



UNIVERSIDADE FEDERAL DE OURO PRETO
BACHARELADO EM ENGENHARIA DE MINAS
ESCOLA DE MINAS



UTILIZAÇÃO DE ALGORITMO K-MEANS PARA DEFINIÇÃO DE DOMÍNIOS LITOLÓGICOS

CAIO SEEBREGTS

Ouro Preto - MG
2022

Caio Seebregts

**UTILIZAÇÃO DE ALGORITMO K-MEANS PARA DEFINIÇÃO DE DOMÍNIOS
LITOLÓGICOS**

Monografia de Conclusão de Curso apresentado ao curso de Engenharia de Minas do Departamento de Engenharia de Minas da Escola de Minas da Universidade Federal de Ouro Preto, como requisito para conclusão do curso.

Área de concentração: Lavra de Minas

Orientador: Prof. Dr. Ivo Eyer Cabral

Coorientador: Dr. Allan Erlikhman Medeiro Santos

Ouro Preto – MG

2022

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S451u Seebregts, Caio.
Utilização de algoritmo k-means para definição de domínios
litológicos. [manuscrito] / Caio Seebregts. - 2022.
69 f.

Orientador: Prof. Dr. Ivo Eyer Cabral.
Coorientador: Prof. Dr. Allan Erlichman Medeiros Santos.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Escola de Minas. Graduação em Engenharia de Minas .

1. Fosfatos - Minas e mineração. 2. Análise de agrupamento. 3.
Mineração de dados (Computação) - Agrupamento k-means. I. Cabral, Ivo
Eyer. II. Santos, Allan Erlichman Medeiros. III. Universidade Federal de
Ouro Preto. IV. Título.

CDU 622.03:553.64

Bibliotecário(a) Responsável: Sione Galvão Rodrigues - CRB6 / 2526



ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Aos sete dias do mês de janeiro de 2022, às 16h00min, foi instalada a sessão pública remota para a defesa de Trabalho de Conclusão de Curso do discente **Caio Seebregts**, Matrícula: 19.1.1029, intitulado: “**UTILIZAÇÃO DE ALGORITMO K-MEANS PARA DEFINIÇÃO DE DOMÍNIOS LITOLÓGICOS**”, perante comissão avaliadora constituída pelo orientador do trabalho, Prof. Dr. Ivo Eyer Cabral, pelo Dr. Allan Erlikhman Medeiros Santos (Coorientador), pelo Prof. Dr. Carlos Enrique Arroyo Ortiz, e o M.Sc. Bruno Lagler. A sessão foi realizada com a participação de todos os membros por meio de videoconferência, com base no regulamento do curso e nas normas que regem as sessões de defesa de TCC. Inicialmente, o presidente da comissão examinadora concedeu ao discente 20 (vinte) minutos para apresentação do seu trabalho. Terminada a exposição, o presidente concedeu, a cada membro, um tempo máximo de 20 (vinte) minutos para perguntas e respostas ao discente sobre o conteúdo do trabalho, na seguinte ordem: primeiro o Prof. Dr. Carlos Enrique Arroyo Ortiz, segundo o M.Sc. Bruno Lagler, terceiro o Dr. Allan Erlikhman Medeiros Santos e, em último, o Prof, Dr. Ivo Eyer Cabral. Dando continuidade, ainda de acordo com as normas que regem a sessão, o presidente solicitou ao discente e aos espectadores que se retirassem da sessão de videoconferência para que a comissão avaliadora procedesse à análise e decisão. Após a reconexão do discente e demais espectadores, anunciou-se, publicamente, que o discente foi aprovado por unanimidade, com a nota 10,0 (dez, vírgula zero). O discente, por sua vez, encaminhará para o Repositório Institucional da UFOP, no prazo máximo de 15 (quinze) dias, uma versão final, contemplando todas as recomendações apresentadas pelos avaliadores. Para constar, foi lavrada a presente ata que, após aprovada, foi assinada pela presidente da comissão.

Ouro Preto, 07 de janeiro de 2022

Ivo Eyer Cabral

Presidente: Prof. Dr. Ivo Eyer Cabral

Membro: Dr. Allan Erlikhman Medeiros Santos

Membro: Prof. Dr. Carlos Enrique Arroyo Ortiz

Membro: M.Sc. Bruno Lagler

Discente: Caio Seebregts

RESUMO

A utilização de técnicas de agrupamento está cada vez mais presente no cotidiano, abrangendo diversas áreas do conhecimento e com os mais diferentes objetivos. No presente trabalho, foi feita uma análise de agrupamento utilizando o algoritmo *k-means* a fim de buscar a eliminação da subjetividade na descrição de testemunhos de sondagem de um banco de dados de uma mina de fosfato buscando a definição de domínios litológicos mais concisos. Além disso, foi realizada a validação das classes criadas utilizando o coeficiente da silhueta e análise de variância multivariada.

Palavras chaves: Análise de agrupamento; *k-means*; domínios litológicos.

ABSTRACT

The use of cluster analysis is increasingly present in everyday life, covering different areas of knowledge and aiming different objectives. In the present work, a cluster analysis was executed using the k-means algorithm in order to eliminate subjectivity in the description of drill hole core samples from a phosphate mine database pursuing more concise lithological domains. In addition, cluster's validation was performed using the silhouette score and multivariate analysis of variance.

Key-words: Cluster analysis, k-means, lithological domains.

LISTA DE FIGURAS

Figura 2.1: Localização da área de estudo na Faixa de Dobramentos Brasília. Extraído de Monteiro (2009).....	11
Figura 2.2: Localização da Mina Domingos, no município de Arraias-TO, a norte da cidade de Campos Belos-GO.....	12
Figura 2.3: Alternância do siltito argiloso (roxo) com o siltito arenoso (bege) que constituem o minério da Mina Domingos.....	13
Figura 3.1: Seção geológica indicando inconsistências no modelo.....	14
Figura 4.1: Localização dos furos de sondagem. Notar o adensamento dos furos na porção norte do depósito.....	17
Figura 5.1: Matriz de dispersão dos óxidos para as amostras descritas como "Arenito".....	21
Figura 5.2: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Arenito".....	22
Figura 5.3: Boxplot das análises dos óxidos para as amostras descritas como "Arenito".....	22
Figura 5.4: Matriz de dispersão dos óxidos para as amostras descritas como "Brecha-Tipo I".	23
Figura 5.5: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Brecha-Tipo I".	24
Figura 5.6: Boxplot das análises dos óxidos para as amostras descritas como "Brecha-Tipo I".	24
Figura 5.7: Matriz de dispersão dos óxidos para as amostras descritas como "Calcário".....	25
Figura 5.8: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Calcário".	26
Figura 5.9: Boxplot das análises dos óxidos para as amostras descritas como "Calcário".	26
Figura 5.10: Matriz de dispersão dos óxidos para as amostras descritas como "Dolomito"....	27
Figura 5.11: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Dolomito".	28
Figura 5.12: Boxplot das análises dos óxidos para as amostras descritas como "Dolomito"....	28
Figura 5.13: Matriz de dispersão dos óxidos para as amostras descritas como "Granito".....	29
Figura 5.14: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Granito".	30
Figura 5.15: Boxplot das análises dos óxidos para as amostras descritas como "Granito".....	30
Figura 5.16: Matriz de dispersão dos óxidos para as amostras descritas como "Marga".....	31

Figura 5.17: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Marga".	32
Figura 5.18: Boxplot das análises dos óxidos para as amostras descritas como "Marga".	32
Figura 5.19: Matriz de dispersão dos óxidos para as amostras descritas como "Silexito".	33
Figura 5.20: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Silexito".	34
Figura 5.21: Boxplot das análises dos óxidos para as amostras descritas como "Silexito".	34
Figura 5.22: Matriz de dispersão dos óxidos para as amostras descritas como "Siltito".	35
Figura 5.23: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Siltito".	36
Figura 5.24: Boxplot das análises dos óxidos para as amostras descritas como "Siltito".	36
Figura 5.25: Matriz de dispersão dos óxidos para as amostras descritas como "Siltito-Brechado".	37
Figura 5.26: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Siltito-Brechado".	38
Figura 5.27: Boxplot das análises dos óxidos para as amostras descritas como "Siltito-Brechado".	38
Figura 5.28: Matriz de dispersão dos óxidos para as amostras descritas como "Solo".	39
Figura 5.29: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Solo".	40
Figura 5.30: Boxplot das análises dos óxidos para as amostras descritas como "Solo".	41
Figura 5.31: Teste da Silhueta para diferentes quantidades de classes.	42
Figura 5.32: Histograma de Frequência das Litologias por Classes.	43
Figura 5.33: Ocorrência de amostras descritas como Siltito em cada Classe, notar a concentração nas Classes 4, 3 e 6.	44
Figura 5.34: Análise de componentes principais para três componentes de acordo com as litologias descritas.	44
Figura 5.35: Análise de componentes principais para três componentes de acordo com as classes geradas.	45
Figura 5.36: Coeficiente da silhueta para as amostras agrupadas na Classe 1.	46
Figura 5.37: <i>Boxplot</i> das análises dos óxidos para a Classe 1. Notar a grande quantidade de <i>outliers</i> e os valores altos de SiO ₂ , Al ₂ O ₃ , Fe ₂ O ₃ e de LOIPer.	47
Figura 5.38: Disposição típica da Classe 1 (em azul) na porção norte do depósito. Notar a concentração na superfície e a ocorrência em menor quantidade em subsuperfície.	47

Figura 5.39: Coeficiente da silhueta para as amostras agrupadas na Classe 2.	48
Figura 5.40: Boxplot das análises dos óxidos para a Classe 2.	49
Figura 5.41: Disposição típica da Classe 2 (em vermelho), notar relação espacial com a Classe 1 em alguns pontos (em azul).	49
Figura 5.42: Coeficiente da silhueta para as amostras agrupadas na Classe 3.	50
Figura 5.43: Boxplot das análises dos óxidos para a Classe 3.	51
Figura 5.44: Disposição típica da Classe 3 (em verde). Classe 1 em azul e Classe 2 em vermelho.	51
Figura 5.45: Coeficiente da silhueta para as amostras agrupadas na Classe 4.	52
Figura 5.46: Boxplot das análises de óxidos para a Classe 4.	52
Figura 5.47: Disposição da Classe 4 (em roxo). Classe 1 em azul, Classe 2 em vermelho e Classe 3 em verde.	53
Figura 5.48: Coeficiente da silhueta para as amostras da Classe 5.	54
Figura 5.49: Boxplot das análises dos óxidos para a Classe 5.	54
Figura 5.50: Correlograma dos óxidos da Classe 5.	55
Figura 5.51: Gráfico de Dispersão de P_2O_5 e CaO para a Classe 5.	55
Figura 5.52: Localização das amostras atribuídas à Classe 5 (Em amarelo). Notar a associação com as amostras da Classe 10 (Em azul claro), relacionada aos carbonatos.	56
Figura 5.53: Coeficiente da silhueta para as amostras atribuídas a Classe 6.	57
Figura 5.54: Boxplot das análises dos óxidos para a Classe 6.	58
Figura 5.55: Gráfico de Dispersão de P_2O_5 e CaO para a Classe 6.	58
Figura 5.56: Localização das amostras atribuídas à Classe 6 (Em rosa). Classe 1 em azul, Classe 2 em vermelho, Classe 3 em verde, Classe 4 em roxo, Classe 5 em amarelo.	59
Figura 5.57: Coeficiente da silhueta para amostras atribuídas a Classe 7.	59
Figura 5.58: Boxplot das análises dos óxidos para a Classe 7.	60
Figura 5.59: Localização da Classe 7 no setor norte do depósito.	60
Figura 5.60: Coeficiente da silhueta para as amostras atribuídas a Classe 8.	61
Figura 5.61: Boxplot das análises dos óxidos para a Classe 8.	62
Figura 5.62: Correlograma entre os valores de óxidos para a Classe 8.	62
Figura 5.63: Localização de algumas amostras atribuídas a Classe 8 na porção central do depósito.	63
Figura 5.64: Coeficiente da silhueta para as amostras atribuídas a Classe 9.	63
Figura 5.65: Boxplot das análises dos óxidos para a Classe 9.	64

Figura 5.66: Localização de algumas das amostras atribuídas a Classe 9 no centro do depósito.	64
Figura 5.67: Coeficiente da silhueta para amostras atribuídas a Classe 10.	65
Figura 5.68: Boxplot das análises dos óxidos para a Classe 10	66
Figura 5.69: Localização das amostras atribuídas a Classe 10 no centro do depósito.	66
Figura 5.70: Seção mostrando a organização das Classes a partir do entendimento de três classes distintas.	67
Figura 5.71: Teor de P_2O_5 em diferentes furos de sondagem. Notar a forte correlação das Classes 5 e 8 com os teores de P_2O_5	67

LISTA DE TABELAS

Tabela 4.1: Disposição do banco de dados com os valores de óxidos e da perda por calcinação.	15
Tabela 5.1: Distribuição das litologias descritas nas diferentes Classes.	43
Tabela 5.2: Resultado da análise de variância multivariada (MANOVA) para as classes geradas.	45

SUMÁRIO

1	INTRODUÇÃO	10
2	CONTEXTUALIZAÇÃO GEOLÓGICA.....	11
3	JUSTIFICATIVA E OBJETIVOS	13
4	MATERIAIS E MÉTODOS	15
5	RESULTADOS E DISCUSSÕES	20
5.1	Análise Exploratória dos Dados.....	20
5.1.1	Arenito	21
5.1.2	Brecha-Tipo I.....	23
5.1.3	Calcário.....	24
5.1.4	Dolomito.....	27
5.1.5	Granito	29
5.1.6	Marga.....	31
5.1.7	Silexito.....	33
5.1.8	Siltito	35
5.1.9	Siltito Brechado	37
5.1.10	Solo.....	39
5.2	Análise de Agrupamentos	41
5.2.1	Classe 1	46
5.2.2	Classe 2.....	48
5.2.3	Classe 3.....	49
5.2.4	Classe 4.....	51
5.2.5	Classe 5.....	53
5.2.6	Classe 6.....	56
5.2.7	Classe 7.....	59
5.2.8	Classe 8.....	61
5.2.9	Classe 9.....	63
5.2.10	Classe 10.....	65
5.2.11	Visão Geral	66
6	CONCLUSÕES.....	68
	REFERÊNCIAS	69

1 INTRODUÇÃO

A aquisição de dados geológicos para a posterior modelagem e análise econômica de um depósito mineral é uma atividade que demanda tempo, dinheiro e bons profissionais. Muitas vezes, a interpretação geológica de um depósito muda conforme mais informações são obtidas ao longo do tempo. Tais mudanças podem mostrar que informações valiosas coletadas, sob um entendimento vigente, não são mais úteis, sendo necessário dispensar ainda mais dinheiro e tempo para obtenção de novos dados. Isso é especialmente verdade na descrição de testemunhos de sondagem, que, além de estar submetida ao entendimento geológico vigente, uma grande quantidade de subjetividade do profissional está presente.

Quando levados a fases posteriores, sem a devida correção, tais dados podem levar a interpretações errôneas que vão desde a superestimação da quantidade de minério presente no depósito, ou desempenho altamente variável da planta de processamento durante as operações, a um aumento do CAPEX. De acordo com Michaux (2020) tais problemas são consequência de um entendimento parcial do que era economicamente viável dentro do depósito mineral.

A utilização de métodos de análise de agrupamento (“*cluster analysis*”) se tornou cotidiana e abrange as mais diferentes áreas: ciências econômicas (Monfort *et al.*, 2013), publicidade (Punj & Stewart, 1983) e diversas áreas da engenharia (Prassas *et al.*, 1996; Dindarloo & Siami-Irdemoosa, 2017; Fouedjio, 2016). Tal metodologia é uma ferramenta que busca encontrar padrões em dados multivariados a partir de diferentes características e propriedades (Koch, 2013).

