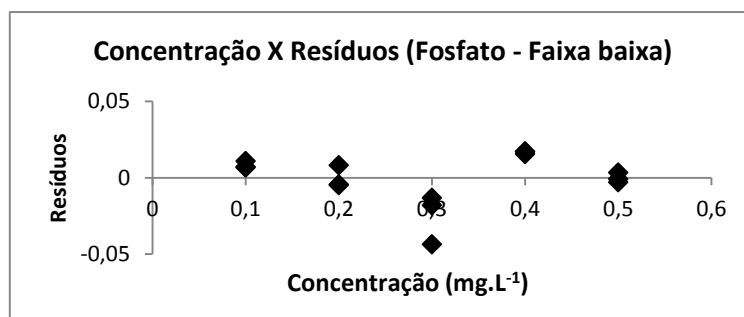


Figura 5.14: Gráfico de resíduos construído para as curvas analíticas do íon fosfato.

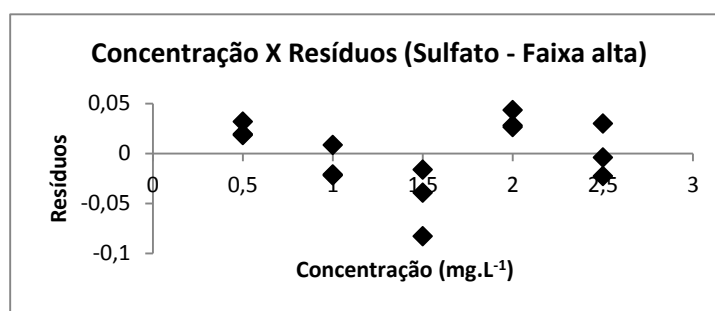


Figura 5.15: Gráfico de resíduos construído para as curvas analíticas do íon sulfato.

Ao observar os gráficos de resíduos, pode-se concluir que, algumas curvas apresentaram pontos que se distanciaram um pouco de suas replicatas (brometo faixa baixa e sulfato faixa alta) ou/e indicaram um pequeno aumento da dispersão dos dados com o aumento da concentração (nitrito faixa baixa e nitrato faixa alta) ou/e mostraram uma pequena curvatura dos resíduos (cloreto, nitrito e nitrato para a faixa baixa).

Apesar dos gráficos de resíduos indicarem uma pequena não conformidade dos dados, os testes estatísticos apontam que as curvas podem ser utilizadas.

As equações das retas construídas para todos os sete íons (fluoreto, cloreto, nitrito, brometo, nitrato, fosfato e sulfato) e seus respectivos R^2 encontram-se no Quadro 5.9.

Quadro 5.9: Resumo das equações das retas para os íons fluoreto, cloreto, nitrito, brometo, nitrato, fosfato e sulfato e seus respectivos R²

Íon	Curva analítica (Faixa baixa)		Curva analítica (Faixa alta)	
	Equação da reta	R ²	Equação da reta	R ²
Fluoreto	$y = 2,6738x - 0,0125$	0,9465	-	-
Cloreto	$y = 1,3016x + 0,0985$	0,9769	$y = 1,8164x - 0,1923$	0,9971
Nitrito	$y = 0,803x + 0,0017$	0,9907	-	-
Brometo	$y = 0,6584x - 0,006$	0,9950	-	-
Nitrato	$y = 0,791x - 0,0004$	0,9895	$y = 0,9289x - 0,0821$	0,9977
Fosfato	$y = 2,2334x - 0,034$	0,9976	-	-
Sulfato	-	-	$y = 1,1917x - 0,0876$	0,9985

5.2. Validação parcial

No decorrer da validação parcial, foi verificado se a qualidade dos dados obtidos pelo método está de acordo com a exigida através da avaliação de alguns parâmetros. Os parâmetros analisados foram linearidade, limites detecção e quantificação por duas abordagens e repetibilidade.

5.2.1. Linearidade

Como foi mencionado anteriormente, a linearidade de um método pode ser avaliada através da análise visual da curva analítica construída após a aplicação dos testes estatísticos e pela verificação do valor do desvio padrão do coeficiente angular.

A observação das curvas analíticas indica que os dados, para todos os íons em questão, têm relação linear aparente e a proposta de descrevê-los através de um modelo linear seria válida.

Para a outra abordagem, o Quadro 5.10 apresenta os valores do desvio padrão relativo do coeficiente angular com os respectivos parâmetros usados para obtê-lo para cada íon.

Quadro 5.10: Desvio padrão relativo do coeficiente angular para cada íon

	F ⁻	Cl ⁻	NO ₂ ⁻	Br ⁻	NO ₃ ⁻¹	PO ₄ ³⁻	SO ₄ ²⁻
S _{y/x}	0,097	0,030	0,012	0,007	0,012	0,017	0,035
s _b	0,176	0,055	0,022	0,013	0,023	0,030	0,007
b	2,674	1,302	0,803	0,658	0,791	2,233	1,192
DPR _{in}	0,066	0,043	0,027	0,020	0,029	0,014	0,005
DPR _{in} *100	6,6	4,3	2,7	2,0	2,9	1,4	0,5

Ao analisar o Quadro 5.10, observa-se que, para maioria dos íons, o valor de RSD está abaixo do valor máximo permitido, indicando uma baixa dispersão dos dados em torno da linha de tendência dos dados.

Apenas o íon fluoreto apresentou um desvio padrão relativo de seu coeficiente ligeiramente maior que o valor estabelecido, o que não impede o uso da curva analítica mesmo que de forma cautelosa.

5.2.2. Limites de detecção (LD) e quantificação (LQ)

Os limites de detecção e quantificação foram avaliados por meio de duas abordagens: relação sinal/ruído e desvio padrão e inclinação da reta.

Em ambas as abordagens é necessário a análise do branco, o Quadro 5.11, então, apresenta os valores obtidos para as áreas dos picos relativas à análise de três replicatas do branco (água ultrapurificada Mili Q).

Quadro 5.11: Valores obtidos para as áreas dos picos para a análise de três replicatas do branco

ÍON	Área (µS.min)		
	Branco 1	Branco 2	Branco 3
Fluoreto	6,08E-03	5,32E-03	8,35E-03
Cloreto	8,19E-03	8,92E-03	1,10E-02
Nitrito	1,76E-03	1,99E-03	2,42E-03
Brometo	6,41E-05	4,42E-04	6,88E-04
Nitrato	1,76E-03	1,99E-03	2,42E-03
Fosfato	9,94E-04	2,55E-03	2,56E-03
Sulfato	1,08E-02	1,17E-02	1,08E-02

No método da relação sinal/ruído o valor de sinal esperado para o limite de quantificação é resultado da multiplicação do sinal apresentado pelo branco por dez.

Desse modo, ao comparar o sinal obtido por sucessivas diluições procurou-se a solução que apresentou um sinal próximo do valor de sinal esperado.

As áreas esperadas para o limite de quantificação, as áreas observadas em que, após sucessivas diluições, foram encontradas a razão sinal/ruído de aproximadamente 10:1, suas respectivas concentrações (LQ) e os limites de detecção obtidos através da Equação 21 são mostrados no Quadro 5.12.

Quadro 5.12: Limites de quantificação e limites de detecção obtidos pela relação sinal/ruído

	Área esperada – 10*sinal branco ($\mu\text{S}\cdot\text{min}$)	Área observada ($\mu\text{S}\cdot\text{min}$) - LQ	LQ ($\text{mg}\cdot\text{L}^{-1}$)	LD ($\text{mg}\cdot\text{L}^{-1}$)
Fluoreto	0,066	0,073	0,03	0,009
Cloreto	0,094	0,100	0,05	0,018
Nitrito	0,021	0,023	0,03	0,01
Brometo	0,004	0,003	0,01	0,002
Nitrato	0,021	0,018	0,02	0,008
Fosfato	0,020	0,022	0,05	0,02
Sulfato	0,111	0,113	0,09	0,03

Os parâmetros utilizados para o cálculo dos limites pela abordagem da regressão linear e os valores resultantes são mostrados no Quadro 5.13:

Quadro 5.13: Limites de detecção e quantificação decorrentes do método do desvio padrão do branco e inclinação da reta

Íon	Desvio padrão (brancos)	Coefficiente angular	LQ ($\text{mg}\cdot\text{L}^{-1}$)	LD ($\text{mg}\cdot\text{L}^{-1}$)
Fluoreto	1,58E-03	2,674	0,006	0,002
Cloreto ^{*1}	1,47E-03	1,302	0,011	0,004
Nitrito	3,37E-04	0,803	0,004	0,001
Brometo	3,14E-04	0,658	0,005	0,002
Nitrato ^{*1}	3,37E-04	0,791	0,004	0,001
Fosfato	9,00E-04	2,233	0,004	0,001
Sulfato ^{*2}	5,34E-04	1,191	0,004	0,001

*1 Curvas analíticas utilizadas: 0,10 – 0,50 $\text{mg}\cdot\text{L}^{-1}$ (faixa baixa)

*2 Curva analítica utilizada: 0,50 – 2,50 $\text{mg}\cdot\text{L}^{-1}$ (faixa alta)

O método da regressão linear apresentou valores mais baixos do que os obtidos para a relação sinal/ruído. Entretanto, ao realizar as análises no

cromatógrafo das soluções com as concentrações dos limites de quantificação indicados (0,004 mg.L⁻¹, 0,005 mg.L⁻¹, 0,006 mg.L⁻¹ e 0,01 mg.L⁻¹), as áreas obtidas apresentaram um valor muito diferente ao previsto pela equação da reta de cada íon, indicando que esse limites não são confiáveis.

Portanto, optou-se por adotar os limites obtidos pela abordagem da relação sinal/ruído.

5.2.3. Repetibilidade

Os desvios padrões e as concentrações médias calculados para os dados das soluções padrões com concentração 0,10 mg.L⁻¹, 0,30 mg.L⁻¹ e 0,50 mg.L⁻¹ de todos os íons em estudo encontram-se no Apêndice D (página 98). A partir desses resultados foi possível calcular os desvios padrões relativos, os quais são representados no Quadro 5.14.

Quadro 5.14: Desvios padrões relativos calculados para as soluções padrões com 0,10 mg.L⁻¹, 0,30 mg.L⁻¹ e 0,50 mg.L⁻¹ dos íons em estudo

Concentração (mg.L ⁻¹)	%Desvio padrão relativo						
	Fluoreto	Cloreto	Nitrito	Brometo	Nitrato	Fosfato	Sulfato
0,10	2,8	4,5	2,3	2,9	7,2	9,9	1,3
0,30	4,3	6,4	7,5	5,6	6,6	2,9	6,3
0,50	10,3	2,9	1,5	3,2	1,8	1,1	1,4

Ao confrontar os valores de %DPR_{rep} com o valor máximo de aceitação para %DPR_{rep} sugerido pela AOAC (analito presente em 0,00001% - %DPR_{rep}<15%), verifica-se que, para todos os desvios padrões relativos, %DPR_{rep}<15%. Com isso, as condições de repetibilidade das medições são consideradas adequadas.

