

**ANA YUMI ASADA JACOMO**

**AGRUPAMENTO DE AMOSTRAS DE DRX A PARTIR DE MAPAS  
AUTO-ORGANIZÁVEIS COMO SUPORTE À CARACTERIZAÇÃO  
MINERAL NA MINA DO SOSSEGO, PA**

**São Paulo**

**2021**

**ANA YUMI ASADA JACOMO**

**AGRUPAMENTO DE AMOSTRAS DE DRX A PARTIR DE MAPAS  
AUTO-ORGANIZÁVEIS COMO SUPORTE À CARACTERIZAÇÃO  
MINERAL NA MINA DO SOSSEGO, PA**

Trabalho de Formatura em Engenharia de  
Minas do curso de graduação do  
Departamento de Engenharia de Minas e de  
Petróleo da Escola Politécnica da  
Universidade de São Paulo.

Área de concentração: Engenharia Mineral

Orientadores: Prof. Dr. Cleyton de Carvalho  
Carneiro/Profa. Dra. Carina Ulsen

Colaboradores: Rodrigo César Teixeira de  
Gouvêa/Thais Bortotti Fagundes

**São Paulo**

**2021**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

### Catlogação-na-publicação

Serviço de Biblioteca e Documentação

Escola Politécnica da Universidade de São Paulo

Jacomo, Ana Yumi Asada

Agrupamento de amostras de DRX a partir de mapas auto-organizáveis como suporte à caracterização mineral na mina do Sossego, PA / A. Y. A. Jacomo -- São Paulo, 2021.

51 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Minas e de Petróleo.

1.Mapas auto-organizáveis 2.Difração de raios-X 3.Índice de Davies Bouldin 4.Análise exploratória de dados I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Minas e de Petróleo II.t.

## **AGRADECIMENTOS**

Primeiramente, gostaria de agradecer a todo o grupo que me auxiliou no desenvolvimento deste trabalho, especialmente ao meu orientador, o professor Cleyton Carneiro, e ao Rodrigo Gouvêa, que me acompanhou de perto todo esse período. Muito obrigada pelos ensinamentos, disponibilidade e paciência ao longo deste último ano.

Não poderia deixar de mencionar também, a minha querida professora Carina Ulsen, uma mulher incrível, que sempre foi e sempre será uma inspiração e exemplo profissional para mim. Muito obrigada por ter sido minha mentora durante toda a minha graduação e por ter me apresentado tantas pessoas maravilhosas que tive a oportunidade de trabalhar em grupo. Minha graduação não teria sido tão completa sem você, sou grata por todas as oportunidades e por estar comigo até meus últimos momentos na Poli.

Thais Bortotti, nos encontramos em diversas oportunidades, obrigada pelo acompanhamento e pela ajuda no desenvolvimento deste trabalho. Rafael Macedo, não tenho palavras para dizer o quanto eu aprendi com você, vou levar comigo seu amor pelo conhecimento e por ensinar o próximo, obrigada pela parceria.

Gostaria de agradecer à Escola Politécnica da Universidade de São Paulo por todos os ensinamentos, não apenas os adquiridos em sala de aula, mas os conhecimentos para levar ao longo de minha vida, com os grupos de extensão, intercâmbio e iniciação científica, além de colegas que viraram verdadeiros amigos.

À minha família, em especial minha mãe, Yumiko, e meu pai, Alfredo, agradeço por terem me mostrado a importância dos estudos e por sempre terem lutado para poder me proporcionar a melhor educação ao longo de minha vida, mesmo em meio a tantas dificuldades. Sem vocês eu não teria chegado até aqui, muito obrigada.

Por fim, meu namorado, João Pedro, um outro presente que a Escola Politécnica me proporcionou, muito obrigada por ter sido tão companheiro ao longo de todos esses anos de graduação.

## **ANEXOS**

**Anexo 01 – Relação das amostras presentes em cada classe**

**Anexo 02 – Amostras das classes comparadas com SHIMIZU (2012) – 7 Grupos**

**Anexo 03 – Amostras das classes comparadas com SHIMIZU (2012) – 4 Grupos**

**Anexo 04 – Posição 2 $\theta$  determinada no método de DRX**

**Anexo 05 – Amostras representativas – 5 clusters**

## RESUMO

Este estudo compara duas técnicas de agrupamento de dados, a Clusterização Hierárquica Aglomerativa (CHA), método mais tradicionalmente utilizado, desenvolvido no trabalho de SHIMIZU (2012), e os Self Organizing Maps (SOM) como suporte à técnica da difratometria de raios-X (DRX), um dos principais métodos de caracterização microestrutural de materiais. O número ideal de clusters também foi comparado com o método tradicional, que depende do conhecimento de um especialista, utilizando o Índice *Davies-Bouldin* (IDB). A interpretação dos difratogramas das amostras mais representativas de cada classe definida pelo IDB foi realizada e comparada com as características presentes nos corpos minerais de Sossego e de Sequeirinho, localizados na Mina do Sossego, no Pará. Os resultados obtidos foram satisfatórios, mostrando que ambos auxiliam positivamente a técnica de difratometria de raios-X, mas consideram critérios de agrupamentos diferentes, sendo o maior diferencial do novo método dos SOM, a determinação do número ideal de clusters por meio do IDB.

**Palavras-chave:** mapas auto-organizáveis, difração de raios-X, análise exploratória de dados, aprendizado não supervisionado.

## ABSTRACT

*This work aims to compare two data analysis techniques: the hierarchical agglomerative clustering (HAC), more commonly and traditionally used, as presented on SHIMIZU (2012), and the Self Organizing Maps (SOM). Both can be used to support the X-ray diffraction method, one of the main techniques for microstructural characterization of materials. The ideal number of clusters were also compared with the traditional method, which depends on specific knowledge of the operator, using the Davies-Bouldin Index (DBI). The most representative sample of each group defined by the DBI were determined, its diffractograms interpreted and compared to the main characteristics of each ore body, Sossego and Sequeirinho, in Sossego Mine, located in Pará, Brazil. The obtained results were satisfactory, showing that both can positively support the X-ray diffraction method, but using different approaches to cluster the samples into each group, also having the main advantage of the SOM related to the determination of the ideal numbers of clusters using the DBI.*

**Keywords:** Self-Organizing map (SOM), X-ray diffraction (XRD), exploratory data analysis, unsupervised learning.

## LISTA DE FIGURAS

Figura 1 - Seções transversais dos corpos de Sequeirinho e Sossego (MONTEIRO et al., 2008).....	16
Figura 2 - Fluxograma do processo da usina do Sossego (ROSA et al., 2007).....	17
Figura 3 - Fenômeno de difração de raios-X - Lei de Bragg (ANTONIASSI, 2010).....	20
Figura 4 - Geometria Bragg-Brentano (JENKINS, 1989).....	20
Figura 5 - Difratoograma do cristal de NaCl (adaptado de SILVA, 2013).....	21
Figura 6- Estrutura dos SOM (quadrada e hexagonal) (BREARD, 2017) .....	22
Figura 7 - Etapa de treinamento dos mapas auto-organizáveis (BREARD, 2017) .....	23
Figura 8 - Matriz-U resultante de um mapa Kohonen (FRASER, 2007; KITANI, 2013).....	24
Figura 9 - Aglomeração dos dados com maior similaridade e formação de classes (clusters) (KUMAR, 2020a) .....	25
Figura 10 - Etapa de junção entre classes mais próximas até a formação final de duas classes (adaptado de KUMAR, 2020a).....	26
Figura 11 - Dendrograma e definição do número ótimo de classes (KUMAR, 2020a).....	26
Figura 12 - Definição da menor distância entre o ponto e os centróides de cada classe (adaptado de KUMAR, 2020b) .....	27
Figura 13 – Difratoogramas mais representativos de cada cluster (SHIMIZU, 2012) .....	28
Figura 14 - Matriz-U de Treinamento .....	32
Figura 15 - Matriz U com Hits.....	32
Figura 16 - Comparação índices <i>Davies-Bouldin</i> .....	33
Figura 17 – Contagem dos mínimos e segundos mínimos globais .....	33
Figura 18 - Divisão da matriz-U em 4 classes .....	34
Figura 19 - Divisão da matriz-U em 5 classes.....	35
Figura 20 - Divisão da matriz-U em 7 classes .....	35
Figura 21 – Difratoogramas das amostras mais representativas de cada cluster .....	37



## LISTA DE TABELAS

Tabela 1- Características dos corpos da Mina do Sossego (MARQUES, 2015)	15
Tabela 2 - Histórico de produção de cobre na Mina do Sossego (elaboração própria com dados da empresa).....	18
Tabela 3 - Principais fases minerais identificadas no DRX .....	29
Tabela 4 - Parâmetros de treinamento dos Mapas Auto-organizáveis.....	30
Tabela 5 - Porcentagem das amostras de SHIMIZU (2012) em cada cluster definido (4 grupos) .....	35
Tabela 6 - Porcentagem das amostras de SHIMIZU (2012) em cada cluster definido (7 grupos) .....	35
Tabela 7 - Amostras por cluster e corpos minerais .....	36
Tabela 8 - Amostra mais representativa de cada classe (5 grupos).....	36
Tabela 9 - Amostra mais representativa de cada classe (7 grupos).....	37
Tabela 10 - Amostras mais representativas de cada cluster determinado por SHIMIZU (2012) .....	39

## LISTA DE ABREVIATURAS E SIGLAS

BMU	<i>Best Matching Unit</i>
CHA	Clusterização Hierárquica Aglomerativa
IDB	Índice <i>Davies e Bouldin</i>
DRX	Difração de raios-X
IA	Inteligência Artificial
IOCG	<i>Iron oxide-copper-gold</i>
ICDD	<i>International Center for Diffraction Data</i>
ROM	<i>Run of Mine</i>
SOM	<i>Self-Organizing Map</i>
VQ	<i>Vector Quantization</i>

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	11
1.1. Objetivo	12
1.2. Relevância do Projeto / Justificativa	12
<b>2. REVISÃO BIBLIOGRÁFICA</b>	14
2.1. Mina do Sossego	14
2.1.1. História	14
2.1.2. Geologia	14
2.1.3. Beneficiamento do minério de cobre	16
2.1.4. Produção de concentrado de cobre	18
2.2. Difração de raios-X	18
2.2.1. Raios-X e o fenômeno da difração de raios-X	19
2.2.2. Difratorômetro de raios-X	20
2.3. Mapas Auto-Organizáveis (SOM)	21
2.4. Índice <i>Davies e Bouldin</i> (IDB)	24
2.5. Análise de agrupamento ( <i>cluster analysis</i> )	24
2.5.1. Clusterização Hierárquica Aglomerativa (CHA)	25
2.5.2. Clusterização por <i>K-means</i>	26
<b>3. MATERIAIS E MÉTODOS</b>	27
3.1. Materiais	27
3.2. Procedimentos Adotados	29
3.2.1. Difractogramas	29
3.2.2. Tratamento dos dados	29
3.2.3. Análise dos Mapas Auto-organizáveis (SOM)	30
3.2.4. Índice <i>Davies e Bouldin</i> (IDB)	30
3.2.5. Algoritmo K-means	31
<b>4. RESULTADOS E DISCUSSÃO</b>	31
4.1. Resultados	31
4.2. Discussão	38
4.2.1. Comparação dos resultados entre classes	38
<b>5. CONCLUSÃO</b>	41
<b>6. REFERÊNCIAS BIBLIOGRÁFICAS</b>	42

## 1. INTRODUÇÃO

A Mina do Sossego, localizada em Canaã dos Carajás, no Pará, iniciou suas operações em 2004 e ainda hoje corresponde a uma significativa produção de cobre no Brasil. De acordo com a Vale, empresa responsável pelas operações, o complexo de mineração dispõe de uma capacidade nominal de 93 mil toneladas por ano de cobre em concentrados (VALE, 2020). Apenas em 2019, a Mina do Sossego representou aproximadamente 15% do valor total de comercialização do cobre produzido no país (ANM, 2020).

Durante a fase de exploração em um complexo de mineração, são realizadas análises para determinar alguns fatores fundamentais que auxiliarão na decisão da viabilidade da mina. Um desses fatores é o conhecimento da variabilidade mineralógica no complexo, sendo comumente determinado a partir do método de difração de raios X.

No entanto, essa técnica amplamente utilizada na mineração possui alguns pontos de atenção. Cada difratograma leva um tempo considerável de análise, não sendo um método muito ágil, além de não apresentar uma análise precisa na determinação das fases mineralógicas, uma vez que cada profissional detém um certo conhecimento e realiza as interpretações de forma individual, podendo ocorrer um viés interpretativo nos dados obtidos.

Os desenvolvimentos na área de Inteligência Artificial (IA), especialmente das suas ramificações em ciência de dados e aprendizado de máquina (*machine learning*), vêm trazendo uma maior eficiência e exatidão na obtenção de resultados, especialmente relacionados a protocolos analíticos repetitivos.

O aprendizado de máquinas é amplamente empregado em tarefas de classificação, regressão e análise exploratória de dados. Nesse contexto, os mapas auto-organizáveis são muito utilizados na exploração automatizada da estrutura interna dos dados, de forma não supervisionada, ou semi-supervisionada. Além de permitir o emprego de técnicas de agrupamento de dados aos *neurons* treinados segundo as informações vetoriais e topológicas dos dados de entrada.

Tradicionalmente, a análise exploratória e o agrupamento de dados dos difratogramas de raios-X é feita utilizando a Clusterização Hierárquica Aglomerativa (CHA). Esse método baseia-se na distância euclidiana como medida de dissimilaridade dos valores de  $2\theta$  entre as amostras. Entretanto, a definição do número ótimo de agrupamentos depende da definição do *cutoff* pelo usuário, ou seja, dependente do conhecimento de especialista e, consequentemente, pode apresentar um viés de interpretação.