Uma análise de agrupamento pode ser entendida como bem sucedida quando observações dentro de um mesmo grupo sejam semelhantes entre si e significativamente distintas de observações pertencentes a outros grupos. O algoritmo de agrupamento *k-means*, proposto por MacQueen (1967), consiste no agrupamento de uma população multidimensional em *k* diferentes grupos, baseado na distância da observação e do centroide do grupo seu grupo.

Dessa maneira, o presente trabalho apresenta a aplicação do algoritmo de agrupamento *k-means* a um banco de dados geológico resultado de diferentes campanhas de sondagem. Além disso, traz a comparação entre o agrupamento dos dados antigos e o obtido a partir da aplicação do algoritmo, utilizando diferentes tipos de validação, desde a visual até a estatística. Finalmente, aponta como as classes encontradas com o uso do algoritmo se relacionam com o contexto geológico local.

2 CONTEXTUALIZAÇÃO GEOLÓGICA

A área está inserida no contexto geotectônico do Cráton do São Francisco, na porção noroeste da Bacia do São Francisco, à leste da Faixa de Dobramentos Brasília (Fig. 2.1) (Dardenne et al., 1986). A Faixa de Dobramentos Brasília é um cinturão de dobras e falhas de empurrão neoproterozoico com vergência em direção ao Cráton do São Francisco (Fuck et al., 1994).

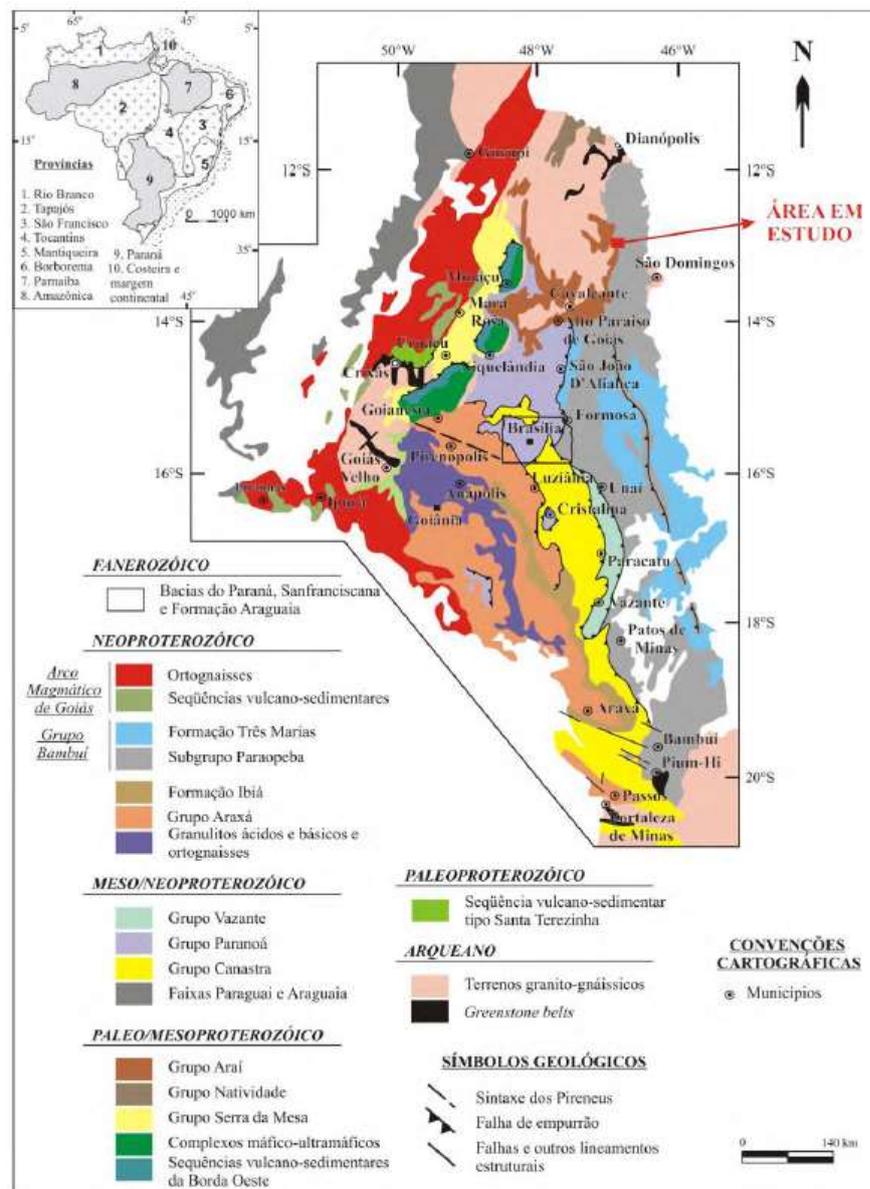


Figura 2.1: Localização da área de estudo na Faixa de Dobramentos Brasília. Extraído de Monteiro (2009)

Dentro desse contexto, encontra-se o Grupo Bambuí com aproximadamente 2000m de espessura e interpretado como tendo gênese durante uma transgressão pós glacial em um mar epicontinental quente (Sial et al., 2009; Drummond et al., 2015). O Grupo Bambuí é formado, da base para o topo, pela Formação Sete Lagoas, Formação Serra de Santa Helena, Formação Serra do Jacaré, Formação Serra da Saudade e Formação Três Marias.

A ocorrência de rochas fosfatadas do Grupo Bambuí está relacionada à Formação Sete Lagoas, distribuída ao longo da borda leste do Cráton São Francisco (Fig. 2.1). Esta formação possui aproximadamente 350m de espessura e é constituída por pelo menos duas sequências sedimentares: a sequência basal, constituída por rochas carbonáticas dobradas; e a sequência de siltitos e arenitos fosfatados (Drummond et al., 2015).



Figura 2.2: Localização da Mina Domingos, no município de Arraias-TO, a norte da cidade de Campos Belos-GO.

Na região do município de Arraias, no Estado do Tocantins, a ocorrência da Formação Sete Lagoas da origem à Mina de Domingos, foco do presente estudo (Fig 2.2). Nessa mina, o minério é constituído principalmente por uma alternância de siltitos argilosos e siltitos arenosos fosfatados que se encontram sobre calcários e dolomitos (Fig 2.3).

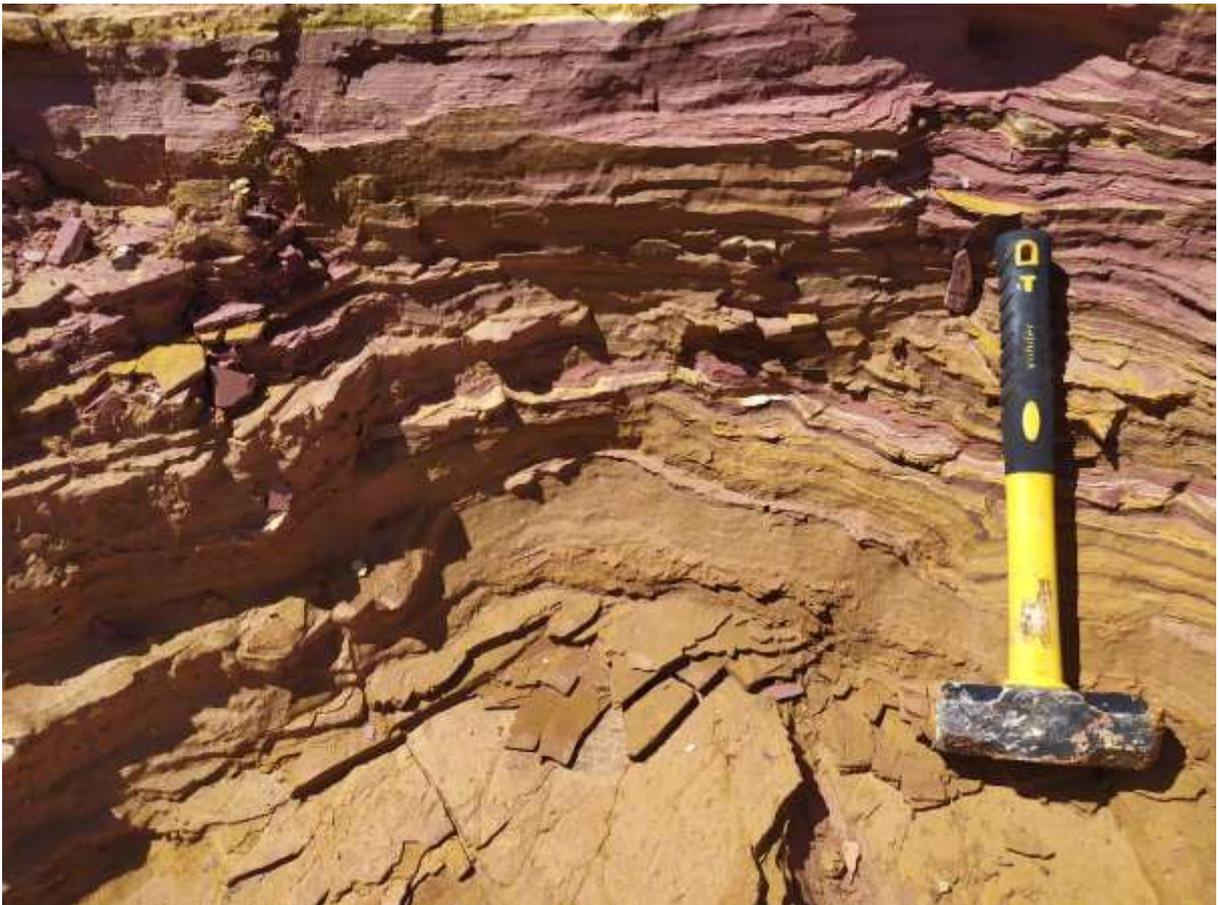


Figura 2.3: Alternância do siltito argiloso (roxo) com o siltito arenoso (bege) que constituem o minério da Mina Domingos.

3 JUSTIFICATIVA E OBJETIVOS

Objetivou-se determinar domínios para a modelagem de recursos minerais a partir da análise de agrupamentos através do uso do algoritmo *k-means* (MacQueen, 1967). Tal estudo é importante pois por se tratar de um banco de dados histórico de um depósito onde ocorreram mudanças no entendimento geológico com o passar do tempo (Drummond, 2015; Choupina & Franco, 2020) e cuja a modelagem foi feita com base na descrição litológica de poucos furos de sondagem diamantada e muitos furos de sondagem de circulação reversa, a qual é de difícil

descrição por conta da variação vertical do corpo mineralizado e pela natureza das amostras obtidas. Dessa forma, é possível, no modelo geológico, encontrar algumas discrepâncias.

A Figura 3.1 apresenta algumas das discrepâncias e pode-se observar a presença de furos de sondagem que terminam ainda no corpo mineralizado e furos de sondagem que estão lado a lado com diferentes litologias, algo que não é de trivial ocorrência em um depósito sedimentar. Dessa maneira, a determinação de domínios para a modelagem de recursos minerais a partir de técnicas de agrupamento possibilita a minimização da subjetividade da descrição litológica feita pelo geólogo a partir do entendimento geológico vigente e garante uma consideração maior das relações estatísticas e espaciais entre as amostras.



Figura 3.1: Seção geológica indicando inconsistências no modelo.

Isto posto, o presente trabalho tem como objetivo a tentativa de utilização de técnicas de agrupamento para a definição de domínios para a modelagem geológica e estimativa de recursos minerais, levando em consideração as relações estatísticas e espaciais das amostras, em detrimento da subjetividade da descrição litológica.

4 MATERIAIS E MÉTODOS

Para o presente estudo, utilizou-se um banco de dados de amostras oriundas de testemunhos de sondagem circulação reversa, todos feitos com direção vertical. As amostras utilizadas foram as que possuíam todos os resultados químicos para 9 óxidos (P₂O₅, Al₂O₃, CaO, Fe₂O₃, K₂O, MgO, MnO₂, Na₂O, SiO₂) e da perda por calcinação (LOIper) e descrição litológica (Tabela 4.1). Desta forma, todas as amostras que não possuíam todos os resultados e a descrição litológica foram descartadas, passando de um banco de dados de 14794 provenientes de sondagem de circulação reversa para um banco de dados com 10493 amostras. O comprimento de todas as amostras utilizadas é de 1 metro, devido à natureza da sondagem de circulação reversa.

Tabela 4.1: Disposição do banco de dados com os valores de óxidos e da perda por calcinação.

HOLE-ID	FROM	TO	HOLE TYPE	P2O5	Al2O3	CaO	Fe2O3	K2O	MgO	MnO2	Na2O	SiO2	LOIper	Litologia
DOMI-RC-0001	0	1	RC	0.09	17.8	0.15	11.05	1.14	0.63	0.04	0.005	61.3	7.63	SOLO
DOMI-RC-0001	1	2	RC	0.06	19.75	0.06	7.65	1.05	0.58	0.04	0.005	62.5	7.54	SILTITO
DOMI-RC-0001	2	3	RC	0.08	17.9	0.04	7.07	1.12	0.62	0.1	0.005	65.7	6.75	SILTITO
DOMI-RC-0001	3	4	RC	0.11	16.8	0.07	6.77	1.66	0.82	0.14	0.005	66.2	6.24	SILTITO
DOMI-RC-0001	4	5	RC	0.12	16.2	0.12	6.25	2.8	2.02	0.1	0.02	65.6	5.56	SILTITO
DOMI-RC-0001	5	6	RC	0.12	15.9	0.1	6.22	3.12	2.68	0.1	0.08	65.4	5.54	SILTITO
DOMI-RC-0001	6	7	RC	0.13	15.95	0.02	6.37	3.27	3.86	0.09	0.16	63.1	5.66	SILTITO
DOMI-RC-0001	7	8	RC	0.14	15.35	0.06	6.63	3.24	2.7	0.16	0.09	63.9	5.45	SILTITO
DOMI-RC-0001	8	9	RC	0.15	15.7	0.02	7.12	3.23	2.95	0.18	0.09	63.7	5.62	SILTITO
DOMI-RC-0001	9	10	RC	0.15	15.6	0.03	7.14	3.22	3.6	0.18	0.14	63.6	5.67	SILTITO
DOMI-RC-0001	10	11	RC	0.14	15.4	0.04	6.43	3.39	3.62	0.17	0.19	64.7	5.34	SILTITO

As litologias são clássicas das sucessões pelito-carbonatadas da Formação Sete Lagoas, sendo elas:

- i. *Siltito*: Rocha sedimentar clástica constituída predominantemente por fragmento de mineral ou rocha muito fino do tamanho silte (diâmetro 1/256 a 1/16 mm).
- ii. *Silexito*: Variedade de *chert* que ocorre principalmente associada a rochas carbonáticas ou sílico-carbonáticas, em nódulos ou camadas originadas, muitas vezes, por silicificação, desde singenética até epigenética. *Cherts*, no sentido *sensu strictu*, são formados a partir de precipitação química de sílica.
- iii. *Arenito*: Rocha sedimentar clástica cujas partículas são predominantemente do tamanho de areia (0,62 a 2,00 mm de diâmetro). O arenito possui frequentemente, matriz fina, síltico-argilosa, e cimento que pode ser de sílica, óxido/hidróxido de ferro e calcita. O

tamanho de areia que predomina pode dar os nomes: arenito grosso, médio ou fino. Muitas vezes apresenta boa estratificação, mas pode se apresentar maciço.

- iv. *Margas*: Rocha sedimentar detrítica-carbonatada. Apresenta frequentemente estruturas sedimentares, como estratificações e concreções, além de fósseis. Componentes: Mistura de calcita (menos frequentemente dolomita) e minerais argilosos, com vestígios de quartzo, micas e resíduos carbonosos. São frequentes nódulos de gesso, calcita e pirita.
- v. *Calcário*: Rocha sedimentar carbonatada composta essencialmente por carbonato de cálcio.
- vi. *Dolomitos*: Rocha carbonatada cujo constituinte principal é a dolomita. O dolomito está relacionado à substituição do Ca de calcários micríticos e outros por Mg por dolomitização.
- vii. *Brecha Tipo 1*: Clastos poliderivados, mal classificados e mal selecionados. Com silicificação.
- viii. *Solo*: engloba a unidade de solo que ainda não apresenta textura/estrutura da rocha pretérita.
- ix. *Granito*: Granitos e gnaisses da Suíte Aurumina.
- x. *Siltito Brechado*: Material silicificado, muito consistente e de alta dureza, geralmente bandado e associado a zonas de deformação estrutural. Os teores geralmente são altos, superiores a 10% P₂O₅.