5.2.4. Recuperação analítica

A recuperação analítica foi calculada através da comparação entre o valor observado para a área do pico de cada íon nas concentrações 0,10 mg.L⁻¹, 0,30 mg.L⁻¹ e 0,50 mg.L⁻¹ e o valor esperado pela equação da reta para as concentrações selecionadas.

Quadro 5.15: Recuperação analítica calculada para as soluções padrão com 0,10 mg.L⁻¹, 0,30 mg.L⁻¹ e 0,50 mg.L⁻¹ dos íons em estudo

Replicata	Concentração (mg.L ⁻¹)	Fluoreto	Cloreto	Nitrito	Brometo	Nitrato	Fosfato	Sulfato
n=1	0,10	109	112*	108	99	122*	104	90
n=2	0,10	113*	112*	112*	102	116*	104	93
n=3	0,10	116*	121*	113*	104	106	106	92
n=1	0,30	88	88	90	92	88	93	86
n=2	0,30	95	98	97	101	94	97	94
n=3	0,30	95	99	104	102	101	98	97
n=1	0,50	93	100	101	99	100	100	98
n=2	0,50	98	104	101	99	103	100	98
n=3	0,50	113*	105	104	105	104	100	101

* Valores fora da faixa sugerida como permitida

Ao confrontar os valores de recuperação analítica com o valor máximo de aceitação para esse parâmetro sugerido pela AOAC (analito presente em 0,00001%, 80% - 110%), verifica-se que, para a maioria dos íons os valores calculados ficaram na faixa.

Ao avaliar os resultados que ficaram fora da faixa, percebe-se que todos ficaram próximos ao valor máximo e que não foram todas as replicatas para o íon em questão que se apresentaram inadequadas, com isso, os resultados obtidos para o parâmetro recuperação analítica foram considerados satisfatórios.

Os parâmetros adotados para a realização da validação parcial do método de análise ânions em soluções aquosas pela técnica de cromatografia de íons indicaram, de uma forma geral, que o método atende aos critérios exigidos, indicando, portanto, que a validação parcial foi bem sucedida.

5.3. Análise das águas minerais

As concentrações calculadas para os íons (fluoreto, cloreto, nitrito, nitrato e sulfato), através das equações das retas propostas para as curvas analíticas para todas as 16 águas minerais, são apresentadas no Quadro 5.16, o qual apresenta cada marca da água mineral com um respectivo código de 1 à 16.

Quadro 5.16: Concentrações calculadas para os íons fluoreto, cloreto, nitrito, nitrato e sulfato para todas as 16 águas minerais

		Concentração (mg.L ⁻¹)				
		FLUORETO	CLORETO	NITRITO	NITRATO	SULFATO
São Lourenço	1	0,14	1,51	0,84	2,32	3,29
	1	0,14	1,49	0,80	2,31	3,29
	1	0,14	1,53	0,83	2,39	3,36
Bonafont	2	0,12	0,27	0,17	1,68	0,67
	2	0,12	0,28	0,17	1,71	0,68
	2	0,12	0,28	0,17	1,66	0,67
Nestlè	3	0,04	1,90	0,08	1,15	1,54
	3	0,05	1,92	0,08	1,13	1,56
	3	0,05	1,89	0,08	1,14	1,54
Pouso Alto	4	0,03	0,08	0,19	0,22	2,69
	4	0,03	0,08	0,17	0,22	2,69
	4	0,03	0,08	0,18	0,21	2,71
Sferrie	5	0,30	0,22	0,04	0,32	0,54
	5	0,31	0,21	0,04	0,33	0,54
	5	0,33	0,20	0,04	0,32	0,53
Minalba Premium	6	0,06	0,08	0,11	0,53	0,19
	6	0,06	0,08	0,11	0,53	0,19
	6	0,06	0,07	0,11	0,53	0,19
Petropolis	7	0,05	1,78	0,03	1,07	1,51
	7	0,05	1,78	0,03	1,07	1,51
	7	0,05	1,78	0,03	1,07	1,51
Qualitá	8	0,11	1,28	0,30	0,05	7,24
	8	0,13	1,28	0,29	0,05	7,20
	8	0,11	1,30	0,30	0,05	7,30
Canção Nova	9	0,10	0,41	0,03	0,11	0,24
	9	0,15	2,43	0,03	0,21	1,28
	9	0,16	2,41	0,03	0,12	1,26
Pouso Alto (gás)	10	0,04	0,30	0,03	0,23	2,39
	10	0,04	0,26	0,03	0,22	2,39
	10	0,04	0,27	0,03	0,23	2,78
Fastio	11	0,03	4,19	0,05	1,63	0,85
	11	0,03	4,25	0,05	1,62	0,85
	11	0,03	4,23	0,05	1,63	0,85
Luso	12	0,06	8,66	0,05	1,54	1,48
	12	0,06	8,71	0,04	1,58	1,48
	12	0,06	8,83	0,05	1,59	1,48
Evian	13	0,03	9,65	0,81	4,06	13,76
	13	0,03	9,45	0,82	3,95	13,46
	13	0,03	9,77	0,86	4,08	13,94
Acqua Panna	14	0,03	7,84	0,03	3,87	22,25
	14	0,03	7,55	0,03	3,83	22,09
	14	0,03	7,63	0,03	3,96	22,42
S.Pellegrino	15	0,66	52,15	0,03	2,95	398,76
	15	0,58	51,94	0,03	3,02	399,53
	15	0,76	51,26	0,03	3,02	442,77
Perrier	16	0,03	20,50	0,03	8,89	26,00
	16	0,03	20,57	0,03	8,92	27,02
	16	0,03	20,54	0,03	8,76	27,08

As águas que apresentaram concentrações abaixo do limite de quantificação, tiveram os valores substituídos por seus respectivos limites de quantificação (destacados em vermelho no Quadro 5.16).

Os íons brometo e fosfato se apresentaram, na maioria das amostras de águas minerais, com teores abaixo dos limites de quantificação e, com isso, não foram utilizados na aplicação do PCA.

Com a finalidade de exemplificar a comparação entre as concentrações obtidas pela análise e as concentrações indicadas nos rótulos, a Figura 5.16 demonstra, para três marcas de águas minerais brasileiras (São Lourenço, Bonafont e Sferriê), os teores de sulfato, nitrato, cloreto e fluoreto apresentados na análise e no rótulo.

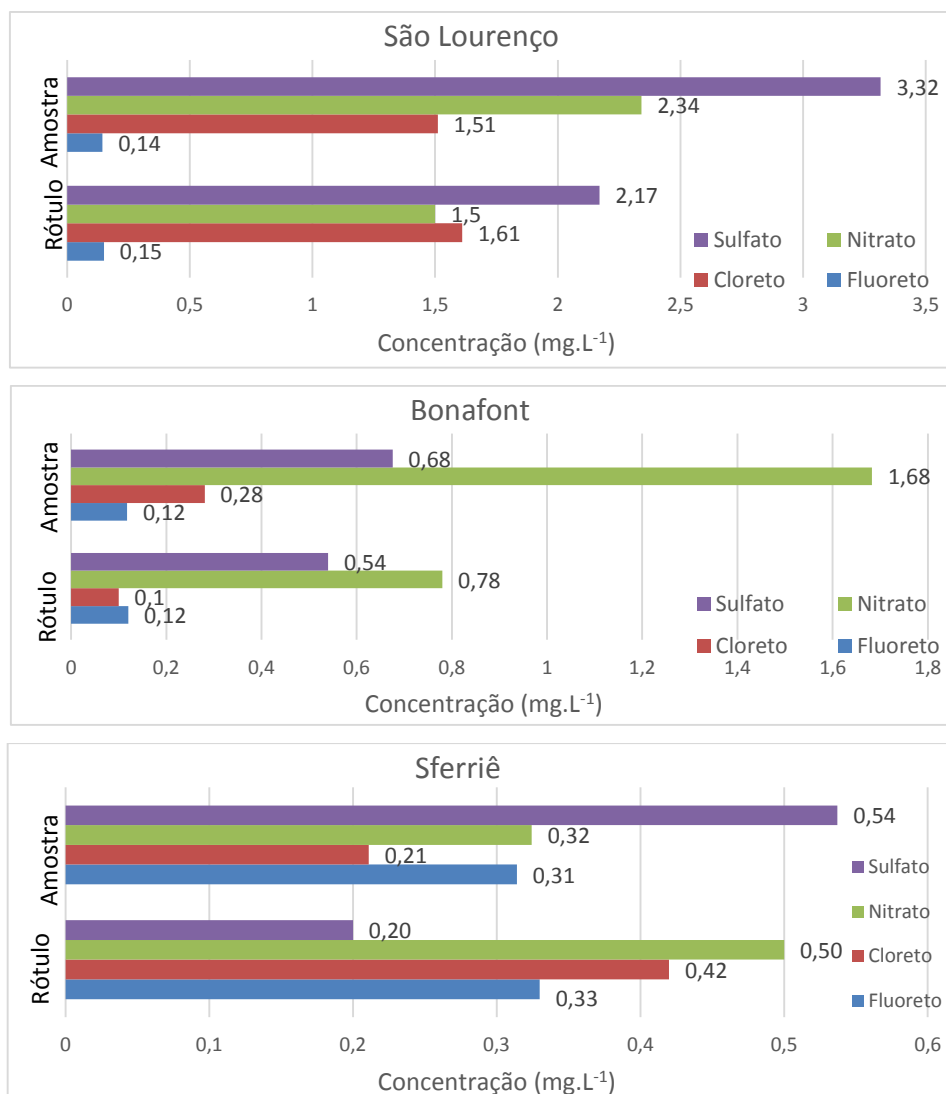


Figura 5. 16: Teores de sulfato, nitrato, cloreto e fluoreto apresentados na análise e no rótulo para três águas minerais brasileiras

Ao analisar a Figura 5.16 e considerando o método adequado para determinação das espécies em estudo, verifica-se que os teores de alguns íons (SO_4^{2-} e NO_3^-) obtidos pelas curvas não foram concordantes, nas três águas, com o indicado no rótulo, indicando uma menor rigidez no controle dos teores pelas empresas responsáveis pela fabricação ou a presença de um fator externo/interno que resulte no aumento da concentração desses íons.

Com a finalidade de realizar a mesma averiguação para as águas minerais estrangeiras das marcas Luso, Evian e Acqua Panna, foram construídos gráficos, os quais são representados na Figura 5.17, comparando os resultados obtidos na análise e os rótulos.