Este estudo de caso irá avaliar um novo modelo de agrupamento de dados associados à difratometria de raios-X. Os resultados obtidos serão comparados com a técnica tradicional, realizado no estudo de SHIMIZU (2012) e, posteriormente, os difratogramas das amostras mais representativas serão analisados.

### **1.1. Objetivo**

Este estudo visa avaliar a aplicabilidade da ferramenta de agrupamento de mapas auto-organizáveis associada à técnica de difração de raios-X (DRX) na definição automatizada de agrupamentos ideais das amostras da Mina do Sossego e a sua interpretação.

Como objetivos específicos, este trabalho se propõe a:

1. Comparar agrupamentos não supervisionados com os agrupamentos definidos por um especialista
2. Analisar a representatividade de amostras em relação aos dois métodos, *Self-Organizing Maps* (SOM) e a Clusterização Hierárquica Aglomerativa (CHA)
3. Definir o número ótimo de *clusters* a partir do Índice *Davies-Bouldin*
4. Interpretar os difratogramas mais representativos de cada grupo

### **1.2. Relevância do Projeto / Justificativa**

A difratometria de raios X é uma das principais técnicas de caracterização microestrutural de materiais, sejam eles orgânicos ou inorgânicos, sendo amplamente aplicado em diversas áreas de conhecimento. O processo de

interpretação das fases mineralógicas nos difratogramas ainda é repetitivo e demorado a depender do volume de análises. Dessa forma, técnicas de agrupamento de dados podem ser empregadas com o objetivo de identificar similaridades entre conjuntos de amostras, permitindo a escolha criteriosa daquelas mais representativas para análise e interpretação, reduzindo, dessa forma, o tempo de trabalho.

Alguns trabalhos como SHIMIZU (2012) e MACCHIAROLA et al. (2007) conduzem a análise de agrupamentos presente nos *softwares* comerciais, como o *HighScore da Malvern Panalytical*, que consiste no agrupamento se baseando no método de Clusterização Hierárquica Aglomerativa. Esse tipo de agrupamento de amostras se baseia em separações consecutivas das amostras em grupos de maior similaridade, utilizando a métrica euclidiana.

Apesar de efetivo para muitos trabalhos, essa abordagem é dependente da decisão de um operador especializado para a definição do número ideal de grupos. Desse modo, tal abordagem é suscetível ao viés interpretativo e dependente da exploração *a priori* dos dados em situações em que se há pouco conhecimento sobre as amostras analisadas.

Considerando esses pontos, uma abordagem matemática, baseada na maximização da similaridade intragrupos e na dissimilaridade entre grupos é importante na medida em que evita esse viés interpretativo e permite uma escolha criteriosa de um número ideal de grupos sem a necessidade de maior exploração dos dados.

Portanto, ambas as técnicas facilitam a caracterização e a separação de fases minerais, reduzindo a quantidade de amostras que necessitam de uma análise individual e, conseqüentemente, o tempo de trabalho. Além da análise do SOM, o uso conjunto do Índice *Davies-Bouldin* pode trazer um maior benefício em comparação com a determinação do *cutoff* pelo próprio especialista na CHA, evitando vieses e uniformizando a definição do número ideal de clusters.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. Mina do Sossego

#### 2.1.1. História

A Mina do Sossego, localizada no Canãa dos Carajás, no estado do Pará, foi a primeira operação de cobre da empresa Vale S.A. (antiga Companhia Vale do Rio Doce). A descoberta do depósito de cobre do Sossego ocorreu em 1997, juntamente com a empresa *Phelps Dodge* (atual *Freeport*), que vendeu suas participações na região em 2001 para a própria Vale. Em 2003, o Projeto Sossego foi implementado com a abertura da cava de Sequeirinho e apenas em 2004 as operações da usina de beneficiamento mineral foram iniciadas. Neste mesmo ano, o primeiro embarque de concentrado de cobre foi realizado, com um total de 269 mil toneladas de concentrado vendidos, sendo 252 mil toneladas destinados ao mercado externo e 17 mil ao mercado interno (CARVALHEIRO, 2019; VALE, 2012).

A abertura da Mina do Sossego contribuiu consideravelmente no desenvolvimento da região, investindo um total de R\$ 200 milhões até 2020 em obras de infraestrutura, principalmente via parcerias com a prefeitura de Canãa dos Carajás, com obras relacionadas à segurança pública e ao saneamento básico, além de ações sociais e culturais no município (VALE, 2012).

#### 2.1.2. Geologia

A Província Mineral de Carajás (PMC), território onde se encontra a Mina do Sossego, contém a maior concentração conhecida de depósitos de óxido-ferro-cobre-ouro (comumente conhecida em inglês como *IOCG*) e são relevantes na exploração mineral devido ao seu grande potencial econômico associado (MARQUES, 2015).

De acordo com Hitzman (1992), esse tipo de depósito é geralmente composto por grandes quantidades de óxidos de ferro e precipitados hidrotermais, contendo magnetita ou hematita, com sulfetos de cobre e ouro associados e podem conter concentrações significativas de carbonatos.

A mineralização da Mina do Sossego é composta por cinco corpos, sendo eles: Pista, Sequeirinho, Baiano, Curral e Sossego. A origem da rocha hospedeira, o tipo e a intensidade hidrotermal de cada depósito afetam as características minerais dos principais corpos, de Sequeirinho e de Sossego e, apesar da proximidade entre eles, diferem-se intensamente entre si (MONTEIRO et al., 2008). As principais diferenças dos corpos de Sequeirinho-Pista-Baiano e Sossego-Curral encontram-se na Tabela 1.

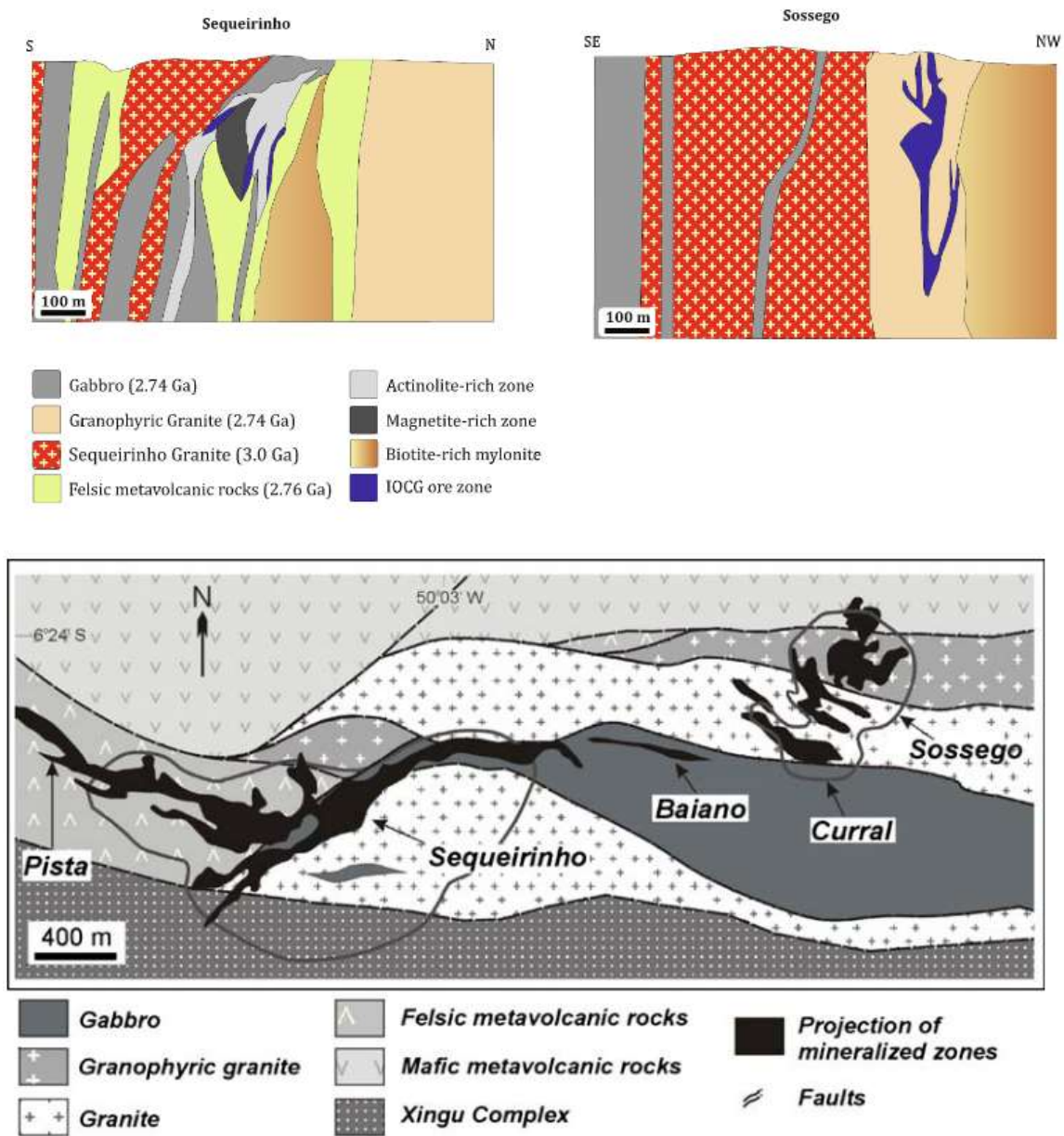
Tabela 1- Características dos corpos da Mina do Sossego (MARQUES, 2015)

	Corpos	
	Sequeirinho-Pista-Baiano	Sossego-Curral
Reserva	85% de 355 Mt com 1,1% Cu e 0,28 g/t Au	15% de 355 Mt com 1,1% Cu e 0,28 g/t de Au
Rochas hospedeiras	Granito Sequeirinho, gabronoritos e rocha vulcânica félsica Pista	Granito granofírico Sossego e gabro
Alteração hidrotermal	Alteração sódica (albita-hematita), alteração sódio-cálcica (rica em actinolita) associada com formação de magnetita-apatita	Ateração potássica e clorítica bem desenvolvida e alteração hidrolítica tardia
Mineralogia do minério	Calcopirita, magnetita, pirrotita, pirita. Menores molibdenita, esfalerita, siegenita, millerita, ouro, Pd-melonita, galena, cassiterita e hessita.	Calcopirita, magnetita e pirita. Menores siegenita, millerita, hessita, Pd-meliorita, molibdenita, ouro e cassiterita
Assinatura geoquímica do minério	Cu-Fe-Au-Ni-Co-Pd-Se-V-P-ETRL (elementos terras raras leves), baixo conteúdo de Ti e U e enriquecido em Co, Ni, Pd, C e Se comparando-se com Sossego-Curral	Cu-Fe-Au-Ni-Co-Pd-Se-V-P-ETRL e relativamente enriquecido em Au, Pb, Sn, Rb, Y e Nb em comparação com Sequeirinho-Pista Baiano

A seção transversal dos dois principais corpos minerais da Mina do Sossego pode ser observada na Figura 1, com destaque à presença de granito e rochas metavulcânicas félsicas em Sequeirinho e granito granofírico e gabro em Sossego.



Figura 1 - Seções transversais dos corpos de Sequeirinho e Sossego (MONTEIRO et al., 2008)



### 2.1.3. Beneficiamento do minério de cobre

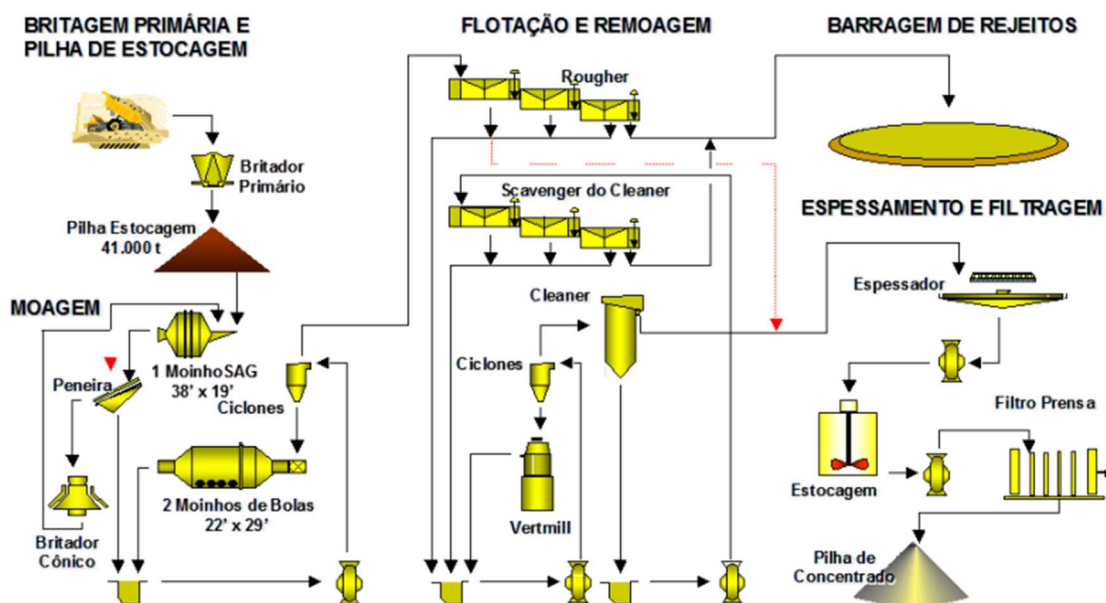
A etapa de cominuição e classificação do minério é um processo determinante na qualidade do concentrado final, uma vez que a eficiência do processo afetará diretamente na recuperação do minério de interesse.

O circuito da usina do Sossego foi testado em diversos projetos pilotos e sofreram algumas otimizações após o início de suas operações. Um ponto de destaque foi a implementação do moinho semi-autógeno (SAG) no

processamento de cobre e ouro, sendo a primeira operação da empresa Vale S.A. a operar com esse tipo de equipamento (MIRANDA et al., 2015).

