



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
DEPARTAMENTO DE QUÍMICA

**APLICAÇÃO DE FERRAMENTAS QUIMIOMÉTRICAS PARA
AVALIAÇÃO DA CONCENTRAÇÃO DE METAIS EM
RESÍDUOS DA INDÚSTRIA DE ÓLEO E GÁS**

THIAGO LEHMKUHL LUCIANO

Florianópolis
Mês/2020

Thiago Lehmkuhl Luciano

**APLICAÇÃO DE FERRAMENTAS QUIMIOMÉTRICAS PARA
AVALIAÇÃO DA CONCENTRAÇÃO DE METAIS EM RESÍDUOS DA
INDÚSTRIA DE ÓLEO E GÁS**

Relatório apresentado ao Departamento de Química
da Universidade Federal de Santa Catarina,
como requisito parcial da disciplina de
Estágio II (QMC 5512)

Prof^a Dr^a Tatiane de Andrade Maranhão

MSc. Klaiani Bez Fontana

Florianópolis
01/2021

AGRADECIMENTOS

Gostaria de agradecer primeiramente à minha orientadora Prof^a. Dr^a. Tatiane de Andrade Maranhão, por me aceitar como orientando durante os últimos anos de graduação e auxiliar em tomadas de decisões que foram essenciais para a realização deste trabalho.

À minha colega Thebny pela paciência, por me ensinar o funcionamento do laboratório, a operar os equipamentos e pela tremenda ajuda na construção do meu primeiro TCC. À Klaiani por todas as chamadas de vídeo que ocorreram durante toda a confecção deste trabalho final.

Meus agradecimentos a todos os professores da UFSC que sempre se mostraram interessados em sanar todas as dúvidas, por ensinar como fazer ciência e por engrandecer o aprendizado que tive durante a graduação.

À todas as pessoas que passaram por mim durante essa fase da minha vida, por todas as amizades, cervejas, festas e perrengues compartilhados. Uma menção especial ao grupo do gramadinho, poucas coisas podem ser tão simples e fazer tanta falta como os encontros que sempre ocorriam ali.

Por fim, gostaria de agradecer meu pai e minha mãe, que sempre fizeram de tudo para que eu pudesse seguir o meu caminho, pelos legítimos sacrifícios e pelos legítimos sorrisos. Me sinto extremamente sortudo quando penso sobre a minha vida e percebo como eles sempre estiveram ao meu lado.

SUMÁRIO

1 INTRODUÇÃO	pag 10
2 REVISÃO DA LITERATURA	pag 12
2.1 Indústria petroquímica	pag 12
2.2. Resíduos Sólidos da Indústria de Petróleo e Gás	pag 13
2.3 ICP-MS	pag 14
2.3.1 Interferências ICP-MS	pag 15
2.4 Quimiometria	pag 17
2.4.1 Pré-tratamento de dados	pag 18
2.4.2 Métodos de decomposição multivariada	pag 19
2.4.3 Análise de cluster (CA)	pag 22
3 OBJETIVOS	pag 26
3.1 Objetivo Geral	pag 26
3.2 Objetivos Específicos	pag 26
4 MATERIAL E MÉTODOS	pag 27
4.1 Instrumentação	pag 27
4.2 Amostras e analitos	pag 28
4.3 Reagentes	pag 28
4.4 Procedimentos analíticos	pag 28
4.5 Segurança no laboratório	pag 30
4.6 Descarte e tratamento de resíduos	pag 30
5 RESULTADOS E DISCUSSÃO	pag 32

5.1 Origem e seleção dos dados	pag 32
5.2 Estatística básica	pag 34
5.3 Ferramentas quimiométricas	pag 35
5.3.1 Autoescalamento dos dados	pag 35
5.3.2 Análise de componentes principais	pag 36
5.3.3 Análise hierárquica de clusters	pag 40
6 CONCLUSÃO	pag 44
7 REFERÊNCIAS	pag 46
8 APÊNDICES	pag 51

LISTA DE FIGURAS

FIGURA 1. Gráfico dos scores de PC1 e PC2 para 88 elementos químicos.....	pag 21
FIGURA 2. Gráfico de pesos da componente principal 1 e para componente principal 2 para oito propriedades.....	pag 22
FIGURA 3. Exemplo onde a PCA não é capaz de diferenciar os grupos das amostras.....	pag 23
FIGURA 4. Dendograma simulado como o gerado pelo HCA para 13 amostras.....	pag 25
FIGURA 5. Concentrações médias e desvios dos analitos A) Ni, Fe, Al e Zn e B) As, Cr, Co e Pb nos diferentes resíduos da indústria de petróleo e gás.....	pag 34
FIGURA 6. Gráfico dos escores de PC1 e PC2 para 34 amostras de resíduos provenientes da indústria de petróleo e gás.....	pag 36
FIGURA 7. Gráfico de pesos (loadings) da componente principal 1 e para componente principal 2 para oito concentrações de analitos.....	pag 37
FIGURA 8. Gráfico das médias das concentrações dos analitos.....	pag 39
FIGURA 9. Dendograma gerado pelo HCA para 4 amostras e 32 casos.....	pag 41
FIGURA 10. Dendograma comparando as similaridades entre as concentrações dos analitos	pag 42

LISTA DE TABELAS

TABELA 1. Primeiro programa potência de microondas para digestão das amostras.....	pag 29
TABELA 2. Segundo programa de potência de microondas para digestão das amostras.....	pag 29
TABELA 3. Tabela 3. Dados de limite de detecção (LOD) e limite de quantificação (LOQ) existentes nos relatórios.....	pag 33
TABELA 4. Dados de concentração dos analitos (em mg/Kg) nas diferentes amostras de resíduos da indústria de petróleo e gás do laboratório LEMA/LARES do Departamento de Química da UFSC.....	pag 33

LISTA DE ABREVIATURAS

PCA. – Análise de componentes principais (Principal component analysis)

HCA. – Análise de cluster hierárquico (Hierarchical cluster analysis)

LEMA. – Laboratório de Espectrometria Atômica

LARES. – Laboratório de Análise de Resíduos

ICP–MS. – Espectrometria de massa com plasma indutivamente acoplado

ABNT. – Associação Brasileira de Normas Técnicas

NBR. – Normas brasileiras

RF. – Radio frequência

Q. – Quadripolar

LIT. – Armadilha linear de íons

QIT. – Armadilha de íons quadripolar

PARAFAC. – Análise de fator paralelo

SBQ. – Sociedade Brasileira de Química

IUPAC. – União Internacional de Química Pura e Aplicada (International Union of Pure and Applied Chemistry)

PC. – Componente principal (Principal componente)

CA. – Análise de cluster (Cluster analysis)

ED. – Distância euclidiana (Euclidean distance)

RESUMO

Com o desenvolvimento da tecnologia aplicada a análise química, a realização de análises com alta frequência analítica em diversas amostras em um tempo cada vez menor, gera-se uma grande quantidade de dados. Esses dados muitas vezes necessitam de tratamentos matemáticos para que haja completa visualização e entendimento do sistema estudado. A área da química responsável pelo estudo e aplicação desses métodos estatísticos é a quimiometria. Esse trabalho teve seu projeto inicial alterado devido a pandemia associada ao coronavírus, que não permitiu por motivo de segurança sanitária a execução da parte experimental prevista, este trabalho então apresenta revisões teóricas e aplicações práticas de ferramentas estatísticas à resultados de determinações de metais em amostras ambientais, assuntos ligados à química analítica e suas vertentes. O trabalho descreve os fundamentos da espectrometria de massa com plasma indutivamente acoplado, aparelho utilizado para gerar os dados das amostras, expondo vários processos quimiométricos, dando ênfase à análise de componentes principais (PCA) e à análise de cluster hierárquico (HCA). Esses dois métodos foram aplicados em dados de concentração de arsênio, cromo, cobalto, níquel, alumínio, chumbo, zinco e ferro, em amostras de resíduos da indústria de petróleo e gás fornecidas pela Petrobras, em diferentes anos de convênio com o grupo do Laboratório de Espectrometria Atômica e Laboratório de Análise de Resíduos (LEMA/LARES) na UFSC. A PCA separou cada tipo de resíduo (4 tipos no total) em diferentes grupos relacionadas às componentes principais. Com essa análise foi possível perceber que as concentrações de As, Cr e Pb foram as principais causas da separação do lodo dos demais resíduos, os dados de Co, Ni e Al foram os principais responsáveis pelo agrupamento dos catalisadores e que as quantidades de Fe e Zn foram determinantes para a separação do cascalho e da borra. Por fim, a HCA agiu como prova real, reafirmando as análises providas da PCA e melhorando a visualização das diferenças entre os grupos dos cascalhos e das borras.

Palavras-chave: Quimiometria, espectrometria de massa com plasma indutivamente acoplado, análise de dados, resíduos petroquímicos.

1 Introdução

No começo de 2020 o planejamento era fazer um projeto que pretendia investigar se existiam mudanças significativas ao se modificar o procedimento para obtenção de extrato lixiviado de resíduos sólidos, descrito pela norma ABNT NBR 10005. Pretendia-se trocar os volumes de solução extratora, mantendo-se a proporção massa/volume preconizada na norma.

Com o começo da pandemia causada pelo novo coronavírus, os laboratórios ficaram fechados para os alunos da graduação. Com isso, o tema do trabalho foi modificado. Neste momento optou-se por continuar com a temática de resíduos da indústria de óleo e gás, mas agora o objetivo de utilizar ferramentas quimiométricas para análises de dados das concentrações de metais totais dessas amostras, previamente determinados em função do convênio celebrado pelo LEMA/LARES da UFSC, possibilitando discussões sobre a indústria de óleo e gás, os resíduos por ela gerados, as metodologias que são aplicadas nesse tipo de análise e a viabilidade do uso de ferramentas quimiométricas para a análise de dados provindos dessas metodologias.

A indústria de óleo e gás tem extrema importância para a vida e economia brasileira e mesmo com os avanços nas pesquisas de obtenção renovável de energia e métodos de reciclagem, esse mercado continuará tendo importante papel no mundo. Porém, esse ramo é responsável pela geração de toneladas de diferentes tipos de resíduos que, em sua composição, podem conter altas concentrações de substâncias que podem ser maléficas tanto para o meio ambiente quanto para a vida das pessoas.

As discussões sobre a preservação da natureza vêm ganhando cada vez mais força com o passar do tempo. Para os químicos, o tratamento e análise de resíduos industriais é uma das principais maneiras de contribuição para a saúde do meio ambiente. Para tanto, necessita-se de equipamentos e metodologias capazes de realizar análises das substâncias classificadas como perigosas, com baixos limites de detecção, alta reprodutibilidade e boa confiança.

O avanço da tecnologia aplicada nos instrumentos, capazes de fazer análises multielementares com alta frequência analítica, como a espectrometria de massa com plasma acoplado indutivamente (ICP-MS), juntamente com o desenvolvimento dos

processos químicos adequados ao preparo de amostras complexas, possibilitou a geração de grandes quantidades de dados em um pequeno espaço de tempo, criando a necessidade de métodos avançados para projetar experimentos, calibrar instrumentos e analisar todos esses dados gerados.

A quimiometria é, em suma, uma interação entre métodos estatísticos e certas matemáticas em medições de processos químicos. Utilizada comumente pela química analítica, ela promove aproximações e visualizações de dados, facilitando a correção de erros nos mais diversos métodos analíticos, auxiliando a interpretação e reconhecimento de dados que não estariam tão evidentes sem o devido tratamento matemático.

Esse trabalho visa discutir resultados de determinação de metais e semimetais por ICP-MS, analitos estes presentes em amostras de resíduos provenientes da indústria de petróleo e gás, sendo estes: Lodo, borra, cascalho de perfuração e catalisadores. Assim como uma revisão bibliográfica para o entendimento da quimiometria aplicada neste estudo. Foram aplicadas duas ferramentas quimiométricas que se destacam nas análises de dados químicos, a análise de componentes principais e a análise de cluster hierárquico. Assim, será possível entender a importância do uso destas ferramentas para melhor compreensão do tema de resíduos industriais, assim como maximizar informações para além de valores de concentrações para expansão do conhecimento para tomadas de decisão quanto ao gerenciamento destes resíduos.

2 Revisão da literatura

2.1 Indústria petroquímica

A indústria de óleo e gás é o ramo que emprega como matérias primas os mais diversos tipos de matéria, como o gás natural, gases liquefeitos de petróleo, naftas, resíduos de petróleo e alguns tipos de petróleo cru. Essa área possui um alto poder germinativo e tem correlação com os mais diversos setores da econômica. Diversos insumos podem ser produzidos para o setor de fertilizantes, pesticidas, plásticos, detergentes, adesivos, solventes, explosivos, produtos farmacêuticos, combustíveis, entre outros.¹

Esse ramo industrial é responsável pela geração de toneladas de resíduos, sendo eles produzidos durante as etapas de estocagem, transporte e refino do petróleo. Em 2013, a indústria de petróleo e gás brasileira produziu cerca de um milhão de toneladas de um dos resíduos decorrentes do processo de extração do óleo bruto, a borra de petróleo.² Sendo assim, novas estratégias e metodologias vêm sendo estudadas e empregadas para caracterizar e reutilizar tais resíduos de forma eficiente, contribuindo para a redução desse passivo.