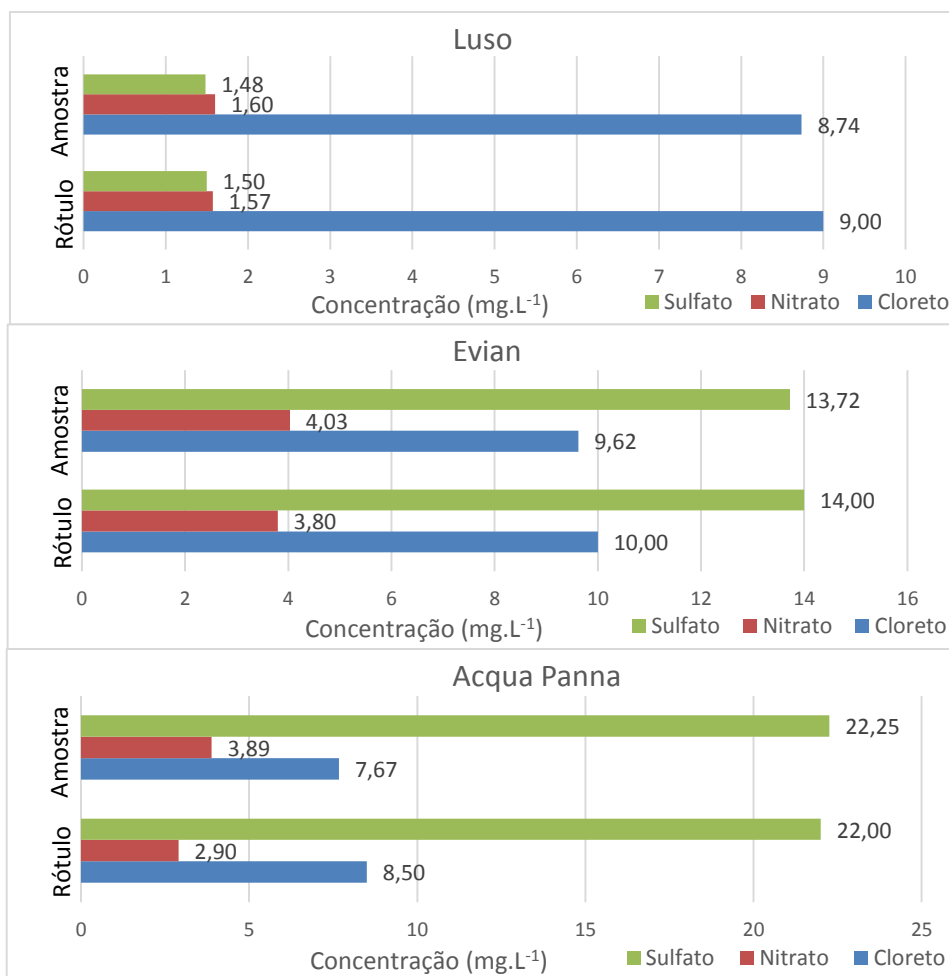


Figura 5.17: Teores de sulfato, nitrato e cloreto apresentados na análise e no rótulo para três águas minerais estrangeiras

Os resultados apresentados para as amostras na Figura 5.17, os quais foram calculados da mesma maneira através das curvas analíticas que nas águas brasileiras, indicam, ao ser comparados, com os valores no rótulo, uma boa concordância.

Assim sendo, as empresas brasileiras responsáveis pela fabricação das três águas citadas, apresentaram um menor controle na qualidade do de seus produtos em relação as empresas estrangeiras responsáveis pelas marcas Luso, Evian e Acqua Panna.

5.4. Aplicação da técnica quimiométrica

Os resultados atingidos de concentrações calculadas pelas regressões foram utilizados como variáveis para discriminação das amostras por PCA. Os comandos

utilizados na plataforma GNU Octave para obter os resultados da aplicação da ACP estão descritos e detalhados no Apêndice E (página 106) na forma de um roteiro. Os resultados das 16 águas minerais gerados pela, serão mostrados de acordo com a ordem empregada nesse anexo.

- ✓ Construção da matriz de dados X ($n \times p$);

Os dados organizados, com as amostras nas linhas e as concentrações de cada íon nas colunas e com códigos representando cada amostra de água (número de 1 à 16) em uma coluna anterior aos resultados, foram copiados no programa e carregados para o ambiente de trabalho do software.

- ✓ Pré-tratamento dos dados;

A matriz dos dados foi dividida em duas partes, códigos e resultados, e os dados foram pré-tratados com o método de autoescalamiento. usando um comando que equivale a Equação 28 (tópico 3.5.1.2. – Pré-tratamento de dados).

- ✓ Determinação dos autovetores, autovalores e *scores*;

O pacote de estatística, o qual contém a função responsável pelo PCA (*princomp*), foi carregado e função *princomp* foi, então, aplicada aos dados.

A partir dessa função foram obtidos os valores de autovetores (Quadro 5.17), que indicam as direções dos novos eixos representados por cada componente, e autovalores (Quadro 5.18), os quais representam a variância explicada por cada componente.

Quadro 5.17: Autovetores relacionados com a matriz de covariância dos dados pré-tratados

Variável	Autovetor				
	Componente 1	Componente 2	Componente 3	Componente 4	Componente 5
Fluoreto	-0.522	0.259	-0.322	-0.742	0.082
Cloreto	-0.578	-0.156	0.098	0.387	0.694
Nitrito	0.105	-0.586	-0.800	0.072	0.028
Nitrato	-0.206	-0.747	0.481	-0.348	-0.214
Sulfato	-0.582	0.082	-0.124	0.417	-0.682

Quadro 5.18: Autovalores relacionados com a matriz de covariância dos dados pré-tratados

	Autovalores
Componente 1	2.806
Componente 2	1.185
Componente 3	0.878
Componente 4	0.119
Componente 5	0.012

Os scores relacionados com cada componente principal estão representados no Apêndice F (página 111).

- ✓ Contribuição de cada componente principal;

A partir dos autovalores calculados foi possível calcular a contribuição de cada componente principal para a explicação da variância total, a qual consiste na normalização em relação à soma dos autovalores.

Para calcular essa contribuição, pode ser explorado com os alunos o conceito dos autovalores e mostrar como sua soma representa a variância total do experimento. Os valores obtidos para cada contribuição estão descritos no Quadro 5.19.

Quadro 5.19: Contribuição de cada componente principal para a explicação da variância total

	Contribuição (%)
PC1	56.116
PC2	23.705
PC3	17.555
PC4	23.823
PC5	0.24081

As percentagens cumulativas ao adicionar cada componente principal no estudo dos dados originais são exibidas no Quadro 5.20.

Quadro 5.20: Percentagens cumulativas ao adicionar cada componente principal na explicação da variação dos dados originais

	% Cumulativa
PC1	56.1
PC2	79.8
PC3	97.4
PC4	99.7
PC5	100.0

Um gráfico para %cumulativa ao adicionar cada componente principal na explicação é retratada na Figura 5.18.

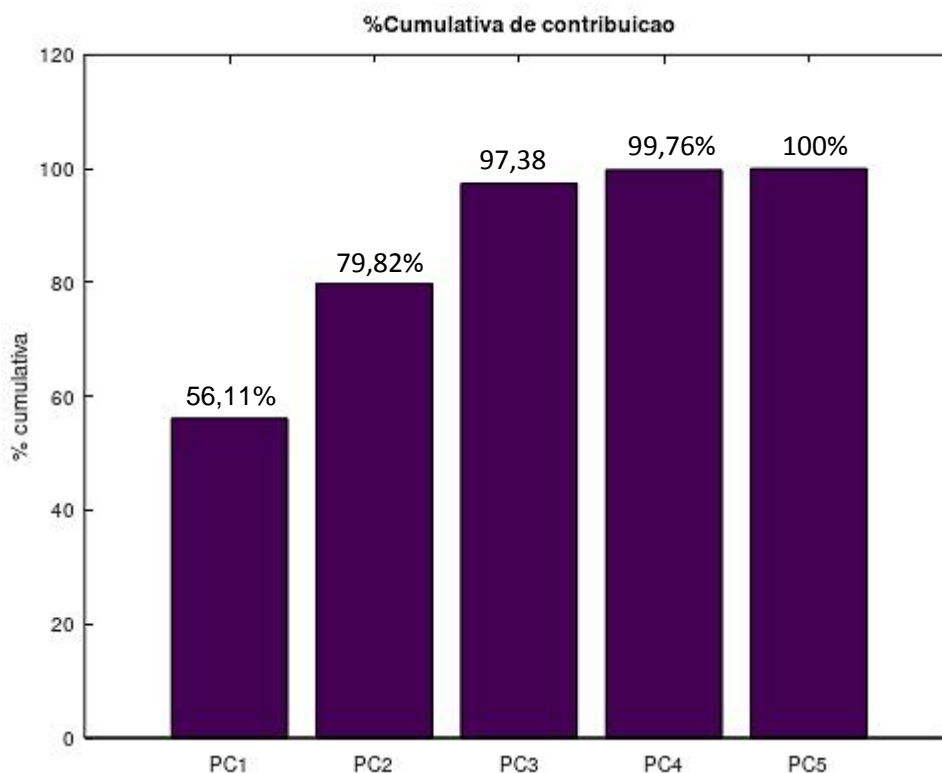


Figura 5.18: Percentagem cumulativa da contribuição dos componentes principais

A primeira componente, de acordo com a Figura 5.18, contém mais de 50% da variação, contendo, assim, mais informação sobre os dados originais do que as outras componentes principais e as variáveis originais.

Verifica-se, também, que, ao usar apenas as componentes PC1 e PC2, aproximadamente 80% da estrutura dos dados originais é explicada. A avaliação desses dois componentes é considerada, então, suficiente para uma análise de componentes principais sem uma perda significativa de informação.

Dessa forma, PC3, PC4 e PC5 não contribuem significativamente e podem ser descartadas, sendo possível, assim, reduzir de 5 dimensões (concentração dos cinco íons) para 2 (PC1 e PC2).

✓ Construção do gráfico dos scores;

As projeções dos dados nos novos eixos (PC1 e PC2) é representada pelos scores. Seus valores, no plano PC1 x PC2, são mostrados na Figura 5.19.