De acordo com o fluxograma de ROSA et al. (2007) ilustrado na Figura 2, o processo se inicia com o transporte do *Run of Mine (ROM)* por caminhões fora-de-estrada até um britador primário do tipo giratório e seu produto britado é encaminhado por um transportador de correia de longa distância até uma pilha pulmão com capacidade útil de 41 mil toneladas.

Figura 2 - Fluxograma do processo da usina do Sossego (ROSA et al., 2007)



O minério é retomado da pilha pulmão e encaminhado até a alimentação do moinho semi-autógeno (SAG). A moagem do SAG opera em circuito fechado e alimentam duas peneiras vibratórias, sendo o material passante bombeado para o circuito de classificação, enquanto o material retido na peneira é levado a um britador cônico para rebitagem antes de retornar ao início desse circuito.

O circuito de classificação também opera em circuito fechado reverso em conjunto com dois moinhos de bolas associados. O *underflow* da ciclonagem irá alimentar o moinho de bolas antes de retornar novamente ao ciclone, enquanto o *overflow* será encaminhado para o processo de flotação *rougher*.

A flotação é composta basicamente por três etapas, o *rougher*, o *scavenger-cleaner* e o *cleaner*, com o estéril encaminhado para barragem de rejeitos e o

material de interesse para uma etapa final de espessamento e filtragem antes de ser encaminhado até a pilha de concentrado.

#### 2.1.4. Produção de concentrado de cobre

A produção de concentrado de cobre na Mina do Sossego representa aproximadamente 30% da produção total do concentrado pela Vale S.A., com exceção no ano de 2019, que houve uma queda considerável na produção devido à uma manutenção não programada na correia transportadora e no moinho de bolas, retomando a operação normalmente apenas no primeiro trimestre de 2020 (VALE, 2020).

Os concentrados de cobre produzidos na Mina do Sossego são vendidos principalmente para refinadoras de cobre (*smelters*) localizados na Europa, Índia e Ásia via contratos de médio e longo prazo (VALE, 2020). O restante da produção de concentrado de cobre pela empresa no Brasil é proveniente da Mina do Salobo, também localizada no Estado do Pará (VALE, 2019). O histórico de produção da empresa desde 2016 pode ser visto na Tabela 2.

Tabela 2 - Histórico de produção de cobre na Mina do Sossego (elaboração própria com dados da empresa)

	Produção de cobre (mil toneladas métricas)				
	2016	2017	2018	2019	2020
Sossego	92,6	99,7	92,2	65,5	87,7
Total produzido pela Vale - Brasil	268,5	293,1	284,8	254,9	260,5
% produção Sossego	34,5	34,0	32,4	25,7	33,7

## 2.2. Difração de raios-X

A difração de raios-X é um método analítico não destrutivo, que permite identificar e quantificar as fases cristalinas de uma amostra. A identificação das fases se dá por meio da comparação dos dados obtidos com uma base de dados de referência, gerida pelo *International Center for Diffraction Data* (ICDD), sediado nos Estados Unidos. Essa base de referência contém informações

cristalográficas básicas e propriedades físicas dos compostos cristalinos (MACCHIAROLA, 2007).

A difração de raios-X possui outras aplicações menos utilizadas, podendo também determinar parâmetros de cela unitária, orientação preferencial e tamanho médio de cristalinos, tensão residual sobre estruturas cristalinas e até mesmo ser utilizada como um dispositivo de controle de temperatura (KAHN, 2007).

#### 2.2.1. Raios-X e o fenômeno da difração de raios-X

Os raios-X são ondas eletromagnéticas situadas no espectro eletromagnético entre as radiações gama e ultravioleta. O comprimento de onda dos raios-X se encontra entre 0,1 e 100 Å (*Angström*), mas são geralmente utilizados na técnica de difratometria de raios X entre 0,5 e 2,5 Å, por corresponderem a mesma ordem de grandeza dos menores espaçamentos atômicos analisados em materiais orgânicos e inorgânicos (MARTINI, 2015).

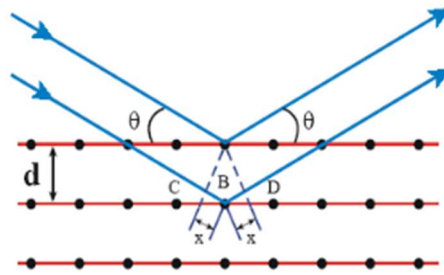
De acordo com KAHN (2007) e ANTONIASSI (2010), os raios-X, quando atingem um material podem ser espalhados elasticamente, sem perda de energia pelos elétrons de um átomo (espalhamento coerente). O fóton de raios-X muda sua trajetória ao colidir com o elétron da amostra, no entanto, mantém o comprimento de onda da radiação incidente.

Se os átomos que geraram esse espalhamento estiverem arranjados de maneira sistemática, como em uma estrutura cristalina, e ao mesmo tempo apresentarem distâncias próximas ao do comprimento de onda do fóton incidente, as relações de fase se tornam periódicas e o fenômeno da difração dos raios-X podem ser observados em vários ângulos de incidência do feixe, satisfazendo a Lei de Bragg, conforme a equação 1 abaixo.

$$n \lambda = 2 d \sin \theta \quad (1)$$

onde, “ $\lambda$ ” corresponde ao comprimento de onda da radiação incidente, “ $n$ ” a ordem de difração ou fase, “ $d$ ” a distância entre os planos atômicos e “ $\theta$ ” o ângulo de incidência dos raios-X. A Figura 3 ilustra o fenômeno da difração de raios-X.

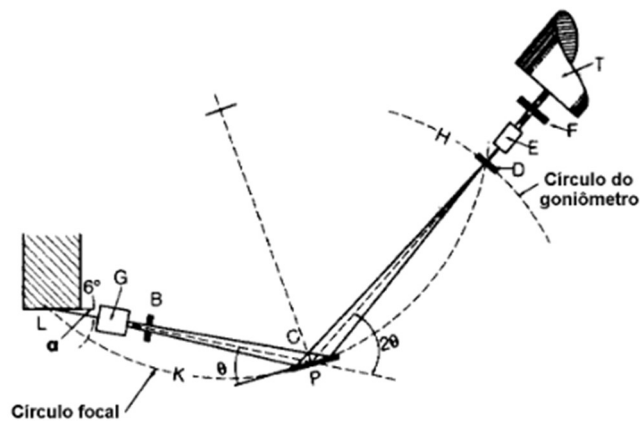
Figura 3 - Fenômeno de difração de raios-X - Lei de Bragg (ANTONIASSI, 2010)



### 2.2.2. Difratorômetro de raios-X

O difratômetro tradicional detecta o feixe difratado por meio da geometria Bragg-Brentano, habilitando a obtenção do ângulo  $2\theta$  (JENKINS, 1989). Conforme ilustrado na Figura 4, a fonte “L” emite os raios-X que atravessam as fendas “G” e “B” e chegam até a superfície da amostra localizada em “C”. Os raios são difratados com um certo ângulo  $2\theta$  e percorrem até a fenda de recepção “D”, responsável por captar a radiação proveniente da amostra e encaminhar até o detector de raios-X em “T”.

Figura 4 - Geometria Bragg-Brentano (JENKINS, 1989)

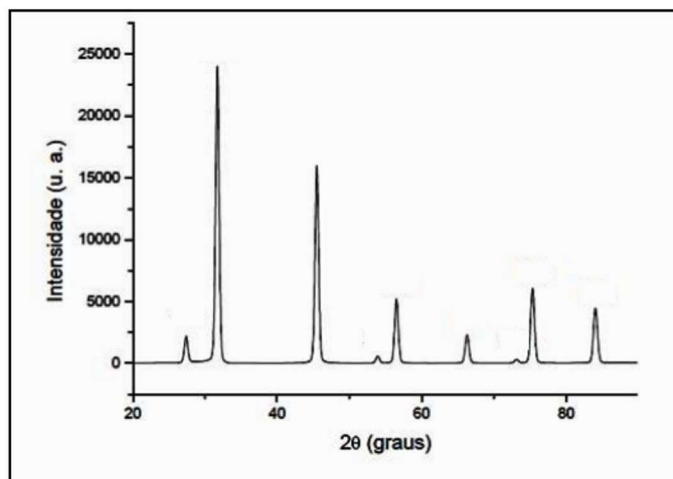


onde, “L” a fonte de raios-X, “G” e “E” as fendas soller, “B” a fenda divergente, “C” a amostra a ser analisada, “D” a fenda receptora, “F” a fenda de espalhamento e “T” o detector de raios-X.

O resultado dessa técnica são os difratogramas, que são representados por picos difratados contendo as seguintes informações: o perfil, a intensidade, a

posição angular ( $2\theta$ ) ou a posição entre planos cristalinos ( $d$ ). A Figura 5 ilustra um difratogramas com os dados obtidos em um cristal de NaCl.

Figura 5 - Difratograma do cristal de NaCl (adaptado de SILVA, 2013)



Como cada material cristalino apresenta um perfil difratométrico característico, é possível identificá-lo conforme as posições angulares e as intensidades dos picos obtidos (MARTINI, 2015). Com isso, as fases cristalinas resultantes da difratometria de raios-X podem ser determinadas a partir da comparação dos padrões difratométricos das fases individuais que ficam localizadas na base de dados *ICDD*.

### 2.3. Mapas Auto-Organizáveis (SOM)

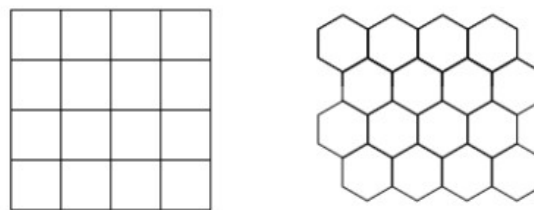
Os mapas auto-organizáveis (SOM) representam uma rede neural com aprendizado não supervisionado, ou seja, buscam identificar padrões a partir dos dados fornecidos, sem a apresentação de rótulos de classificação fornecidos pelos dados de treinamento do algoritmo.

Essa técnica foi desenvolvida por Teuvo Kohonen em 1982 e permite analisar, visualizar e interpretar dados multidimensionais a partir dos princípios do algoritmo de Quantização Vetorial (VQ – *Vector Quantization*), mapeando uma relação não linear entre o espaço de entrada e o espaço de saída (KOHONEN, 2001). São amplamente utilizados em pesquisas geofísicas e geoquímicas, as quais envolvem uma ampla quantidade de dados complexos a serem analisados (KITANI, 2013; FRASER, 2007).

Na análise dos SOM, as amostras inseridas nos dados de entrada são tratadas como vetores  $n$ -dimensionais em um espaço  $n$ -dimensional definido pela quantidade de variáveis analisadas. Os dados de saída, são projetados em um mapa 2D que representa todos esses parâmetros complexos inseridos inicialmente.

O tamanho e a tipologia do mapa são definidos a partir do número de variáveis consideradas e essa estrutura básica do mapa é um fator importante, uma vez que determina a quantidade de neurônios semeados, os quais são tipicamente gerados de modo aleatório (PERILLO, 2019). A geometria de vizinhança entre os neurônios forma uma representação de mapa e assume tipicamente uma grade quadrada ou hexagonal, conforme ilustrado na Figura 6.

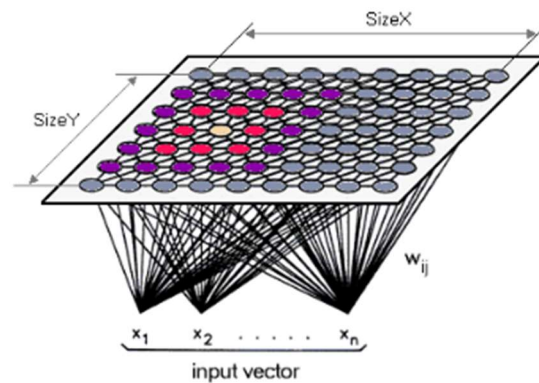
Figura 6- Estrutura dos SOM (quadrada e hexagonal) (BREARD, 2017)



A análise dos SOM se inicia com um processo de treinamento dos neurônios, o qual consiste em duas etapas: a competitiva e a cooperativa. Na etapa competitiva, as amostras inseridas são comparadas com todos os neurônios semeados por meio de medidas de similaridade, normalmente utilizando-se a distância Euclidiana. A menor distância determinada entre os dois vetores é denominada como *Best Matching Unit* (BMU). Já na etapa de cooperação, os *BMUs* definidos anteriormente são considerados os centros de uma certa região e uma vizinhança é definida (FRASER, 2007).

Após a fase de treinamento, os neurônios semeados que representam os dados de entrada podem ser visualizados em um mapa 2D por meio dos nós que agrupam os vetores com maior similaridade entre si, preservando a topologia entre eles. Os mapas que representam os dados de saída em um espaço 2D a partir dos dados de entrada em um espaço  $n$ -dimensional são denominados mapas Kohonen ou mapas auto-organizáveis (SOM). Uma visualização desse processo de redução dimensional pode ser vista na Figura 7.