Ter consciência ecológica é algo de extrema importância dada a realidade atual do mundo em que nos encontramos. A cada dia que passa, novas notícias sobre desastres ambientais, má gestão de resíduos sólidos, poluição do ar e de águas são apresentadas. Em função de sua elevada importância, o setor petroquímico é um dos principais geradores de poluentes, e este está sob contínua fiscalização das agências de proteção ambiental que, quando descumprem as leis e normas propostas pelo governo, são passíveis de punição pelos órgãos competentes.³

É preciso ter muita atenção sobre os resíduos produzidos por essa indústria, uma vez que muitos estudos comprovam a existência de diversas substâncias que, em altas concentrações, podem afetar o meio ambiente e a saúde das pessoas que entram em contato com tais descartes.^{4,5,6} Alumínio, arsênio, bário, chumbo, cobre, ferro, mercúrio, manganês, prata e zinco são exemplos de metais e semimetais que, em determinadas concentrações, são classificados como perigosos pela Norma de classificação de resíduos sólidos industriais, ABNT NBR 10004:2004⁷ e podem estar presentes nesses descartes. Vale ressaltar que diversas matrizes distintas compõem

o leque de resíduos gerados pelo setor industrial de petróleo e gás, cabendo observar as características únicas de cada resíduo como um problema individual.

2.2. Resíduos Sólidos da Indústria de Petróleo e Gás

O cascalho de perfuração é um resíduo produzido pela indústria de petróleo e gás, sendo proveniente do processo de extração do óleo cru. A perfuração do solo libera fragmentos de rochas que são cortadas pela broca. Esses fragmentos produzidos durante a extração são posteriormente separados do fluido de perfuração usado na escavação e então são recolhidos e analisados para serem devidamente tratados, descartados ou reutilizados.⁸ A composição química desses rejeitos vai depender da origem deste óleo cru e dos processos de extração e refinação do mesmo.⁹

Sendo um dos resíduos mais significativos da indústria de petróleo, a borra oleosa é uma emulsão estável e complexa constituída tipicamente de óleo (30-50%), água (30-50%), e sólidos (10-12%) composta de diferentes hidrocarbonetos de petróleo, água, metais pesados e partículas sólidas.^{2,10} Uma quantia significativa desse resíduo é gerado durante a exploração, produção, estocagem, transporte e processo de refinação do petróleo.¹¹ Estima-se que em 2013, só Brasil foi responsável por gerar 2940 toneladas por dia de borra, logo, métodos para lidar e analisar esse resíduo vem ganhando interesse.²

O petróleo é uma mistura muito complexa de diferentes hidrocarbonetos, juntamente com menores quantidades de compostos contendo nitrogênio, oxigênio e enxofre. Os produtos originados dessa matéria prima não são constituídos do óleo cru, existe uma variedade de processos que são requeridos para a transformação do mesmo, entre eles vale citar os processos de catálise, rejeição de carbono, rotas de adição de hidrogênio e quebra de correntes carbônicas.¹²

A catálise ganhou um importante papel na indústria de óleo e gás. O primeiro método catalítico utilizado no refinamento do petróleo para a quebra de ligações de hidrogênio foi nos anos 80 e utilizou argila modificada, mas nos dias atuais os processos refinadores utilizam zeólitos e aluminossilicatos.¹² Após seu uso prolongado, os catalisadores perdem sua atividade devido a sinterização, deterioração

e acúmulo de metais, logo, eles se transformam em resíduos que precisam ser analisados, descartados ou reutilizados de maneira correta.¹³

As diferenças em composição possíveis nesses diferentes resíduos fazem do monitoramento e gerenciamento destes resíduos algo nada trivial. Por ser capaz de analisar a maior parte dos elementos presentes na tabela periódica, principalmente para análise de metais em nível traço (ppb e ppt), a espectrometria de massa com plasma indutivamente acoplado é uma ótima opção para análises de metais e semimetais em diferentes matrizes.¹⁴ Diversos trabalhos na literatura fazem uso dessa técnica para desvendar a composição de petróleo e seus resíduos.^{5,6,15}

2.3 ICP-MS

Um dos principais desafios da química atualmente é a necessidade de aumentar a sensibilidade dos métodos de análise, pois cada vez mais as legislações impõem regras que exigem a detecção de compostos com potenciais tóxicos em baixíssimas concentrações, chegando, atualmente, nos níveis de pg L^{-1} .¹⁴

Inicialmente utilizado para a quantificação de elementos baixas concentrações em amostras líquidas, a espectrometria de massa com plasma indutivamente acoplado, ou ICP-MS, desde sua introdução nos anos 80, se transformou em uma incrível técnica para análise de diversos compostos.¹⁶ Muitos fabricantes produzem equipamentos robustos e confiáveis, capazes de fazer detecção multielementar, com alta resolução espectral e limites de detecção muito baixos.^{17,18}

Tal técnica tem a capacidade de analisar a maior parte dos elementos presentes na tabela periódica,¹⁴ juntamente com compostos orgânicos,¹⁹ organometálicos²⁰ e biomoléculas complexas.²¹ Logo, tal equipamento foi importante para análises químicas, ajudando a estudar, por exemplo, a degradação do sistema nervoso pelo consumo de mercúrio, envenenamentos por arsênio em água e degeneração de órgãos sexuais de animais por compostos organoestênicos.²²

Em suma, o ICP-MS possui várias características que o fazem ser uma ótima técnica de análise elementar, com alta sensibilidade e seletividade para grande parte dos elementos, um espectrômetro simples, permite a determinação de isótopos de um

mesmo analito e ainda possibilita o uso de padrões externos em sua calibração para prevenir flutuações instrumentais.^{17,18,23}

A metodologia baseia-se na separação, identificação e quantificação dos elementos, levando em consideração a sua razão carga/massa (m/z)¹⁷. O plasma indutivamente acoplado, em alta temperatura, é a fonte de íons padrão utilizada na maioria dos equipamentos comerciais.²² Ele utiliza gases nobres como argônio para gerar o plasma, dando uma alta eficiência para a vaporização, dissociação ou atomização, excitação e ionização das amostras a serem analisadas.¹⁴

O plasma é gerado por uma descarga no gás, com a energia sendo mantida com o auxílio de um gerador de radiofrequência (RF).¹⁷ A alta temperatura, cerca de 5773K, podendo chegar a 7000K, permite a completa fragmentação de cada molécula presente.¹⁴

O campo magnético é induzido pela corrente de RF através da bobina, produzindo uma região de indução no gás que conduz. O campo magnético produzido oscila em função da frequência do gerador, em torno de 27MHz, fazendo o campo permanecer dentro da tocha. Produzidos pela faísca da bobina, os elétrons livres giram em torno do campo magnético, transformando a energia provinda da bobina em energia cinética para os elétrons. Os elétrons em alta velocidade colidem com o gás em fluxo, essas colisões geram resistência que por sua vez gera aquecimento por efeito Joule. Uma vez em altas temperaturas, o sistema é auto-sustentável.¹⁷

Depois de passar pelo plasma, os analitos ionizados são separados de acordo com a sua razão m/z em um analisador de massas. Os analisadores de massas mais utilizados são: Quadripolar (Q), armadilha linear de íons (LIT), armadilha de íons quadripolar (QIT) e Orbitrap. O analisador quadripolar usa campos de radiofrequência (RF) quadripolar para filtrar seletivamente os íons de interesse baseado em sua razão carga/massa, semelhantemente ao QIT, onde os íons são criados e aprisionados em um campo de RF quadripolar 3D para que então sejam ejetados e separados.²⁴

2.3.1 Interferências ICP-MS

Por ter uma grande gama de interferências, existem diversos estudos que são voltados exclusivamente para melhoramento e retirada desses intermédios na técnica

de ICP-MS.²⁵ A evolução da tecnologia fez com que muitos problemas fossem corrigidos, mas interferências espectrais e não espectrais ainda são muito frequentes¹⁷.

As interferências espectrais são aquelas ligadas a sobreposição de massas que possuam a mesma massa nominal do analito, causando sinais largos e incorretos. Essa sobreposição pode ocorrer através de isótopos de elementos diferentes.²⁵ Elas são facilmente evitadas utilizando diferentes isótopos ou equações elementares para análise. Íons poliatômicos formados a partir dos gases do plasma, dos gases de arraste, da água, dos ácidos utilizados no preparo de amostra e da própria matriz também são causas de interferências espectrais.²⁵

Os íons de óxidos são outras fontes desse tipo de interferência. Eles são criados pela dissociação incompleta da amostra ou da recombinação dos íons no plasma. A quantidade de vapor de água também pode formar tais interferentes, pois a quantidade de água introduzida no plasma afeta o equilíbrio do plasma. Aumentar a temperatura do plasma ou diminuir a vazão do gás carregador são maneiras de otimização de parâmetros que podem solucionar esse caso.¹⁷

Já os íons dupla carga diminuem o sinal do analito e geram isótopos com metade da massa do íon de origem, uma vez que o analisador de massa pode apenas contar isótopos monovalente, resultando em interferências em massas semelhantes. Ao aumentar a vazão do gás, resfriando o plasma, ou diminuir a potência RF, diminuindo a temperatura do plasma, podem ser soluções simples para diminuir tal interferência.¹⁷

Alguns métodos para superar esses desafios espectrais podem ser citados. O uso de ácido nítrico, por exemplo, é indicado para ser usado no preparo da amostra para análises de vanádio e arsênio, uma vez que o ácido hidrocloreídrico pode acarretar na aparição de $^{35}\text{Cl}^{16}\text{O}^+$ e $^{40}\text{Ar}^{35}\text{Cl}^+$.²⁵ A precipitação e a extração de solvente são outras ferramentas capazes de ajudar nessa tarefa. Lyon et al.²⁶ reduziu as interferências precipitando os cloretos com prata, prevenindo a formação de $^{12}\text{C}^{35}\text{Cl}_2^+$, $^{40}\text{Ar}^{37}\text{Cl}^+$ e $^{35}\text{Cl}^{16}\text{O}^+$. Gases de plasma alternativos²⁶, utilização de nebulizadores microconcentricos²⁷, ablação a laser,²⁸ geração de hidretos²⁹ e vaporização eletrotérmica³⁰ também são algumas das técnicas documentadas.

Já as interferências não-espectrais estão caracterizadas pela redução ou aprimoramento no sinal dos analitos, dado a fatores que exercem influência no transporte da amostra, na ionização no plasma, na extração de íons ou na taxa de transferência de íons. Essas distorções são a segunda maior causa de interferências no ICP-MS e estão diretamente ligadas à natureza e concentração da amostra.²⁵

Muitas vezes, durante e após a coleta dos dados desse tipo de equipamento, é interessante utilizar certas ferramentas estatísticas, seja para prevenir possíveis interferências já previstas ou para tratar a grande quantidade de informações coletadas sobre o sistema. A quimiometria é o ramo da química responsável por estudar diferentes mecanismos matemáticos capazes de ajudar os pesquisadores em suas análises e seus entendimentos sobre o sistema investigado.³¹

2.4 Quimiometria

O primeiro cientista a usar o termo "kemometri", que em português é o equivalente a "quimiometria", foi o sueco Svante Wold em 1971. Os primeiros jornais desta área foram publicados em 1986-1987, sendo eles o "Chemometrics and Intelligent Laboratory Systems" e o "Journal of Chemometrics". Essa ciência pode ser brevemente descrita como a interação entre métodos estatísticos e certeza matemática em medições de processos químicos. A quimiometria foi criada devido à mudança dos tipos e tamanhos dos dados obtidos pelas técnicas analíticas, assim como o avanço dos microprocessadores capazes de fazer contas cada vez mais complexas.³¹

Normalmente utilizada pela química analítica, ela engloba matemática, estatística e lógica formal para poder criar teorias e métodos para medições químicas. Ela é capaz de promover novas aproximações e percepções sobre análises de diferentes tipos de espectroscopia e medições de dados químicos.³² A grande virada dessa ciência veio no século XXI, com o desenvolvimento de diversos softwares capazes de fazer o processamento dos dados, dando espaço para construção de equipamentos que analisam em várias dimensões, criando assim novas opções para melhorar os métodos de análise de dados.³¹

Uma boa descrição desse assunto é dada por Massarte et al.³³ que diz: "Quimiometria é a disciplina química que usa matemática, estatística e lógica formal

(1) para projetar ou selecionar procedimentos experimentais ideais; (2) para prover o máximo de informação química relevante por análise de dados químicos e (3) para obter conhecimento sobre o sistema químico". As necessidades que fizeram que a quimiometria fosse criada e estudada são: (1) A imensa quantidade de dados produzidos por modernos analisadores multi-elementares e multi-componentes, que necessitam dos métodos quimiométricos para extrair informações essenciais; e (2) o avanço da computação, que permite processar essa grande massa de dados complexos.²

É importante ressaltar a importância do desenvolvimento da computação, pois é através dos computadores que temos a capacidade de análise de grandes quantidades de dados. Assim, pode-se dizer que a criação de bibliotecas espectrais, as análises de dados em várias dimensões, as redes neurais artificiais, a pesquisa de inteligência artificial para química e os sistemas super inteligentes impulsionaram todas as formas de análises de dados, englobando assim a quimiometria.³²

Análise de dados multivariada envolve a análise de dados existentes nas diversas variáveis medidas em um dado número de amostras. O objetivo principal aqui é determinar todas as variações nos dados da matriz estudadas. As ferramentas quimiométricas tentam encontrar relações entre as amostras e as variáveis em um certo conjunto de dados, convertendo-os em novas variáveis latentes, ou seja, variáveis que antes estavam "escondidas".³¹

2.4.1 Pré-tratamento de dados

Os dados que se pretende avaliar quimiometricamente são comumente observados como uma matriz de dados em que as linhas representam n amostras e as colunas representam m variáveis, que em química estas variáveis podem ser diferentes analitos, diferentes propriedades físicas ou químicas.