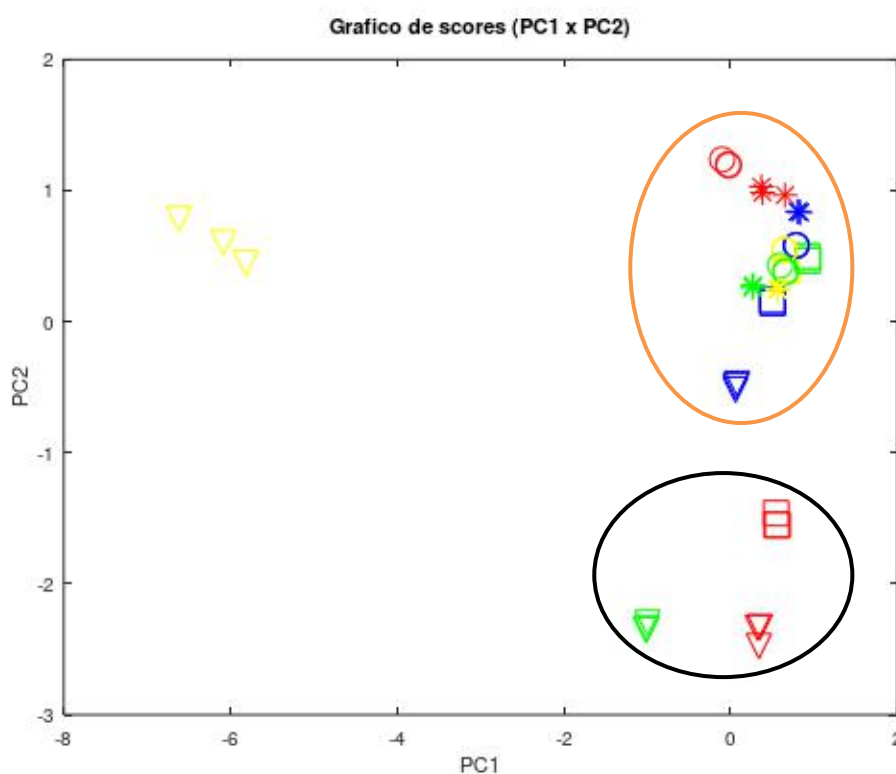


Figura 5.19: Gráfico de scores (PC1 x PC2) *

*Legenda

□ São Lourenço	○ Sferriê	* Canção Nova (gás)	▽ Evian
□ Bonafont	○ Minalba Premium	* Pouso alto (gás)	▽ Acqua Panna
□ Nestlé	○ Petrópolis	* Fastio	▽ S. Pellegrino (gás)
□ Pouso alto	○ Qualitá	* Luso	▽ Perrier (gás)

Ao observar a Figura 5.19, pode-se observar que apenas uma água (da marca S.Pellegrino) se distanciou expressivamente das outras no eixo da componente PC1. Ao avaliar os dados originais do rótulo (Anexo L), verifica-se que a água mineral em questão apresentou, em comparação as outras marcas, o maior teor de sulfato com uma concentração em torno de 410 mg.L⁻¹.

É importante que, nesse tipo de análise, seja discutido com os alunos a influência de cada variável original nas novas variáveis. A PC1 foi capaz de distanciar a água S.Pellegrino das outras. Uma vez que essa água tem um alto teor de sulfato, provavelmente a influência desta variável original é muito grande para a PC1. Isso auxilia como recurso didático para que os alunos entendam os diferentes tipos de análises possíveis que podem ser conseguidas a partir de uma ACP.

Em relação à separação no eixo da PC2, dois grupos de águas minerais são constatados. O grupo com três águas minerais (elipse preta) com as marcas São Lourenço, Evian e Perrier têm como diferencial as concentrações de nitrito e nitrato.

São Lourenço e Evian apresentam as maiores concentrações de nitrito e Perrier a maior concentração de nitrato. Novamente, uma correlação das variáveis originais através de uma análise dos scores pode ser feita, indicando possíveis influências nos novos eixos.

✓ *Determinação dos loadings;*

Tendo discutido a análise das amostras pelos scores, pode-se introduzir a relação das variáveis originais às novas variáveis. O peso de cada variável original nas duas componentes principais (PC1 e PC2), ou seja, os *loadings* encontrados são mostrados no Quadro 5.21.

Quadro 5.21: *Loadings* nas duas componentes principais (PC1 e PC2)

Variável original	PC1	PC2
Fluoreto	-0.87411	0.28162
Cloreto	-0.969	-0.17033
Nitrito	0.17642	-0.63763
Nitrato	-0.34564	-0.81382
Sulfato	-0.9758	0.089851

Ao observar apenas a magnitude dos valores dos *loadings*, ignorando os sinais, verifica-se que as concentrações de fluoreto, cloreto e sulfato tem os maiores pesos no cálculo para a primeira componente principal. Já, as concentrações de nitrito e nitrato, são as variáveis que mais influenciaram para a componente PC2.

O gráfico dos *loadings* no plano PC1 x PC2 encontra-se na Figura 5.20.

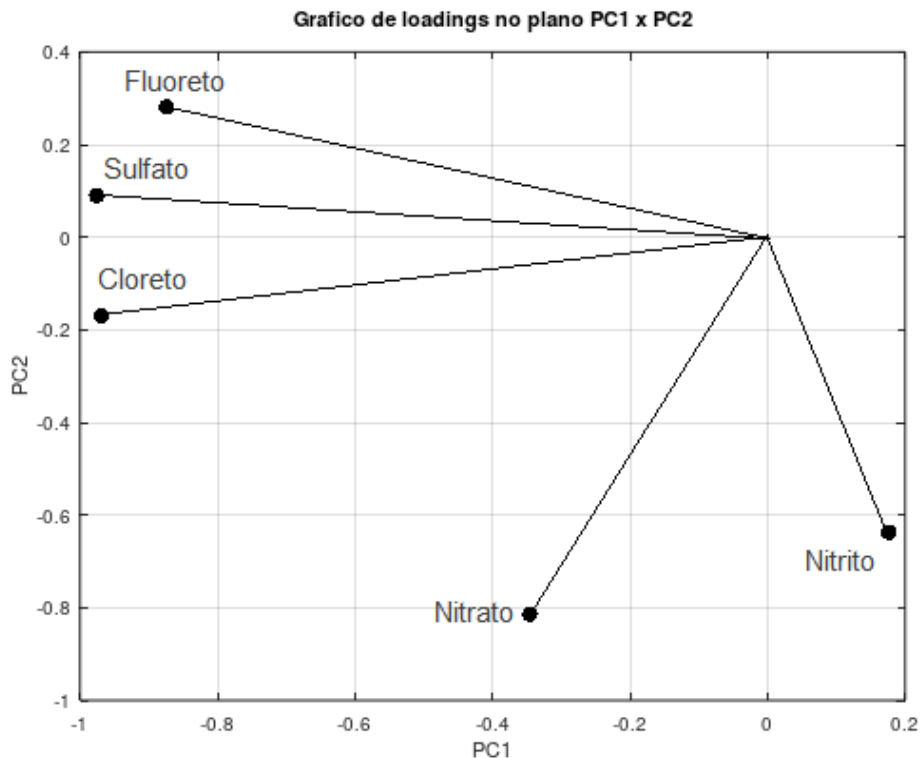


Figura 5.20: Gráfico dos *loadings* no plano PC1 x PC2

Ao examinar os ângulos entre os vetores criados pelo *loading* de cada variável, é possível inferir que as concentrações de fluoreto, cloreto e sulfato são correlacionadas significativamente entre si, diferentemente dos outros íons (nitrito e nitrato).

Dessa forma é possível confirmar as observações feitas para o gráfico de *scores* com relação as variáveis originais que mais influenciaram nas distinções observadas em cada componente principal. Isso permite aos alunos entenderem conceitualmente a aplicação de cada análise, de modo que a discussão sobre estas naturaliza sua interpretação e aceitação.

5.5. Elaboração do roteiro de aula e proposta de implementação no curso

Com base no procedimento realizado para a análise de 16 águas minerais de marcas diferentes com a técnica de cromatografia de íons foi proposto um roteiro para a aula de Análise Instrumental I, do Instituto de Química da UFRJ.

A realização da prática proposta é viável considerando um tempo de aula com duração de 4 horas, incluindo as injeções das soluções padrões (1,25 h), das

amostras (0,75 h), e da explicação da técnica utilizada pelo equipamento e da ferramenta quimiométrica PCA e sem contar com a realização de replicatas.

Determinação de ânions utilizando a técnica de cromatografia de íons com detecção condutimétrica

01 – Fundamentos (adaptado do roteiro atual)

A cromatografia de íons (CI) é uma técnica que engloba todos métodos cromatográficos que separam substâncias iônicas e substâncias que se ionizam facilmente na fase móvel. Estes métodos são a cromatografia de par iônico, a cromatografia de troca iônica e a cromatografia por exclusão iônica. Em princípio, tanto fase móvel como os analitos são substâncias polares e/ou iônicas. A configuração básica de um cromatógrafo de íons é semelhante a um sistema de cromatografia de alta eficiência (HPLC) convencional: uma bomba conduz a fase móvel através de uma válvula de injeção. Os volumes típicos de injeção variam de 5 à 100 µL.

O coração de um sistema cromatográfico é a coluna de separação. A seleção da fase estacionária adequada e das condições cromatográficas determina a qualidade da análise. Em geral, a separação é obtida à temperatura ambiente. Em poucos casos, tais como nas determinações de ácidos graxos de cadeia longa são utilizadas maiores temperaturas. Os analitos são detectados e quantificados por um sistema de detecção. O desempenho do detector é verificado de acordo com os seguintes critérios: sensibilidade, linearidade, resolução e limite de detecção. O detector comumente empregado em cromatografia iônica é o de condutividade, que é usado com ou sem um sistema supressor de condutividade de fase móvel. Além desse, são usados detectores UV/Vis, amperométrico e de fluorescência.

A CI é baseada em uma interação entre os íons de uma solução e uma substância sólida contendo os grupos funcionais, que podem fixar íons como resultado de forças eletrostáticas. No caso mais simples, na cromatografia de cátions, são grupos de ácido sulfônico; enquanto na cromatografia de ânions, são grupos de amônio quaternário. Teoricamente, íons com a mesma carga podem ser completa e reversivelmente trocados entre duas fases. O processo de troca iônica leva a uma condição de equilíbrio. O deslocamento de equilíbrio depende da

afinidade dos íons participantes em relação aos grupos funcionais da fase estacionária.

02 – Aparelhagem

- Cromatógrafo de íons
- Coluna de separação de ânions
- Alça de amostragem: 200 μL

03 - Reagentes e soluções:

- ✓ Eluente: Na_2CO_3 25 mmol.L^{-1}
- ✓ Solução de padrão misto contendo:

Ânion	Concentração (mg.L^{-1})
Fluoreto	20
Cloreto	30
Nitrito	100
Nitrato	100
Sulfato	150

- ✓ Amostras: águas minerais brasileiras e estrangeiras

Preparo dos padrões para a cada curva analítica:

Diluir para balão de 10,00 mL os seguintes volumes da solução padrão misto: 40 μL , 60 μL , 80 μL e 100 μL . Completar o volume com água ultrapurificada do tipo I Milli Q.

04 - Procedimento:

- ✓ Curva analítica:

Lavar e preencher a alça de amostragem manualmente com cerca de 1 mL dos padrões preparados no equipamento. O tempo de corrida será de aproximadamente 15 min. Integrar os sinais gerados no cromatograma e através do software construir a curva analítica. O procedimento deverá ser consultado anteriormente com o professor da disciplina ou o monitor disponível.