Figura 7 - Etapa de treinamento dos mapas auto-organizáveis (BREARD, 2017)



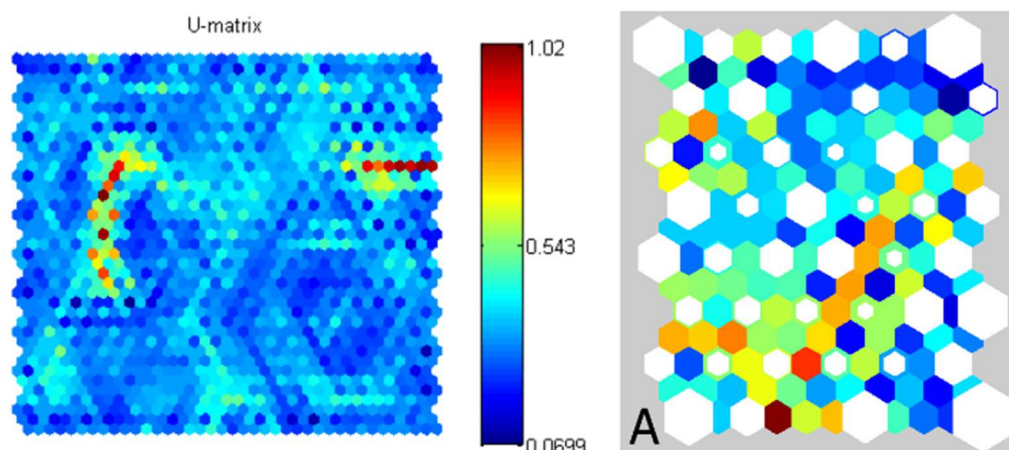
Outra forma de visualização dos mapas auto-organizáveis se dá por meio da Matriz-U, que foi criada no início dos anos 90 com o propósito de trabalhar em conjunto com os mapas 2D de Kohonen, visando auxiliar na visualização dos mapas representativos de um espaço de alta dimensão.

A Matriz-U indica a proximidade entre as amostras representadas nos mapas 2D. Essa distância entre as amostras, calculada pela distância Euclidiana, é definida e representada por cores em uma escala de valores. As cores mais frias, próximas da tonalidade de azul, representam uma menor distância entre os neurônios, possuindo uma maior similaridade entre si, e as cores mais quentes, próximas da tonalidade do vermelho, representam uma maior distância entre os neurônios, possuindo uma menor similaridade entre si (FRASER, 2007).

Uma outra forma de visualização é a partir de hexágonos brancos contidos na Matriz-U, que indicam a quantidade de amostras representadas em cada nó, sendo o tamanho do hexágono proporcional à quantidade de amostras. As representações da Matriz-U podem ser vistas na Figura 8 abaixo.



Figura 8 - Matriz-U resultante de um mapa Kohonen (FRASER, 2007; KITANI, 2013)



#### 2.4. Índice *Davies e Bouldin* (IDB)

O Índice de *Davies e Bouldin* (IDB) é um indicador que considera as similaridades entre os dados, ajudando a separar e definir um valor de classes representativas. O propósito do uso desse índice é de determinar o número ótimo de *clusters* maximizando a distância *intra-cluster* (homogeneidade e compacidade), ou seja, a similaridade *intra-cluster*, e tentando reduzir a distância entre as amostras no *cluster*, ou seja, a dissimilaridade entre as classes (SINGH, 2020).

As principais determinações adotadas na formulação do índice, consideram a aplicação em bases hierárquicas, a compatibilidade com extensas bases de dados, o processamento de dados com dimensionalidade arbitrária e sem que haja a necessidade de interação do usuário na definição de parâmetros (*Davies*, 1979). A definição do parâmetro pelo usuário é o número máximo de *clusters* e, caso não seja definido, o melhor Índice *Davies-Bouldin* será o de mesma ordem do número de amostras.

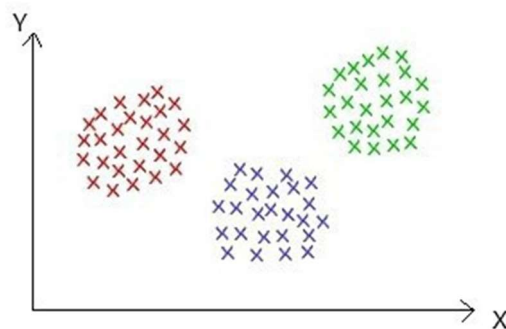
#### 2.5. Análise de agrupamento (*cluster analysis*)

A seguir serão descritas as duas técnicas de análise por agrupamento que serão relevantes para esse trabalho, a Clusterização Hierárquica Aglomerativa e a clusterização por K-means.

### 2.5.1. Clusterização Hierárquica Aglomerativa (CHA)

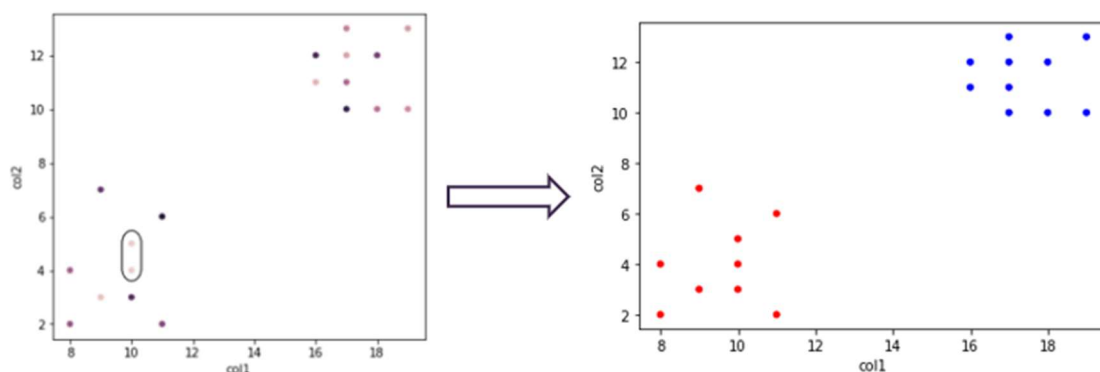
A Clusterização Hierárquica Aglomerativa é considerada uma técnica não supervisionada de aprendizagem de máquinas que visa dividir uma determinada população em diferentes classes a partir da similaridade entre cada amostra (KUMAR, 2020a). Esses algoritmos são amplamente utilizados em análises estatísticas a fim de determinar um número ideal de classes em relação aos dados analisados (MINGOTI, 2005). Na Figura 9, é possível visualizar a formação de 3 classes, mostrando uma maior similaridade entre as amostras dentro de cada classe.

Figura 9 - Aglomeração dos dados com maior similaridade e formação de classes (*clusters*) (KUMAR, 2020a)



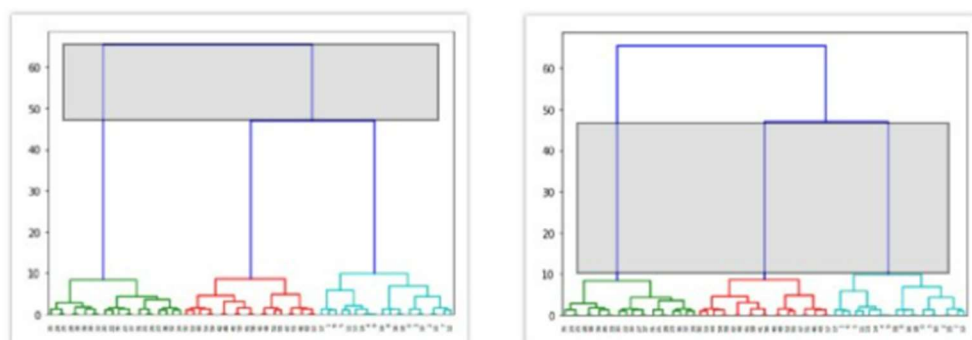
Ao descrever o método, pode-se dizer que, inicialmente, cada dado é considerado uma única classe. Em seguida, esses dados começam a ser comparados entre si, dois a dois, a partir de uma medida de distância, geralmente por meio da distância Euclidiana. Outras formas de medidas podem ser utilizadas como a distância entre os pontos mais distantes entre duas classes, a distância entre os centróides entre duas classes ou pela distância média entre todos os pontos nas duas classes. As duas classes mais próximas entre si, ou seja, mais similares entre si, são unidas formando uma única classe. Esse último processo descrito é repetido diversas vezes até se atingir o número ideal de classes, que será definido posteriormente pelo usuário. Apenas duas classes podem ser agrupadas em cada etapa, formando uma nova classe, e não podem se separar depois de aglomerados (KUMAR, 2020a; MINGOTI, 2005). A Figura 10 abaixo ilustra as etapas do método.

Figura 10 - Etapa de junção entre classes mais próximas até a formação final de duas classes (adaptado de KUMAR, 2020a)



Uma das grandes vantagens do uso desse método é a possibilidade de analisar visualmente a estrutura formada a partir da similaridade entre os dados e a formação de classes a partir de um dendrograma. O dendrograma é uma representação gráfica que mostra as etapas do método e onde pode ser definido o número ideal de classes a partir da análise da maior altura do retângulo formado, uma vez que a maior distância Euclidiana entre o número ótimo de classes, conforme ilustrado na Figura 11 (KUMAR, 2020a). Nesse caso, pode-se afirmar que o número ideal seria de três classes.

Figura 11 - Dendrograma e definição do número ótimo de classes (KUMAR, 2020a)

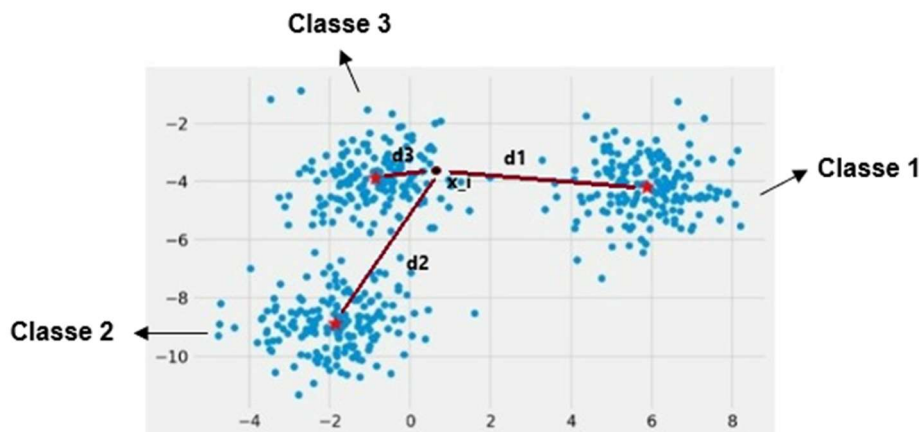


### 2.5.2. Clusterização por *K-means*

A clusterização por *K-means* é um método não supervisionado, sendo um algoritmo bastante conhecido e utilizado em diversas áreas pela sua fácil implementação e habilidade de classificação. Essa técnica necessita de uma definição prévia de um número de grupos,  $k$ , para iniciar o processo. Em seguida, encontram-se os  $k$ -pontos, sendo esses os centroides de cada um dos grupos,

e as distâncias entre cada um dos pontos e os centroides são comparadas. Aquele que apresentar a menor distância Euclidiana calculada passa a pertencer à essa classe (BRENTAN, 2018; KUMAR, 2020b).

Figura 12 - Definição da menor distância entre o ponto e os centróides de cada classe (adaptado de KUMAR, 2020b)



A Figura 12 ilustra essa etapa de comparação entre as distâncias e, nesse caso, o  $d3$  é a que representa a menor distância Euclidiana calculada a partir do ponto  $x_i$ . Portanto, esse ponto passará a pertencer à classe 3.

Apesar de ser um método simples de ser empregado, existe uma dificuldade em determinar o número ideal de classes para ser inserido como dado de entrada, podendo afetar diretamente a performance da técnica de clusterização por K-means. Com isso, é aconselhado o uso de outros algoritmos que possam auxiliar na definição do número ideal de *clusters* (MINGOTI, 2006; BRENTAN, 2018). Nesse trabalho, o Índice de *Davies-Bouldin* (IDB) foi utilizado para definir esse número e que será descrito ao longo do texto.

### 3. MATERIAIS E MÉTODOS

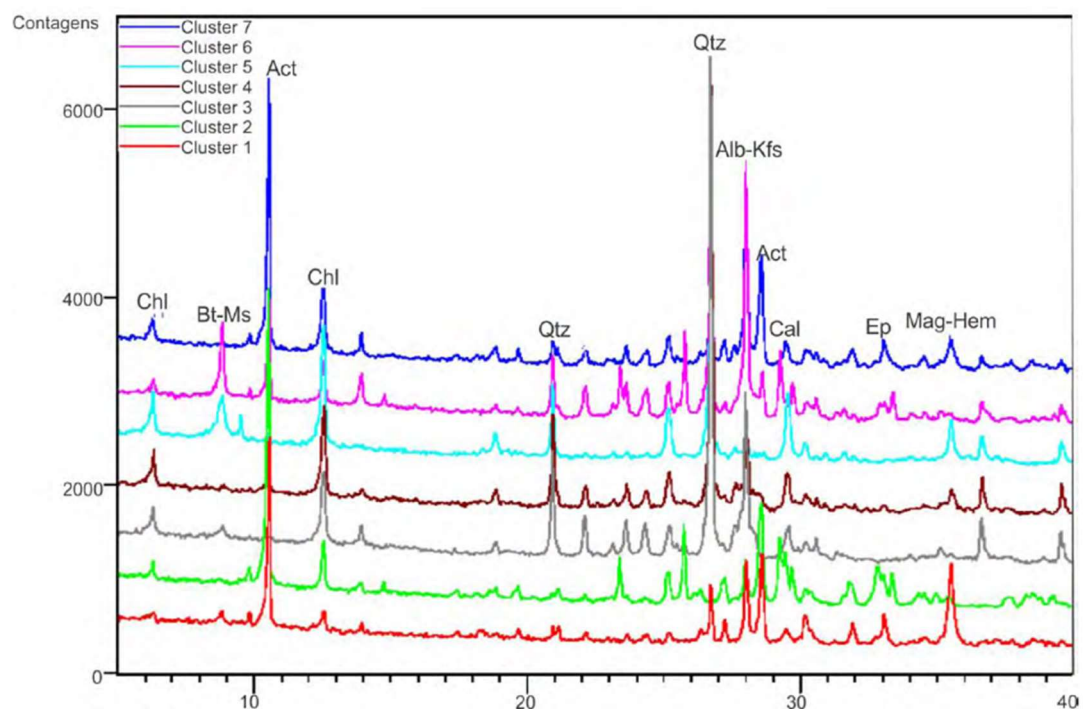
#### 3.1. Materiais

As amostras utilizadas neste trabalho são provenientes da Mina do Sossego, no Pará, e pertencem ao estudo de SHIMIZU (2012), que analisa a aplicabilidade do método de Clusterização Hierárquica Aglomerativa, visando classificar os diferentes tipos de minérios presentes nos corpos de Sequeirinho e de Sossego.