Devido as diferentes possibilidades de complexidade dos dados que deseja-se trabalhar o pré-processamento dos dados é uma etapa crucial. Os dados podem ser centrados na média e/ou autoescalados.³⁴ Centrar os dados na média faz com que cada variável tenha média zero através da seguinte equação:

$$X_{(i,j)CM} = X_{(i,j)} - \bar{X}_j \quad \text{Equação (1) De forma que:}$$

$X_{(i,j)CM}$ = valor centrado na média para variável j, na amostra i

$X_{(i,j)}$ = valor da variável j na amostra i

\bar{X}_j = média dos valores da coluna j calculados conforme equação (2), abaixo:

$$\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{i,j} \quad \text{Equação (2)}$$

Sendo n o número de amostras.

O procedimento de autoescalamento de dados é comum quando as variáveis se encontram em unidades diferentes, para assim dar o mesmo peso a todas as variáveis.³⁵ Autoescalar os dados significa dividir os dados centrados na média pelo desvio padrão, de forma que terão média igual a zero e desvio padrão igual a 1 na matriz de dados.

2.4.2 Métodos de decomposição multivariada

Essa técnica é aplicada para quantificar contribuições, agrupando as diferenças nas previsões médias de modelos multivariados. Utiliza-se modelos regressivos para particionar os componentes da diferença de grupo, (significado ou uma proporção) em dois componentes, um deles atribuível à diferentes composições entre grupos (diferença em características ou atribuições), e um componente atribuível às diferenças nos efeitos das características (diferenças nas respostas, coeficientes ou comportamentos).³⁶

Esses métodos são muito úteis quando se quer estudar dados multivariados reduzindo-os a um certo número de variáveis correlacionadas, chamadas de componentes principais, decompondo os dados em uma dimensão menor.³¹ Aqui, existem dois métodos normalmente utilizados pelos químicos, onde um é a análise de componentes principais, ou Principal Components Analysis (PCA), e a análise de fator paralelo, ou parallel factor analysis (PARAFAC).

O objetivo da PCA é encontrar o subespaço no espaço das variáveis onde os dados mais variam. As variáveis originais, normalmente correlacionadas, são

linearmente transformadas em um número menor de variáveis não relacionadas.³⁷ Ela reduz a dimensionalidade dos dados considerando um amplo número de variáveis relacionadas, enquanto retém o máximo da variação no conjunto de dados.³¹ Isso pode ser útil quando grandes quantidades de dados podem ser aproximadas por um modelo estrutural de complexidade moderada.³⁸ Usa-se os dados não correlacionados e roda-os de uma maneira onde os primeiros componentes possuem a maior variação em todos as variáveis originais.³¹

Existem inúmeras aplicações para a PCA, normalmente utiliza-se essa ferramenta em análises de misturas biológicas.³¹ Essa análise pode ser usada como uma ferramenta redutora de dimensões para outras análises computacionais mais aprofundadas. Ela é aplicada para identificar as relações entre variáveis com alta variância, ajudando a entender o conjunto de dados de uma análise. Muitas vezes as relações encontradas pela PCA permitem recuperar dados perdidos. Essa capacidade pode ser aplicada para redução de ruídos das medições e desenvolvimento de sensores sensíveis.³⁷

Segundo Ferreira et al.³⁹ a ideia por trás da PCA é projetar dados multivariados em um espaço de menor dimensão, reduzindo a dimensionalidade do espaço original dos dados, sem que a relação entre os dados seja prejudicada. No livro, o autor fala que a metodologia possibilita descobrir, visualizar e interpretar diferenças existentes entre variáveis e examinar as relações que podem existir entre as amostras, podendo detectar amostras com componentes atípicos.

A PCA pode ser realizada por decomposição em autovalores (valores próprios) de uma matriz de covariância, geralmente depois de centralizar e normalizar a matriz de dados. Os resultados de PCA são avaliados em termos de pontuações (*scores*) de componentes ou pontuações de fatores (os valores de variável transformados correspondem a um ponto), e carregamentos (*loadings*), o peso pelo qual cada variável normalizada é multiplicada para se obter a componente. Com frequência a PCA pode ser reveladora da estrutura interna dos dados, de uma forma que melhor explica a variância nos dados. Quase todas as matrizes de dados podem ser simplificadas pela análise de componentes principais. Ela pode ser utilizada junto com um conjunto bem selecionado de variáveis para criar um modelo para entender como um sistema químico se comporta, podendo ser utilizada como uma previsão para

quando um novo dado é medido para o mesmo sistema.³⁸ Importante ressaltar que se as variáveis originais não forem correlacionadas, a análise não faz nada relevante, apenas ordena os dados de acordo com suas variâncias.³¹

Lyra et al,³⁵ em um trabalho com fins didáticos, criou um conjunto de dados de propriedades químicas dos elementos e utilizou a PCA nesses dados. Ele verificou a classificação dos elementos contidos na tabela periódica. A classificação de metais, semimetais e não metais surgiu naturalmente com a PCA e os dados encontrados são compatíveis com a tabela periódica oficial da SBQ/IUPAC. As propriedades que foram utilizadas para a análise dos elementos foram a primeira energia de ionização, o raio atômico, a afinidade eletrônica, a eletronegatividade de Pauling, a densidade, o calor específico, a entropia e a condutividade térmica.

Na Figura 1 temos escores das duas primeiras componentes principais. As novas coordenadas dos objetos nas novas variáveis são as componentes principais (PC), e os scores são as projeções no espaço dos objetos originais. É nesse gráfico que se deve procurar as relações entre os objetos analisados, no caso desse exemplo, os elementos de uma mesma classe ficam agrupados dentro de uma mesma região no gráfico de escores.³⁵ Nota-se que a PC1 modela 34% da variância da matriz de dados e a PC2 24%. Ou seja, 58% da variância está nas duas primeiras componentes principais. Logo, podemos simplificar a análise cortando a base em um número de variáveis menor, sem perda significativa de informações.³⁵

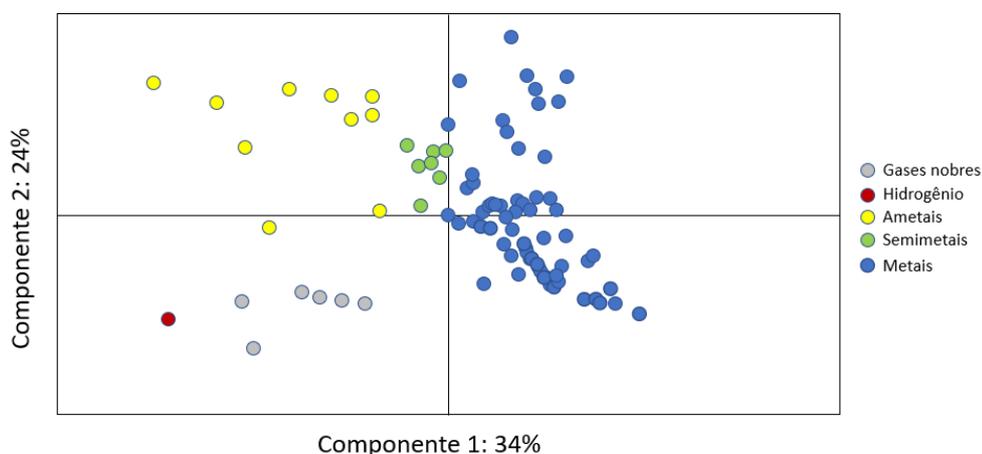


Figura 1. Gráfico dos scores de PC1 e PC2 para 88 elementos químicos. Fonte: adaptado de Lyra et al.³⁵

Na Figura 2 apresenta-se os pesos para a primeira e a segunda componente principal. Esses pesos correspondem, geometricamente, aos cossenos dos ângulos que as componentes fazem com as variáveis originais. A combinação linear dos pesos das variáveis que definem cada componente. Observando a PC1, nota-se que o raio atômico tem sinal contrário ao da eletronegatividade, da afinidade eletrônica e da primeira energia de ionização. O que faz sentido, uma vez que o raio atômico varia de maneira inversa com a variação dessas outras propriedades.³⁵ A entropia, que é uma propriedade termodinâmica, tem peso negativo elevado, o que ajuda a discriminar os metais, que possuem entropias mais baixas que os não metais.³⁵

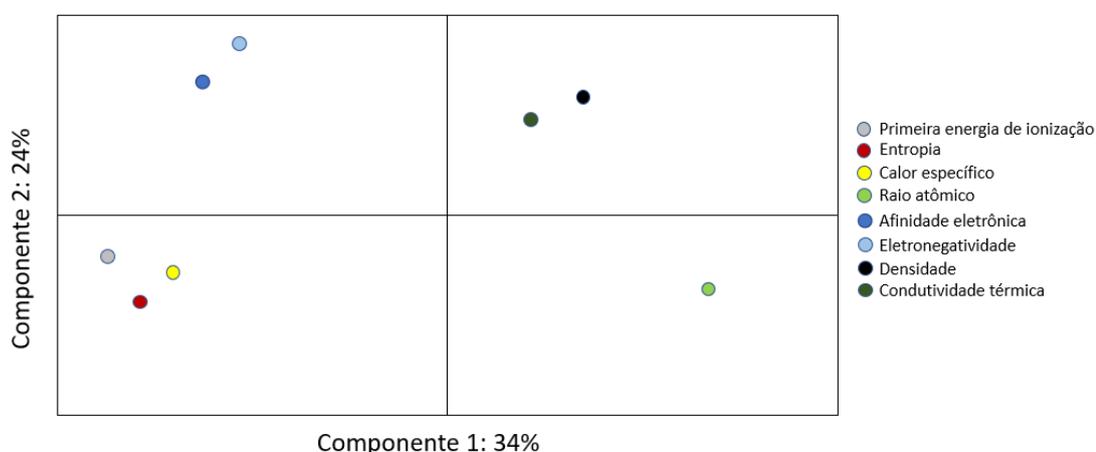


Figura 2. Gráfico de pesos da componente principal 1 (eixo x) e para componente principal 2 (eixo y) para oito propriedades. Fonte: adaptado de Lyra et al.³⁵

A PARAFAC, resumidamente, é um método decompositor para uma matriz de terceira ordem que pode ser vista como uma simplificação de um PCA bilinear para uma matriz de maior ordem. Aqui a decomposição dos dados é feita em componentes trilineares. Cada componente consiste em um vetor de informações e dois vetores de carregamento em vez de um vetor de informação e um de carregamento como no PCA.³¹

2.4.3 Análise de cluster (CA)

Enquanto a PCA é aplicada na impressão digital e outros conjuntos de dados para exploração da análise dos dados, o que reflete nas medições originais ao reconhecer os fatores dominantes, excluindo possíveis fatores de interferência, a análise de clusters (CA) foca em classificar os objetos baseando-se nas

características quantitativas.³² Ela classifica os componentes no espaço, com base nas similaridades. O resultado disso é um cluster de alta homogeneidade nos intergrupos e alta heterogeneidade para os diferentes grupos.³¹

Em muitos dos casos, a PCA consegue prover grupos de objetos parecidos, porém, nem sempre essa análise consegue fazer isso de maneira satisfatória, como mostrado na Figura 3. A CA é explicitamente utilizada para encontrar tais grupos. Nessa análise divide-se os grupos de objetos em classes onde objetos similares são da mesma classe, ela procura os objetos que estão próximos no espaço.⁴⁰

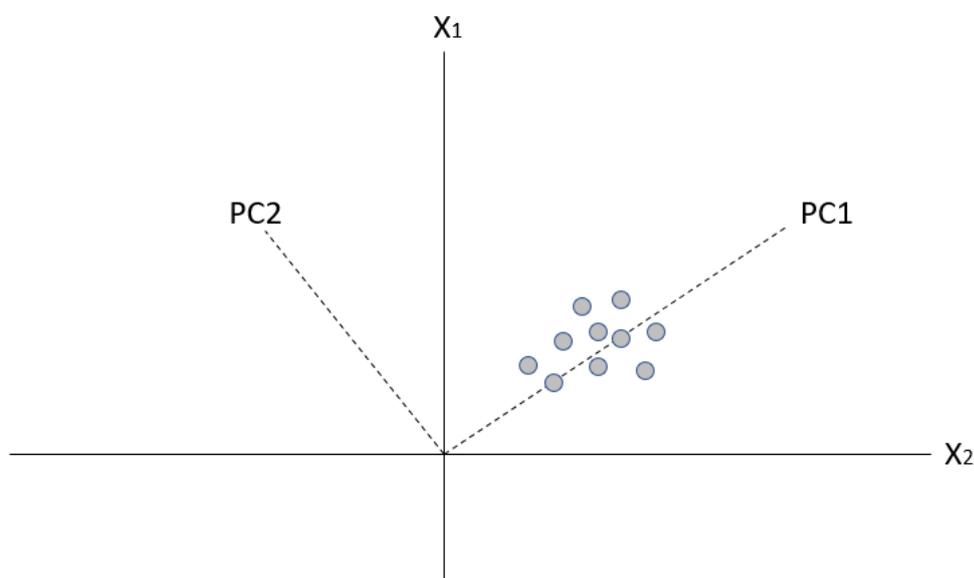


Figura 3. Exemplo onde a PCA não é capaz de diferenciar os grupos das amostras.