✓ *Amostra de água:*

Lavar e preencher a alça de amostragem manualmente com cerca 1 mL das amostras que serão analisadas.

Observação: Caso os rótulos das águas minerais indiquem a presença dos ânions com uma concentração acima do último ponto da curva, diluir as amostras com o intuito adequar a concentração de todos os íons as faixas de trabalho selecionadas.

05 - Resultados:

✓ *Construir as curvas analíticas para cada ânion analisado. Determinar as concentrações dos ânions estudados nas amostras. Na amostra de água mineral, comparar os resultados obtidos com os fornecidos no rótulo, expressando o erro relativo.*

✓ *Com os resultados para cada água seguir o roteiro para a aplicação da ACP no GNU Octave detalhada no Apêndice E (página 106).*

6. Conclusão

A validação parcial do método utilizado para a análise de ânions na técnica de cromatografia de íons, demonstrou, de uma forma geral, que o método é adequado, apresentando, por exemplo, resultados para os parâmetros linearidade e repetibilidade abaixo do permitido e limites de detecção e quantificação indicando que, a maioria dos íons estudados, são detectados pelo equipamento com valores abaixo dos apresentados nos rótulos das águas minerais.

A partir da ferramenta quimiométrica (ACP) aplicada no GNU Octave 5.1.0, foi possível analisar as concentrações de cinco íons em diferentes marcas de águas minerais apenas por 2 variáveis, PC1 e PC2, as quais explicaram aproximadamente 80% da estrutura dos dados originais.

A interpretação desses dados viabilizou a distinção da água mineral da marca S.Pellegrino devido ao seu alto teor de sulfato em relação ao restante e a distinção das marcas São Lourenço e Evian, por apresentarem as maiores concentrações de nitrito e a Perrier, a maior concentração de nitrato.

Ao longo da realização dos experimentos, tratamento de dados e uso do software foram feitas anotações, as quais auxiliaram, junto com o procedimento atual descrito na apostila da disciplina, na elaboração do roteiro proposto para a aula.

Tal roteiro sugere a aplicação da ACP com uma rotina do GNU Octave para águas minerais brasileiras e estrangeiras como uma forma de iniciar a implementação destas ferramentas no currículo do curso de Química.

7. Referências bibliográficas

ANVISA. **Guia para tratamento estatístico da validação analítica - Medicamentos**. GUIA nº 10, versão 1, de 30 de agosto de 2017.

AOAC International. **Official methods of analysis of AOAC International, in Guidelines for Standard Method Performance Requirements**. Gaithersburg, Apêndice F, 2016.

BERTOLO, R; HIRATA, R.; FERNANDES, A. **Hidrogeoquímica das águas minerais envasadas do Brasil**. Revista Brasileira de Geociências. 37(3): 515-529, 2007.

ANVISA. **Resolução RDC nº 274**, Diário Oficial, 23 de setembro de 2005.

COLLINS, C.H., BRAGA, G.L., BONATO, P.S. **Fundamentos de cromatografia**. Editora da UNICAMP, Campinas - SP, 2006. 452p.

CORREIA, P. R. M.; FERREIRA, M. M. C. **Reconhecimento de padrões por métodos não supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos**. Quim. Nova, Vol. 30, No. 2, 481-487, 2007.

EATON, J. W. **GNU Octave and reproducible research**. Journal of Process Control 22 (2012) 1433– 1438

EITH, C.; KOLB, M.; RUMI, A.; SEUBERT, A.; VIEHWEGER, K. H. **Práticas em Cromatografia de íons**. 2º edição, Metrohm, Herisau – Suíça, 2007.

FREITAS, M. B.; BRILHANTE, O. M.; ALMEIDA, L. M. **Importância da análise de água para a saúde pública em duas regiões do Estado do Rio de Janeiro: enfoque para coliformes fecais, nitrato e alumínio**. Cad. Saúde Pública, Rio de Janeiro, 17(3):651-660, 2001.

GONZÁLEZ, A. G.; HERRADOR, M. A. **A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles**. Trends in Analytical Chemistry, Vol. 26, No. 3, 2007.

GUEDES, L. F. M. **Separação em fase sólida para a determinação de ânions por cromatografia de íons em amostra salinas, ambientais e da indústria de petróleo**. Dissertação de mestrado em Geociências – Universidade Federal Fluminense (UFF), 2010.

GUIZELLINI, A. T.; ROSSI, V. V. **Quimiometria na graduação em Química no Brasil**. In: Encontro Nacional de Ensino de Química, 16.; ENCONTRO DE EDUCAÇÃO QUÍMICA NA BAHIA, 5., 2012, Salvador. Anais eletrônicos.

Disponível

em:

<http://www.portalseer.ufba.br/index.php/anaiseneq2012/article/view/7113>.

Acesso em: 28 agosto 2019.

HARRIS, D. C. **Análise Química Quantitativa**. 8 ed. Rio de Janeiro: LTC, 2012, p. 898

Hongyu, K., Sandanielo, V. L. M., Junior, G. J. O. **Análise de Componentes Principais: resumo teórico, aplicação e interpretação**. E&S - Engineering and Science, vol. 1, ed. 5, 2016.

HOTELLING H. **Analysis of a complex of statistical variables into principal components**. J. Educ. Psychol. 24, 498–520, 1933.

INMETRO - Instituto Nacional de Metrologia, Normalização e Qualidade Industrial **Orientações sobre Validação de Métodos de Ensaio Químicos**, DOQCGCRE-008, 2018.

JEBER, A.; PROFETA, A. L. **Recursos minerais de Minas Gerais On line: síntese do conhecimento sobre as riquezas minerais, história geológica, e meio ambiente e mineração de Minas Gerais - Águas minerais**. Belo Horizonte: Companhia de Desenvolvimento de Minas Gerais (CODEMGE) – 2018. Disponível em: <http://recursomineralmg.codemge.com.br/wp-content/uploads/2018/10/AguasMinerais.pdf> . Acessado em: 22/12/2019.

JOLLIFFE, I. T.; CADIMA, J. **Principal component analysis: a review and recent developments**. Phil. Trans. R. Soc. A 374:20150202.

KAPPEL, S. **When HPLC fails : IC in food, water, and pharmaceutical analysis**. Metrohm International Headquarters; Herisau, Switzerland, 2019

LYRA, W. S.; SILVA, E. C.; ARAÚJO, M. C. U.; FRAGOSO, W. D. **Classificação periódica: Um exemplo didático para ensinar análise de componentes principais**. Quim. Nova, Vol. 33, No. 7, 1594-1597, 2010.

MASSART, D.L. et al. **Handbook of Chemometrics and Qualimetrics: Part A – Data Handling in Science and Technology** –, Elsevier, v.20, 1997.

MILANEZ, T. V.; SOUZA, A.; BERNARDO, P. E. M. **Nitrato e nitrito em água mineral envasada comercializada na cidade de São Paulo**. Bol. Inst Adolfo Lutz. 2015; 25(1):12-14.

MILLER, J. N.; MILLER, J. C. **Statistics and Chemometrics for Analytical Chemistry**. Editora Pearson. Ed.6, pg:56 – 59, 2010.

MONTALVÃO, V. T. K. **Modelagem cinética da liberação de calor de um sistema de geração de nitrogênio (sgn) através de quimiometria e desenvolvimento de programas em Matlab para sua predição** .Dissertação de mestrado em Química – Universidade Federal do Rio de Janeiro,2014.

MOURA, L. R. C.; PORTO, G. D.; CUNHA, N. R. S.; MOURA, L. E. L. de; VEIGA, R. T. **O comportamento de compra e a percepção dos atributos da água mineral pelos consumidores**. PERSPECTIVA, Erechim. v.35, n.130, p. 97-112, 2011.

NAZARIO, C. E. D.; LANÇAS, F. M. S. **Suportes cromatográficos e fases estacionárias para cromatografia líquida: preparo, evolução e tendências.** Scientia Chromatographica 2013; 5(2):111-135.

NETO, J. M. M. **Estatística multivariada – Uma visão didática-metodológica.** Revista Crítica, Filosofia na ciência, 2004.

OLIVEIRA, E.C. de. **Comparação das diferentes técnicas para a exclusão de “outliers”.** In: ENQUALAB 2008 – CONGRESSO DA QUALIDADE EM METROLOGIA, São Paulo, 2008.

PEARSON K. 1901 ***On lines and planes of closest fit to systems of points in space.*** Phil. Mag. 2, 559–572. (doi:10.1080/14786440109462720)

PEREIRA, J. S. F. **Determinação de cloreto por cromatografia de íons em petróleo extrapesado, coque e resíduo de vácuo após decomposição em sistema fechado por combustão iniciada por microondas.** Dissertação de mestrado – Universidade Federal de Santa Maria, RS, 2007.

PEREIRA, P. C. S. *et al.* **A quimiometria nos cursos de graduação em química: proposta do uso da análise multivariada na determinação de pK_a .** Quím. Nova, vol.37, no.8, São Paulo, 2014.

RODRIGUES, E. C. **Modelos de Regressão Linear Simples – Análise de Resíduos. Slides** - UFOP, 27 de setembro de 2016.

TEÓFILO, R.F.; FERREIRA, M.M.C. **Quimiometria II: Planilhas eletrônicas para cálculos de planejamentos experimentais, um tutorial.** Química Nova, v. 29, n.2, p.338-350, 2006.

TORRES, L. **Vendas de água mineral aumentaram 150% em supermercado no Rio.** G1, 2020. Disponível em: <https://g1.globo.com/rj/rio-de-janeiro/noticia/2020/01/27/vendas-de-agua-mineral-aumentaram-em-supermercado-no-rio.ghtml>. Acessado em: 18/02/2020.

VALDERRAMA, P.; PAIVA, V. B.; MARÇO, P. H.; VALDERRAMA, L. **Proposta Experimental didática para o ensino de análise de componentes principais.** Química Nova, Vol. XY, No. 00, p. 1-5, 26 outubro de 2015.

8. ANEXO

ANEXO A – Passo a passo da aplicação do teste de Grubbs

No caso do Teste de Grubbs para um valor suspeito há as seguintes etapas:

- 1) Colocar os valores em ordem decrescente ou crescente;
- 2) Calcular o valor de G (G_{calc}) para o valor que se mais distancia da média dos valores – Equação 39:

$$G_{calc.} = \frac{|(x_i - \bar{x})|}{s} \quad (\text{Equação 39})$$

Sendo:

x_i = valor suspeito de ser aberrante

\bar{x} = média dos valores dos dados

s = desvio padrão dos valores dados (Equação 40)

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (\text{Equação 40})$$

Sendo:

n = números de análises

- 3) Comparar com valor de G tabelado ($G_{tab.}$) - Anexo B

Se $G_{calc.} > G_{tab.}$, o valor suspeito é aberrante e é retirado do conjunto.