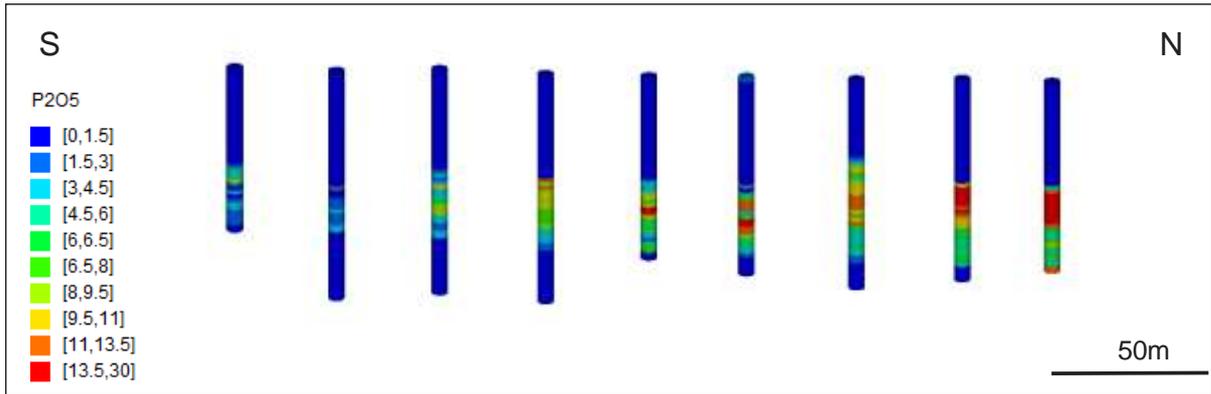


Figura 4.2: Seção mostrando os teores de P₂O₅ para uma seção típica do depósito.

A finalidade da utilização de uma técnica de agrupamentos é a criação de grupos de maneira que as observações amostrais dentro do mesmo grupo possuam um bom grau de similaridade, enquanto observações de grupos distintos possuam um bom grau de diferenças. Além disso, por se tratarem de dados geológicos, com localização espacial definida, é necessário também que os grupos encontrados possuam uma boa correlação espacial, possuindo sentido geológico.

Dessa maneira, optou-se o algoritmo *k-means* para a realização do agrupamento devido a sua extensa aplicação nas mais diversas situações (Hennig et al. 2016). Foi proposto por MacQueen (1967) para agrupar dados numéricos de maneira que cada grupo possuí um centroide chamado média (*mean*, em inglês). Antes da utilização do algoritmo, é necessário que se especifique o número de grupos a serem gerados de maneira que a se minimizar a função erro. Seja X o conjunto de dados a ser agrupado, e $C = \{C_1, \dots, C_n\}$ os diferentes agrupamentos distintos k de X , dessa maneira, a função erro pode ser definida como (Gan et al. 2017):

$$E = \sum_{i=1}^k \sum_{x_n \in C_i} d(x_n, c_i)$$

Onde x_n é uma observação pertencente ao agrupamento C_i , que possui como centroide c_i , e $d(x_n, c_i)$ é a distância entre x_n e c_i . Inicialmente, as observações são colocadas em k agrupamentos iniciais, que podem ser selecionados aleatoriamente. Posteriormente, o centro de cada grupo é calculado como um vetor médio de todas as observações encontradas no determinado grupo. A distância entre o centroide do agrupamento e a observação é calculado e essa é atribuída ao agrupamento mais próximo. O processo é repetido até que o valor da função erro seja minimizado e não mude significativamente com o aumento de grupos. A realização

do algoritmo *k-means* no presente trabalho se deu com o uso do *software* Minitab, que utiliza a distância euclidiana para a geração dos agrupamentos.

Para validação dos agrupamentos criados utilizou-se a metodologia análise multivariada de variância (MANOVA, em inglês), tal metodologia surge da necessidade de comparação de uma ou mais populações e é usada para entender relações entre variáveis categóricas independentes e duas ou mais variáveis dependentes métricas (Johnson, 2018; Hair Jr. et al., 2009 *apud* Nazareth, 2020).

A MANOVA baseia-se em quatro testes multivariados principais, sendo eles: De Wilk, Lawley-Hotelling, Pillai e maior raiz de Roy. De acordo com Nazareth (2020), as interpretações de cada teste para rejeitar ou não a hipótese nula (quando todos os grupos pertencem a mesma população) são as seguintes:

1. Teste de Wilk: No caso do Teste de Wilk, se os grupos se assemelharem, o valor do teste será próximo de 1, e, em contrapartida, quanto mais próximo de 0 o valor do teste, mais diferentes são os grupos.
2. Teste de Lawley-Hotelling: É um teste de valor positivo, onde o aumento do valor indica efeitos que contribuem para o modelo.
3. Teste de Pillai: De maneira análoga ao Teste de Lawley-Hotelling, é um teste estatístico de valor positivo, onde, também, o aumento do valor do teste contribui para o modelo.
4. Teste de maior raiz de Roy: De maneira análoga aos testes de Lawley-Hotelling e de Pillai, maiores valores para esse teste apontam para diferenças significativas entre os grupos.

No teste MANOVA, o p-valor é a probabilidade de se obter resultados ao menos tão extremos quanto os resultados realmente observados, levando em conta que a hipótese nula é verdadeira (Greenland *et al.*, 2016). Dessa forma, um valor de p-valor pequeno significa que o resultado obtido seria improvável levando em conta a veracidade da hipótese nula (Greenland *et al.*, 2016).

Além do teste MANOVA, realizou-se o teste das silhuetas para cada agrupamento criado. Esse teste consiste em um método gráfico que apresenta em sua largura o coeficiente da silhueta. Esse coeficiente é calculado para cada amostra pertencente ao grupo, usando a distância média entre ela e as outras amostras do grupo e a distância entre ela e as amostras do grupo mais próximo de acordo com a fórmula: (Rousseeuw, 1987 *apud* Moreira, 2020).

$$s = \frac{(b - a)}{\max(a, b)}$$

onde b é a distância da amostra e as amostras do mesmo grupo e a , a distância média entre a amostra e cada uma das amostras do grupo mais próximo. Dessa maneira, os valores de s se encontram dentro do intervalo $[-1,+1]$, sendo que valores mais próximos de 1 indicam um grupo bem definido, enquanto valores próximos de 0 indicam sobreposições entre os grupos e valores negativos podem indicar que a amostra foi indevidamente atribuída ao respectivo grupo (Moreira, 2020). A altura da silhueta indica a quantidade de amostras inseridas no determinado grupo.

Ademais, com as classes definidas e validadas tanto pelo teste da silhueta como pela análise multivariada de variância, foram realizadas seções geológicas para entender a relação espacial dos grupos, além da sua aderência com a geologia local. As seções foram posicionadas de modo a observar o comportamento espacial de todos os agrupamentos encontrados com o uso do *k-means*.

5 RESULTADOS E DISCUSSÕES

5.1 Análise Exploratória dos Dados

Neste capítulo, serão dispostas as informações estatísticas das diferentes litologias presentes no banco de dados. A Figura 5.1.1 mostra o histograma de frequência das amostras descritas como cada litologia.

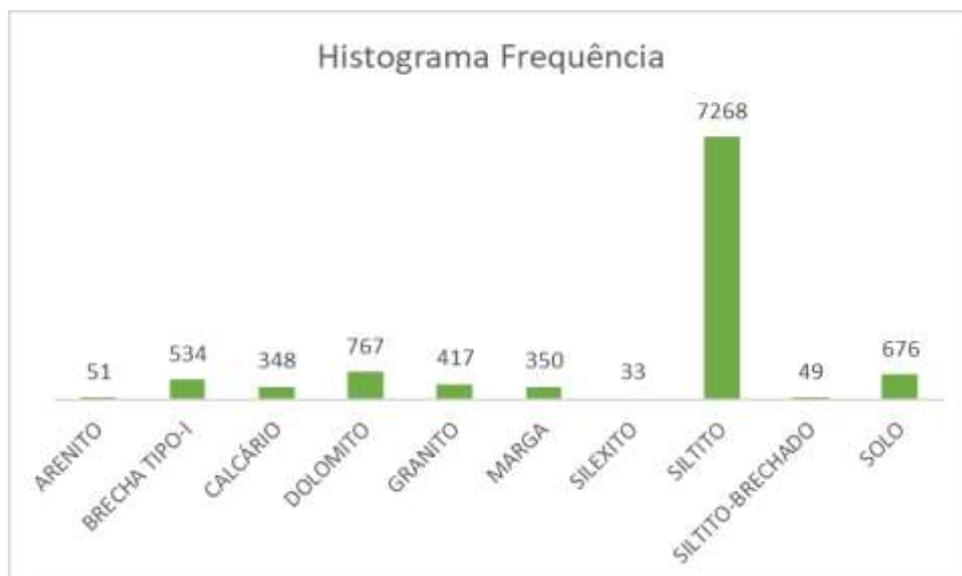


Figura 5.1.1: Frequência de amostras descritas conforme cada litologia do banco de dados.

5.1.1 Arenito

A litologia “Arenito” foi descrita 51 vezes no banco de dados. A Figura 5.1 apresenta a matriz de dispersão dos óxidos dessas amostras. Nota-se correlações positivas ótimas entre CaO e P₂O₅ (Coeficiente de correlação de Pearson = 0,93), K₂O e MgO (Coeficiente de correlação de Pearson = 0,98), MgO e Al₂O₃ (Coeficiente de correlação de Pearson = 0,84) e K₂O e Al₂O₃ (Coeficiente de Correlação de Pearson = 0,89) (Fig. 5.2). Já a correlação entre o SiO₂ e os outros óxidos é claramente negativa, sendo maior para o Fe₂O₃ (Coeficiente de correlação de Pearson = -0,86) e para o Al₂O₃ (Coeficiente de correlação de Pearson = -0,71) (Fig. 5.2).

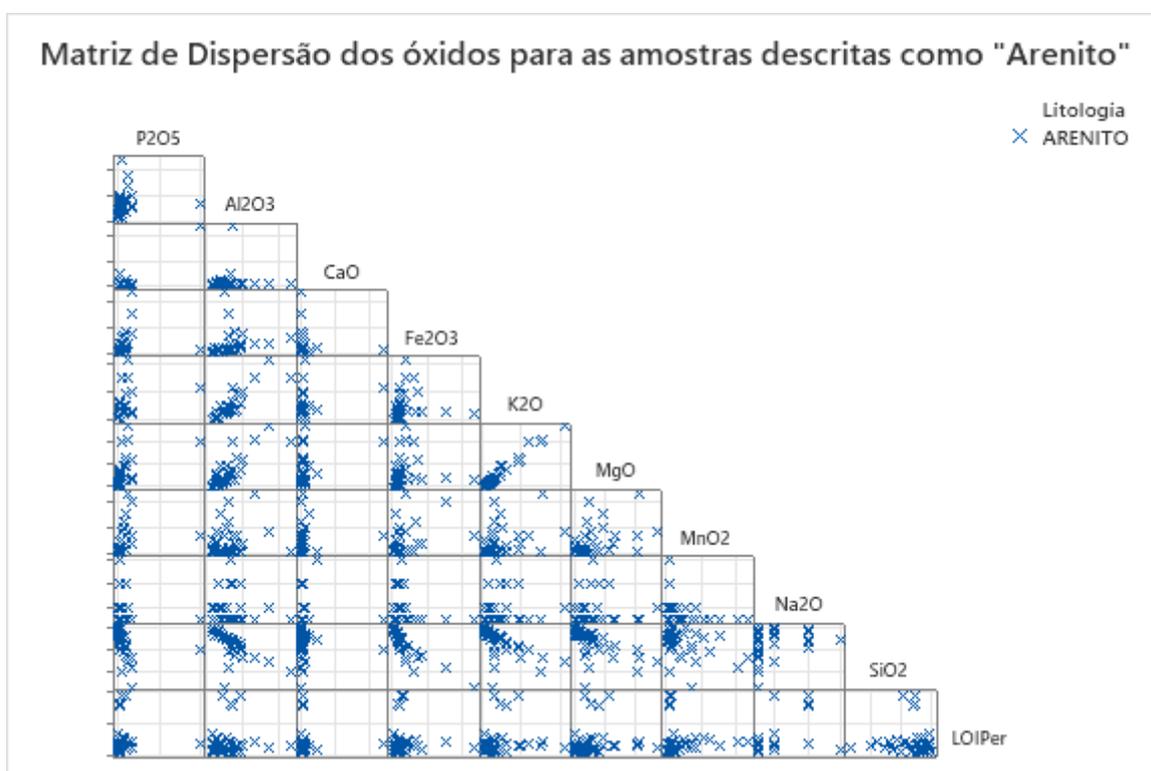


Figura 5.1: Matriz de dispersão dos óxidos para as amostras descritas como "Arenito".

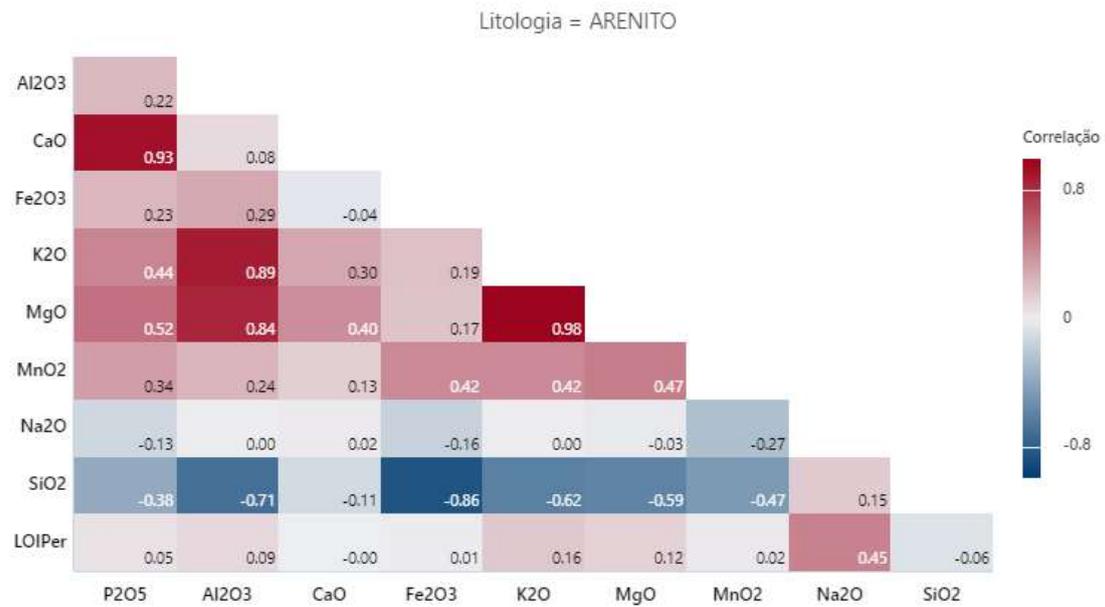


Figura 5.2: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Arenito".

A Figura 5.3 mostra o *boxplot* dos resultados das análises químicas das amostras de Arenito. É possível notar um alto valor de SiO₂, condizente com um arenito quartzoso, e valores expressivos de Al₂O₃, Fe₂O₃ e de PPC (LOIPer), relacionados com a matriz siltico-argilosa e com a cimentação com óxidos de ferro.