Adaptado⁴⁰

Existem dois tipos de análises de clusters, a hierárquica (HCA) e a não-hierárquica. Na química, a análise de clusters hierárquicos é um tratamento bastante recorrente para análise de matrizes complexas. A vantagem do HCA está na flexibilidade de alterar os critérios que o pesquisador pretende analisar, assim como a sua aplicação em métodos de vinculação para diferentes aplicações.³²

O HCA mostra as similaridades nos dados de uma amostra em relação a quaisquer outras amostras. Isso é expresso graficamente com um dendrograma, uma figura com diferentes grupos e sua proximidade por alta redução em dimensão aos dados originais. O gráfico é feito assumindo que as amostras comparadas, as quais têm valores numéricos parecidos, possuem alta proximidade nos valores espaciais.³¹

Uma maneira de entender o processo matemático é utilizando um número X de conjuntos de dados. Cada um desses conjuntos irá virar um cluster, logo, temos X clusters na análise. As distâncias de cada cluster representa a similaridade entre os dados. Cada par de clusters que estão mais próximos são emergidos em um único cluster e esse novo terá sua distância calculada em relação a todos os outros dados.³¹

As distâncias entre os objetos, que posteriormente irão criar os clusters, são normalmente calculadas através da distância euclidiana (ED),³¹ mas existem outras formas menos utilizadas. ⁴¹ ED é a distância entre dois pontos em n dimensões no espaço com coordenadas (x1, x2, ..., xn) e (y1, y2, ..., yn), sendo definida como a Equação 3.⁴⁰

O método de Ward, utilizado para calcular a distância entre os clusters, faz a soma dos quadrados dentro dos agrupamentos calculada sobre todas as variáveis. No procedimento de Ward, a seleção dos pares de agrupamentos a combinar é feita na combinação de agregados que minimizam a soma interna de quadrados no conjunto completo de agrupamentos separados. Em cada passo, os dois agrupamentos combinados são aqueles que minimizam o aumento na soma total de quadrados em todas as variáveis em todos os agrupamentos. Esse procedimento tende a combinar agrupamentos com um pequeno número de observações, pois a soma de quadrados está diretamente relacionada com o número de observações envolvidas.³⁴

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

Equação (3)

A Figura 4 mostra um exemplo de dendograma simplificado de um HCA. Pode-se perceber pequenos grupos de amostras com alta similaridade, como por exemplo as amostras 1, 2, 3 e as amostras 4, 5, 6. O primeiro exemplo possui maior similaridade entre os integrantes quando comparados com o segundo exemplo. Nessa análise hipotética, pode-se dividir essas amostras em dois grandes grupos que possuem alta semelhança, o primeiro, que vai da amostra 1 até a 6, e o segundo grupo, 7 a 13. Os dois grupos acabam se ligando em distâncias maiores, o que significa que existe baixa similaridade entre os dados desses dois grandes grupos.

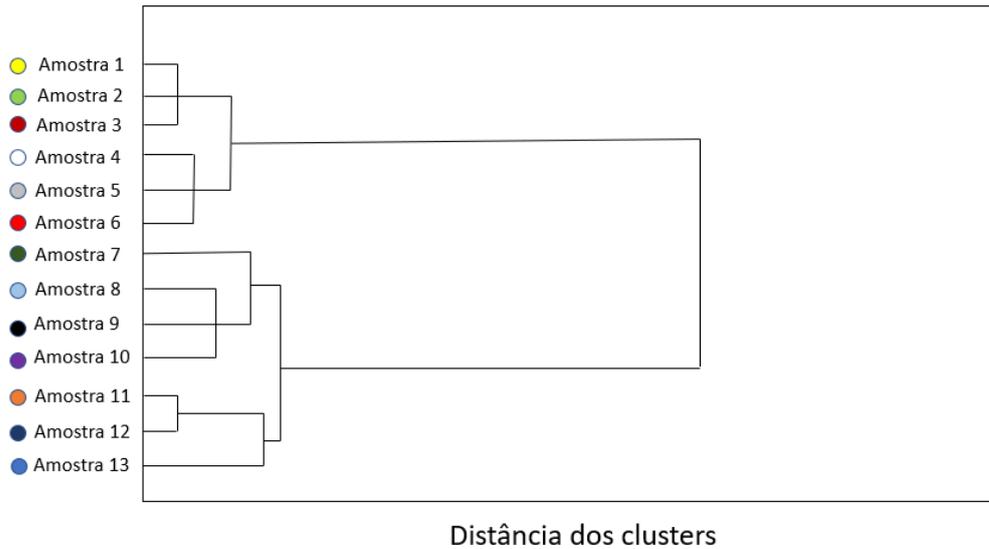


Figura 4. Dendrograma simulado como o gerado pelo HCA para 13 amostras

As similaridades são o critério para a separação ou agrupamentos das análises. Quanto mais similares são os dados, menor será a distância entre os clusters. As distâncias podem ser baseadas em uma única dimensão ou em várias, cada uma representa a condição de agrupamentos dos dados. Cabe ao pesquisador decidir qual dos métodos ele irá escolher, uma vez que o algoritmo não é capaz de saber se as relações são realmente próximas, ele apenas representa qual dessas similaridades são mais ou menos significativas.⁴¹ Essa análise não é estática, ou seja, quando se retira ou adiciona um novo conjunto de dados, todas as distâncias podem ser mudadas, o sistema todo pode gerar dados bem diferentes, e relações que antes eram grandes podem diminuir e vice-versa.³¹

3 Objetivos

3.1 Objetivo Geral

Utilizar ferramentas quimiométricas, mais especificamente a análise de componentes principais e análise de cluster hierárquico, para o tratamento de dados de concentração total de metais e semimetais de resíduos da indústria de óleo e gás determinados por espectrometria de massa com plasma indutivamente acoplado, visando encontrar similaridades e diferenças entre as amostras.

3.2 Objetivos Específicos

- Fazer o tratamento e uma organização prévia dos dados obtidos, selecionando as concentrações de As, Cr, Co, Ni, Fe, Al, Pb e Zn em distintas amostras de resíduos da indústria de óleo e gás para serem analisadas.
- Fazer o pré-processamento dos dados com o autoescalamento.
- Aplicar a análise de componentes principais nos dados.
- Utilizar a análise de cluster hierárquico nas informações.
- Correlacionar as duas análises, investigando se ambas as informações obtidas são concordantes entre si.

4. Materiais e métodos

Como este trabalho utiliza dados de relatórios analíticos associados à rotina do laboratório em convênio com a Petrobrás, nos anos de 2012 a 2018, reuniu-se todas as informações dos relatórios para melhor descrever esse tópico. Por falta de dados, muitos dos procedimentos e parâmetros são descritos de maneira genérica.

4.1 Instrumentação

Os dados foram obtidos com a utilização de um ICP-MS, modelo ELAN 6000 (Perkin Elmer, Thornhill, Canadá), equipado com nebulizador de fluxo cruzado e câmara de nebulização Scott, o argônio com pureza de 99,996% (White Martins, São Paulo, Brasil) como gás para geração do plasma e o forno digestor modelo Ethos Plus, (Milestone, Sorisole, Itália).

A análise de componentes principais e a análise hierárquica de clusters foram feitas utilizando o programa Statistica, versão 13.5.0.17 (StatSoft, Tulsa, USA) oferecido pela UFSC. Os dados foram pré-processados utilizando o autoescalamento. Durante a aplicação da PCA, utilizou-se a distância euclidiana e o método Ward para analisar as similaridades entre as amostras. Para criação de gráficos de concentração utilizou-se o Microsoft Excel 2019, versão 16.0 (Microsoft Corporation, Redmond, USA)

Para o uso de um ICP-MS, avaliações diárias faz-se necessárias para conferência do desempenho do instrumento. Os resultados devem ser comparados diariamente com parâmetros recomendados no manual do instrumento. Para essas avaliações utiliza-se uma solução padrão $10 \mu\text{g L}^{-1}$ de Ce, Ba, Mg, Pb e Rh em HNO_3 1% v/v.

Um bom desempenho do equipamento, é confirmado quando: As contagens maior ou igual a 140000 para Rh, 24000 Mg, 180000 para Pb e as razões CeO^+/Ce^+ e $\text{Ba}^{++}/\text{Ba}^+$ são menores que 3% ($< 0,03$). Quando necessário algum ajuste para atingir esses valores pode-se otimizar a radiofrequência e vazão do gás nebulizador, não sendo suficiente uma otimização completa pode ser requerida, envolvendo uma série de otimizações. Sendo assim, as condições de daily do instrumento Elan 6000 alocado no Departamento de Química foram: Potência da RF de 1100 W, vazão do gás nebulizador de 1.1 Lmin^{-1} , voltagem da lente iônica de 9.5 V, 40 leituras por replicata,

dwell time de 50 ms, cone amostrador/skimmier de Pt. O procedimento de avaliação da performance do ICP-MS deve ser feito diariamente ou antes da determinação.

4.2 Amostras e analitos

Os relatórios analíticos selecionados para esta avaliação, são de determinação de concentração total de metais e semi-metais em amostras de lodo, cascalho, catalisadores e borra, por espectrometria de massa com plasma indutivamente acoplado. Utilizou-se as concentrações de As, Cr, Co, Ni, Al, Pb, Zn e Fe, para as análises quimiométricas propostas. Ao total foram 32 amostras selecionadas, 12 amostras de cascalho, 4 amostras de lodo, 4 amostras de borra e 12 de amostras catalisadoras. Estas amostras foram analisadas nos anos de 2012 a 2018.

4.3 Reagentes

Todos os reagentes utilizados foram de no mínimo grau analítico de pureza. A água ultra-pura obtida no sistema Milli-Q (Millipore, Bedford, MA, EUA) a uma resistividade de 18,2 M Ω cm. O ácido nítrico 65% m/m (HNO₃, Vetec, Rio de Janeiro, Brasil) e o ácido clorídrico 37% m/m (HCl, Neon, São Paulo, Brasil) foram bidestilados abaixo da temperatura de ebulição em destilador de quartzo (Kürner Analysentechnik, Rosenheim, Alemanha). Ácido hidrofúrico 50% m/m (HF, Química Moderna, São Paulo, Brasil) foi destilado abaixo da temperatura de ebulição em um sistema de poli(tetra-fluoroetileno) (Kürner Analysentechnik, Rosenheim, Itália). O peróxido de hidrogênio 30% m/m (H₂O₂, Neon, São Paulo, Brasil) foi utilizado para a digestão das amostras. O ácido bórico (H₃BO₃, Neon, São Paulo, Brasil) foi utilizado após a digestão das amostras para complexar com o fluoreto, evitando danos aos acessórios de quartzo do ICP-MS. Por fim, os padrões monoelementares dos elementos de interesse são de diferentes fabricantes (Aldrich, Steinhein, Suíça e Spex Industries, Edison, USA)

4.4 Procedimentos analíticos

Sem tratamento prévio, as amostras foram homogeneizadas e, quando necessário, maceradas com um almofariz de ágata com pistilo, até se obter uma mistura pastosa e uniforme. Elas foram, então, sujeitas à digestão ácida auxiliada por microondas. Para a digestão, uma massa de 50 mg a 150 mg de cada amostra foi

separada, adicionando então 4 mL de HNO₃ 65% m/m, 1 mL de HCL 37% e 1 mL de H₂O₂ 30%. Deixou-se a mistura em repouso por 1 hora.

Após esse tempo, adicionou-se 0,5 mL de HF 50% para que então os resíduos fossem submetidos a dois programas de digestão. O primeiro programa é descrito pela Tabela 1. Após a realização do primeiro programa, esperou-se resfriamento completo. Adicionou-se 0,3 g de H₃BO₃ para que então as amostras fossem submetidas ao segundo programa de digestão, como descrito na Tabela 2.