Se $G_{calc.} < G_{tab}$ o valor suspeito não é aberrante.

- 4) Caso o valor suspeito seja discrepante, deve-se repetir os passos 1,2 e 3 até que na não constem mais valores discrepantes;
- 5) Caso o valor suspeito não seja aberrante, verifica-se a existência de dois valores discrepantes em cada extremidade do conjunto.

No segundo teste, para dois valores suspeitos, há as seguintes etapas:

- 1) Calcular a diferença quadrática no conjunto com todos os valores (s_0^2 – Equação 41).

$$s_0^2 = \sum_i^n (x_i - \bar{x})^2 \quad (\text{Equação 41})$$

- 2) Calcular a diferença quadrática no conjunto, sem os valores suspeitos, os quais são os dois menores ou os dois maiores valores ($s_{1,2}^2$ e $s_{n-1,n}^2$, respectivamente).

$$s_{1,2}^2 = \sum_{i=3}^n (x_i - \bar{x}_{1,2})^2 \quad (\text{Equação 42})$$

Sendo: $\bar{x}_{1,2}$ = média do conjunto, sem os dois menores valores.

$$\bar{x}_{1,2} = \frac{\sum_{i=3}^n x_i}{n-2} \quad (\text{Equação 43})$$

$$s_{n-1,n}^2 = \sum_{i=1}^{n-2} (x_i - \bar{x}_{n-1,n})^2 \quad (\text{Equação 44})$$

Sendo: $\bar{x}_{n-1,n}$ = média do conjunto, sem os dois maiores valores

$$\bar{x}_{n-1,n} = \frac{\sum_{i=1}^{n-2} x_i}{n-2} \quad (\text{Equação 45})$$

3) Calcular o valor de G (G_{calc}):

$$G_{calc} = \frac{s_{1,2}^2}{s_0^2} \quad (\text{Equação 46})$$

$$G_{calc} = \frac{s_{n-1,n}^2}{s_0^2} \quad (\text{Equação 47})$$

4) Comparar G calculado (G_{calc}) com G tabelado (G_{tab} – Anexo B), em um nível de significância (α) escolhido.

Se $G_{calc} < G_{tab}$, o valor par suspeito é aberrante e é retirado do conjunto.

Se, $G_{calc} > G_{tab}$, o valor par suspeito não é aberrante.

5) Caso o valor par suspeito seja aberrante, deve-se repetir os passos 1, 2, 3 e 4 até que não constem mais valores aberrantes.

ANEXO B - Quadro com os valores críticos de G

P	Um valor		Dois valores	
	1%	5%	1%	5%
-	-	-	-	-
3	1,155	1,155	-	-
4	1,496	1,481	0	0,0002
5	1,764	1,715	0,0018	0,009
6	1,973	1,887	0,0116	0,0349
7	2,139	2,02	0,0308	0,0708
8	2,274	2,126	0,0563	0,1101
9	2,387	2,215	0,0851	0,1492
10	2,482	2,29	0,115	0,1864
11	2,564	2,355	0,1448	0,2213
12	2,636	2,412	0,1738	0,2537
13	2,699	2,462	0,2016	0,2836
14	2,755	2,507	0,228	0,3112
15	2,806	2,549	0,253	0,3367
16	2,852	2,585	0,2767	0,3603
17	2,894	2,62	0,299	0,3822
18	2,932	2,651	0,32	0,4025
19	2,968	2,681	0,3398	0,4214
20	3,001	2,709	0,3585	0,4391
21	3,031	2,733	0,3761	0,4556
22	3,06	2,758	0,3927	0,4711
23	3,087	2,781	0,4085	0,4857
24	3,112	2,802	0,4234	0,4994
25	3,135	2,822	0,4376	0,5123
26	3,157	2,841	0,451	0,5245
27	3,178	2,859	0,4638	0,536
28	3,199	2,876	0,4759	0,547
29	3,218	2,893	0,4875	0,5574
30	3,236	2,908	0,4985	0,5672
31	3,253	2,924	0,5091	0,5766
32	3,27	2,938	0,5192	0,5856
33	3,286	2,952	0,5288	0,5941
34	3,301	2,965	0,5381	0,6023
35	3,316	2,979	0,5469	0,6101

ANEXO C – Passo a passo a para a aplicação do teste de Cochran

Para a aplicação do teste de Cochran os seguintes passos são necessários:

1) Calcular o valor de C ($C_{calc.}$);

$$C_{calc.} = \frac{S_{Máx}^2}{\sum_{i=1}^n s_i^2} \quad (\text{Equação 48})$$

Onde:

$S_{Máx}^2$ = maior valor de variância (variância máxima)

$\sum_{i=1}^n s_i^2$ = somatório de todas as variâncias de cada nível da curva analítica

A variância de um conjunto de dados experimentais é dada por:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (\text{Equação 49})$$

Sendo:

\bar{x}_i = valor da média de todos os valores

n = número de replicatas

2) Comparar com o valor de C tabelado ($C_{tab.}$) – Anexo D.

Quando $C_{calc.} < C_{tab.}$, pode-se dizer que as variâncias são estatisticamente equivalentes e, dessa forma, os dados têm um comportamento homocedástico.

Quando $C_{calc.} > C_{tab.}$, diz-se que as variâncias não são estatisticamente equivalentes e aumentam (ou diminuem) conforme o aumento da concentração da espécie analisada, sendo a condição de homocedasticidade violada. Nesse caso, é preciso realizar uma transformação ponderada dos dados, de forma a transformar a heterocedasticidade em homocedasticidade antes da utilização da Regressão Linear.

ANEXO D - Quadro com os valores críticos de C

Número de Pontos	N° de medidas (réplicas) para a variável y			
	2	3	4	5
5	0,841	0,684	0,598	0,544
6	0,781	0,616	0,532	0,48
7	0,727	0,561	0,48	0,431
8	0,68	0,516	0,438	0,391
9	0,638	0,478	0,403	0,358
10	0,602	0,445	0,373	0,331
11	0,57	0,417	0,348	0,308
12	0,541	0,392	0,326	0,288
13	0,515	0,371	0,307	0,271
14	0,492	0,352	0,291	0,255
15	0,471	0,335	0,276	0,242
16	0,452	0,319	0,262	0,23
17	0,434	0,305	0,25	0,219
18	0,418	0,293	0,24	0,209
19	0,403	0,281	0,23	0,2
20	0,389	0,27	0,22	0,192

ANEXO E – Quadro com os valores críticos de F*

Graus de liberdade (Denominador)	Graus de Liberdade (Numerador)							
	2	3	4	5	6	10	12	∞
2	19,00	19,16	19,25	19,30	19,33	19,40	19,41	19,50
3	9,55	9,28	9,12	9,01	8,94	8,79	8,74	8,53
4	6,94	6,59	6,39	6,26	6,16	5,96	5,91	5,63
5	5,79	5,41	5,19	5,05	4,95	4,74	4,68	4,36
6	5,14	4,76	4,53	4,39	4,28	4,06	4,00	3,67
10	4,10	3,71	3,48	3,33	3,22	2,98	2,91	2,54
12	3,89	3,49	3,26	3,11	3,00	2,75	2,69	2,30
20	3,49	3,10	2,87	2,71	2,60	2,35	2,28	1,84
∞	3,00	2,60	2,37	2,21	2,10	1,83	1,75	1,00

ANEXO F - Nomes, fontes e concentrações dos íons estudados descritos nos rótulos de cada marca.

	Fonte	Concentração (mg.L ⁻¹)			
		F ⁻	Cl ⁻	NO ₃ ⁻	SO ₄ ²⁻
São Lourenço	São Lourenço - MG	0,15	1,61	1,50	2,17
Bonafont	Jundáí - SP	0,12	0,10	0,78	0,54
Nestlé	Petrópolis - RJ	N.D.	1,81	1,35	0,83
Pouso Alto	Pouso Alto - MG	0,02	0,10	0,48	2,06
Sferrie	Toledo - PR	0,33	N.D.	0,50	0,20
Minalba Premium	Campos de Jordão - SP	0,04	0,18	0,74	0,12
Petrópolis	Petrópolis - RJ	N.D.	1,74	0,35	0,40
Qualitá	Barra Mansa - SP	0,11	0,88	0,79	3,49
Canção Nova	São Bento	N.D.	1,61	0,01	0,78
Pouso Alto (gás)	Pouso Alto - MG	0,03	0,05	N.D.	3,94
Fastio	Chamoim - Portugal	N.D.	4,20	N.D.	1,00
Luso	Luso - Portugal	N.D.	9,0	1,60	1,50
Evian	Haute Savoie - França	0,06	10	3,8	14
Acqua Panna	Toscana - Itália	N.D.	8,5	2,9	22,0
S.Pellegrino	Lombardia - Itália	N.D.	49,4	2,9	402
Perrier	Vergeze - França	N.D.	22	7,8	33

* N.D. – Não Declarado

9. Apêndice

Apêndice A - Modo de utilização do software GNU Octave

Modo de utilização:

A sintaxe do Octave é amplamente compatível com o Matlab e, através do Octave Forge, pode-se encontrar pacotes desenvolvidos para o GNU Octave, semelhante às caixas de ferramentas do Matlab. A seguir é mostrada a tela do programa (Figura 21) onde, na esquerda, tem-se o “Navegador de Arquivos” (retângulo amarelo), onde é possível carregar os dados a serem utilizados para o “Ambiente de Trabalho” (retângulo vermelho).