De acordo com BERGERMAN (2009), as amostras coletadas são provenientes de testemunhos de sondagem e referem-se aos planos de lavra dos anos de 2007 e 2008.

De acordo com SHIMIZU (2012), foram tratadas 110 amostras em laboratório e analisadas a partir dos resultados obtidos pelo método de difratometria de raios-X. Os difratogramas mais representativos de cada grupo determinado por SHIMIZU (2012) podem ser visualizados na Figura 13.

Figura 13 – Difratogramas mais representativos de cada *cluster* (SHIMIZU, 2012)



Legenda: Qtz = quartzo, Ab-Kfs = albita-feldspato potássico, Act = actinolita, Mag-Hem = magnetita-hematita, Chl = clorita, Bt-Ms = biotita-muscovita, Cal = calcita, Ep = epidoto.

A partir dos difratogramas encontrados, foram determinadas as seguintes fases minerais em cada cluster, conforme descrito na Tabela 3.

**Tabela 3 - Principais fases minerais identificadas no DRX**

Identificação de fases minerais - DRX	
Cluster 1	Elevada presença de actinolita e menor quantidade de quartzo e feldspato. Presença de magnetita e hematita.
Cluster 2	Elevada presença de actinolita e menor quantidade de quartzo e feldspato.
Cluster 3	Rico em quartzo e presença de feldspato e clorita.
Cluster 4	Rico em quartzo e presença de feldspato.
Cluster 5	Rico em quartzo e presença reduzida de feldspato em relação ao cluster 3 e 4. Presença de mica, clorita e calcita.
Cluster 6	Actinolita, quartzo e feldspato. Presença de mica.
Cluster 7	Actinolita, quartzo e feldspato. Presença de clorita.

## 3.2. Procedimentos Adotados

### 3.2.1. Difrátogramas

De acordo com as especificações descritas no trabalho de SHIMIZU (2012), foi realizado inicialmente o processo de britagem e moagem das amostras tais quais e posteriormente foram pulverizadas. Em seguida, esse material foi preparado e encaminhado ao equipamento *X'Pert Pro* da marca *PANalytical* para a análise de difratometria de raios-X.

### 3.2.2. Tratamento dos dados

Foram coletados dados dos 110 difratogramas disponíveis e organizados em uma única base de dados utilizando linguagem *Python* e a interface do *Jupyter Lab*. Os arquivos de formato *.xrdlm* (*Panalytical XPert X-ray diffraction*) foram abertos utilizando a biblioteca *Xrdtools* (PARUCH, 2021) e foi criada uma base

dados única para o treinamento do algoritmo de mapas auto-organizáveis, onde cada linha represa uma análise de DRX e cada coluna descreve o valor do passo ( $2\theta$ ) conforme ilustrado no Anexo 4.

### 3.2.3. Análise dos Mapas Auto-organizáveis (SOM)

As análises dos Mapas Auto-organizáveis foram realizadas por meio do programa *SiroSOM*®. Inicialmente, foi desenvolvida uma etapa de inicialização, sendo gerada uma malha de 20x20 nós, dispostas em superfície toroidal e com células hexagonais.

Posteriormente, foi realizada uma etapa de treinamento em todas as análises considerando os parâmetros dispostos na Tabela 4.

Tabela 4 - Parâmetros de treinamento dos Mapas Auto-organizáveis

Etapa de treinamento	Quantidade de neurônios na direção X	Quantidade de neurônios na direção Y	Tipo de cela	Raio de vizinhança inicial	Raio de vizinhança final	Épocas de treinamento
Grosso	20	20	Hexagonal	20	8	200
Fino	20	20	Hexagonal	8	1	2000

Fonte: Elaboração Própria

### 3.2.4. Índice *Davies e Bouldin* (IDB)

Os resultados da etapa de treinamento, relacionados aos vetores quantizados pelos *neurons* treinados no algoritmo de Mapas Auto-Organizáveis, também chamados de *Best Matching Units* (BMUs), foram adquiridos e utilizados numa análise automatizada do Índice de *Davies-Bouldin* (IDB).

Essa análise automatizada do IDB foi desenvolvida em *script* de *Jupyter Notebook* em ambiente *Python* utilizando as funções da biblioteca *Sklearn* (PEDREGOSA et al., 2011). A análise foi desenvolvida rodando o IDB em 500 iterações de 2 a 30 grupos, calculando o índice para cada configuração de agrupamento utilizando o algoritmo *K means* para os agrupamentos analisados.

Uma análise estatística foi então desenvolvida analisando os mínimos globais e segundos mínimos globais (menores valores do conjunto global de resultados) de cada uma das iterações de IDB de 2 a 30 grupos. O número ótimo de

agrupamentos foi definido pela moda dos valores de mínimos e segundos mínimos globais dessa análise.

As repetições de iterações de IDB são importantes para garantir a estabilidade dos resultados uma vez que a inicialização do algoritmo *K-Means* é aleatória e alguns agrupamentos podem se coalescer devido a inicialização em algumas iterações, mas não em outras.

### 3.2.5. Algoritmo K-means

Com a definição do valor ideal de *clusters* calculados a partir do Índice *Davies e Bouldin*, foi realizado um agrupamento dos vetores quantizados nos *BMUs* do mapa Kohonen segundo o algoritmo *K-means* para o número ótimo de *clusters* definidos pela análise modal dos IDB. Os vetores de entrada (amostras) relacionadas a esses *BMUs* foram então classificadas segundo esse agrupamento. Os conjuntos de amostras relacionadas a cada grupo foram então comparadas aos obtidos pelo método de agrupamento hierárquico aglomerativo do trabalho de Shimizu (2012).

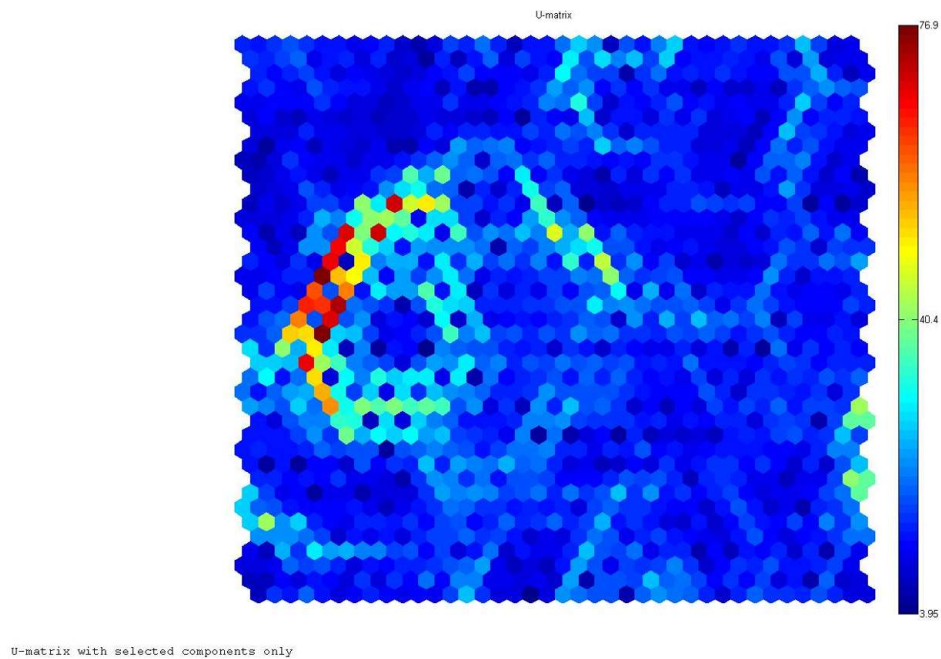
## 4. RESULTADOS E DISCUSSÃO

### 4.1. Resultados

A partir dos dados de entrada e os parâmetros inseridos, a Matriz-U de Treinamento e a Matriz-U com *Hits* foram obtidas. A Matriz-U ilustra a proximidade entre as amostras a partir das cores em escala de valores, sendo as cores mais quentes indicando uma maior distância entre elas e as cores frias indicando uma maior proximidade entre elas no espaço n-dimensional, onde n é o número de *features* treinados.

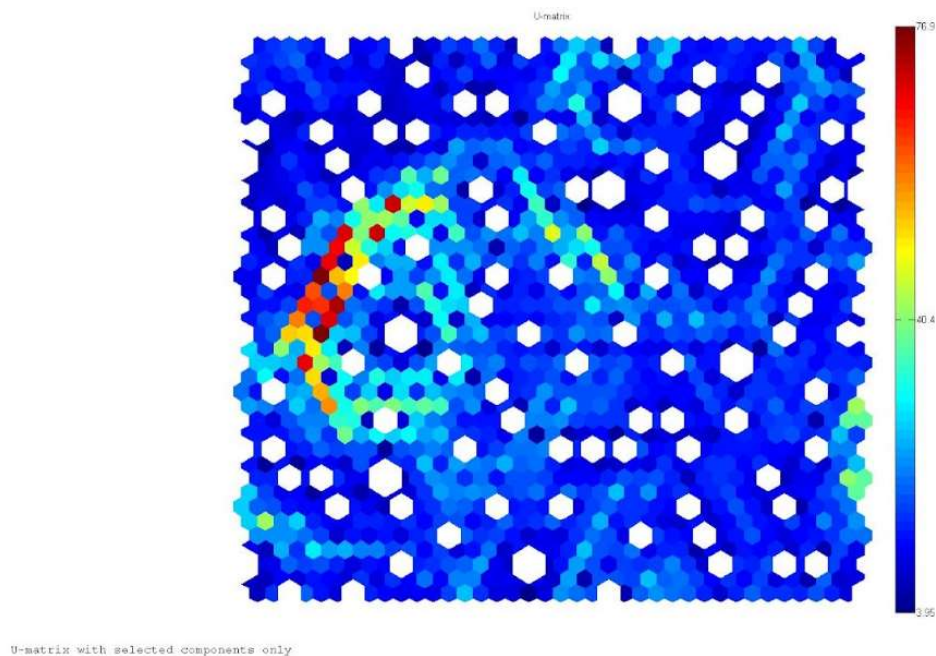


Figura 14 - Matriz-U de Treinamento



Já a Matriz-U com *Hits*, ilustra tanto a proximidade entre os dados de entrada (amostras) a partir das cores em escala de valores, quanto a quantidade de amostras presentes conforme a variação do tamanho dos hexágonos brancos.

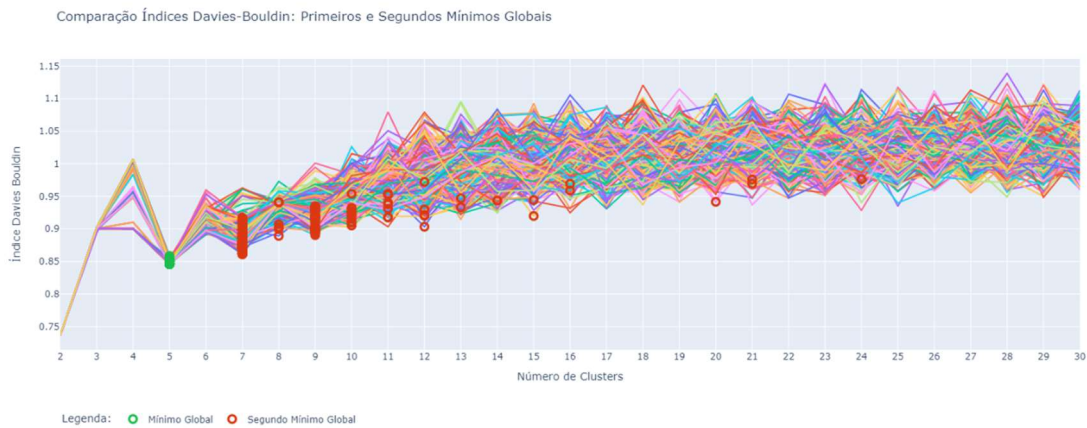
Figura 15 - Matriz U com Hits



Depois de realizada a análise do SOM, foram feitas 500 iterações, a fim de determinar o número ideal de classes a partir do Índice *Davies-Bouldin*. Na

Figura 16, a partir do mínimo global encontrado, determinou-se 5 como sendo o número ideal de classes.

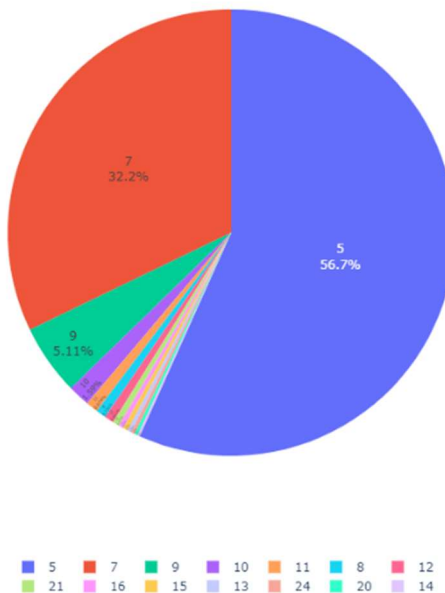
Figura 16 - Comparação Índices *Davies-Bouldin*



Todas as contagens modais de mínimos e segundos mínimos globais determinados a partir da análise de comparação dos Índices *Davies-Bouldin* (IDB) podem ser visualizadas no gráfico (Figura 17).