Tabela 1. Primeiro programa potência de microondas para digestão das amostras

Etapas	Potência (W)	Tempo (min)
Primeira	250	2
Segunda	0	2
Terceira	250	6
Quarta	400	6
Quinta	650	6
Sexta	250	6

Tabela 2. Segundo programa de potência de microondas para digestão das amostras

Etapas	Potência (W)	Tempo (min)
Primeira	250	3
Segunda	0	2
Terceira	400	2
Quarta	500	2

Após os processos de digestão, as soluções foram transferidas para um frasco de polipropileno e avolumadas para 50 mL com água deionizada. Todas as amostras de resíduos, assim como as soluções branco, foram preparadas em triplicata. Por fim, todas as soluções foram diluídas adequadamente e as concentrações dos elementos estudados foram determinadas pelo ICP-MS.

4.5 Segurança no laboratório

Para segurança de todos, faz-se necessário o conhecimento das normas de segurança e boa conduta de laboratório. Ao entrar no laboratório deve-se procurar onde ficam as saídas de emergência, alarmes de incêndio, extintores, chuveiros e lava-olhos de emergência.

Equipamentos de proteção individual se fazem necessários quando os equipamentos de proteção coletiva não são o suficiente para garantir a integridade do químico. Equipamentos de proteção dos olhos e face também devem ser utilizados quando necessário, como o uso de óculos de segurança, óculos de proteção contra radiação, protetores faciais e proteção respiratória.

O uso de jaleco é inerente e obrigatório ao se entrar em um laboratório, uma vez que ele protege boa parte do corpo, formando um nível a mais de proteção contra derramamento de produtos químicos, pequenas explosões das vidrarias e contaminação por contato com equipamentos e objetos do laboratório. O Laboratório em que este trabalho foi realizado possui todos os EPI's narrados e foram usados corretamente.

4.6 Descarte e tratamento de resíduos

Resíduos ácidos são coletados em separado dos demais para posterior tratamento por empresa especializada, sempre que possível devem ser neutralizados, diluídos e descartados em recipientes devidamente identificados para posterior tratamento. Solventes orgânicos são descartados de forma separada dos resíduos ácidos para evitar reações secundárias, armazenados em containers devidamente etiquetados, separando solventes clorados de não clorados, até o dia da coleta. Soluções contendo halogênios são coletados em separado dos metais e solventes orgânicos, descartados em containers apropriados para soluções aquosas, ter cuidado ao misturar soluções contendo metais não inertes, compostos de alta reatividade e com pH's não neutro.

Os resíduos sólidos possuem algumas especificações a serem seguidas, deve-se ter disponível um recipiente forrado com saco plástico para armazenagem.

Segundo o site gestaoderesiduos.ufsc.br a coleta dos resíduos químicos da Universidade Federal de Santa Catarina é programada junto à empresa Proactiva, que possui contrato vigente com a universidade. A coleta acontece uma vez ao mês, com calendário pré estabelecido. A maioria dos resíduos, sendo eles sólidos ou líquidos, tem como destino final um aterro industrial controlado. A incineração é prevista para objetos perfurocortantes e solventes halogenados. As lâmpadas são descontaminadas, para solventes não halogenados faz-se a blendagem para coprocessamento.

5 Resultados e discussão

5.1 Origem e seleção dos dados

Primeiramente, foram reunidos 12 relatórios que o LEMA fez devido ao convenio entre o laboratório e a Petrobras, do período de 2012 a 2018. Aqui, diferentes resíduos e procedimentos eram feitos em cada um desses laudos. Após uma breve análise, percebeu-se que a metade dos relatórios eram de análise quantitativa da concentração total dos elementos e faziam uso do ICP-MS. Os demais eram de diferentes metodologias e utilizavam diferentes aparelhos.

Optou-se pelos laudos de concentração total dos elementos com ICP-MS, uma vez que eu já possuía uma familiaridade com tal equipamento e esses laudos possuíam a maior quantidade de dados passíveis de implementar análises quimiométricas.

Dentre os seis trabalhos pré-selecionados, fez-se mais uma vez uma análise superficial dos dados. Percebeu-se que em alguns deles não existiam valores para alguns dos analitos ou muitos dos metais analisados possuíam concentrações abaixo dos limites de detecção da técnica utilizada. Ao final, optou-se por quatro relatórios que analisaram lodo prensado, cascalho de perfuração, catalisadores e borras. Os limites de detecção para os analitos, retirados dos laudos encontram-se mostrados na Tabela 3. Os resultados de LOQ para catalisadores e lodos não são apresentados pois os relatórios analíticos não possuíam tais informações. Os resultados de concentração para os analitos: As, Cr, Co, Ni, Fe, Al, Pb e Zn, nas diferentes matrizes encontram-se mostradas na Tabela 4.

Tabela 3. Dados de limite de detecção (LOD) e limite de quantificação (LOQ) existentes nos relatórios

	Catalisadores	Lodo	Borra		Cascalho	
	LOD ($\mu\text{g g}^{-1}$)	LOD (mg kg^{-1})	LOD ($\mu\text{g L}^{-1}$)	LOQ ($\mu\text{g L}^{-1}$)	LOD (mg kg^{-1})	LOQ (mg kg^{-1})
As	0,1	2	0,09	0,1	0,03	0,1
Cr	0,2	0,5	0,09	0,3	0,1	0,34
Co	0,011	0,03	0,02	0,07	0,01	0,03
Ni	0,032	0,7	0,03	0,1	0,33	1,09

Tabela 3. Continuação: Dados de limite de detecção (LOD) e limite de quantificação (LOQ) existentes nos relatórios

Fe	3	18	1,08	3,56	6,06	20
Al	0,6	0,7	0,18	0,6	10,1	33,4
Pb	0,016	0,09	0,03	0,1	0,01	0,03
Zn	0,049	0,7	0,15	0,5	0,07	0,23

Tabela 4. Dados de concentração dos analitos (em mg kg⁻¹) nas diferentes amostras de resíduos da indústria de petróleo e gás do laboratório LEMA/LARES do Departamento de Química da UFSC.

	As	Cr	Co	Ni	Fe	Al	Pb	Zn
Cat 1	1,19	36,20	148,00	2780,00	3459,00	223001	16,60	48,54
Cat 2	0,58	27,00	54,00	1141,00	3039,00	202773	17,20	49,40
Cat 3	4,40	42,00	161,00	5476,00	3586,00	190183	22,20	133,00
Cat 4	1,40	42,00	120,00	2184,00	4714,00	285097	20,80	86,00
Cat 5	0,53	41,00	49,59	1133,00	4150,00	237196	14,30	37,40
Cat 6	1,50	41,00	169,00	3959,00	5228,00	218130	29,80	85,00
Cat 7	0,51	77,00	35,24	1173,00	2778,00	196384	10,10	20,00
Cat 8	1,59	28,90	109,40	1542,00	4381,00	190703	21,60	150,00
Cat 9	0,39	33,00	32,00	845,00	3475,00	209838	12,00	29,00
Cat 10	1,20	30,90	17,30	429,00	3629,00	205154	12,40	25,00
Cat 11	0,32	29,20	17,71	364,00	2805,00	182626	12,10	24,20
Cat 12	0,41	29,00	47,00	602,00	3707,00	160931	19,00	89,00
Lodo 1	339,00	216,00	36,50	313,00	2,73	2,94	154,00	0,36
Lodo 2	309,00	232,00	38,30	346,00	3,04	3,32	169,00	0,35
Lodo 3	310,00	181,00	39,40	341,00	2,80	3,22	162,00	0,40
Lodo 4	309,00	206,00	35,20	300,00	2,76	3,03	161,00	0,38
Borra 1	31,10	27,11	5,76	38,74	17742,31	62691,83	13,04	501,25
Borra 2	61,38	35,32	7,62	45,05	23032,15	76655,53	14,95	595,67
Borra 3	115,07	44,80	15,33	99,83	19428,07	38376,65	31,17	1813,96
Borra 4	82,30	42,34	14,04	102,41	17400,54	33726,96	26,23	1818,53

Tabela 4. Continuação: Dados de concentração dos analitos (em mg kg⁻¹) nas diferentes amostras de resíduos da indústria de petróleo e gás do laboratório LEMA/LARES do Departamento de Química da UFSC.

Casc.1	8,70	40,30	4,00	0,33	6,06	7887,00	11,80	406,80
Casc.2	9,80	36,10	3,60	0,33	6,06	7260,00	19,10	414,50
Casc.3	10,60	63,00	3,54	0,33	6,06	2879,00	30,20	607,00
Casc.4	8,40	57,80	3,64	0,33	6,06	2688,00	96,60	1085,70
Casc.5	9,70	70,20	3,36	0,33	6,06	2692,00	67,50	588,70
Casc.6	10,60	41,00	3,72	0,33	6,06	7113,00	16,00	462,00
Casc.7	7,20	43,40	5,47	0,33	6,06	11885,00	6,61	682,00
Casc.8	9,40	64,80	6,65	0,33	6,06	14883,00	5,82	570,50
Casc.9	5,60	38,90	1,19	0,33	6,06	1772,00	4,14	462,08
Casc.10	10,00	42,40	1,60	0,33	6,06	2706,00	3,44	706,70
Casc.11	6,50	68,20	3,95	0,33	6,06	7270,00	8,72	950,80
Casc.12	7,20	85,10	6,04	0,33	6,06	14582,00	11,52	1288,30

5.2 Estatística básica

Através da média e desvio das concentrações dos analitos nas quatro diferentes matrizes pode-se obter valores de concentrações bem diferentes entre os analitos e matrizes, com a heterogeneidade das matrizes desvios altos foram observados. Pouca informação elucidativa é possível desta forma, conforme pode-se observar na Figura 5.

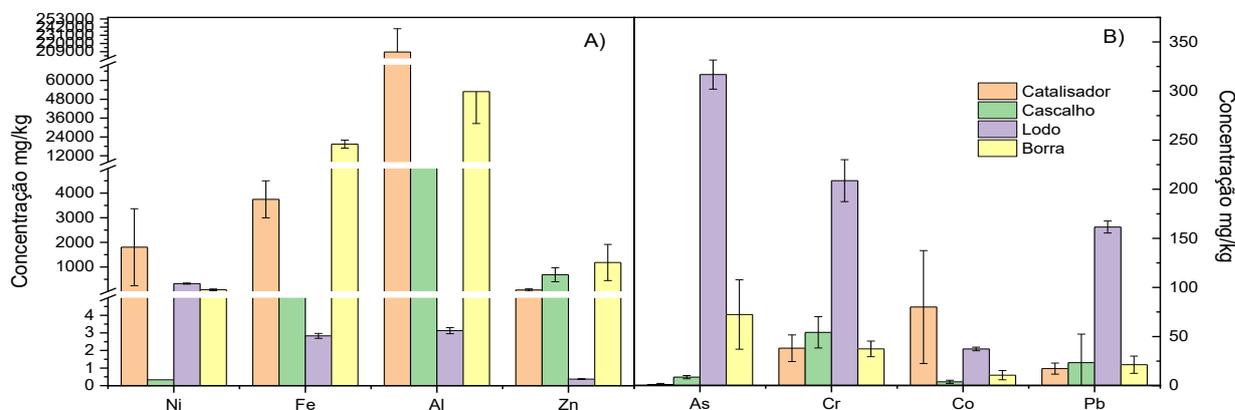


Figura 5. Concentrações médias e desvios dos analitos A) Ni, Fe, Al e Zn e B) As, Cr, Co e Pb nos diferentes resíduos da indústria de petróleo e gás.

Observando a Figura 5 é possível retirar algumas informações. Percebe-se, por exemplo, que os catalisadores possuem uma grande quantidade de alumínio e concentração baixa de arsênio. Para as borras são observadas altas concentrações de alumínio e ferro e quantidades moderadas dos outros analitos. As amostras de lodos têm quantidades de zinco muito pequenas e concentrações baixas de alumínio e ferro. Por fim o cascalho que possui alumínio e zinco como analitos em maiores concentrações, juntamente com baixas quantidades de cobalto, níquel e ferro.

Com os dados no gráfico poucas informações são possíveis de serem concluídas do sistema estudado, é muito difícil fazer qualquer tipo de relação entre os analitos e os resíduos. Sabemos quais são os elementos em maiores e menores concentrações, percebemos que alguns deles possuem altos e baixos desvios padrões, mas não conseguimos ver muitas tendências, similaridades ou fazer previsões sobre os dados.

5.3 Ferramentas quimiométricas

Para um estudo mais aprofundado desses dados, escolheu-se aplicar duas ferramentas quimiométricas muito utilizadas pela química e outras áreas que trabalham com dados, a PCA e a HCA. Com essas análises espera-se encontrar tendências e similaridades entre os resíduos, explicando como essas ferramentas matemáticas chegaram a essas informações.

5.3.1 Autoescalamento dos dados

Como muitas vezes as propriedades estudadas apresentam variáveis com significados físicos e unidades diferentes, é necessário fazer um pré-processamento dos dados para utilizá-los em PCA e HCA. O autoescalamento permite que as variáveis exerçam influências igualitárias e as torna adimensionais.³⁵ O Apêndice I apresenta os dados após o autoescalamento automaticamente calculados pelo software Statistica.