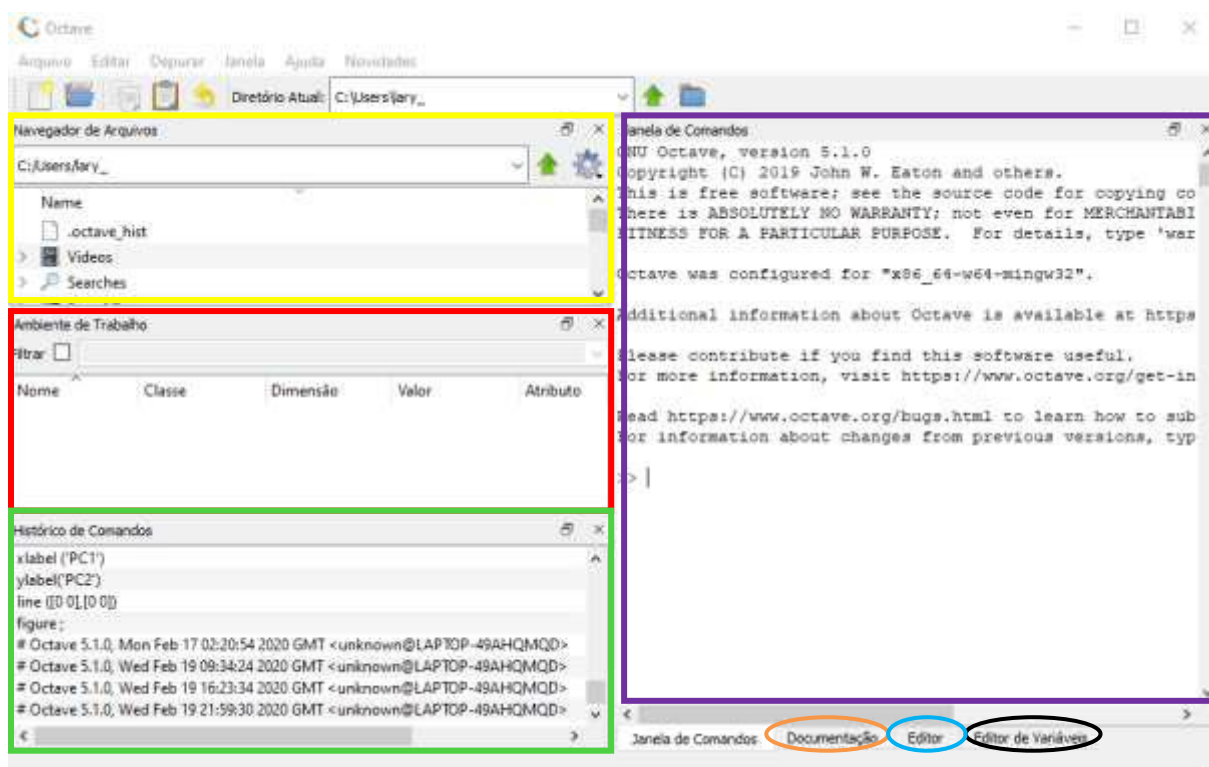


Figura 5.21: Tela do programa com as principais funções destacadas

Ainda na esquerda, há o “Histórico de Comandos” (retângulo verde), o qual facilita a visualização e o uso das funções previamente utilizadas.

À direita, tem-se a “Janela de Comandos” (retângulo roxo) onde são colocadas as funções para calcular os resultados.

Os dados gerados e variáveis criadas são encontrados no ambiente de trabalho e, ao clicar duas vezes em cima de alguns deles, os dados são mostrados

em uma planilha no “Editor de Variáveis” (elipse preta), possibilitando a exportação desses dados para, por exemplo o Excel.

A aba “Documentação” (elipse laranja), possui o índice de funções, o qual fornece os comandos de todas as funções no GNU e suas utilidades. A aba “Editor” (elipse azul), por sua vez, permite que os dados a serem utilizados sejam salvos no formato que possibilite o seu carregamento e, com isso, seu uso pelo programa.

Apêndice B - Áreas dos picos referentes aos íons fluoreto, cloreto, nitrito, brometo, nitrato, fosfato e sulfato nas soluções padrões das curvas analíticas.

		Concentração (em mg.L-1)									
		0,10	0,20	0,30	0,40	0,50	1,00	1,50	2,00	2,50	
Área do pico (em $\mu\text{S} \cdot \text{min}$)	FLUORETO	n=1	0,279	0,512	0,751	0,975	1,494	-	-	-	-
		n=2	0,287	0,499	0,696	0,976	1,231	-	-	-	-
		n=3	0,295	0,521	0,749	1,289	1,292	-	-	-	-
	CLORETO	n=1	0,255	0,330	0,480	0,605	0,746	1,571	2,368	3,467	4,339
		n=2	0,255	0,329	0,432	0,620	0,786	1,583	2,429	3,492	4,339
		n=3	0,276	0,335	0,485	0,619	0,782	1,655	2,478	3,532	4,418
	NITRITO	n=1	0,089	0,157	0,252	0,330	0,408	-	-	-	-
		n=2	0,092	0,154	0,217	0,316	0,419	-	-	-	-
		n=3	0,093	0,159	0,236	0,307	0,409	-	-	-	-
	BROMETO	n=1	0,063	0,122	0,193	0,252	0,339	-	-	-	-
		n=2	0,061	0,132	0,176	0,261	0,321	-	-	-	-
		n=3	0,059	0,129	0,194	0,251	0,320	-	-	-	-
	NITRATO	n=1	0,096	0,157	0,239	0,312	0,407	0,834	1,239	1,808	2,234
		n=2	0,091	0,153	0,209	0,314	0,410	0,832	1,260	1,810	2,231
		n=3	0,083	0,147	0,222	0,318	0,396	0,850	1,284	1,819	2,253
	FOSFATO	n=1	0,036	0,078	0,116	0,155	0,196	-	-	-	-
		n=2	0,043	0,082	0,110	0,154	0,200	-	-	-	-
		n=3	0,040	0,084	0,111	0,151	0,197	-	-	-	-
	SULFATO	n=1	-	-	-	-	0,540	1,082	1,617	2,322	2,888
		n=2	-	-	-	-	0,527	1,083	1,661	2,324	2,869
		n=3	-	-	-	-	0,528	1,113	1,684	2,339	2,922

Apêndice C - Tabelas ANOVA das curvas analíticas dos íons cloreto, nitrito, brometo, nitrato, fosfato e sulfato

		Fonte de variação	SQ	Nº de graus de liberdade	MQ	F
FAIXA BAIXA (0,10 – 0,50 mg.L⁻¹)	CLORETO	Regressão	0,517097	4	0,129274	406,929
		Resíduos	0,003177	10	0,000318	
		Total	0,520274	14		
	NITRITO	Regressão	0,194296	4	0,048574	494,9534
		Resíduos	0,000981	10	9,81E-05	
		Total	0,195278	14		
	BROMETO	Regressão	0,130174	4	0,032543	598,2394
		Resíduos	0,000544	10	5,44E-05	
		Total	0,130718	14		
	NITRATO	Regressão	0,18898773	4	0,047247	672,7303
		Resíduos	0,000702316	10	7,02E-05	
		Total	0,189690046	14		
	FOSFATO	Regressão	0,045465	4	0,011366	1272,239
		Resíduos	8,93E-05	10	8,93E-06	
		Total	0,045554	14		
FAIXA ALTA (0,50 – 2,50 mg.L⁻¹)	CLORETO	Regressão	24,80039	4	6,200098	3520,774
		Resíduos	0,01761	10	0,001761	
		Total	24,818	14		
	NITRATO	Regressão	6,484502833	4	1,621126	9989,313
		Resíduos	0,00162286	10	0,000162	
		Total	6,486125693	14		
	SULFATO	Regressão	10,66246	4	2,665615	5800,278
		Resíduos	0,004596	10	0,00046	
		Total	10,66705	14		

Apêndice D – Desvios padrões (DP) e concentrações médias determinadas (CMD) das soluções padrões 0,10 mg.L⁻¹, 0,30 mg.L⁻¹ e 0,50 mg.L⁻¹

	Concentração (mg.L ⁻¹)	DP	CMD
Fluoreto	0,10	0,008	0,287
	0,30	0,031	0,732
	0,50	0,138	1,339
Cloreto	0,10	0,012	0,262
	0,30	0,030	0,465
	0,50	0,022	0,771
Nitrito	0,10	0,002	0,091
	0,30	0,018	0,235
	0,50	0,006	0,412
Brometo	0,10	0,002	0,061
	0,30	0,010	0,188
	0,50	0,010	0,327
Nitrato	0,10	0,007	0,090
	0,30	0,015	0,223
	0,50	0,007	0,404
Fosfato	0,10	0,004	0,039
	0,30	0,003	0,112
	0,50	0,002	0,198
Sulfato	0,10	0,002	0,122
	0,30	0,020	0,309
	0,50	0,007	0,532

Apêndice E – Comandos utilizados na plataforma do GNU Octave para a aplicação do ACP nos dados gerados na análise das 16 águas minerais

Os comandos utilizados tiveram como base funções realizadas no Matlab e funções encontradas na plataforma do GNU.

- 1) Construção da matriz de dados X ($n \times p$);
 - ✓ Organizar os dados de tal forma que as amostras fiquem nas linhas da tabela e as concentrações de cada íon nas colunas;
 - ✓ Para dados em que o valor ficar abaixo do limite de quantificação, preencher o elemento da matriz com o próprio LQ;
 - ✓ Colocar uma coluna anterior às concentrações enumerando cada amostra;
 - ✓ Substituir as vírgulas por pontos nos dados e copiá-los no “Editor” (indicado em vermelho);
 - ✓ Nomear e salvar os dados;
 - ✓ Carregar os dados através do navegador de arquivos (indicado em azul);

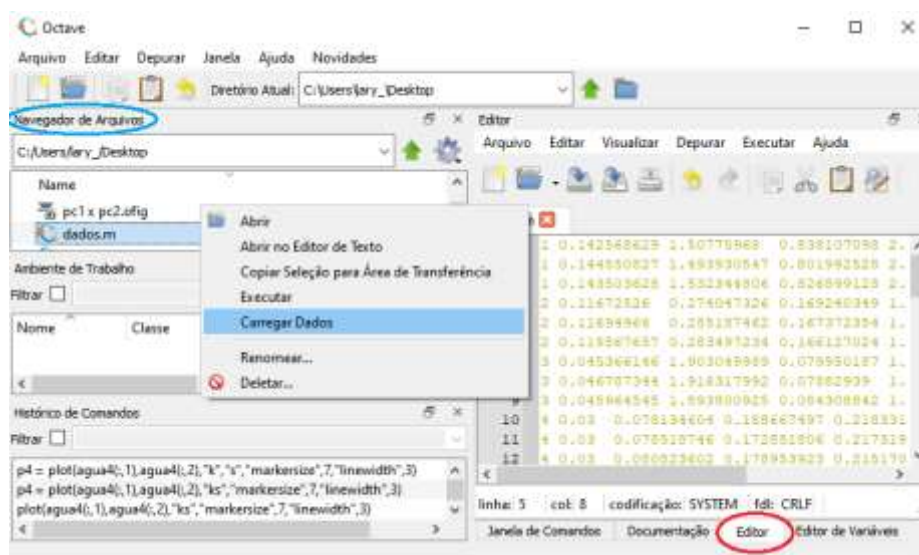


Figura 5.22: Tela do programa com as funções para carregar os dados destacadas

- 2) Pré-tratamento de dados;
 - ✓ Dividir, na “Janela de comandos”, a matriz criada no intuito de delimitar a parte correspondente as classes das amostras () e parte dos resultados

das amostras. No exemplo utilizado a primeira coluna com códigos de 1 à 16 corresponde as classes (matriz 48x1) e a parte dos resultados corresponde ao restante da matriz (matriz 48x5).

```
>> y = dados(:,1);  
>> X = dados(:,2:end);
```

- ✓ Normalizar os dados, através do autoescalamento, usando um comando que equivale a Equação 19 (tópico 3.5.1.2. – Pré-tratamento de dados).

```
>> Xm = bsxfun(@rdivide, bsxfun(@minus,X,mean(X)),std(X));
```