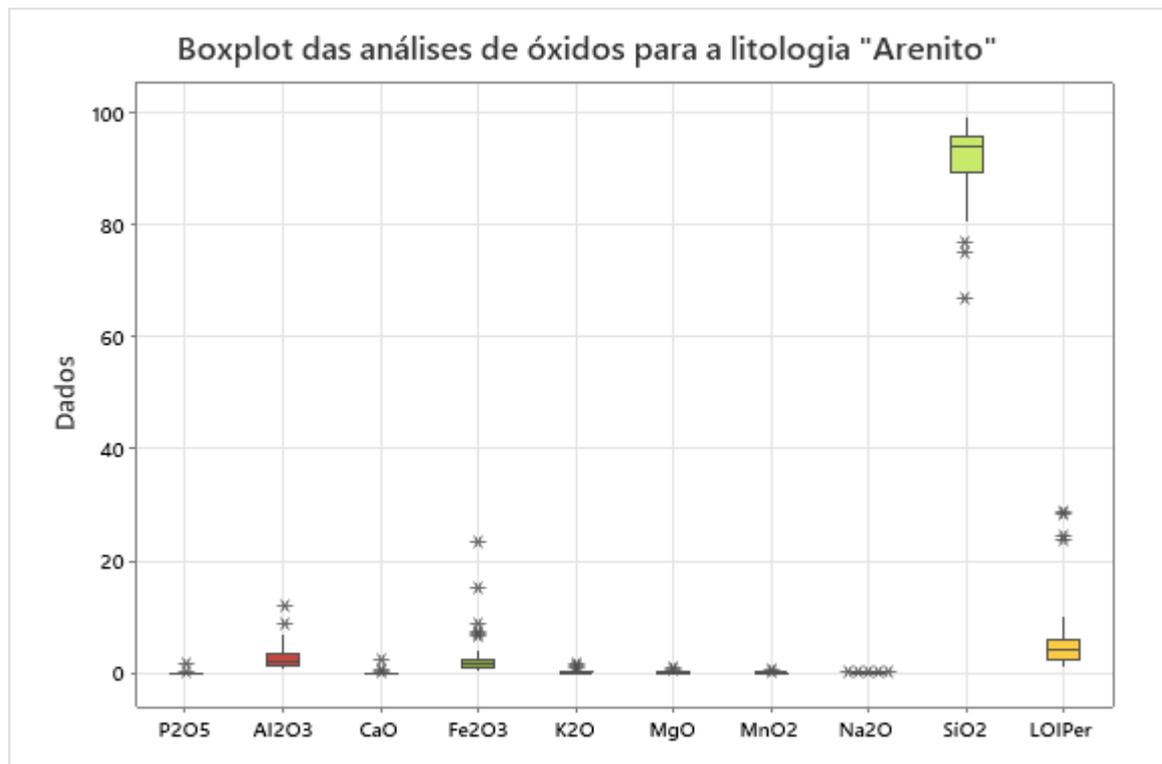


Figura 5.3: Boxplot das análises dos óxidos para as amostras descritas como "Arenito".

5.1.2 Brecha-Tipo I

A litologia Brecha-Tipo I foi descrita 541 vezes. A matriz de dispersão dos óxidos (Fig. 5.4) e o correlograma (Fig. 5.5) mostra excelente correlação entre CaO e P₂O₅ (Coeficiente de correlação de Pearson = 0,99) e correlações boas entre Fe₂O₃ e Al₂O₃ (Coeficiente de correlação de Pearson = 0,75) e K₂O e Al₂O₃ (Coeficiente de correlação de Pearson = 0,79), além de excelentes correlações negativas entre SiO₂ e P₂O₅ (Coeficiente de correlação de Pearson = -0,90) e SiO₂ e CaO (Coeficiente de correlação de Pearson = -0,91). Tais correlações são esperadas para uma litologia rica em apatita (Ca₅(PO₄)₃(OH,F,Cl)).

O *boxplot* dos resultados das análises químicas da litologia Brecha-Tipo I (Fig. 5.6) mostra valores altos de P₂O₅, CaO e SiO₂. Tais resultados são condizentes com um minério rico em apatita que passou por um processo de silicificação.

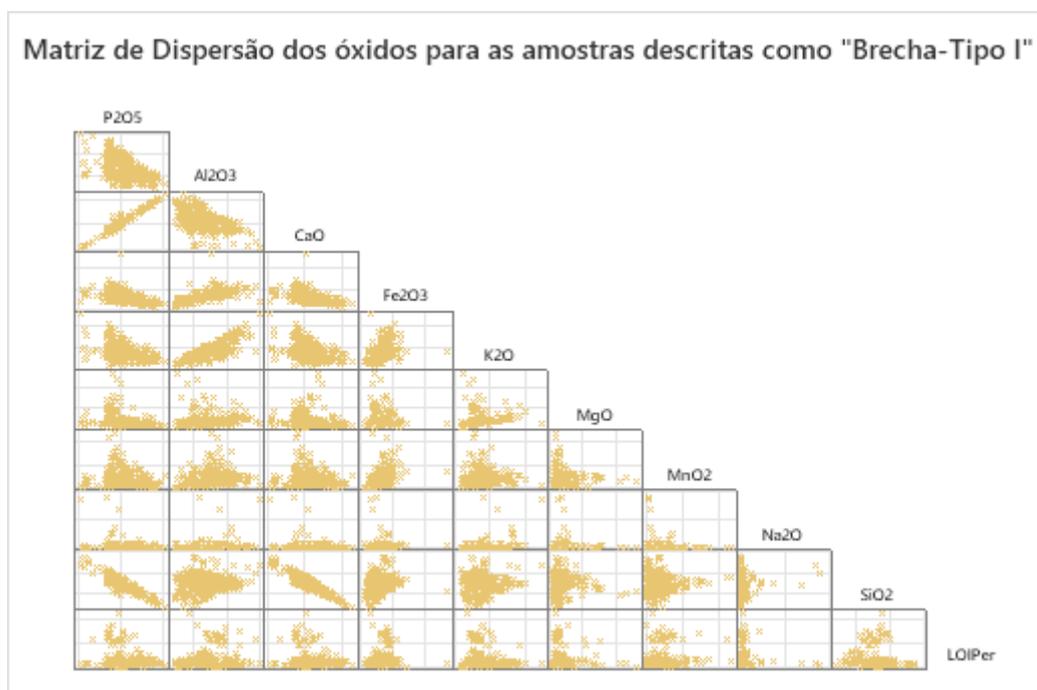


Figura 5.4: Matriz de dispersão dos óxidos para as amostras descritas como "Brecha-Tipo I".

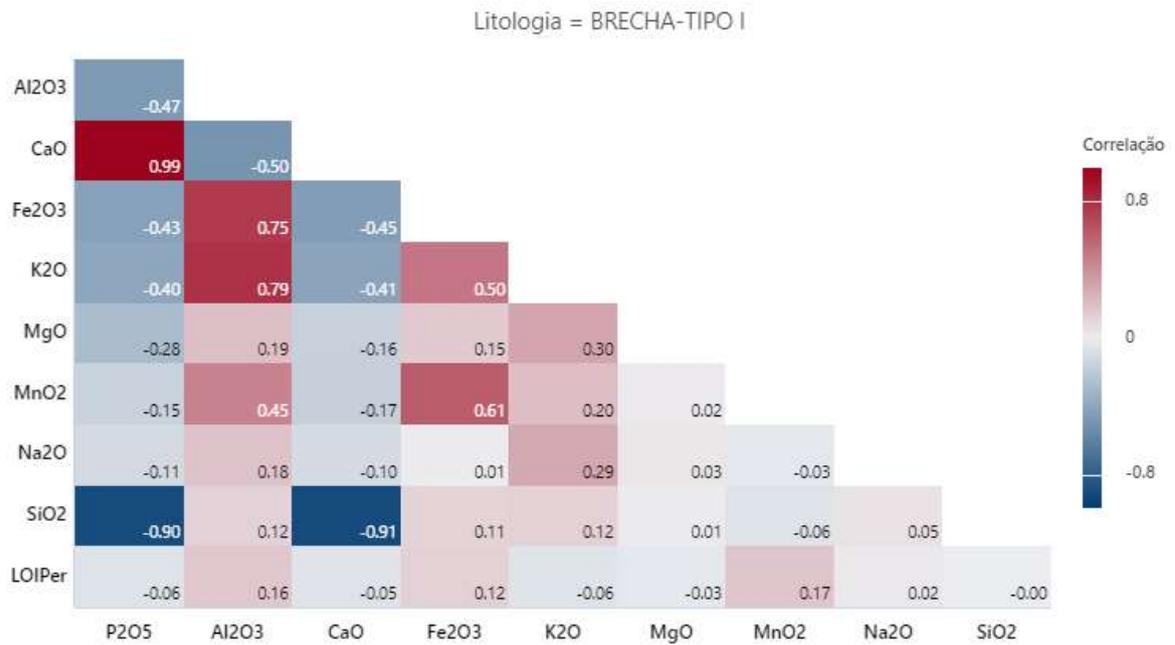


Figura 5.5: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Brecha-Tipo I".

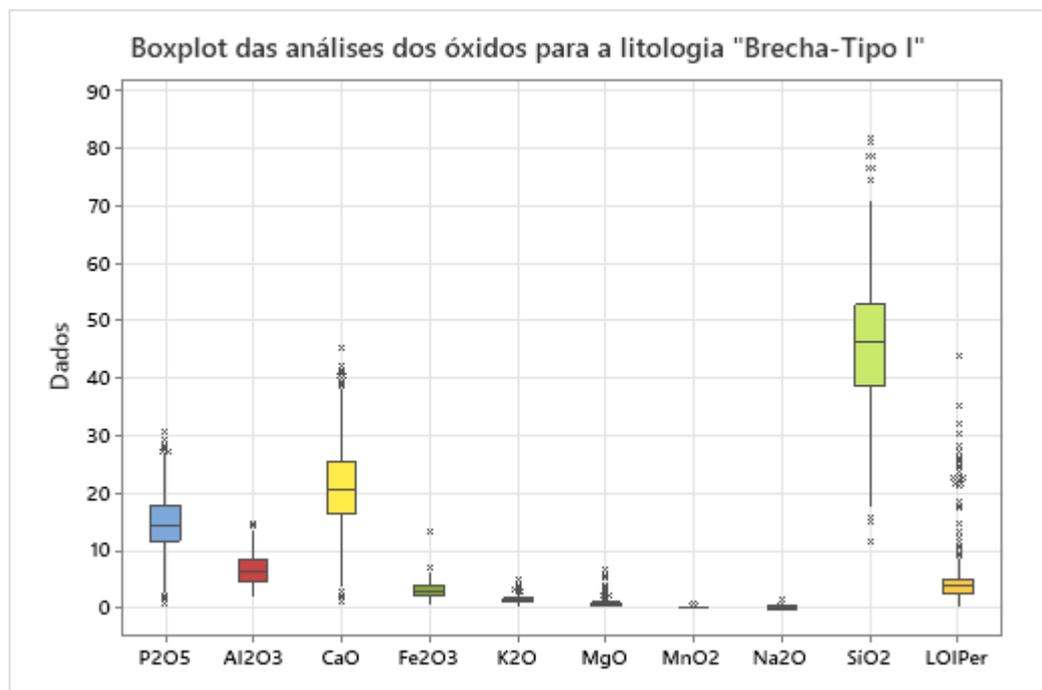


Figura 5.6: Boxplot das análises dos óxidos para as amostras descritas como "Brecha-Tipo I".

5.1.3 Calcário

A litologia Calcário foi descrita 348 vezes no banco de dados. A matriz de dispersão de óxidos e o correlograma mostram correlações positivas ótimas entre os óxidos Fe₂O₃ e Al₂O₃

(Coeficiente de correlação de Pearson = 0,84) e K_2O e Al_2O_3 (Coeficiente de correlação de Pearson = 0,87) e correlação negativa ótima entre SiO_2 e CaO (Figuras 5.7 e 5.8).

O *boxplot* dos resultados das análises químicas dos óxidos das amostras descritas como calcário mostram uma quantidade alta de CaO e MgO , provavelmente provenientes da calcita e dolomita, além de uma alta quantidade de SiO_2 , proveniente do quartzo (Fig. 5.9). Entretanto, existem teores anômalos de Al_2O_3 e K_2O , não muito comuns nos minerais típicos de um calcário, indicando uma contribuição terrígena ou a descrição de margas como calcários. A presença de Fe_2O_3 pode ser explicada pela presença de pirita em alguns intervalos da litologia.

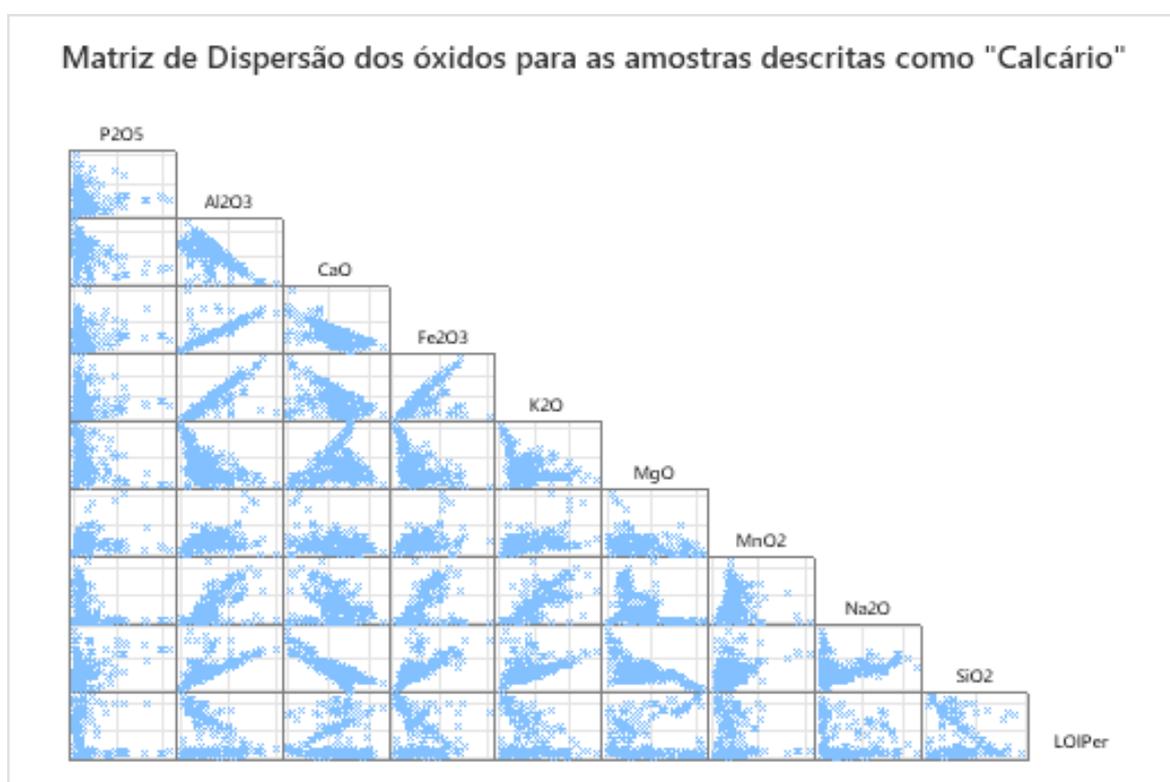


Figura 5.7: Matriz de dispersão dos óxidos para as amostras descritas como "Calcário".