Após o autoescalamento, os valores estão em uma mesma escala e a eles foram atribuídos pesos relacionados aos valores de concentração de todos os analitos e de todos os resíduos. O autoescalamento foi efetuado com todos os dados dos resíduos juntos.

5.3.2 Análise de componentes principais

O primeiro método quimiométrico aplicado a tais informações foi a PCA. O objetivo aqui é encontrar inter-relações entre essas variáveis e explicar em termos de suas dimensões inerentes comuns. Aqui temos um meio de condensar as informações presentes nos dados originais em um conjunto menor de variáveis estatísticas, perdendo o mínimo de informação relevante. A Figura 6 mostra o gráfico de score da PCA para as amostras estudadas.

Observando a Figura 6, pode-se perceber as novas coordenadas (as componentes principais ou “fatores”) e as projeções no espaço dos objetos originais. A primeira informação relevante é que os componentes principais 1 e 2 modelam 42,32% e 32,66% da variância da matriz de dados, respectivamente. Ou seja, quase 75% da variância dos dados está contida nas PC 1 e 2, mostrando que temos a maioria das informações contidas nessas duas dimensões.

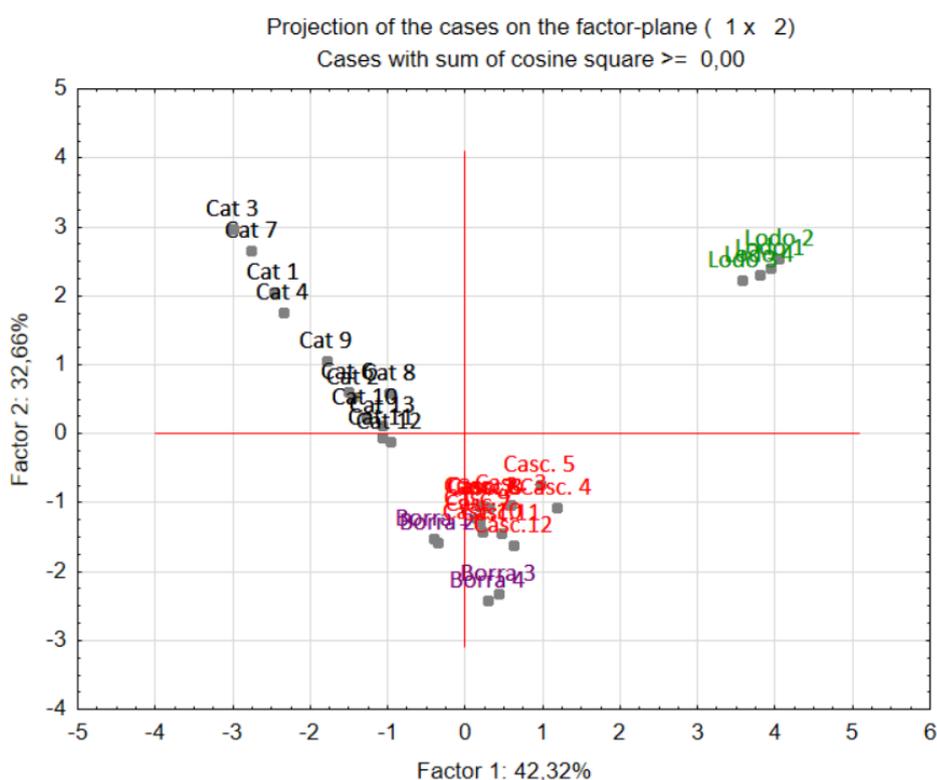


Figura 6. Gráfico dos escores de PC1 e PC2 para 34 amostras de resíduos provenientes da indústria de petróleo e gás.

Conforme observado, os dados de um mesmo resíduo tendem a se agrupar, ocupando a mesma região no gráfico de escores. Enquanto os catalisadores estão no

quadrante negativo da PC1 e positivo da PC2, os lodos estão mais à direita, com valores positivos na PC1. As borras e os cascalhos se apresentam com valores de PC1 quase nulos e levemente negativos para a PC2.

É importante fazer uma rotulagem dos objetos no gráfico de escores, como a mudança de cores nos nomes de cada tipo de amostra, pois muitas vezes a rotulagem adequada permite certos padrões aparecerem.³⁵ Esse passo foi importante para tais dados, já que os cascalhos e as borras se confundiam antes desse procedimento. Com a rotulagem, pode-se observar um padrão entre essas duas amostras, mesmo que esse seja menos evidente do que os vistos para os catalisadores e os cascalhos.

Deve-se inferir algumas interpretações para as componentes principais no gráfico de pesos na Figura 7, onde observa-se a relação entre as variáveis que permitiram a separação dos grupos na PCA. Nota-se que cromo, chumbo e arsênio estão com pesos bem parecidos e bem separados aos demais analitos, percebe-se que o lodo é a amostra com maior quantidade desses analitos, sendo estes os responsáveis pela posição do agrupamento das amostras de lodo. Da mesma forma, Ni, Co e Al caracterizam de forma específica as amostras de catalisadores.

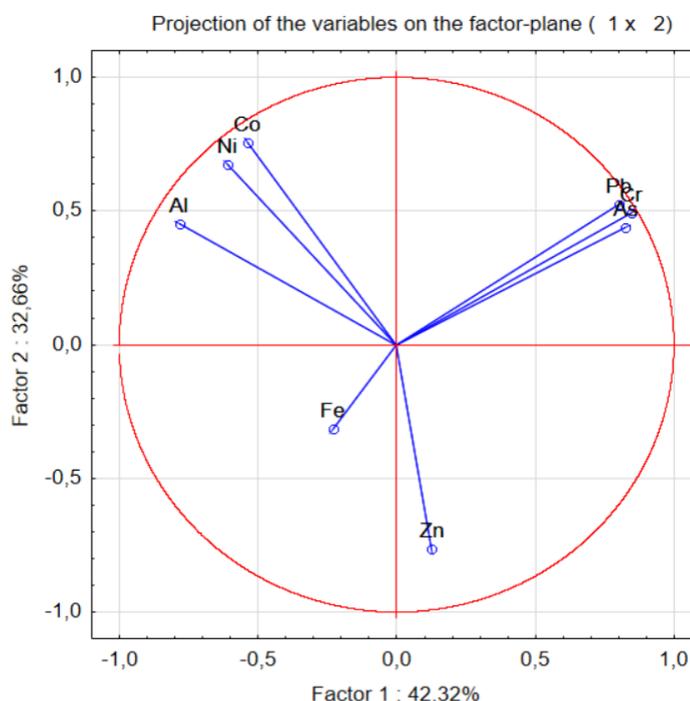


Figura 7. Gráfico de pesos (loadings) da componente principal 1 e para componente principal 2 para oito concentrações de analitos.

O alumínio se mostra como o maior contribuinte para a localização nos escores dos catalisadores. Esses resíduos têm quantidades muito maiores de Al do que qualquer outro resíduo. Percebe-se também que os catalisadores estão bem espalhados pelo gráfico de escores, pode-se relacionar isso aos altos desvios padrões de Al, Fe e Ni. Por exemplo, a quantidade de Ni no catalisador 3 é de 5476 ug g^{-1} , enquanto no catalisador 12 é de 364 ug g^{-1} . Percebe-se que a amostra “Cat 3” está bem mais próxima do peso do Ni enquanto a “Cat 12” deslocou-se em direção do peso do Fe, uma vez que a diferença entre as concentrações de ferro é bem menor. O Fe está exercendo um peso muito maior ao catalisador 12 do que ao 3.

Como catalisadores possuem uma concentração muito maior de Al do que qualquer outro resíduo, o peso desse elemento nos outros resíduos acaba não sendo tão relevante. Por isso a posição do cascalho tem muita influência do Zn, uma vez que esse elemento está presente em maior quantidade nessa amostra, logo depois do Al. Isso estende-se à borra, ela possui como analito em maior concentração o Al, mas como os catalisadores acabaram diminuindo a influência desse metal nos demais analitos, o ferro acaba sendo o fator mais decisivo para a localização da borra no gráfico de escores.

Como as concentrações de Fe, para o cascalho, foram muito baixas, abaixo do limite de detecção, sendo estes os valores colocados nos cálculos de autoescalamento, a separação entre o cascalho e a borra pode ter sido afetada. Se fosse possível ir ao laboratório, fazer um processo analítico diferente, tratando a amostra, a pré-concentrando ou utilizando outra técnica, como a espectrometria de absorção atômica de alta resolução com o forno de grafite, poderia-se inferir valores mais adequados dessas concentrações de Fe. Isso possivelmente faria o cascalho ter ainda mais influência do zinco e menos do ferro, separando de maneira mais satisfatória os dois resíduos.

Com as informações obtidas com a PCA fica muito mais fácil olhar novamente para os dados das médias das concentrações e encontrar tendências e similaridades entre tais informações. A PCA possibilita, de maneira simples, perceber quais analitos possuem tendências parecidas e, agora que se sabe quais são os dados semelhantes, fica muito mais fácil tentar encontrar essas tendências em um gráfico simples como o

da Figura 8 para buscar entender como essa análise quimiométrica chegou à essas conclusões.

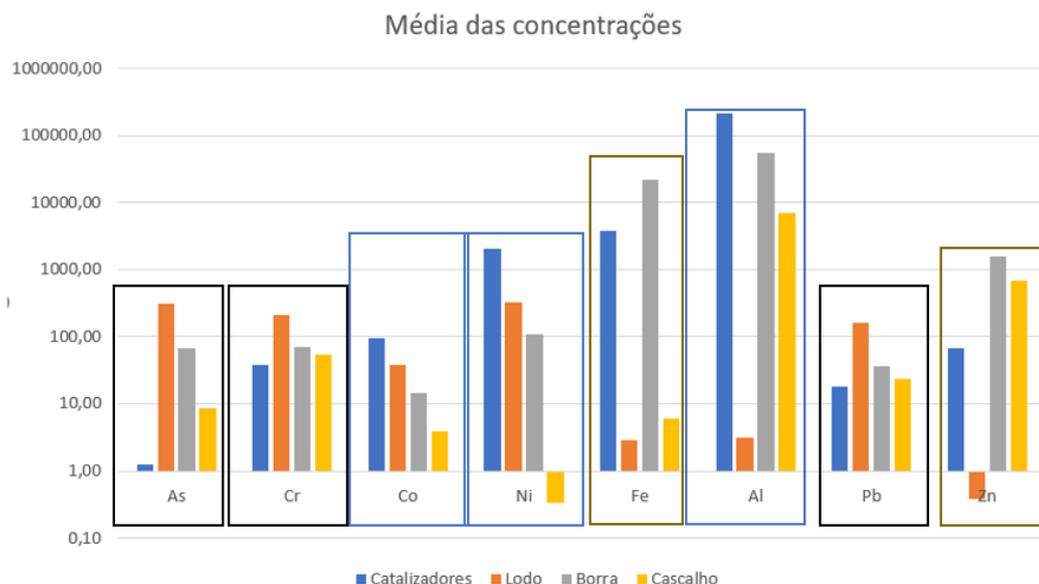


Figura 8. Gráfico das médias das concentrações (mg.kg⁻¹) dos analitos.

Para alumínio, níquel e cobalto também se observa um padrão parecido com o descrito acima, onde os catalisadores são as amostras com maior quantidade desses metais, logo depois o lodo, a borra e o cascalho. Porém, percebe-se que o Al quebra essa tendência para o lodo, isso pode explicar o porquê deste elemento estar levemente separado dos dois outros analitos.

Ferro e zinco estão bem separados dos demais analitos e ambos estão na parte de baixo do gráfico de pesos. Ao se analisar esses dois analitos no gráfico com as tendências, percebe-se que eles não possuem uma tendência tão evidente quanto às outras comparações. Pode-se notar também que nos desvios padrões (Figura 5), o desvio de Fe no cascalho é zero, enquanto no Zn ele é significativo. Esses pontos explicam por que ambos estão separados entre si e entre os demais analitos.

Observando novamente o gráfico de escores (Figura 6) e comparando com os demais gráficos, outras interpretações podem ser feitas. Vê-se que o lodo está em uma parte bem distinta no gráfico de escores. Essa posição pode ser explicada pela quantidade de Pb, Cr e As que, nesse resíduo, são próximas e juntas representam a maior parte dos analitos presentes nessas amostras.

5.3.3 Análise hierárquica de clusters

Muitas vezes a PCA consegue prover grupos de objetos parecidos, porém, nem sempre essa análise consegue fazer isso de maneira satisfatória, como no caso do cascalho e da borra. A análise de clusters hierárquica é explicitamente utilizada para encontrar tais grupos. Nessa análise divide-se os grupos de objetos em classes onde objetos similares são da mesma classe, ela faz isso juntando os objetos que estão próximos no espaço.⁴⁰

A Figura 9 apresenta o dendrograma formado pela HCA quando aplicada aos dados autoescalados. O método de Ward foi utilizado para calcular a distância entre os clusters e a distância euclidiana foi escolhida para calcular as distâncias dos objetos (amostras). No eixo y temos a distância entre os clusters. O $Dlink/Dmax*100$ está dizendo que a distância entre clusters está em porcentagem e que quanto mais perto de zero, menor é a distância entre clusters, logo, maior é a similaridade entre eles. Valores perto de 100 significam que pouca/nenhuma similaridade é encontrada.