Figura 17 – Contagem dos mínimos e segundos mínimos globais

Davies-Bouldin: Primeiros e Segundos Mínimos Globais



Com o número ideal de *clusters* identificado pela análise dos IDB como 5 e 7 clusters, foi realizado um agrupamento dos vetores quantizados nos *BMUs* determinados no SOM por meio do algoritmo *K-means*. As análises foram realizadas com três valores distintos de classes, com quatro, cinco e sete grupos, com a finalidade de uma melhor discussão entre eles.

Figura 18 - Divisão da matriz-U em 4 classes

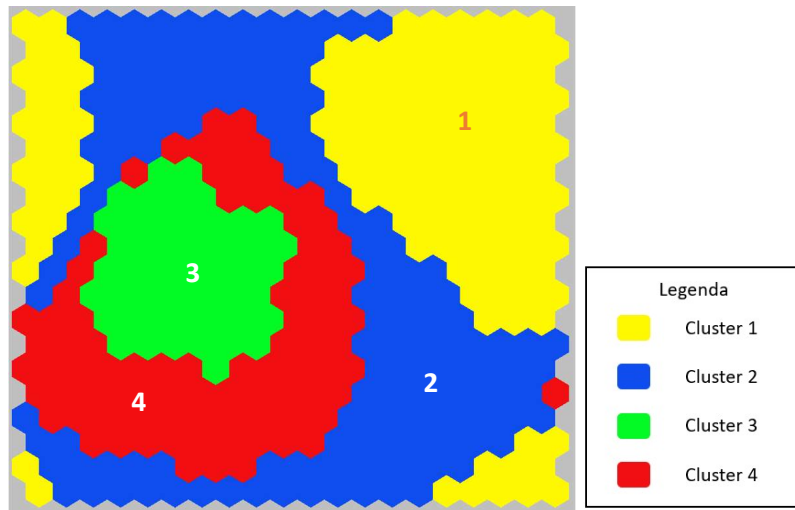


Figura 19 - Divisão da matriz-U em 5 classes

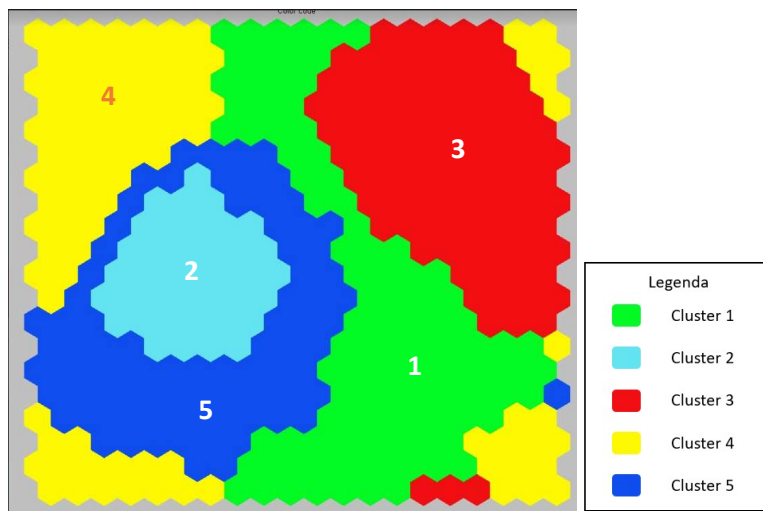
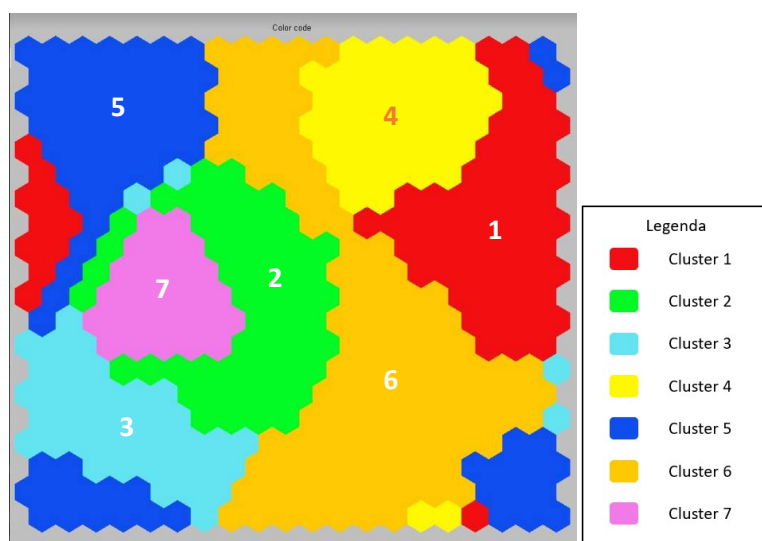


Figura 20 - Divisão da matriz-U em 7 classes



A partir dos resultados coletados do SOM, foram comparados os conjuntos de amostras pertencentes a cada grupo com os obtidos pelo método de agrupamento hierárquico aglomerativo (CHA) no trabalho de SHIMIZU (2012).

Tabela 5 - Porcentagem das amostras de SHIMIZU (2012) em cada *cluster* definido (4 grupos)

SOM/CHA	Cluster 1_CHA	Cluster 2_CHA	Cluster 3_CHA	Cluster 4_CHA
Cluster 1_SOM	0%	76,5%	0%	23,5%
Cluster 2_SOM	9,4%	37,5%	6,3%	46,9%
Cluster 3_SOM	91,7%	0%	8,3%	0%
Cluster 4_SOM	22,2%	0%	44,4%	33,3%

Tabela 6 - Porcentagem das amostras de SHIMIZU (2012) em cada *cluster* definido (7 grupos)

SOM/CHA	Cluster 1_CHA	Cluster 2_CHA	Cluster 3_CHA	Cluster 4_CHA	Cluster 5_CHA	Cluster 6_CHA	Cluster 7_CHA
Cluster 1_SOM	0%	0%	66,7%	16,7%	0%	16,7%	0%
Cluster 2_SOM	54,5%	0%	0%	0%	9,1%	0%	36,4%
Cluster 3_SOM	0%	0%	0%	11,1%	88,9%	0%	0%
Cluster 4_SOM	0%	14,3%	0%	0%	0%	71,4%	14,3%
Cluster 5_SOM	0%	0%	42,9%	47,6%	9,5%	0%	0%
Cluster 6_SOM	12,5%	8,3%	0%	12,5%	0%	4,2%	62,5%
Cluster 7_SOM	100%	0%	0%	0%	0%	0%	0%

Com a definição do número ótimo de *clusters* por meio do Índice *Davies-Bouldin*, foi realizado um estudo mais aprofundado considerando este cenário, dispondo da quantidade de amostras presentes em cada *cluster* e a qual corpo pertencem.

Tabela 7 - Amostras por *cluster* e corpos minerais

	Quantidade de Amostras			
	Sequeirinho	Sossego	Sem identificação	Total
<i>Cluster 1_SOM</i>	22	1	3	26
<i>Cluster 2_SOM</i>	11	1	1	13
<i>Cluster 3_SOM</i>	11	7	7	25
<i>Cluster 4_SOM</i>	2	25	1	28
<i>Cluster 5_SOM</i>	7	9	2	18
Total	53	43	14	110

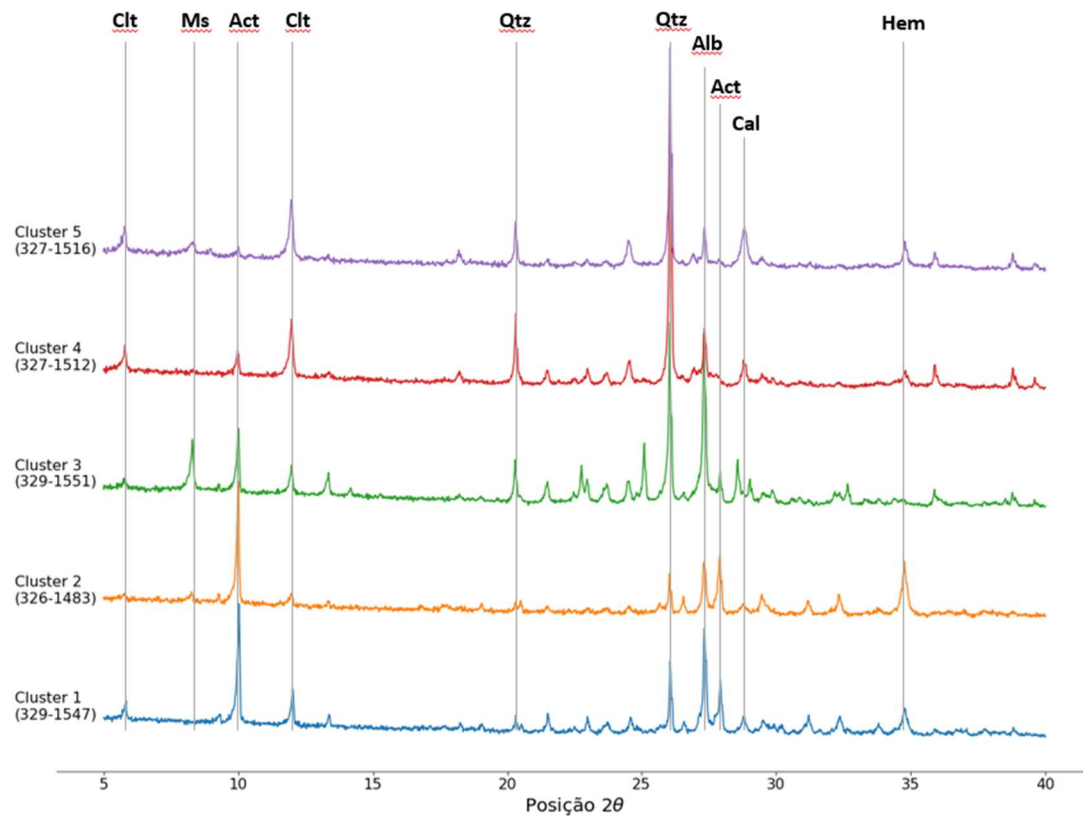
A amostra mais representativa de cada um dos *clusters*, ou seja, a amostra que se encontra mais próxima do centróide de cada grupo foi então calculada. Com isso, obteve-se o seguinte resultado conforme ilustrado na Tabela 8.

Tabela 8 - Amostra mais representativa de cada classe (5 grupos)

5 Classes	Amostras mais representativas
<i>Cluster 1_SOM</i>	1547
<i>Cluster 2_SOM</i>	1483
<i>Cluster 3_SOM</i>	1551
<i>Cluster 4_SOM</i>	1512
<i>Cluster 5_SOM</i>	1516

A partir das amostras mais representativas de cada um dos 5 *clusters*, foi possível determinar os principais minerais contidos em cada um deles, conforme os difratogramas dispostos na Figura 21.

Figura 21 – Difratomogramas das amostras mais representativas de cada *cluster*



Legenda: Clt – Clorita; Ms – Muscovita; Act – Actinolita; Qtz – Quartzo; Alb – Albita; Cal – Calcita e Hem - Hematita

Os difratogramas e as posições  $2\theta$  em relação as amostras também podem ser vistos no ANEXO 04 e 05.

Por fim, a fim de comparação com o estudo de SHIMIZU (2012), as amostras mais representativas para os 7 *clusters* nesse estudo foram determinadas, conforme Tabela 9.

Tabela 9 - Amostra mais representativa de cada classe (7 grupos)

7 Classes	Amostras mais representativas
Cluster 1_SOM	1463
Cluster 2_SOM	1522
Cluster 3_SOM	1516
Cluster 4_SOM	1438
Cluster 5_SOM	1512
Cluster 6_SOM	1547
Cluster 7_SOM	1483

## 4.2. Discussão

### 4.2.1. Comparação dos resultados entre classes

As análises foram desenvolvidas em três cenários distintos, com as amostras aglomeradas em 4, 5 e 7 classes. O trabalho de SHIMIZU (2012) analisou dois cenários distintos, de 4 e 7 *clusters*, sendo esses valores definidos pelo próprio usuário baseado em seus conhecimentos na área. Assim, as divisões em 4 e 7 *clusters* foram realizadas neste trabalho, a fim de comparar os resultados obtidos por um especialista por meio da Clusterização Hierárquica Aglomerativa e os resultados obtidos a partir da aglomeração dos vetores treinados pelos mapas auto-organizáveis.

Ademais, a divisão em 5 *clusters* foi considerada neste estudo, uma vez que o resultado do Índice *Davies-Bouldin*, conforme ilustrado na Figura 16, indicou qual seria o número ideal de classes. Esse número ideal é determinado pela maior quantidade de mínimos globais presentes em cada classe e, conforme a Figura 17, o número de *clusters* que apresentou o maior valor foi o de 5 classes, englobando 57% dos mínimos.

A primeira parte das análises, contabilizam o percentual de amostras definidas pela Clusterização Hierárquica Aglomerativa (CHA) de SHIMIZU (2012) presentes em cada classe determinada neste trabalho. Como exemplo, conforme ilustrado na Tabela 6, no *Cluster 1\_SOM* (*cluster* determinado neste estudo) contém 66,7% das amostras contidas no *Cluster 3\_CHA* (*cluster* determinado por SHIMIZU), 17% no *Cluster 4\_CHA* e 17% no *Cluster 6\_CHA*.

### Cenário 1 - 7 Clusters

O primeiro cenário considera as 7 classes. A partir dos resultados obtidos na Tabela 6, a classe que selecionou o total de amostras de um único *cluster* de SHIMIZU (2012) foi o *Cluster 7\_SOM*, contendo 100% de amostras do *Cluster 1\_CHA*.

Já o *Cluster 1\_SOM* predomina a quantidade de amostras do *Cluster 3\_CHA*, o *Cluster 3\_SOM* predomina a quantidade de amostras do *Cluster 5\_CHA*, o *Cluster 4\_SOM* predomina a quantidade de amostras do *Cluster 6\_CHA* e o *Cluster 6\_SOM* predomina a quantidade de amostras do *Cluster 7\_CHA*.