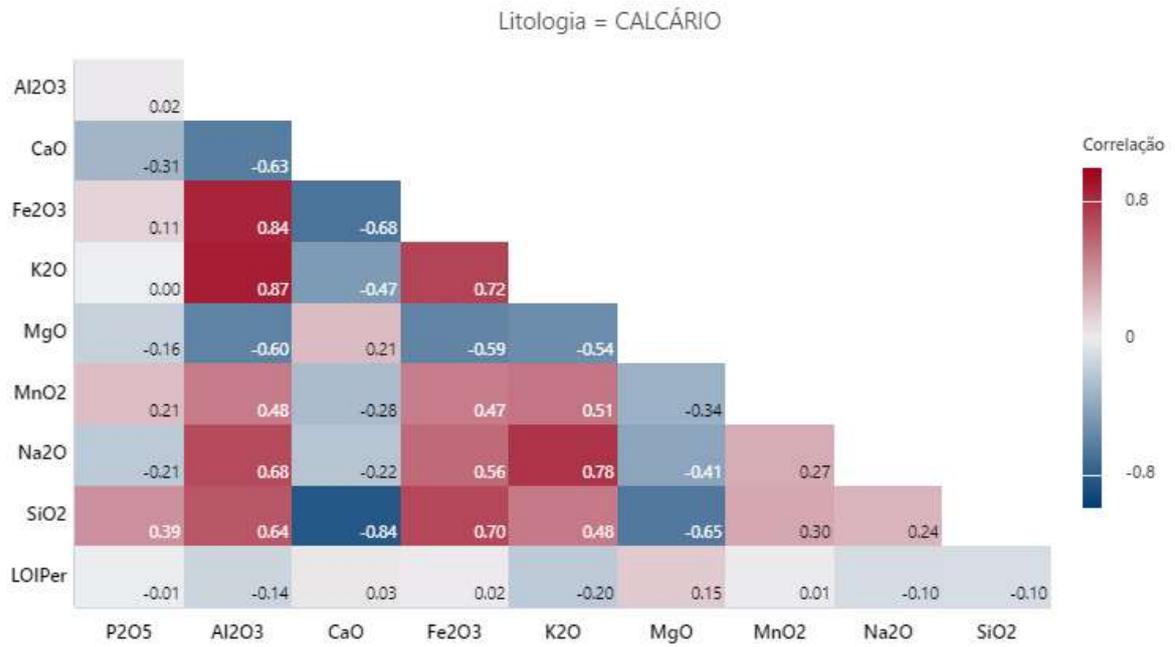


Figura 5.8: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Calcário".

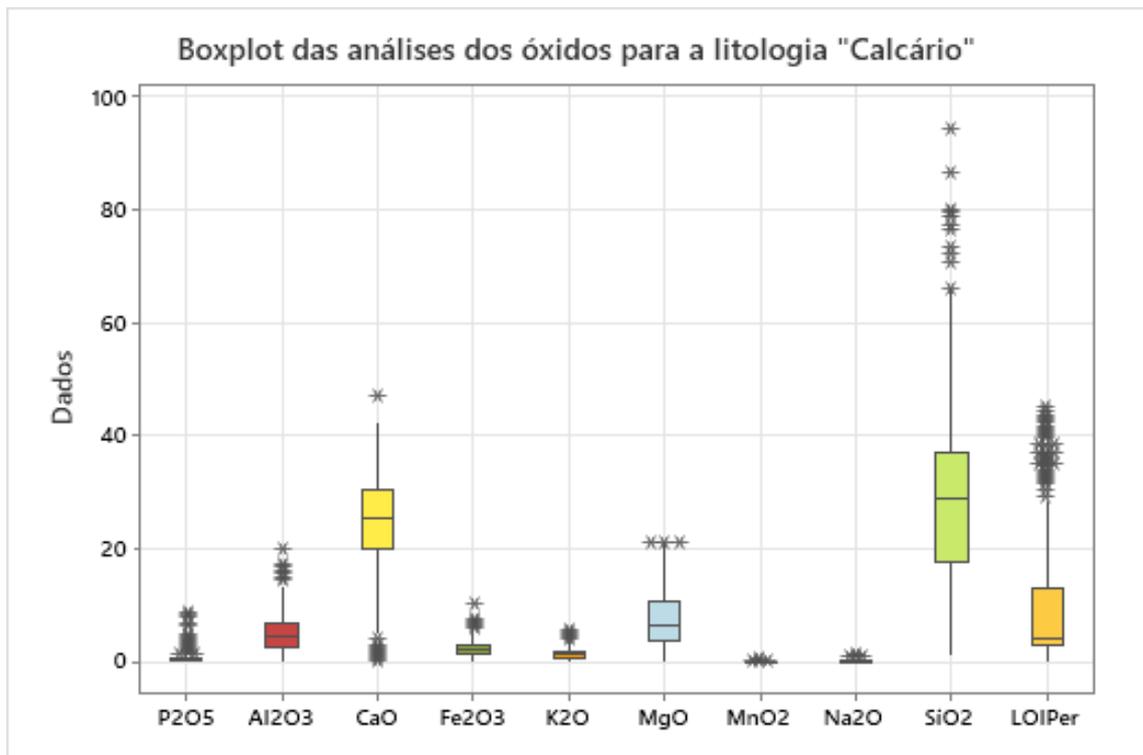


Figura 5.9: Boxplot das análises dos óxidos para as amostras descritas como "Calcário".

5.1.4 Dolomito

A litologia dolomito foi descrita 767 vezes no banco de dados. A matriz de dispersão dos óxidos e o correlograma mostram ótima correlação positiva entre K_2O e Al_2O_3 (Coeficiente de correlação de Pearson = 0,88) e ótima correlação negativa entre os óxidos SiO_2 e MgO (Coeficiente de correlação de Pearson = -0,83), SiO_2 e CaO (Coeficiente de correlação de Pearson = -0,86) e de CaO e Al_2O_3 (Coeficiente de correlação de Pearson = -0,80) (Figuras 5.10 e 5.11).

O *boxplot* das amostras descritas como dolomito mostra altos teores de CaO e MgO , condizentes com a mineralogia composta por dolomita e calcita (Fig. 5.12). Além desses, é possível notar teores altos de Al_2O_3 e de K_2O , óxidos não comuns na mineralogia dos dolomitos. Além disso, os valores de outliers de SiO_2 indicam a presença de amostras com mais de 80% de SiO_2 , condizentes, litologicamente, com um arenito/quartzito ou com um *chert* (Fig. 5.12).

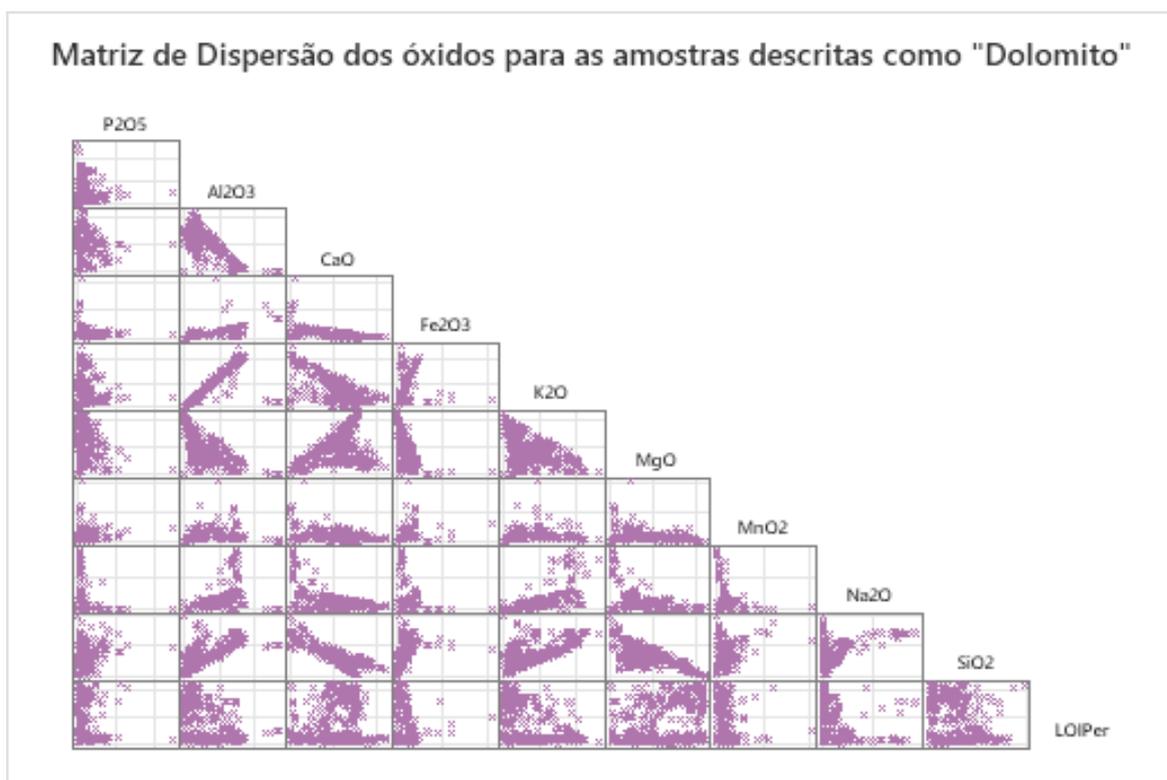


Figura 5.10: Matriz de dispersão dos óxidos para as amostras descritas como "Dolomito".

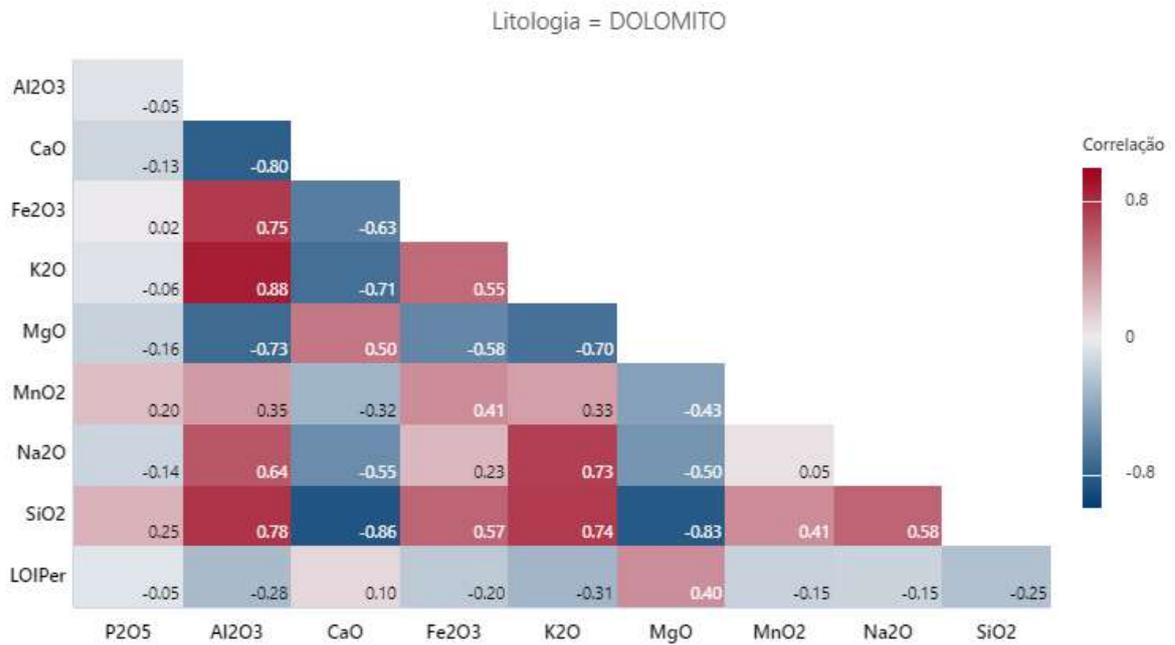


Figura 5.11: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Dolomito".

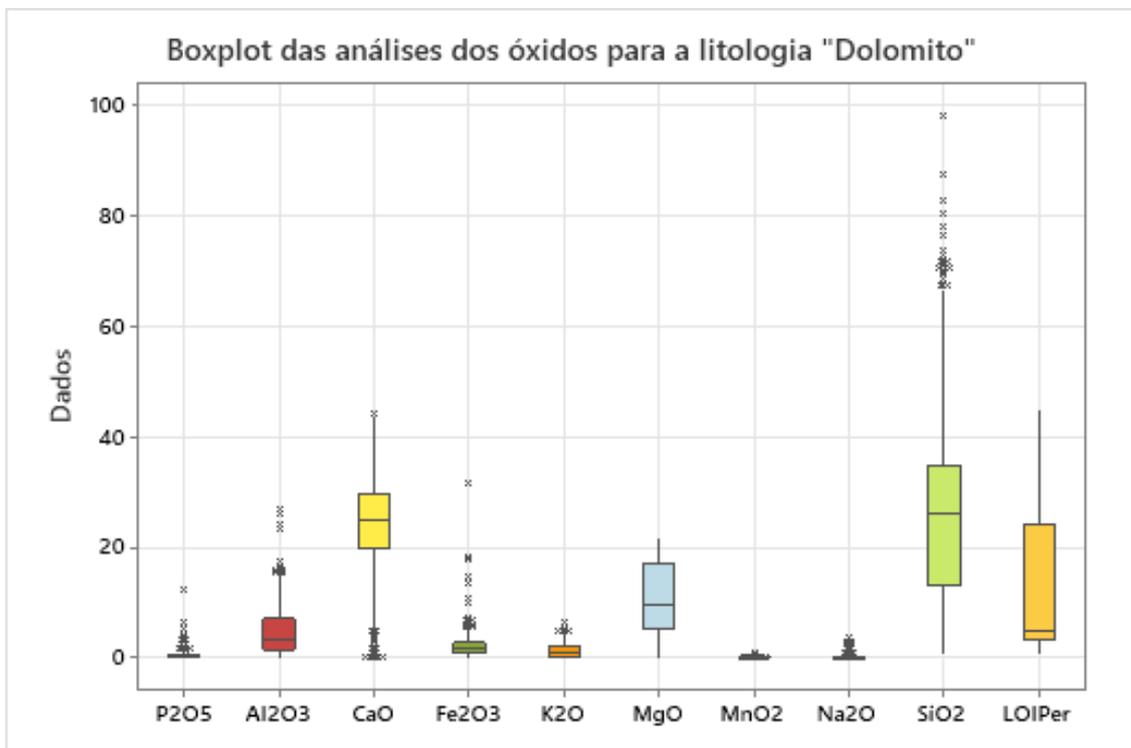


Figura 5.12: Boxplot das análises dos óxidos para as amostras descritas como "Dolomito".

5.1.5 Granito

A litologia granito foi descrita 417 vezes no banco de dados. A matriz de dispersão dos óxidos e o correlograma mostram correlação positiva excelente entre MgO e CaO (Coeficiente de correlação de Pearson = 0,95) e correlação negativa ótima entre SiO₂ e CaO (Coeficiente de correlação de Pearson = -0,84) e entre SiO₂ e MgO (Coeficiente de correlação de Pearson = -0,81) (Figuras 5.13 e 5.14).

O *boxplot* das amostras descritas como granito mostra altos teores de SiO₂, condizente com a mineralogia amplamente silicática de um granito, além de valores consideráveis de K₂O, Na₂O e Al₂O₃, condizentes com a química dos feldspatos (Fig. 5.15).