Observando o dendrograma, pode-se notar uma separação bem evidente entre as amostras de resíduos. Com cerca de 61% de similaridade (tracejado azul) têm-se a separação das quatro distintas matrizes de resíduos. Percebe-se, que o cluster o qual engloba os lodos possui uma similaridade de aproximadamente 97%, marcado pelo tracejado verde, demonstrando que essas amostras apresentaram uma melhor homogeneidade que as demais matrizes. Os cascalhos e as borras apresentam um percentual alto de similaridade entre si. Os cascalhos até se separam em distância muito próximas de zero, mas o cluster formado por todos os cascalhos tem uma similaridade entre os dados de aproximadamente 87%, sinalizado pelo tracejado roxo. As borras também se separam um pouco em pequenas distâncias, mas o cluster formado por elas ainda mostra alta similaridade entre elas, cerca de 83% (tracejado cinza). Por fim, o cluster formado por todos catalisadores mostra uma similaridade de 60%, demarcado com o tracejado vermelho escuro, um valor menor do que os outros, porém ainda é um valor muito satisfatório.

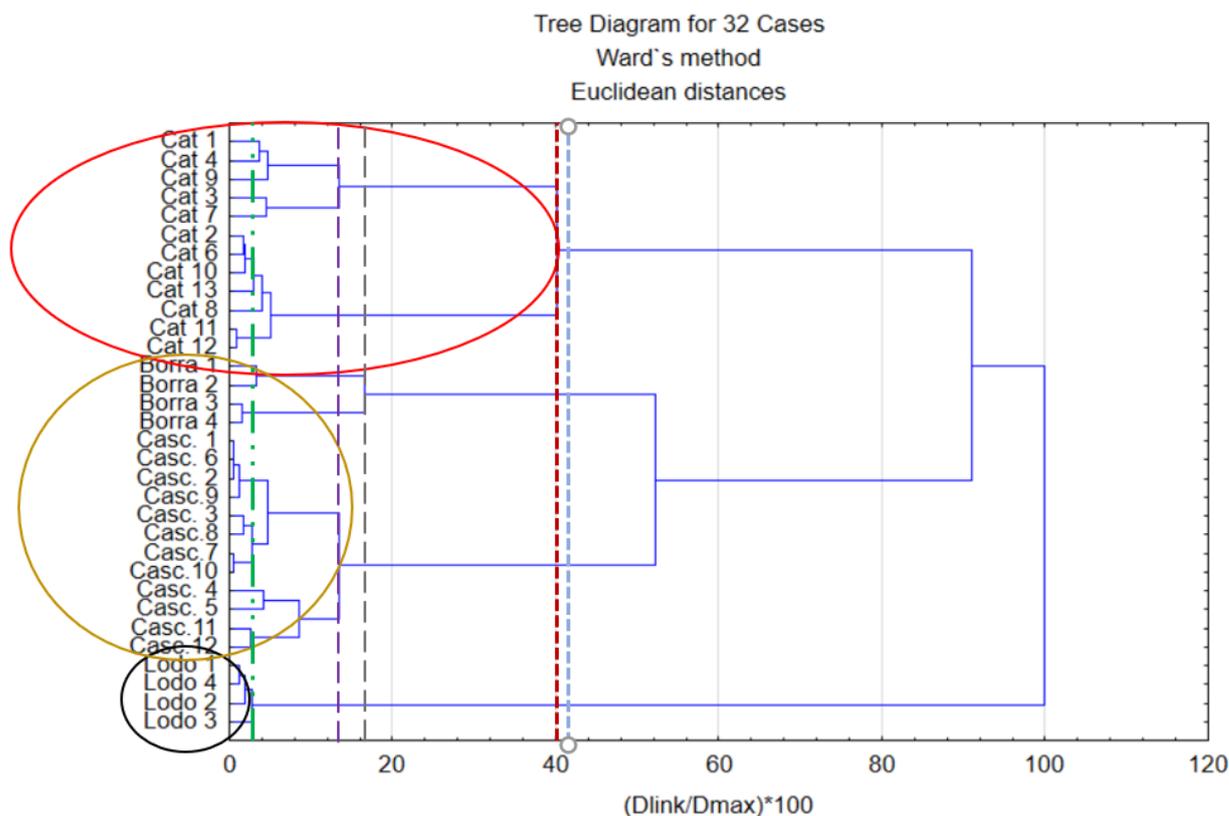


Figura 9. Dendrograma gerado pelo HCA para 4 amostras e 32 casos.

Normalmente, os químicos utilizam a HCA como complemento à PCA. Por isso, comparações podem e devem ser feitas entre a Figura 6 e a Figura 9. Percebe-se que em ambos os casos temos grupos definidos que englobam amostras dos mesmos resíduos, porém o HCA conseguiu distinguir muito melhor o grupo das borras do grupo dos cascalhos, mesmo com os clusters desses dois resíduos mostrando similaridade de mais ou menos 50%. Pode-se ainda perceber como os lodos são demonstrados pelos dois métodos como os resíduos que mais se diferenciam entre os outros. Como na PCA, a HCA também mostra que, mesmo dentro de um grupo bem definido, os catalisadores possuem certas diferenças entre si, evidenciando esta amostra como heterogênea.

Por fim, observou-se que em 53% de similaridade, todos os resíduos estão em seus respectivos clusters, mostrando e reafirmando o que a PCA havia mostrado: é possível diferenciar cada conjunto de amostras apenas tratando matematicamente os dados de concentrações e analisando suas tendências.

Analisando a Figura 10, pode-se ver como a HCA classificou as semelhanças entre as variáveis. Comparando-a com a Figura 7 pode-se fazer algumas observações reafirmando mais tendências encontradas pela PCA anteriormente.

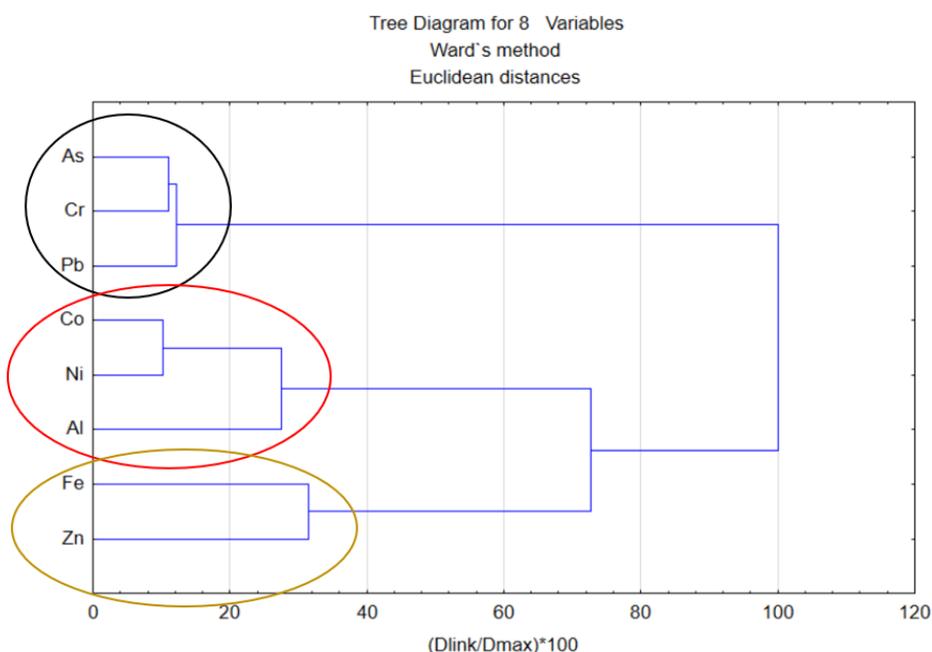


Figura 10. Dendrograma comparando as similaridades entre as concentrações dos analitos.

Nota-se aqui que, igual na PCA, os metais As, Cr e Pb estão com dados semelhantes entre si. Com o auxílio da Figura 10, olhando para o círculo preto, pode-se perceber que esses três metais são responsáveis pela separação do lodo aos demais resíduos, uma vez que o cluster formado pelos lodos (Figura 9) tem basicamente o mesmo comportamento que o cluster formado pelos três metais.

Co, Ni e Al são também muito semelhantes, com o Al mostrando uma diferença um pouco maior do que os demais metais. Percebe-se que o cluster formado por todos os catalisadores, representado no círculo vermelho, é muito parecido com o cluster formado pelos três metais da Figura 10, também identificado pelo círculo vermelho. Aqui pode-se confirmar que o alto desvio padrão dos três metais, em especial o do Al, é um dos motivos das diferenças observadas nos catalisadores.

Por fim, Fe e Zn, demarcados pelo círculo amarelo, estão em um mesmo cluster, porém a similaridade deles é menor do que as dos outros grupos formados pelos metais, exatamente como foi visto na PCA. Novamente percebe-se que esse

par de analitos exerce grande influência ao cascalho e à borra, e que a são as diferenças entre os metais que possibilitam a separação entre borra e cascalho.

Deve-se notar que muito pouco foi possível descrever sobre as semelhanças e tendências das amostras apenas com os gráficos das médias e dos desvios padrões. Após a aplicação da PCA, foi possível ter pontos mais específicos a serem analisados nos dois gráficos de estatística básica, auxiliando em muito a percepção sobre algumas possíveis tendências. Por fim, a HCA vem como uma prova real, melhorando a separação de alguns grupos e reafirmando as percepções feitas com a PCA.

6 Conclusão

Com uso de ferramentas quimiométricas, foi possível fazer um estudo detalhado sobre as informações de relatórios analíticos que continham informações sobre a concentração de As, Cr, Co, Ni, Fe, Al, Pb e Zn em quatro resíduos da indústria de petróleo e gás.

Com o autoescalamento, pode-se padronizar todos os dados obtidos a partir da espectrometria de massa com plasma indutivamente acoplado, permitindo que as variáveis exerçam influências igualitárias e as tornando adimensionais. Com as informações pré-processadas, aplicou-se a análise de componentes principais e a análise de cluster hierárquico, a fim de descobrir, visualizar e interpretar diferenças existentes entre as variáveis e examinar as relações que podem existir entre as amostras.

A PCA pode modelar 75% da variância dos dados nas componentes principais 1 e 2, ou seja, pode-se observar a maioria das informações contidas em todos os dados em duas dimensões. Com isso, constatou-se que cada tipo de resíduo se agrupou em diferentes posições das novas coordenadas criadas (PC1 e PC2).

Com o gráfico de pesos, foi possível identificar a influência que as quantidades de Pb, Cr e As tiveram sobre a separação do lodo das demais amostras. Evidenciou-se que o alumínio foi o maior contribuinte para a localização dos catalisadores, e que os altos desvios padrões nas concentrações de Al, Fe e Ni explicam o espalhamento desse resíduo pelo gráfico de escores. Por fim, mostrou-se que as posições do cascalho e das borras foram influenciadas fortemente pelos dados de Fe e Zn e como dados mais aprofundados das concentrações de ferro poderiam contribuir para uma melhor separação entre os grupos dos cascalhos e os grupos das borras.

O dendrograma formado pela HCA mostrou uma clara divisão entre as amostras estudadas, onde cada tipo de resíduo formou um cluster de alta similaridade. Essa análise ressalta as evidências descritas a partir da PCA e conseguiu distinguir de maneira mais satisfatória o grupo das borras do grupo dos cascalhos. O dendrograma comparando as similaridades entre as concentrações dos analitos melhorou a maneira de relacionar quais analitos foram mais relevantes para as

tendências apresentadas, reforçando mais uma vez as conclusões feitas a partir da PCA.

Com esse estudo foi possível mostrar como a quimiometria pode auxiliar nas compreensões e interpretações que um conjunto de dados pode conter. Apenas com tratamento matemático e dados de concentrações de 8 analitos foi possível separar quatro tipos diferentes amostras em seus respectivos grupos, identificando quais analitos tiveram o maior impacto na formação de cada um desses grupos.

Estudos com a aplicação destas ferramentas quimiométricas permitem modelar o comportamento das diferentes matrizes continuamente analisadas no laboratório de forma a possibilitar uma avaliação uma amostra desconhecida e visualizar uma forma de rastrear a qual tipo de resíduo ela pertence. Este tipo de ferramenta é comumente utilizado em estudos de identificação de origem, localização, ou numa outra esfera de aplicação estudos de identificação de adulteração em amostras de diversas matrizes, demonstrando a potencialidade das ferramentas e a importância de aprendê-las.

7 Referências

¹TORRES, E. M. A evolução da indústria petroquímica brasileira. **Química Nova**, 20, 1997.