O *Cluster 2\_SOM* e o *Cluster 5\_SOM* apresentaram uma quantidade bem dividida entre as classes de SHIMIZU (2012), sendo o primeiro, a maioria das amostras pertencentes ao *Cluster 1\_CHA* e ao *Cluster 7\_CHA* e, o último, das amostras pertencentes ao *Cluster 3\_CHA* e *Cluster 4\_CHA*. Os dados comparados podem ser vistos no ANEXO 02.

Portanto, percebe-se que há uma boa correlação entre os *clusters* comparados, mostrando uma certa semelhança no agrupamento por Clusterização Hierárquica Aglomerativa e pelos *Self-Organizing Maps*. No entanto, é possível notar que os critérios de seleção das amostras em cada grupo não são os mesmos, uma vez que a maioria dos grupos determinados pelo SOM possuem uma mistura entre dois grupos determinados pela CHA, exceto pelo *Cluster 7\_SOM* que contém 100% das amostras do *Cluster 1\_CHA*.

No estudo de SHIMIZU (2012), do método de CHA, foram determinadas as amostras mais representativas para cada um dos 7 *clusters* analisados, com o auxílio do programa *Highscore da Malvern Panalytical*, conforme ilustrado na Tabela 10.

Tabela 10 - Amostras mais representativas de cada *cluster* determinado por SHIMIZU (2012)

Estudo de SHIMIZU (2012)	
7 Classes	Amostras mais representativas
<i>Cluster 1_CHA</i>	1483
<i>Cluster 2_CHA</i>	1495
<i>Cluster 3_CHA</i>	1446
<i>Cluster 4_CHA</i>	1512
<i>Cluster 5_CHA</i>	1526
<i>Cluster 6_CHA</i>	1551
<i>Cluster 7_CHA</i>	1538

Em comparação com as amostras mais representativas determinadas na Tabela 9, pode-se notar que apenas duas amostras são iguais, a do *Cluster 1\_CHA* e a do *Cluster 4\_CHA* com a do *Cluster 5\_SOM* e a do *Cluster 7\_CHA*, respectivamente. Esse resultado pode estar relacionado com a correlação encontrada entre os *clusters* na Tabela 6, podendo inferir que os métodos realizam um agrupamento das amostras baseado em diferentes características.



## **Cenário 2 - 4 Clusters**

Com relação à divisão em 4 classes, a partir dos resultados obtidos na Tabela 5, pode-se observar que o *Cluster 3\_SOM* foi o que mais apresentou uma quantidade de amostras de uma mesma classe que SHIMIZU (2012), neste caso, do *Cluster 1\_CHA*. O *Cluster 1\_SOM* teve a predominância de amostras do *Cluster 2\_CHA* e, por outro lado, o *Cluster 2\_SOM* e o *Cluster 4\_SOM* apresentaram uma quantidade de amostras bem dividida em comparação com as classes de SHIMIZU (2012). No caso, o *Cluster 2\_SOM* apresenta uma combinação de amostras do *Cluster 1\_CHA*, do *Cluster 2\_CHA* e do *Cluster 4\_CHA* e o *Cluster 4\_SOM* apresenta uma combinação de amostras do *Cluster 1\_CHA*, do *Cluster 3\_CHA* e do *Cluster 4\_CHA*. Os dados comparados podem ser vistos no ANEXO 03.

Portanto, percebe-se que há uma boa correlação entre os *clusters* comparados, mostrando uma certa semelhança no agrupamento por Clusterização Hierárquica Aglomerativa (CHA) e pelos *Self-Organizing Maps*, assim como no agrupamento em 7 classes. Também, é possível notar que os critérios de seleção das amostras em cada grupo não são os mesmos, uma vez que a maioria dos grupos determinados pelo SOM possuem uma mistura entre dois grupos determinados pela CHA.

## **Cenário 3 - 5 Clusters**

A divisão das amostras em 5 grupos foi definida como o número ótimo de *clusters* por meio do Índice *Davies-Bouldin* e, a partir disso, foi realizado um estudo mais aprofundado considerando este cenário.

Na Tabela 7, é possível observar a quantidade de amostras presentes em cada um dos grupos e sua respectiva localização, referente ao corpo de Sequeirinho ou ao corpo de Sossego. Há uma predominância de amostras do corpo de Sossego no *cluster 4* e do corpo de Sequeirinho nos *clusters 1* e *2*. Já os *clusters 3* e *5* apresentam amostras dos dois corpos, sem predominância de um deles, o que pode indicar que as amostras pertencem a uma zona de transição.

Uma análise mais aprofundada foi realizada por meio da interpretação dos difratogramas das amostras mais representativas de cada um dos grupos,

conforme ilustrado na Figura 21. No grupo 1, pode ser observado um pico relevante para a presença de actinolita, hematita e albita. Já o grupo 2, apresenta intensidades dos picos semelhantes aos do grupo 1, contendo actinolita, albita e hematita, no entanto, com menor intensidade para o último. O *cluster 4*, diferentemente dos dois grupos previamente mencionados, apresenta picos de intensidade relevantes para clorita e quartzo. Por fim, os cluster 3 e 5 apresentam picos com uma maior intensidade para muscovita, actinolita e quartzo no primeiro e clorita, quartzo, calcita e hematita no último.

Ao comparar esses resultados com os minerais presentes em cada um dos corpos, de Sossego e de Sequeirinho, a relação que foi analisada na Tabela 7 se justifica, uma vez que o primeiro apresenta o quartzo e a clorita como minerais relevantes, assim como determinado no *cluster 4*, e o último apresenta actinolita, hematita e albita como minerais relevantes, assim como observado nos grupos 1 e 2. Como o *cluster 3* e 5 apresentam picos de intensidades de minerais presentes em ambos os corpos, é possível entender que essas amostras são, portanto, representativas de uma zona de transição.

## 5. CONCLUSÃO

A comparação entre os métodos de Clusterização Hierárquica Aglomerativa (CHA) dos *Self Organizing Maps* (SOM) mostra que ambos auxiliam positivamente a técnica de difratometria de raios X, mas consideram critérios de agrupamentos diferentes. Essa diferença também faz com que as amostras mais representativas não sejam as mesmas, a depender dos critérios considerados no momento do agrupamento.

O maior diferencial entre os métodos foi a determinação do número ideal de *clusters*. O método do SOM em conjunto com o IDB possui alguns benefícios em comparação com o método de CHA, uma vez que evita viés de usuário no momento da determinação do *cutoff* e consegue garantir uma uniformidade ou padronização de escolha do número de classes.

Os resultados das análises utilizando o número ideal de clusters, a partir do IDB, com 5 classes, mostrou um resultado satisfatório, demonstrando um bom agrupamento entre as amostras mais semelhantes entre si e sendo possível

classificar o grupo de amostras pertencentes a cada corpo mineral e as fases de transição entre eles.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ANTONIASSI, J. **A difração de raios-X com o método de Rietveld aplicada a bauxitas de Porto Trombetas, PA.** Dissertação (Mestrado). Departamento de Engenharia de Minas e de Petróleo, Escola Politécnica da Universidade de São Paulo, São Paulo, 2010.

BERGERMAN, M. **Modelagem e simulação do circuito de moagem do Sossego.** 2009. Dissertação (Mestrado). Departamento de Engenharia de Minas e de Petróleo, Escola Politécnica da Universidade de São Paulo, São Paulo, 2009.

BREARD, G. T. ***Evaluating Self-Organizing Map Quality Measures as Convergence Criteria.*** 2017. Open Access Master's Theses. Paper 1033. Disponível em: <https://digitalcommons.uri.edu/theses/1033>. Acesso em: 15 set. 2021.

BRENTAN, B. et al. **Hybrid SOM+ *k-means* clustering to improve planning, operation and management in water distribution systems.** Environmental Modelling & Software. Vandoeuvre-les-Nancy: Centre de Recherche en Controle et Automatique de Nancy, 2018. v.106. p.77-88.

CARVALHEIRO, M. I. G. **Trajetória de acumulação de capacidades tecnológicas no setor mineral: a experiência de uma empresa brasileira no beneficiamento de cobre.** 2019. Dissertação (Mestrado). Escola Brasileira de Administração Pública e de Empresas, Centro de Formação Acadêmica e Pesquisa, Rio de Janeiro, 2019.

DAVIES, D. L.; BOULDIN, D. W. **A cluster separation measure.** IEEE transactions on pattern analysis and machine intelligence, v. 1, n. 2, p. 224–227, 1979.

FRASER S.J.; DICKSON. B. L. **A new method for data integration and integrated data interpretation: Self-Organising Maps**. Proceedings of Exploration 07: Fifth Decennial International Conference of Mineral Exploration, 2007.

JENKINS, R. Instrumentation. In: BISH, D.L.; POST, J.E. (Ed.). **Modern powder diffraction**. Washington: Mineralogy Society of America, 1989. p. 19-43. (Reviews in Mineralogy, v. 20).

KAHN, H. **Difração de raios-X**. São Paulo: Escola Politécnica, Universidade de São Paulo, 2007. 17 p. Curso apostilado. (Não publicado)

KITANI, E. C. **Mapeamento e visualização de dados em alta dimensão com mapas auto-organizados**. 2013. Dissertação (Doutorado). Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica da Universidade de São Paulo, São Paulo, 2013.

KOHONEN T. **Self-Organising Maps**. Third Ed. Helsinki University of Technology Neural Networks Research Centre, Finland, 2001.

KUMAR, S. **Agglomerative clustering and dendrograms**. 2020a. Disponível em: <https://towardsdatascience.com/agglomerative-clustering-and-dendrograms-explained-29fc12b85f23>. Acesso em: 05 out. 2021.

KUMAR, S. **Understanding K-Means, K-Means++ and K-Medoids clustering algorithms**. 2020b. Disponível em: <https://towardsdatascience.com/understanding-k-means-k-means-and-k-medoids-clustering-algorithms-ad9c9fbf47ca>. Acesso em: 05 out. 2021.

MACCHIAROLA K. et al. **Modern X-ray diffraction techniques for exploration and analysis of ore bodies**. In: DECENNIAL INTERNATIONAL CONFERENCE ON MINERAL EXPLORATION, 5, 2007, Toronto, Canadá.

MARQUES, I. P. **O depósito de óxido de ferro-cobre-ouro de Bacuri, Província Carajás, Pará: implicações de processos metassomáticos em rochas metaultramáficas para a evolução do sistema hidrotermal.** Dissertação (Mestrado). Instituto de Geociências, Universidade de São Paulo. São Paulo, 2015.

MARTINI, M. **Avaliação do potencial de expansão dos agregados siderúrgicos através da Difractometria de Raios-X e Método de Rietvelt.** Dissertação (Mestrado). Departamento de Engenharia de Materiais, Escola de Engenharia de Lorena, Universidade de São Paulo. Lorena, 2015.

MINGOTI, S. A.; LIMA, J.O. **Comparing SOM network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms.** European Journal of Operational Research. Belo Horizonte: Universidade Federal de Minas Gerais, 2006. v.174. p. 1742-1759

MIRANDA A. et. Al. **Moagem Semi-Autógena da usina do Sossego: Histórico de 10 anos de operação e otimizações.** In: HOLOS ISSN 1807-1600. Ano 31. 2015. v.7, p.43-50.

MONTEIRO, L. V. S. et al. **Spacial and temporal zoning of hydrothermal alteration and mineralization in the Sossego iron oxide-copper-gold deposit, Carajás Mineral Province, Brazil: paragenesis and stable isótopo constrains.** Mineralium Deposita, v. 43, p. 129-159, 2008.

PARUCH, P. **Xrdtools.** Disponível em: <https://github.com/paruch-group/xrdtools>. Acesso em: 05 set. de 2021.

PEDREGOSA F. et al. (2011). **Scikit-learn: Machine Learning in {P}ython.** *Journal of Machine Learning Research*, 12.

PERILLO, B. M. **Imputação e incremento de resolução espacial em dados geofísicos de bacias sedimentares *onshore*: abordagem a partir da técnica *Self Organizing Maps***. Dissertação (Mestrado). Departamento de Engenharia de Minas e de Petróleo, Escola Politécnica da Universidade de São Paulo, São Paulo, 2019.

ROSA, M.A.N. et al. **Controle operacional da usina do Sossego**. In: Meeting of the southern hemisphere on mineral technology e Encontro nacional de tratamento de minérios e metalurgia extrativa. Proceedings do VII Meeting of the southern hemisphere on mineral technology e XXII Encontro nacional de tratamento de minérios e metalurgia extrativa. Ouro Preto: Universidade Federal de Ouro Preto, 2007. v. III. p. 505-512

SHIMIZU, V. **Classificação e caracterização de tipos de minérios de cobre da mina do Sossego – Região de Carajás**. 2012. Dissertação (Mestrado). Departamento de Engenharia de Minas e de Petróleo, Escola Politécnica da Universidade de São Paulo, São Paulo, 2012.

SILVA, A. L. **Caracterização mineralógica por difração de raios-X e determinação de terras raras por ICP-MS de rochas da região sul da Bahia**. Dissertação (Mestrado). Departamento de Engenharia Nuclear, Escola de Engenharia, Universidade Federal de Minas Gerais. Belo Horizonte, 2013.

SINGH, A. K. et al. **Clustering evaluation by Davies-Bouldin Index (DBI) in cereal data using K-means**. In: Fourth International Conference on Computing Methodologies and Communication (ICCMC). IEEE Xplore Part Number. Mathura, 2020. p.306-310.

VALE S.A. **Vale: Nossa história**. Rio de Janeiro: Verso Brasil, 2012. 420 p.

VALE S.A. **Relatório de produção e vendas da Vale em 2019**. Disponível em: [http://www.vale.com/pt/investors/information-market/quarterly-results/resultadostrimestrais/preport4t19\\_p.pdf](http://www.vale.com/pt/investors/information-market/quarterly-results/resultadostrimestrais/preport4t19_p.pdf). Acesso em: 25 set. 2021.