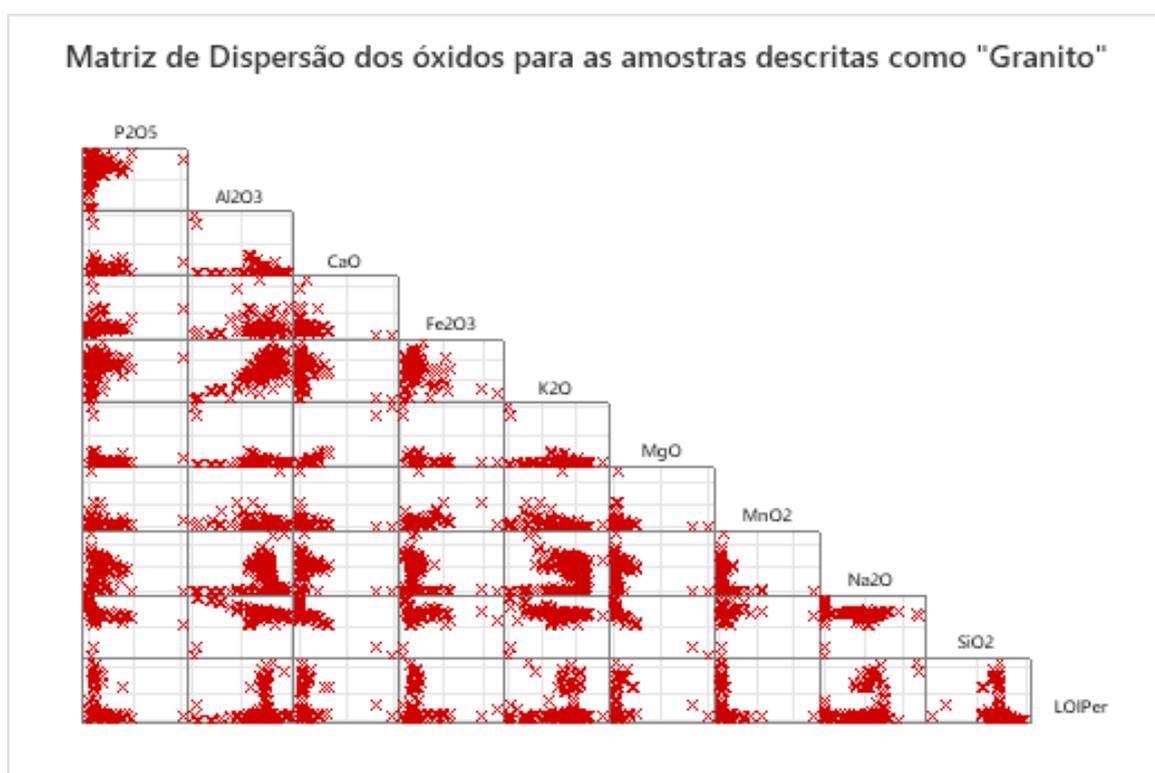


Figura 5.13: Matriz de dispersão dos óxidos para as amostras descritas como "Granito".

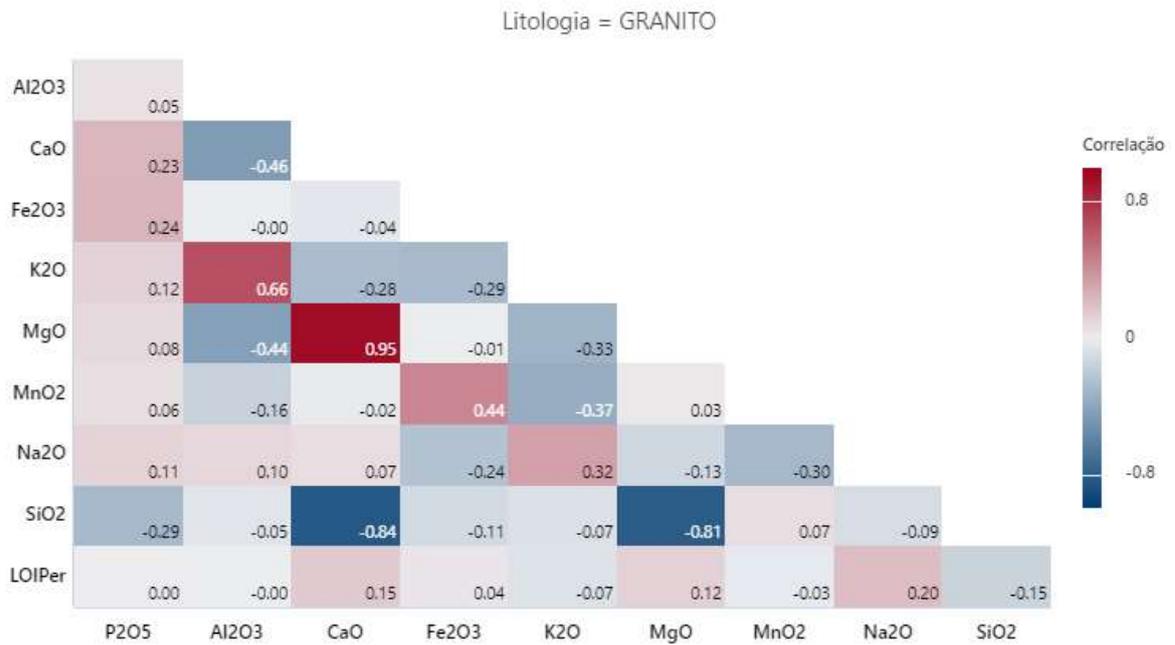


Figura 5.14: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Granito".

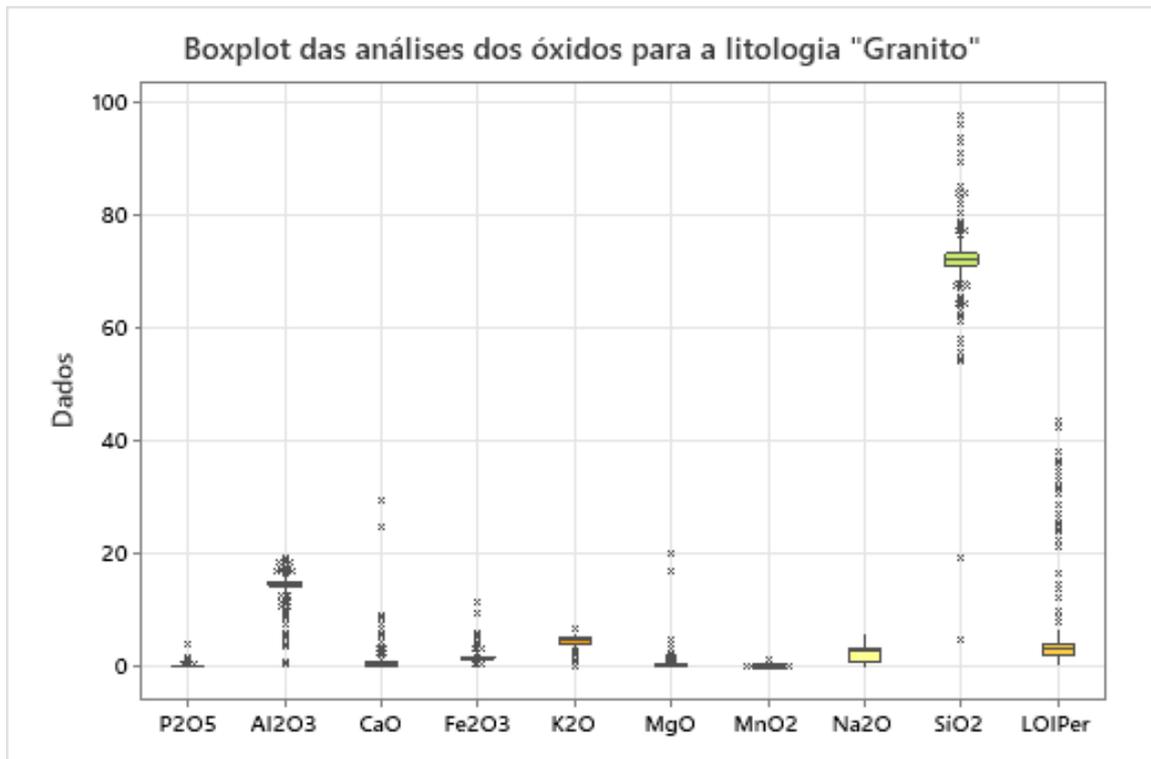


Figura 5.15: Boxplot das análises dos óxidos para as amostras descritas como "Granito".

5.1.6 Marga

A litologia marga foi descrita 350 vezes no banco de dados. A matriz de dispersão e o correlograma mostram correlações positivas excelentes entre os óxidos Fe_2O_3 e Al_2O_3 (Coeficiente de correlação de Pearson = 0,92) e K_2O e Al_2O_3 (Coeficiente de correlação de Pearson = 0,94), e uma correlação positiva ótima entre K_2O e Fe_2O_3 (Coeficiente de correlação de Pearson = 0,85), além de correlações negativas ótimas entre CaO e Al_2O_3 (Coeficiente de correlação de Pearson = -0,80), Fe_2O_3 e CaO (Coeficiente de correlação de Pearson = -0,82), SiO_2 e CaO (Coeficiente de correlação de Pearson = -0,85) e SiO_2 e MgO (Coeficiente de correlação de Pearson = -0,80) (Figuras 5.16 e 5.17).

O *boxplot* dos resultados das análises da litologia marga mostra valores consideráveis de SiO_2 , CaO e MgO , condizentes com a mineralogia de uma marga, além de valores altos de Al_2O_3 e K_2O , usuais em margas com contribuição terrígena relacionada a argilominerais. Além disso, nota-se a presença de P_2O_5 , condizente com a presença de apatita (Fig. 5.18).

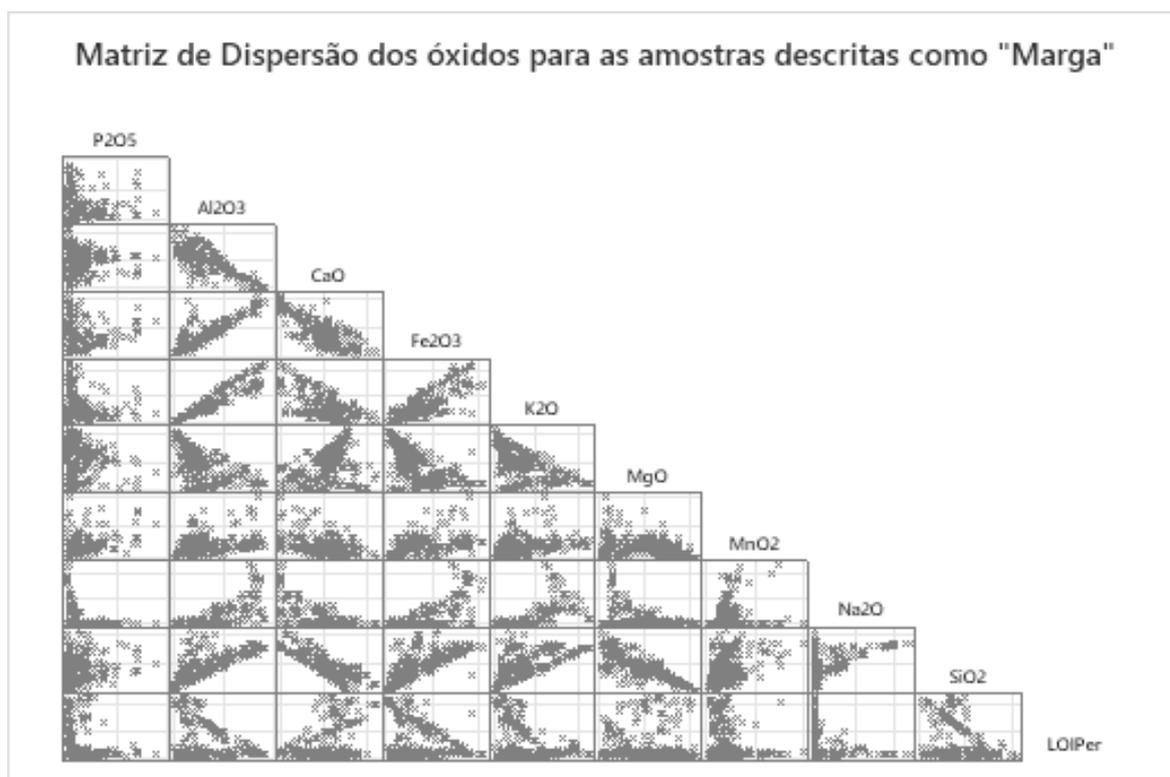


Figura 5.16: Matriz de dispersão dos óxidos para as amostras descritas como "Marga".

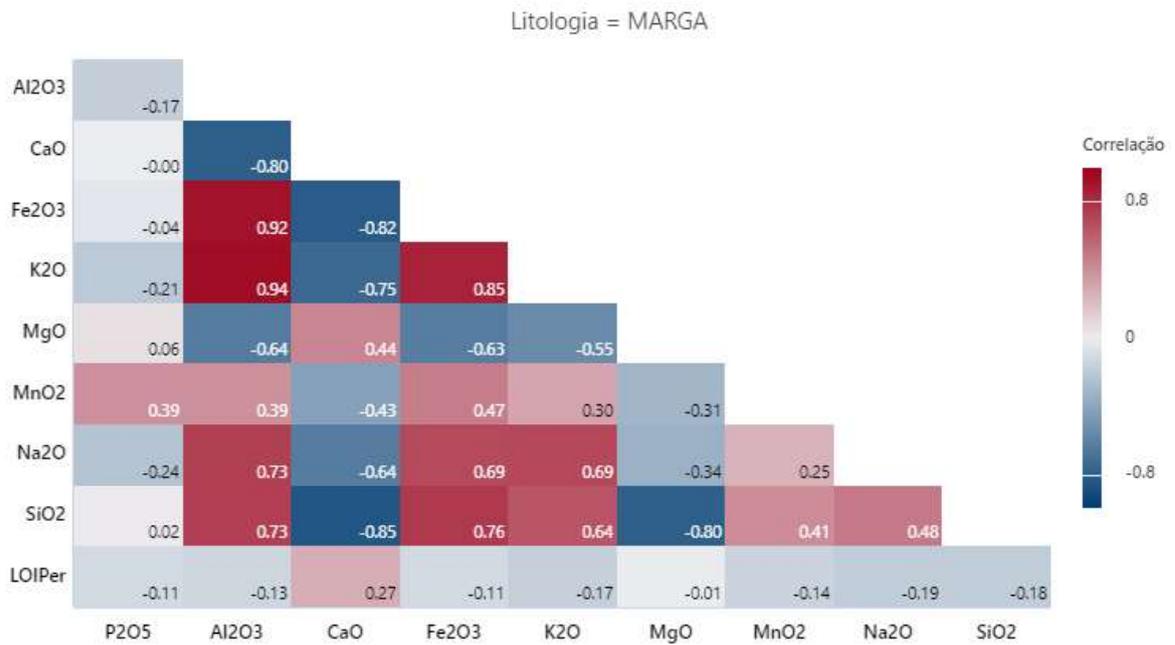


Figura 5.17: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Marga".

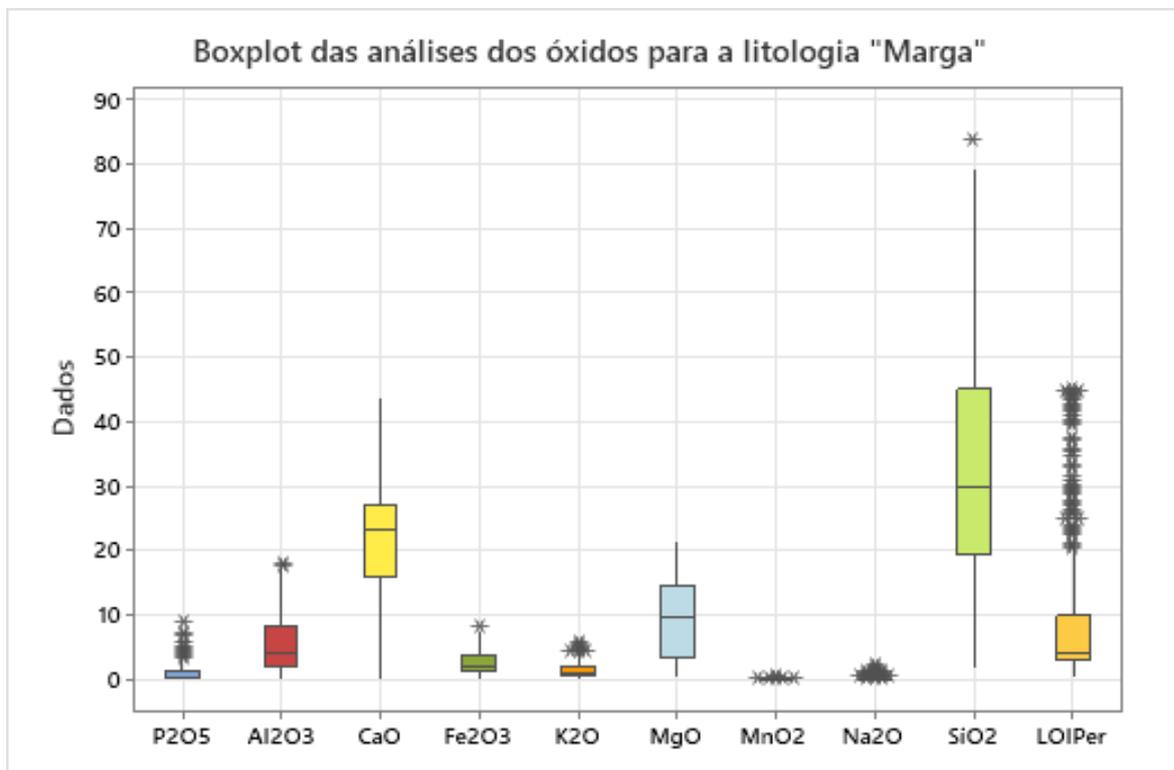


Figura 5.18: Boxplot das análises dos óxidos para as amostras descritas como "Marga".