²GUIMARAES, A. K. V.; et al. Estudo da caracterização da borra de petróleo e processo de extração do óleo. **Eng. Sanit. Ambient.**, 21, 265-274, 2016.

³JARDIM, W. F. Gerenciamento de resíduos químicos: Biossegurança Aplicada a Laboratórios e Serviços de Saúde. **Atheneu**, 167-175, 2004.

⁴DRUZIAN, T. G.; GIESBRECHT, C. K. Determinação simultânea de metais e enxofre por icp-oes em resíduos de destilação de petróleo. **Quim. Nova**, 39, 1065-1070, 2016.

⁵LIENEMANN, C. P.; DREYFUS, S.; PECHEYRAN, C.; DONARD, O. F. X. Trace Metal Analysis in Petroleum Products: Sample Introduction Evaluation in ICP-OES and Comparison with an ICP-MS Approach. **Oil & Gas Science and Tech.**, 62, 69-77, 2007.

⁶SOIN, A. V.; MARYUTIMA, T. A.; SPIVAKOV, B. A. Sample Preparation in the Determination of Metals in Oil and Petroleum Products by ICP MS. **Jornal of Anal. Chem.**, 65, 571-576, 2010.

⁷ABNT, NBR. 10004: 2004. **Resíduos sólidos: Classificação. Associação Brasileira de Normas**, 2004.

⁸FIALHO, P. F. **Cascalho de perfuração de poços de petróleo e gás. estudo do potencial de aplicação em concreto**. 2012. Dissertação (Mestrado em Engenharia Civil) - Centro Tecnológico, Universidade Federal do Espírito Santo, ES.

⁹CORREIA, D.; REGIS, C.; BORGES, H.; JUVINIANO, D. M.; CARLENISE, M.; ALENCAR, P. D et al. Analysis of the use of microemulsion systems to treat petroleum sludge from a water flotation unit. **Jornal of Environ. Chem. Eng.**, 7, 2019.

¹⁰HU, G.; LI, J.; ZENG, G. Recent development in the treatment of oily sludge from petroleum industry: A review. **Journal of Hazardous Materials**, 261, 470–490. 2013.

¹¹XU, N.; WANG, W.; HAN, P.; LU, X. Effects of ultrasound on oily sludge deoiling, **J. Hazard. Mater.**, 171, 914-917, 2009.

¹²EMAM, E. A. Clays as catalysts in petroleum refining industry. **J. of Science and Tech.**, 3, 4, 2013.

¹³ANABTAWI, J. A.; REDWAN, D. S.; AL-JARALLAH, A. M.; AITANI, A. M. Advances in the chemistry of catalytic reforming of naphtha. **Fuel Science and Technology International**, 9, 1–23, 1991.

¹⁴PRÖFROCK, D.; PRANGE, A. Inductively coupled plasma–mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: A review of challenges, solutions, and trends. **Applied Spectroscopy**, 66, 843-868, 2012.

¹⁵FONTANA, K. B.; ARAUJO, R. G. O.; OLIVEIRA, F. J. S. et al. Rare earth elements in drill cutting samples from off-shore oil and gas exploration activities in ultradeep Waters. **Chemosphere**, 263, 127984, 2021.

¹⁶HOUK, R. S. Mass spectrometry of inductively coupled plasmas. **Anal. Chem.**, 58, 97-105, 1986.

¹⁷JARVIS, K. E.; GRAY, A. L.; HOUK, R. S. **Handbook of Inductively Coupled Plasma Mass Spectrometry**. 1 ed., London: Blackie, 1992.

¹⁸MONTASER, A. **Inductively Coupled Plasma Mass Spectrometry**. 1 ed., New York: Wiley-VCH, 1998.

¹⁹PROFROCK, D.; LEONHARD, P.; WILBUR, S.; PRANGE, A. Sensitive, simultaneous determination of P, S, Cl, Br and I containing pesticides in environmental samples by GC hyphenated with collision-cell ICP-MS. **J. Anal. At. Spectrom.**, 19, 623-631, 2004.

²⁰POINT, D.; DAVIS, W.C.; CHRISTOPHER, S. J.; ELLISOR, M. B. et al. Development and application of an ultratrace method for speciation of organotin compounds in cryogenically archived and homogenized biological materials. **Anal. Bioanal. Chem.**, 387, 2343-2355, 2007.

²¹EDLER, M.; JAKUBOWSKI, M.; LINSCHIED, M. Quantitative determination of melphalan DNA adducts using HPLC - inductively coupled mass spectrometry. **J. Mass. Spectrom.**, 41, 507-516, 2006.

²²AMMANN, A. A. Inductively coupled plasma mass spectrometry (ICP MS): a versatile tool. **Jornal of Mass Spectrometry**, 42, 419-427, 2007.

²³BECKER, J. S. **Inorganic mass spectrometry, principles and applications**. 1 ed., West Sussex: Wiley, 2007.

²⁴FENG, X.; LIU, X.; LUO, Q.; LIU, B. S. Mass spectrometry in systems biology: an overview. **Mass Spectrometry reviews**, 27, 635-660, 2008.

²⁵EVANS, E. H.; GIGLIO, J. J. Interferences in inductively coupled plasma mass spectrometry. A review. **Journal of Analytical Atomic Spectrometry**, 8, 1. 1993.

²⁶LYON, T. D. B.; FELL, G. S.; HUTTON, R. C.; EATON, A. N. J. Evaluation of inductively coupled argon plasma mass spectrometry (ICP-MS) for simultaneous multi-element trace analysis in clinical chemistry. **Journal of Analytical Atomic Spectrometry**, 3, 265, 1988.

²⁶MONTASER, A.; CHAN, S.; KOPPENAAL, D. W. Inductively coupled helium plasma as an ion source for mass spectrometry. **Anal. Chem.**, 59, 1240-1242, 1987.

²⁷TORMEN, L.; CHAVES, E. S.; Saint'Pierre, T. D. Determination of trace elements in fuel ethanol by ICP-MS using direct sample introduction by a microcentric nebulizer. **J. Anal. At. Spectrom.**, 23, 1300-1304, 2008.

- ²⁸POZEBON, D.; DRESSLER, V. L. Monitoring of platinum in a single hair by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) after cisplatin treatment for câncer. **Intern. J. of Mass Spectrom.**, 272, 57-62, 2008.
- ²⁹VIERA M.A.; RIBEIRO, A. S. Determination of As, Ge, Hg, Pb, Sb, Se and Sn in coal slurries by CVG-ETV-ICP-MS using external or isotopic dilution calibration. **Microchem. J.**, 82, 127-136, 2006.
- ³⁰CHAVES, E. S.; LEPRI, F. G. Determination of Co, Cu, Fe, Mn, Ni and V in diesel and biodiesel samples by ETV-ICP-MS. **J. Environ. Monit.**, 10, 1211-1216, 2008.
- ³¹KUMAR, N.; BANSAL, A.; SARMA, G. S.; RAWAL, R. K. Chemometrics tools used in analytical chemistry: An overview. **Talanta**, 123, 186–199, 2014.
- ³²HOPLE, P. K. Application of chemometrics in quality evaluation of medicinal plants. **Analytica Chimica Acta**, 500, 365-377, 2003.
- ³³MASSART, D. L.; VANDEGINSTE, B. G. M. et al. **Handbook of Chemometrics and Qualimetrics, Part A**. 1 ed., Amsterdam: Elsevier Science, 1997.
- ³⁴HAIR, J. F.; BLACK, W. C.; BABIN, B. J. et al. **Análise Multivariada de Dados**. 6 ed., Porto Alegre: Bookman, 2009.
- ³⁵LYRA, W. S.; DA SILVA, E. C.; et al. Classificação periódica: um exemplo didático para ensinar análise de componentes principais. **Quim. Nova**, 33, 1594-1597, 2010.
- ³⁶POWERS, D. A.; YOSHIOKA, H.; YUN, M. mvdcmp: Multivariate decomposition for nonlinear response models. **The Stata Journal**, 11, 556-576, 2011.
- ³⁷CAMACHO, J.; PICÓ, J.; FERRER, A. Data understanding with PCA: Structural and Variance Information plots. **Chemometrics and Intelligent Laboratory Systems**, 100, 48-56, 2010.

³⁸WOLD, S.; ESBENSEN, K.; GELADI, P. Principal Component Analysis. **Chemometrics and Intelligent Laboratory Systems**, 2, 37-52, 1987.

³⁹FERREIRA, M. M. C. **Quimiometria - Conceitos, métodos e Aplicações**. 1 ed., Campinas: Editora da Unicamp, 2015.

⁴⁰MILLER, J, N.; MILLER, J, C. **Statistics and Chemometrics for Analytical Chemistry**. 6 ed., Harlow: Pearson Education Limited, 2010.

⁴¹SARAÇLI, S.; DOGMAN, N.; DOGAN, I. Comparison of hierarchical cluster analysis methods by cophenetic correlation. **Jornal of Inequalities and Applications**, 203, 2013.

8 Apêndices

Apêndice I: Tabela com o autoescalonamento das concentrações dos analitos.

	As	Cr	Co	Ni	Fe	Al	Pb	Zn
Cat 1	-0,4883	-0,5088	2,2530	1,6259	-0,0593	1,3756	-0,4250	-0,7543
Cat 2	-0,4941	-0,6691	0,3369	0,3288	-0,1260	1,1704	-0,4131	-0,7525
Cat 3	-0,4576	-0,4078	2,5180	3,7596	-0,0391	1,0426	-0,3146	-0,5869
Cat 4	-0,4863	-0,4078	1,6823	1,1542	0,1399	2,0056	-0,3422	-0,6800
Cat 5	-0,4946	-0,4252	0,2470	0,3225	0,0504	1,5196	-0,4703	-0,7763
Cat 6	-0,4853	-0,4252	2,6811	2,5590	0,2215	1,3262	-0,1647	-0,6820
Cat 7	-0,4948	0,2020	-0,0455	0,3541	-0,1674	1,1056	-0,5531	-0,8108
Cat 8	-0,4845	-0,6360	1,4662	0,6462	0,0871	1,0479	-0,3264	-0,5532
Cat 9	-0,4959	-0,5646	-0,1116	0,0945	-0,0567	1,2421	-0,5157	-0,7930
Cat 10	-0,4882	-0,6012	-0,4112	-0,2347	-0,0323	1,1945	-0,5078	-0,8009
Cat 11	-0,4966	-0,6308	-0,4029	-0,2861	-0,1631	0,9660	-0,5137	-0,8025
Cat 12	-0,4957	-0,6343	0,1942	-0,0978	-0,0199	0,7459	-0,3776	-0,6741
Casc.1	-0,4166	-0,4374	-0,6823	-0,5740	-0,6074	-0,8068	-0,5196	-0,0443
Casc.2	-0,4061	-0,5106	-0,6905	-0,5740	-0,6074	-0,8131	-0,3757	-0,0290
Casc.3	-0,3984	-0,0419	-0,6917	-0,5740	-0,6074	-0,8576	-0,1568	0,3525
Casc.4	-0,4194	-0,1325	-0,6897	-0,5740	-0,6074	-0,8595	1,1523	1,3011
Casc.5	-0,4070	0,0835	-0,6954	-0,5740	-0,6074	-0,8595	0,5786	0,3162
Casc.6	-0,3984	-0,4252	-0,6880	-0,5740	-0,6074	-0,8146	-0,4368	0,0651
Casc.7	-0,4309	-0,3834	-0,6524	-0,5740	-0,6074	-0,7662	-0,6219	0,5011
Casc.8	-0,4099	-0,0105	-0,6283	-0,5740	-0,6074	-0,7358	-0,6375	0,2801
Casc.9	-0,4462	-0,4618	-0,7396	-0,5740	-0,6074	-0,8688	-0,6706	0,0653
Casc.10	-0,4042	-0,4008	-0,7313	-0,5740	-0,6074	-0,8593	-0,6844	0,5500
Casc.11	-0,4376	0,0487	-0,6834	-0,5740	-0,6074	-0,8130	-0,5803	1,0338
Casc.12	-0,4309	0,3431	-0,6408	-0,5740	-0,6074	-0,7388	-0,5251	1,7026
Lodo 1	2,7371	2,6237	-0,0198	-0,3265	-0,6080	-0,8867	2,2840	-0,8497
Lodo 2	2,4506	2,9024	0,0168	-0,3004	-0,6079	-0,8867	2,5797	-0,8497
Lodo 3	2,4602	2,0139	0,0393	-0,3044	-0,6080	-0,8867	2,4417	-0,8496
Lodo 4	2,4506	2,4495	-0,0463	-0,3368	-0,6080	-0,8867	2,4220	-0,8497
Borra 1	-0,2027	-0,6672	-0,6465	-0,5436	2,2082	-0,2508	-0,4952	0,1429
Borra 2	0,0864	-0,5241	-0,6085	-0,5386	3,0479	-0,1091	-0,4575	0,3300
Borra 3	0,5990	-0,3590	-0,4514	-0,4952	2,4758	-0,4974	-0,1377	2,7443
Borra 4	0,2861	-0,4018	-0,4777	-0,4932	2,1539	-0,5446	-0,2351	2,7534