3) Determinação dos autovalores, autovetores e *scores*;

- ✓ A função responsável pela aplicação do PCA no GNU pertence ao pacote de estatística, o qual deve ser baixado na plataforma do GNU. Para utilizar essa função, deve-se carregar o pacote.

```
>> pkg load statistics;
```

- ✓ Aplicar a função *princomp* para os dados tratados (X_m);

```
>> [COEFF, SCORE, latent] = princomp(Xm)
```

Sendo: *COEFF* = autovetores

SCORE = *score*

latent = autovalores

4) Cálculo da contribuição de cada componente principal para a explicação da variância total;

- ✓ Determinar a contribuição de cada componente principal usando um comando que equivale a Equação 27 (tópico 3.5.1.7. - Cálculo da contribuição de cada componente principal).

```
>> Contribuicao = (latent/sum(latent))*100
```

- ✓ Construir um gráfico para melhor visualização das percentagens.

```
>> figure (); bar (Contribuicao(1:5, 1)); title ("Contribuicao de cada  
componente principal (%)");
```

```
>> Componentes = {'PC1','PC2','PC3','PC4','PC5'}
```

```
Componentes =
```

```
{  
 [1,1] = PC1  
 [1,2] = PC2  
 [1,3] = PC3  
 [1,4] = PC4  
 [1,5] = PC5  
 }
```

```
>> set(gca,'XTickLabel',Componentes)
```

```
>> ylabel('Contribuicao %')
```

- ✓ Construir um gráfico que demonstre a percentagem cumulativa ao adicionar cada componente principal na explicação.

```
>> cumsum(Contribuicao)
```

```
ans =
```

```
56.116
```

```
79.822
```

```
97.377
```

```
99.759
```

```
100.000
```

```
>> figure (); bar (cumsum(Contribuicao(1:5,1))); title ("% cumulativa de
contribuicao de cada componente principal");
>> set(gca,'XTickLabel',Componentes)
>> ylabel('% cumulativa de contribuicao')
```

5) Construção dos gráficos dos scores;

- ✓ Definir cada amostra de acordo com a linha da matriz "SCORE" e da matriz de classes definida por y.

```
>> agua1 = SCORE(find(y==1),:);
>> agua2 = SCORE(find(y==2),:);
>> agua3 = SCORE(find(y==3),:);
>> agua4 = SCORE(find(y==4),:);
>> agua5 = SCORE(find(y==5),:);
>> agua6 = SCORE(find(y==6),:);
>> agua7 = SCORE(find(y==7),:);
>> agua8 = SCORE(find(y==8),:);
>> agua9 = SCORE(find(y==9),:);
>> agua10 = SCORE(find(y==10),:);
>> agua11 = SCORE(find(y==11),:);
>> agua12 = SCORE(find(y==12),:);
>> agua13 = SCORE(find(y==13),:);
>> agua14 = SCORE(find(y==14),:);
>> agua15 = SCORE(find(y==15),:);
>> agua16 = SCORE(find(y==16),:);
```

- ✓ Construir o gráfico de scores (PC1 x PC2), relacionando, para cada amostra, cada elemento da coluna 1 com o respectivo elemento da coluna 2 da matriz "SCORE", definindo a cor, tamanho e espessura da linha do marcador.

```
>> figure;
```

```

>> plot(agua1(:,1), agua1(:,2),"rs", "markersize", 10, "linewidth", 2); hold
on;
>> plot(agua2(:,1), agua2(:,2),"bs", "markersize", 10, "linewidth", 2);
>> plot(agua3(:,1), agua3(:,2),"ys", "markersize", 10, "linewidth", 2);
>> plot(agua4(:,1), agua4(:,2),"gs", "markersize", 10, "linewidth", 2);
>> plot(agua5(:,1), agua5(:,2),"ro", "markersize", 10, "linewidth", 2);
>> plot(agua6(:,1), agua6(:,2),"bo", "markersize", 10, "linewidth", 2);
>> plot(agua7(:,1), agua7(:,2),"yo", "markersize", 10, "linewidth", 2);
>> plot(agua8(:,1), agua8(:,2),"go", "markersize", 10, "linewidth", 2);
>> plot(agua9(:,1), agua9(:,2),"r*", "markersize", 10, "linewidth", 2);
>> plot(agua10(:,1), agua10(:,2),"b*", "markersize", 10, "linewidth", 2);
>> plot(agua11(:,1), agua11(:,2),"y*", "markersize", 10, "linewidth", 2);
>> plot(agua12(:,1), agua12(:,2),"g*", "markersize", 10, "linewidth", 2);
>> plot(agua13(:,1), agua13(:,2),"rv", "markersize", 10, "linewidth", 2);
>> plot(agua14(:,1), agua14(:,2),"bv", "markersize", 10, "linewidth", 2);
>> plot(agua15(:,1), agua15(:,2),"yv", "markersize", 10, "linewidth", 2);
>> plot(agua16(:,1), agua16(:,2),"gv", "markersize", 10, "linewidth", 2);
>> title ('Grafico de scores (PC1 x PC2)')
>> xlabel('PC1')
>> ylabel('PC2')

```

- ✓ Repetir os passos acima para os gráficos entre as componentes principais que contribuem significativamente para a explicação da variância total.

6) Determinação e construção do gráfico dos *Loadings*;

- ✓ Determinar os *loadings*, usando um comando que equivale a Equação 29 (tópico 3.5.1.6. - *Scores e loadings*).

```

>> for i=1:size(latent)
P(:,i) = COEFF(:,i) * sqrt(latent(i,1))
End

```

- ✓ Definir cada variável de acordo com a linha da matriz “P”.

```
>>fluoreto = P(1,:);  
>>cloreto = P(2,:);  
>>nitrito = P(3,:);  
>>nitrato = P(4,:);  
>>sulfato = P(5,:);
```

✓ Construir o gráfico de *loadings* relacionando, para cada amostra, cada elemento da coluna 1 com o respectivo elemento da coluna 2 da matriz "P", definindo a cor, tamanho e espessura da linha do marcador.

```
>> figure;  
>> plot(fluoreto(:,1), fluoreto(:,2), "k.", "markersize",25,"linewidth",1); hold on;  
>> plot(cloreto(:,1), cloreto(:,2), "k.", "markersize",25,"linewidth",1); hold on;  
>> plot(nitrito(:,1), nitrito(:,2), "k.", "markersize",25,"linewidth",1); hold on;  
>> plot(nitrato(:,1), nitrato(:,2), "k.", "markersize",25,"linewidth",1); hold on;  
>> plot(sulfato(:,1), sulfato(:,2), "k.", "markersize",25,"linewidth",1); hold on;  
>> title ('Gráfico de loadings no plano PC1 x PC2')  
>> xlabel ('PC1')  
>> ylabel('PC2')
```

✓ O gráfico construído pode ser exportado e ser manipulado para adicionar, a cada ponto, um rótulo com a respectiva variável a qual se relaciona e uma reta da origem até o ponto.

Apêndice F – Scores calculados para cada componente principal

		Scores (PC1)	Scores (PC2)	Scores (PC3)	Scores (PC4)	Scores (PC5)
São Lourenço	1	0.56377	-15.517	-20.115	-0.27016	-0.060159
	1	0.54483	-14.614	-19.075	-0.28725	-0.062272
	1	0.54788	-15.498	-19.628	-0.28801	-0.067121
Bonafont	2	0.50659	0.14972	-0.042375	-0.28359	-0.13275
	2	0.50157	0.14345	-0.030084	-0.28974	-0.13539
	2	0.497	0.16582	-0.041341	-0.29507	-0.12979
Nestlé	3	0.67122	0.39484	0.273	0.1544	-0.046096
	3	0.66792	0.40826	0.26808	0.15208	-0.042346
	3	0.67297	0.39139	0.25486	0.15504	-0.044249
Pouso Alto	4	0.92818	0.46517	-0.25159	0.3528	-0.057471
	4	0.92186	0.50124	-0.20298	0.34857	-0.059056
	4	0.9243	0.48817	-0.2223	0.35076	-0.058207
Sferrie	5	-0.023442	11.929	-0.32796	-0.9728	0.078968
	5	-0.030736	12.004	-0.32519	-0.98337	0.079211
	5	-0.10257	12.376	-0.36957	-10.864	0.090245
Minalba Premium	6	0.79022	0.57928	0.0034936	0.14193	-0.064369
	6	0.78635	0.58225	0.0050651	0.13592	-0.064659
	6	0.7837	0.58397	0.0049025	0.1319	-0.064607
Petropolis	7	0.64296	0.54733	0.39539	0.11933	-0.046691
	7	0.64307	0.54748	0.39529	0.11934	-0.04668
	7	0.6447	0.54663	0.39634	0.1218	-0.046865
Qualitá	8	0.65544	0.37962	-0.79634	0.10442	0.045842
	8	0.59308	0.42845	-0.80394	0.018172	0.054322
	8	0.65889	0.3826	-0.78604	0.11257	0.045464
Canção Nova	9	0.65074	0.96562	0.082122	0.0094113	0.0056419
	9	0.38319	0.98574	0.020109	-0.17062	0.12271
	9	0.36912	10.296	-0.015146	-0.18998	0.13485
Pouso Alto (gás)	10	0.81027	0.83977	0.21473	0.24949	-0.054679
	10	0.83472	0.83281	0.22573	0.28108	-0.059156
	10	0.80538	0.8403	0.21315	0.24492	-0.058634
Fastio	11	0.56582	0.24132	0.51156	0.20757	0.024635
	11	0.56312	0.25221	0.51868	0.21065	0.028946
	11	0.56441	0.24233	0.51091	0.20948	0.02726
Luso	12	0.27201	0.28025	0.48564	0.22298	0.28528
	12	0.26471	0.27056	0.4972	0.21683	0.28452
	12	0.26164	0.25912	0.49753	0.22322	0.28965
Evian	13	0.3244	-23.377	-12.594	0.25268	0.079012
	13	0.35041	-23.314	-13.254	0.26558	0.081824
	13	0.33895	-24.758	-14.308	0.26974	0.088799
Acqua Panna	14	0.058735	-0.49214	10.697	0.047732	-0.14293
	14	0.076895	-0.47401	10.582	0.045094	-0.1535
	14	0.058808	-0.52001	10.875	0.027835	-0.16435
S.Pellegrino	15	-60.864	0.60267	-0.5202	0.18224	0.10204
	15	-58.152	0.4472	-0.33941	0.5556	0.035589
	15	-6.62	0.78331	-0.76024	-0.11496	-0.20214
Perrier	16	-10.083	-23.419	22.542	-0.34247	0.021826
	16	-10.149	-23.551	22.626	-0.34612	0.021658
	16	-0.99829	-22.985	22.262	-0.32068	0.035879