5.1.7 Silexito

A litologia silexito é a que menos aparece no banco de dados, sendo que apenas 33 amostras foram descritas dessa maneira. A matriz de dispersão e o correlograma dos teores de óxidos mostram correlação positiva perfeita entre MgO e CaO (Coeficiente de correlação de Pearson = 1,00) e correlação negativa excelente entre MgO e SiO₂ (Coeficiente de correlação de Pearson = -0,95) e CaO e SiO₂ (Coeficiente de correlação de Pearson = -0,95) (Figuras 5.19 e 5.20).

O *boxplot* dos teores de óxidos do silexito mostra uma grande quantidade de SiO₂, condizente com a mineralogia do silexito, além de quantidades significativas de Al₂O₃, Fe₂O₃ provavelmente provenientes de algum mineral acessório (Figura 5.21). A baixa quantidade de amostras impede uma delimitação da composição química de maneira acurada.

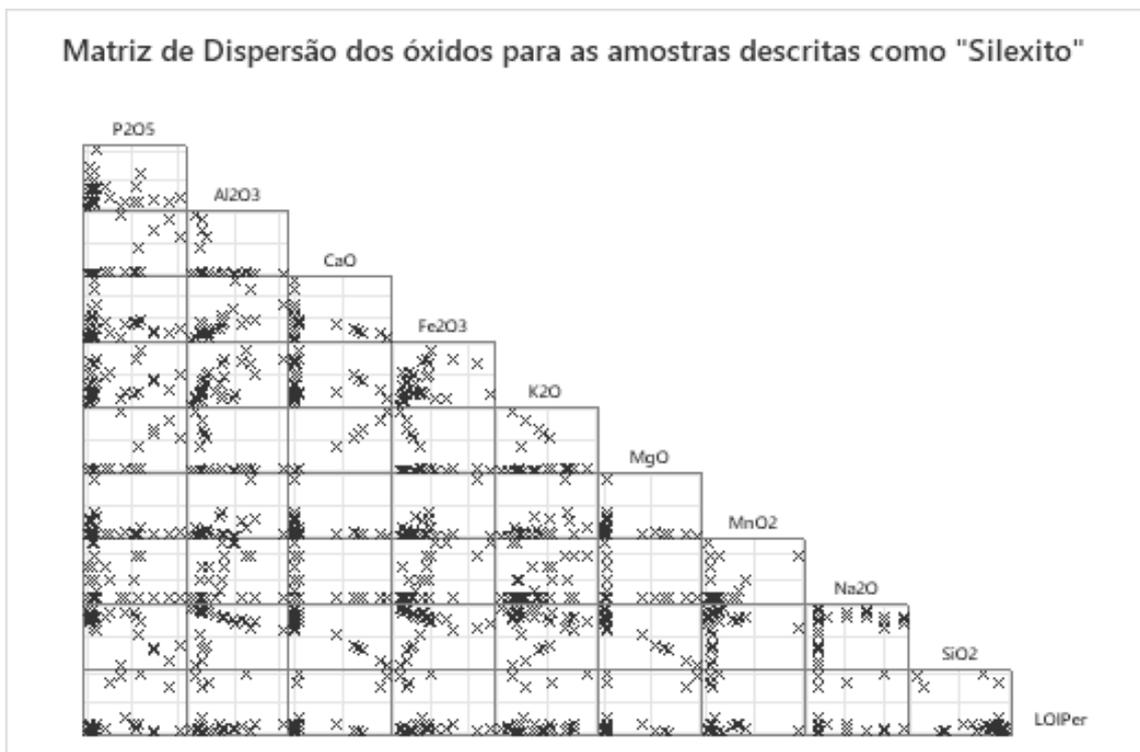


Figura 5.19: Matriz de dispersão dos óxidos para as amostras descritas como "Silexito".

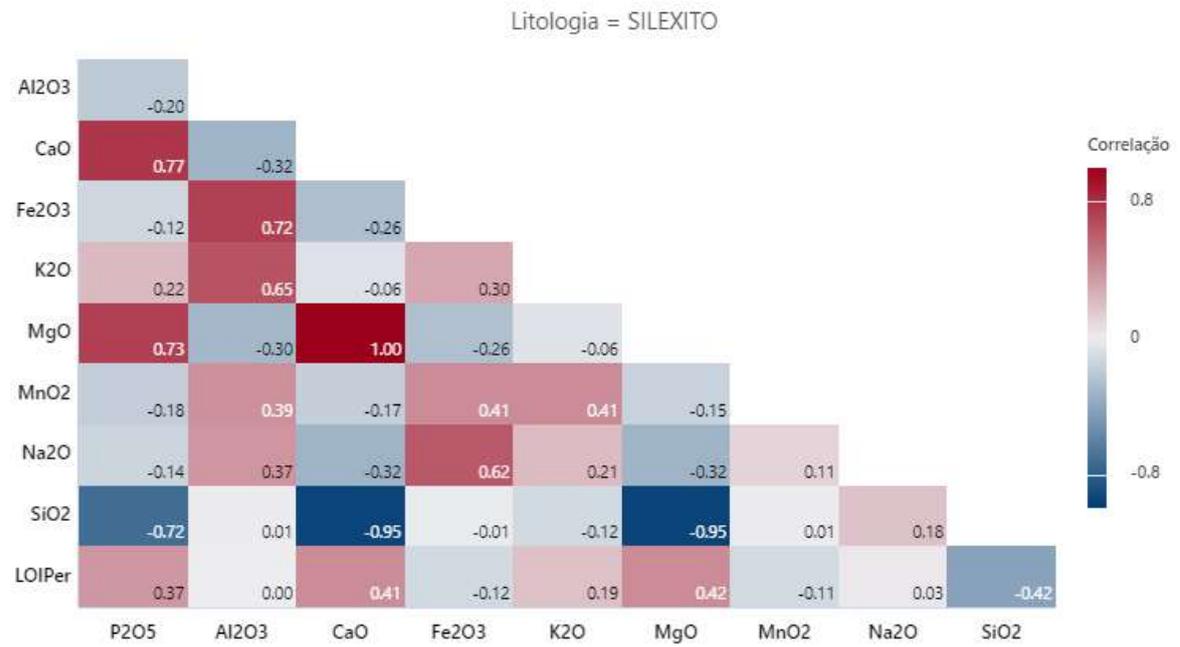


Figura 5.20: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Silexito".

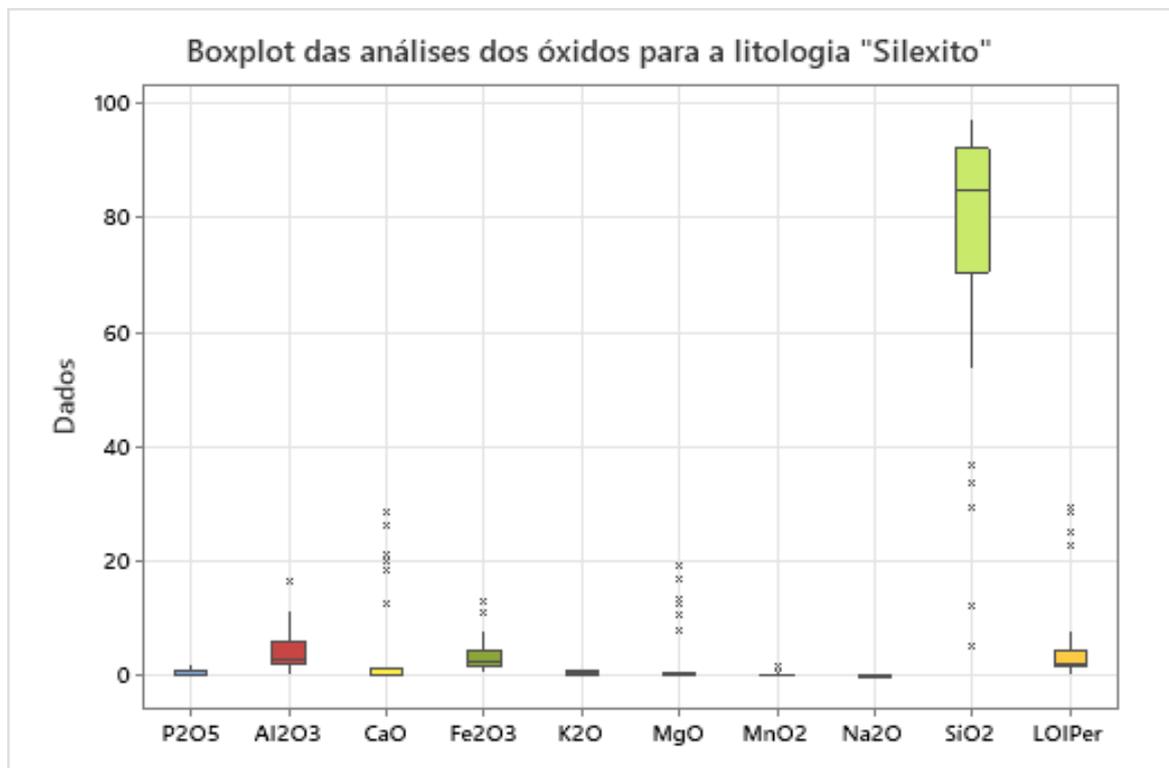


Figura 5.21: Boxplot das análises dos óxidos para as amostras descritas como "Silexito".

5.1.8 Siltito

A litologia siltito é a que mais aparece no banco de dados, sendo descrita em 7268 amostras. É a litologia que corresponde ao minério fosfático da Mina Domingos. A matriz de dispersão dos teores de óxidos e o correlograma mostram baixa correlação entre a maioria dos teores dos óxidos. Correlações positivas boas só aparecem entre os teores de MgO e CaO (Coeficiente de correlação de Pearson = 0,75), CaO e P₂O₅ (Coeficiente de correlação de Pearson = 0,65) e de K₂O e Al₂O₃ (Coeficiente de correlação de Pearson = 0,65) (Figuras 5.22 e 5.23). Correlações negativas boas somente entre os teores de SiO₂ e CaO (Coeficiente de correlação de Pearson = -0,69) e entre SiO₂ e MgO (Coeficiente de correlação de Pearson = -0,72) (Figuras 5.22 e 5.23). O *boxplot* dos teores de óxidos para as amostras de siltito mostra uma grande presença de *outliers* (Figura 5.24). Tal fato, aliado com a grande variação dos teores de P₂O₅ e dos outros óxidos, além da variação visual do siltito (Figura 3.3) levam ao entendimento que talvez a litologia descrita como siltito deveria ser subdividida em mais litologias, ou que outros tipos de classificação deveriam ter sido adotadas.

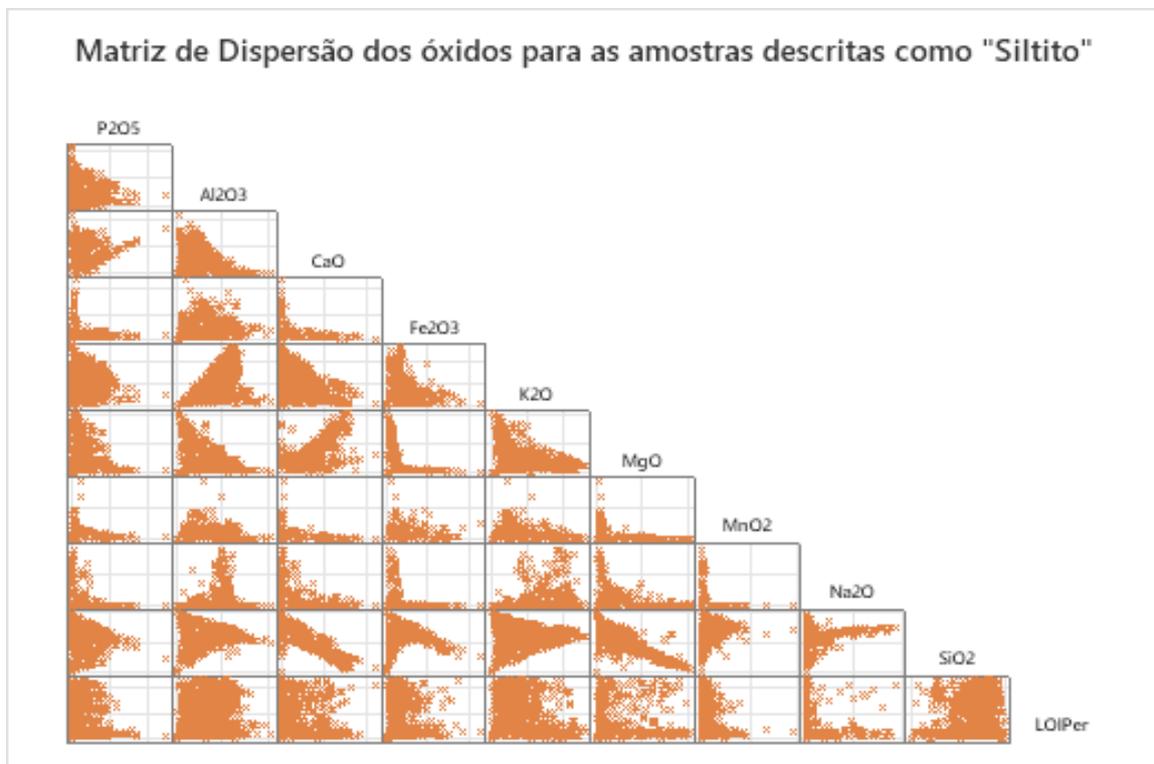


Figura 5.22: Matriz de dispersão dos óxidos para as amostras descritas como "Siltito".