VALE S.A. **Relatório Anual - Vale 2020.** Disponível em:  
[http://www.vale.com/PT/investors/information-market/annual-reports/20f/20FDocs/Vale%2020-F%20FY2020%20-%20Final%20Version\\_pt.pdf](http://www.vale.com/PT/investors/information-market/annual-reports/20f/20FDocs/Vale%2020-F%20FY2020%20-%20Final%20Version_pt.pdf). Acesso em: 01 out. 2021.

VALE S.A. **Website.** Disponível em:  
<http://www.vale.com/brasil/PT/sustainability/vale-nas-comunidades/para/Paginas/canaa-dos-carajas.aspx>. Acesso em: 25 set. 2021.

Agência Nacional de Mineração (ANM). **Anuário Mineral Brasileiro: principais substâncias metálicas** – Brasília: ANM, 2020. 30 p. Disponível em:  
[https://www.gov.br/anm/pt-br/centrais-de-conteudo/publicacoes/serie-estatisticas-e-economia-mineral/anuario-mineral/anuario-mineral-brasileiro/amb\\_2020\\_ano\\_base\\_2019\\_revisada2\\_28\\_09.pdf](https://www.gov.br/anm/pt-br/centrais-de-conteudo/publicacoes/serie-estatisticas-e-economia-mineral/anuario-mineral/anuario-mineral-brasileiro/amb_2020_ano_base_2019_revisada2_28_09.pdf). Acesso em: 05 set. 2021.

# ANEXO 01

## ANEXO 01 - IDENTIFICAÇÃO DE AMOSTRAS POR CLUSTERS

Identificação das amostras - 4 clusters					Identificação das amostras - 5 clusters					Identificação das amostras - 7 clusters						
Cluster 1	Cluster 2	Cluster 3	Cluster 4		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
323-1434	323-1439	323-1448	323-1437		323-1439	323-1448	323-1438	323-1434	323-1437	323-1434	323-1448	323-1437	323-1438	323-1435	323-1439	323-1452
323-1435	323-1450	323-1451	323-1441		323-1450	323-1451	323-1443	323-1435	323-1441	323-1443	323-1451	323-1441	323-1449	323-1436	323-1450	323-1453
323-1436	325-1466	323-1452	323-1447		325-1466	323-1452	323-1445	323-1436	323-1447	323-1444	325-1465	323-1447	325-1461	323-1440	325-1466	325-1476
323-1438	325-1487	323-1453	325-1465		325-1467	323-1453	323-1446	323-1440	325-1465	323-1445	325-1480	327-1514	326-1486	323-1442	325-1467	326-1483
323-1440	325-1468	325-1476	325-1477		325-1468	325-1476	323-1449	323-1442	325-1477	323-1446	326-1484	327-1515	326-1489	325-1470	325-1468	328-1534
323-1442	325-1472	326-1483	325-1480		325-1474	326-1483	325-1461	323-1444	325-1480	325-1462	326-1490	327-1516	326-1493	325-1471	325-1474	328-1536
323-1443	325-1474	326-1484	326-1490		325-1478	326-1484	325-1462	325-1463	326-1490	325-1463	327-1505	327-1517	326-1499	325-1472	325-1477	
323-1444	325-1478	327-1505	326-1491		325-1479	327-1505	325-1464	325-1470	326-1491	325-1464	328-1522	328-1525	326-1500	326-1487	325-1478	
323-1445	325-1479	328-1533	327-1504		326-1481	328-1533	325-1469	325-1471	327-1504	325-1469	328-1533	328-1526	327-1501	326-1497	325-1479	
323-1446	326-1481	328-1534	327-1515		326-1482	328-1534	325-1475	325-1472	327-1515	325-1473	328-1535	328-1530	327-1502	326-1498	326-1481	
323-1449	326-1482	328-1535	327-1516		326-1488	328-1535	326-1485	325-1473	327-1516	325-1475	328-1538	329-1542	327-1503	327-1508	326-1482	
325-1461	326-1488	328-1536	327-1517		326-1492	328-1536	326-1486	326-1487	327-1517	326-1485	328-1551	327-1509	326-1488	327-1509	326-1488	
325-1462	326-1492	329-1541	328-1522		326-1493	329-1541	326-1489	326-1487	328-1522	327-1506	328-1485	329-1541	329-1552	327-1510	326-1491	
325-1463	326-1493		328-1525		326-1494		326-1499	326-1488	328-1525	327-1507			327-1511	326-1492		
325-1464	326-1494		328-1526		326-1495		326-1500	327-1508	328-1526	328-1521			327-1512	326-1494		
325-1469	326-1495		328-1530		326-1496		327-1501	327-1509	328-1530	328-1523			327-1513	326-1495		
325-1470	326-1496		328-1538		327-1520		327-1502	327-1511	328-1538	328-1527			327-1518	326-1496		
325-1471	326-1497		329-1542		328-1524		327-1503	327-1510	329-1542	329-1549			327-1519	327-1504		
325-1473	327-1508		329-1546		328-1532		327-1506	327-1512		329-1550			328-1528	327-1520		
325-1475	327-1509		329-1547		328-1537		327-1507	327-1513					328-1529	328-1524		
326-1485	327-1510				328-1539		328-1521	327-1514					328-1531	328-1532		
326-1486	327-1511				329-1543		328-1527	327-1518					328-1539	328-1537		
326-1487	327-1512				329-1545		329-1549	327-1519					329-1543			
326-1489	327-1513				329-1546		329-1551	328-1523					329-1545			
326-1498	327-1514				329-1547		329-1552	328-1528					329-1546			
326-1499	327-1518				329-1550			328-1529					329-1547			
326-1500	327-1519							328-1531								
327-1501	327-1520							328-1530								
327-1502	328-1524															
327-1503	328-1532															
327-1506	328-1537															
327-1507	328-1539															
328-1521	328-1539															
328-1523	329-1543															
328-1525	329-1545															
328-1527	329-1547															
328-1528	329-1550															
328-1529																
328-1531																
329-1549																
329-1551																
329-1552																



## ANEXO 02

N LCT DRX	Cluster	Cluster SHIMIZU (7 grupos)	N LCT DRX	Cluster	Cluster SHIMIZU (7 grupos)
1434	1	3	1435	5	3
1443	1	3	1436	5	3
1444	1	3	1440	5	3
1445	1	3	1470	5	3
1446	1	3	1471	5	3
1462	1	3	1472	5	3
1469	1	3	1528	5	3
1473	1	3	1529	5	3
1475	1	3	1531	5	3
1506	1	3	1442	5	4
1507	1	3	1487	5	4
1527	1	3	1498	5	4
1463	1	4	1508	5	4
1521	1	4	1510	5	4
1523	1	4	1511	5	4
1464	1	6	1512	5	4
1549	1	6	1518	5	4
1550	1	6	1519	5	4
1448	2	1	1540	5	4
1451	2	1	1497	5	5
1465	2	1	1509	5	5
1484	2	1	1479	6	1
1533	2	1	1491	6	1
1535	2	1	1546	6	1
1541	2	5	1494	6	2
1480	2	7	1504	6	2
1490	2	7	1468	6	4
1522	2	7	1539	6	4
1538	2	7	1545	6	4
1514	3	4	1481	6	6
1441	3	5	1439	6	7
1515	3	5	1450	6	7
1516	3	5	1466	6	7
1517	3	5	1467	6	7
1525	3	5	1477	6	7
1526	3	5	1478	6	7
1530	3	5	1482	6	7
1542	3	5	1488	6	7
1493	4	2	1495	6	7
1438	4	6	1496	6	7
1489	4	6	1520	6	7
1501	4	6	1524	6	7
1551	4	6	1537	6	7
1552	4	6	1543	6	7
1486	4	7	1547	6	7
			1452	7	1
			1453	7	1
			1476	7	1
			1483	7	1
			1534	7	1
			1536	7	1

### ANEXO 03

N LCT DRX	Cluster	Cluster Viviane (4 grupos)	N LCT DRX	Cluster	Cluster Viviane (4 grupos)
1434	1	B	1439	2	D
1435	1	B	1450	2	D
1436	1	B	1466	2	D
1440	1	B	1467	2	D
1442	1	B	1478	2	D
1443	1	B	1481	2	D
1444	1	B	1482	2	D
1445	1	B	1488	2	D
1446	1	B	1495	2	D
1462	1	B	1496	2	D
1463	1	B	1520	2	D
1469	1	B	1524	2	D
1470	1	B	1537	2	D
1471	1	B	1543	2	D
1473	1	B	1550	2	D
1475	1	B	1448	3	A
1487	1	B	1451	3	A
1498	1	B	1452	3	A
1506	1	B	1453	3	A
1507	1	B	1476	3	A
1521	1	B	1483	3	A
1523	1	B	1484	3	A
1527	1	B	1533	3	A
1528	1	B	1534	3	A
1529	1	B	1535	3	A
1531	1	B	1536	3	A
1438	1	D	1541	3	C
1464	1	D	1465	4	A
1486	1	D	1491	4	A
1489	1	D	1504	4	A
1501	1	D	1546	4	A
1549	1	D	1441	4	C
1551	1	D	1515	4	C
1552	1	D	1516	4	C
1479	2	A	1517	4	C
1493	2	A	1525	4	C
1494	2	A	1526	4	C
1468	2	B	1530	4	C
1472	2	B	1542	4	C
1508	2	B	1477	4	D
1510	2	B	1480	4	D
1511	2	B	1490	4	D
1512	2	B	1522	4	D
1514	2	B	1538	4	D
1518	2	B	1547	4	D
1519	2	B			
1539	2	B			
1540	2	B			
1545	2	B			
1497	2	C			
1509	2	C			

# ANEXO 04

Anexo 04 - Posição 20 determinada no método de DRX

DRX-Amostra	Posição														
	a 2.0	a 2.02	a 2.03	a 2.05	a 2.07	a 2.09	a 2.1	a 2.12	a 2.14	a 2.16	a 2.17	a 2.19			
323-1434	1412,697	1325,984	1262,894	1169,882	1105,61	1012,205	957,677	886,811	832,087	774,016	705,512	643,898			
323-1435	1434,941	1327,461	1260,827	1180,61	1138,78	1055,118	998,13	908,366	848,13	771,752	717,126	652,461			
323-1436	1433,169	1344,685	1258,268	1168,406	1114,567	1030,413	956,004	913,091	815,846	783,858	704,823	651,87			
323-1437	1426,083	1341,535	1267,618	1182,283	1118,701	1056,398	984,154	930,217	854,626	798,13	731,299	673,031			
323-1438	1434,843	1364,567	1266,831	1194,291	1154,035	1059,252	1004,626	931,988	865,551	790,157	733,268	663,189			
323-1439	1441,043	1358,268	1241,929	1196,752	1106,398	1049,606	1001,673	921,161	866,339	793,701	710,433	655,02			
323-1440	1427,854	1355,61	1286,811	1190,945	1103,74	1040,157	990,256	870,276	822,835	770,276	715,354	668,209			
323-1441	1424,409	1329,528	1265,354	1162,106	1104,232	1036,811	944,291	895,571	835,63	771,26	719,882	639,567			
323-1442	1404,429	1321,358	1243,701	1166,043	1094,685	1023,819	963,78	899,803	821,26	774,803	684,449	639,173			
323-1443	1424,508	1316,732	1261,713	1181,988	1100,295	1036,909	958,268	897,933	832,283	759,843	693,996	653,051			
323-1444	1395,866	1325,394	1241,043	1166,831	1095,079	1041,339	978,15	915,748	848,819	787,303	731,201	652,264			
323-1445	1407,972	1324,508	1237,697	1183,465	1093,504	1023,524	963,976	890,945	842,224	782,283	711,22	652,854			
323-1446	1447,441	1327,953	1255,02	1177,854	1120,669	1041,142	954,331	897,146	864,665	783,858	721,85	669,783			
323-1447	1416,732	1326,378	1238,189	1184,055	1103,445	1056,201	993,11	910,925	845,866	786,811	712,008	652,362			
323-1448	1426,673	1349,902	1297,441	1205,709	1126,083	1065,846	982,48	931,89	843,504	787,303	724,016	656,102			
323-1449	1432,874	1375,197	1286,516	1228,839	1138,78	1066,929	1007,874	956,004	884,843	812,205	735,925	691,634			
323-1450	1440,256	1344,783	1266,437	1203,346	1115,256	1077,461	969,488	919,094	846,654	801,476	724,705	666,831			
323-1451	1435,63	1348,819	1262,008	1204,035	1116,437	1034,449	985,236	910,827	854,134	787,795	727,264	677,264			
323-1452	1428,248	1349,016	1266,634	1221,358	1144,39	1047,835	971,063	939,665	869,587	787,106	737,008	677,264			
323-1453	1430,906	1342,913	1280,512	1187,402	1127,854	1055,61	965,65	919,488	857,382	780,512	744,587	667,421			
325-1461	1400,197	1340,65	1254,823	1158,366	1090,846	1025,591	973,228	925,295	851,575	788,189	722,441	665,65			
325-1462	1419,39	1320,472	1228,445	1167,224	1094,98	1042,618	969,882	885,433	851,772	769,98	710,433	668,701			
325-1463	1425,886	1385,433	1270,079	1183,957	1106,398	1047,146	973,819	904,724	845,374	785,039	713,681	646,555			
325-1464	1408,071	1343,406	1279,232	1212,402	1126,575	1043,11	965,157	923,622	843,996	789,764	708,76	651,378			
325-1465	1425,098	1350,591	1264,665	1204,724	1122,736	1053,051	1004,035	936,516	860,531	807,874	747,736	673,524			
...	...	...	...	...	...	...	...	...	...	...	...	...			

## ANEXO 05

### Anexo 05 - Amostras Representativas: 5 Clusters