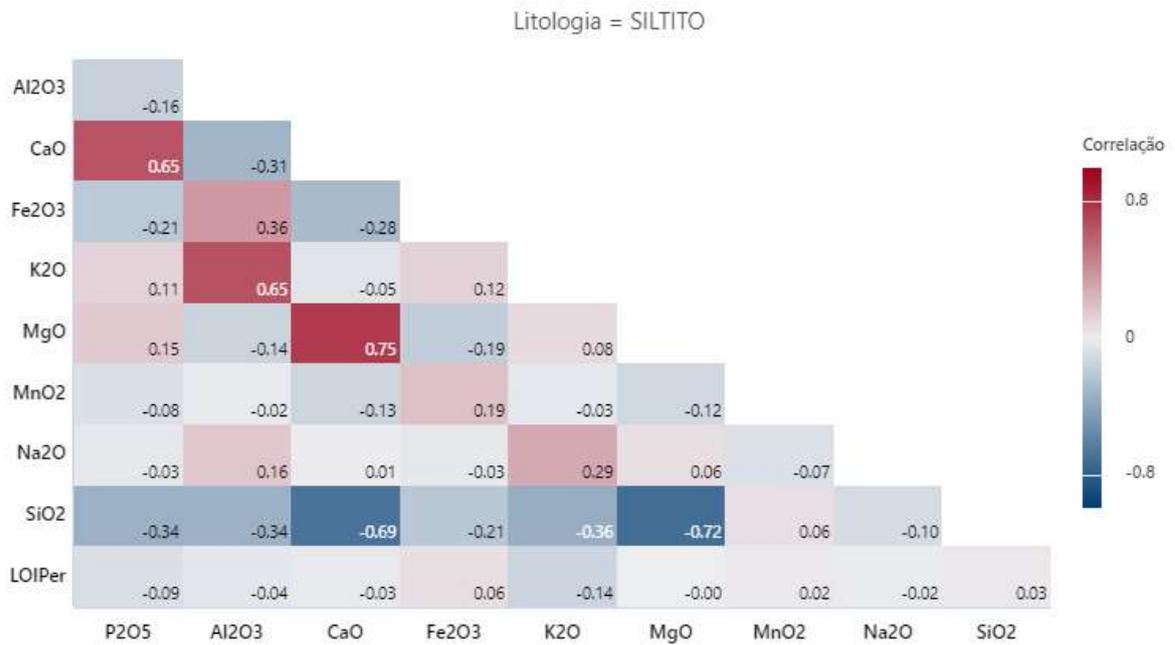


Figura 5.23: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Siltito".

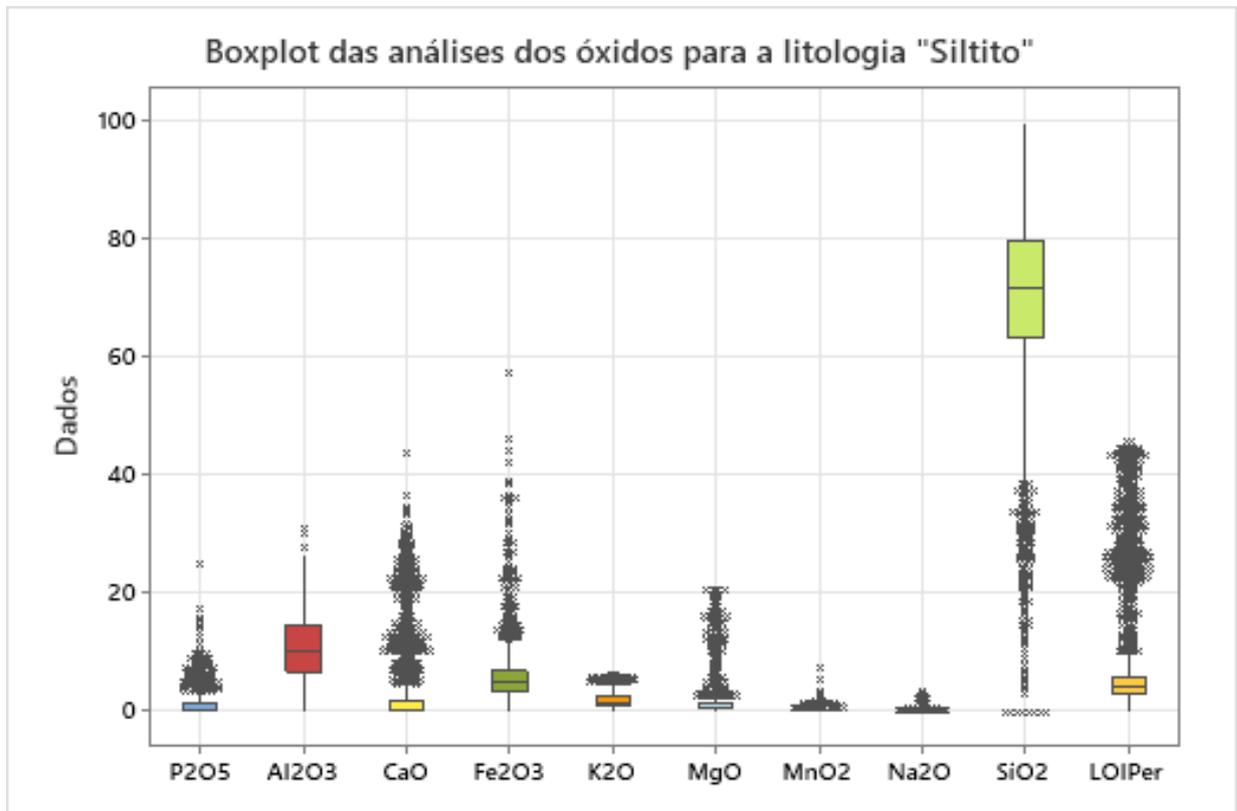


Figura 5.24: Boxplot das análises dos óxidos para as amostras descritas como "Siltito".

Dessa maneira, pode-se inferir que parte dos problemas operacionais enfrentados durante a lavra advém do fato de que o modelo geológico não compreende essa variação litológica do minério, tratando-o como uniforme para a modelagem. Assim sendo, destacam-se:

1. A provável superestimação da quantidade de minério disponível no depósito, uma vez que não foi levada em conta a variação do minério, com diferentes recuperações e teores.
2. Problemas relacionados ao *design* da planta de beneficiamento, por considerar o siltito como sendo apenas uma litologia.
3. Problemas relacionados ao *ramp up* da produção na planta, relacionados a diferenças entre o minério previsto a partir do modelo e o que de fato foi lavrado.
4. A grande variação química pode apresentar diferentes respostas no processamento mineral, especialmente na flotação, o que pode levar a problemas de recuperação.
5. Aumento do uso de reagentes durante a acidulação, o que pode levar a um aumento do custo operacional e a inviabilização da lavra.

5.1.9 Siltito Brechado

A litologia “Siltito Brechado” é descrita em 49 amostras do banco de dados. A matriz de dispersão e o correlograma (Figuras 5.25 e 5.26) mostram ótima correlação positiva entre os teores de MnO_2 e Fe_2O_3 (Coeficiente de correlação de Pearson = 0,80) e entre os teores de K_2O e Al_2O_3 (Coeficiente de correlação de Pearson = 0,87).

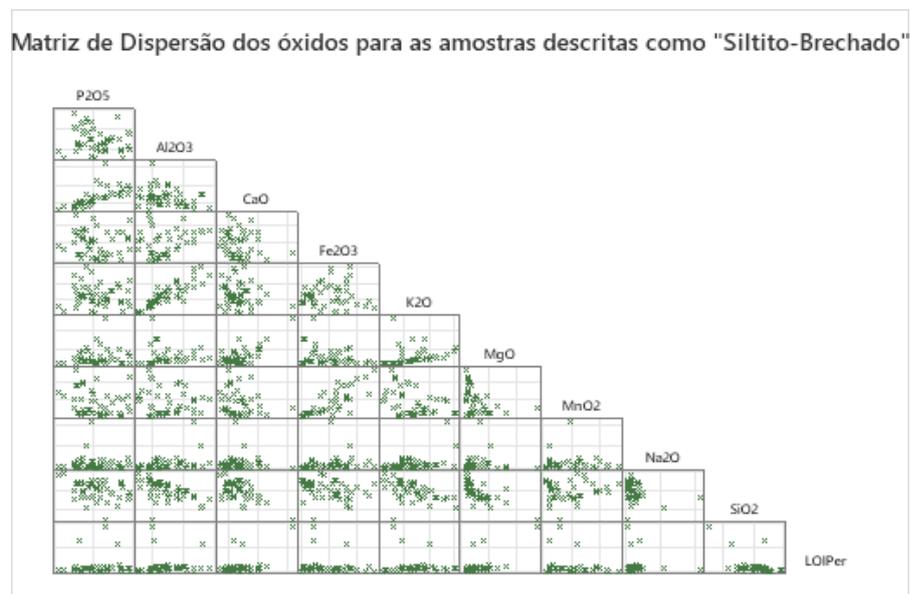


Figura 5.25: Matriz de dispersão dos óxidos para as amostras descritas como "Siltito-Brechado".

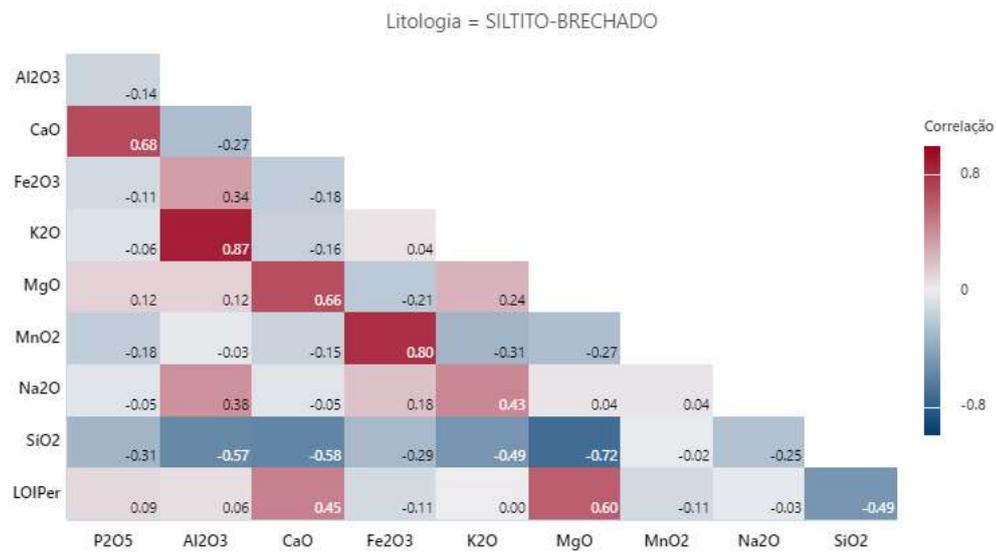


Figura 5.26: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Siltito-Brechado".

O *boxplot* com os resultados das análises químicas das amostras descritas como siltito-brechado (Fig. 5.27) mostra teores elevados de P_2O_5 e CaO , condizentes com a descrição litológica e com uma rocha rica em apatita. Além disso, estão presentes teores significativos de SiO_2 , congruente com a brechcação por fluidos hidrotermais ricos em sílica.

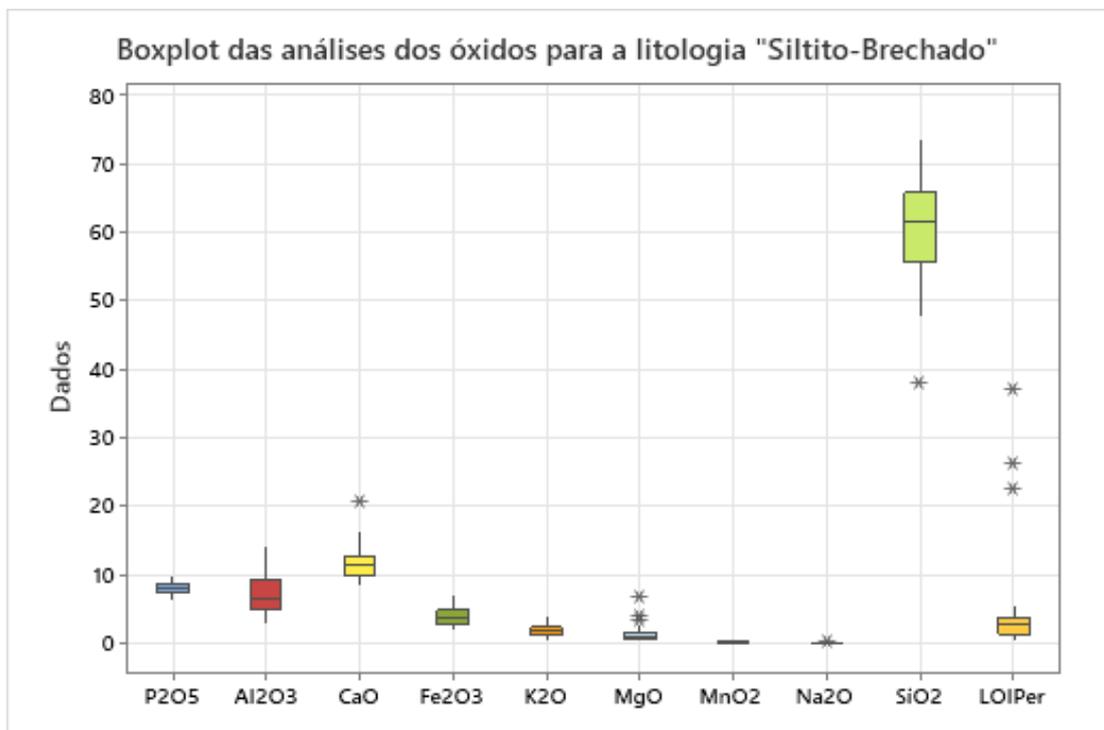


Figura 5.27: Boxplot das análises dos óxidos para as amostras descritas como "Siltito-Brechado".

5.1.10 Solo

A litologia solo aparece na descrição de 676 amostras no banco de dados. A matriz de dispersão e o correlograma mostram uma baixa correlação para todos os óxidos, com exceção de uma correlação positiva ótima para os teores de CaO e P₂O₅ (Coeficiente de correlação de Pearson = 0,80) (Figuras 5.28 e 5.29).

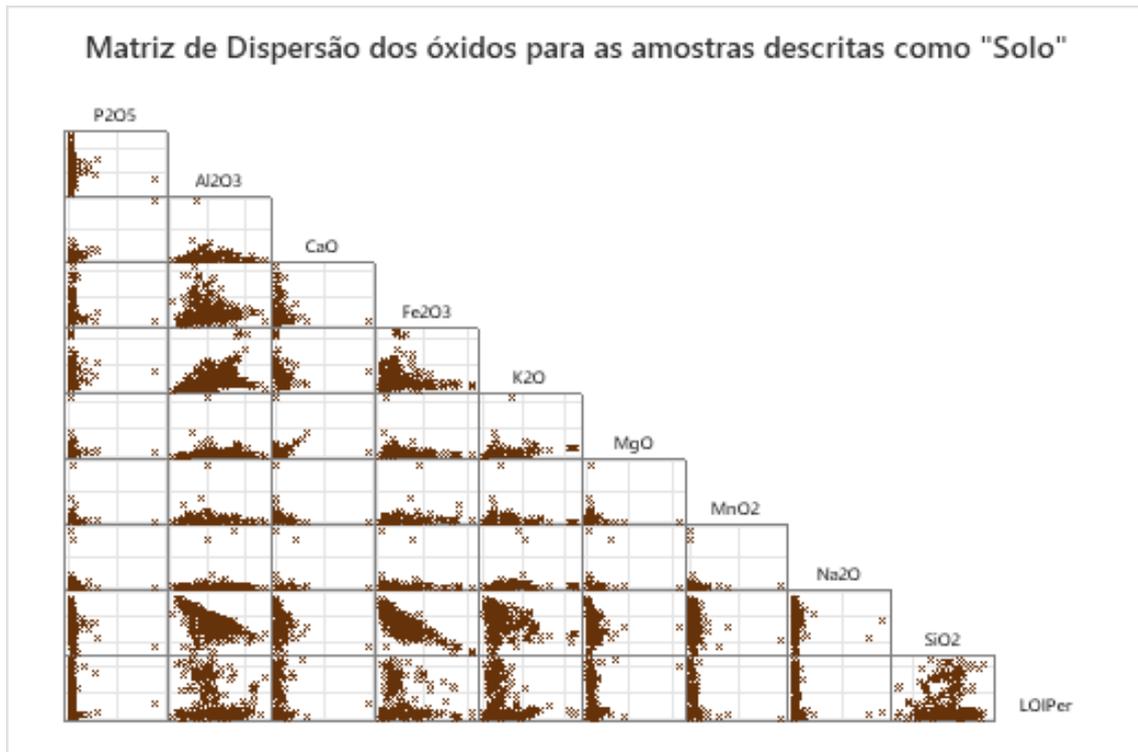


Figura 5.28: Matriz de dispersão dos óxidos para as amostras descritas como "Solo".

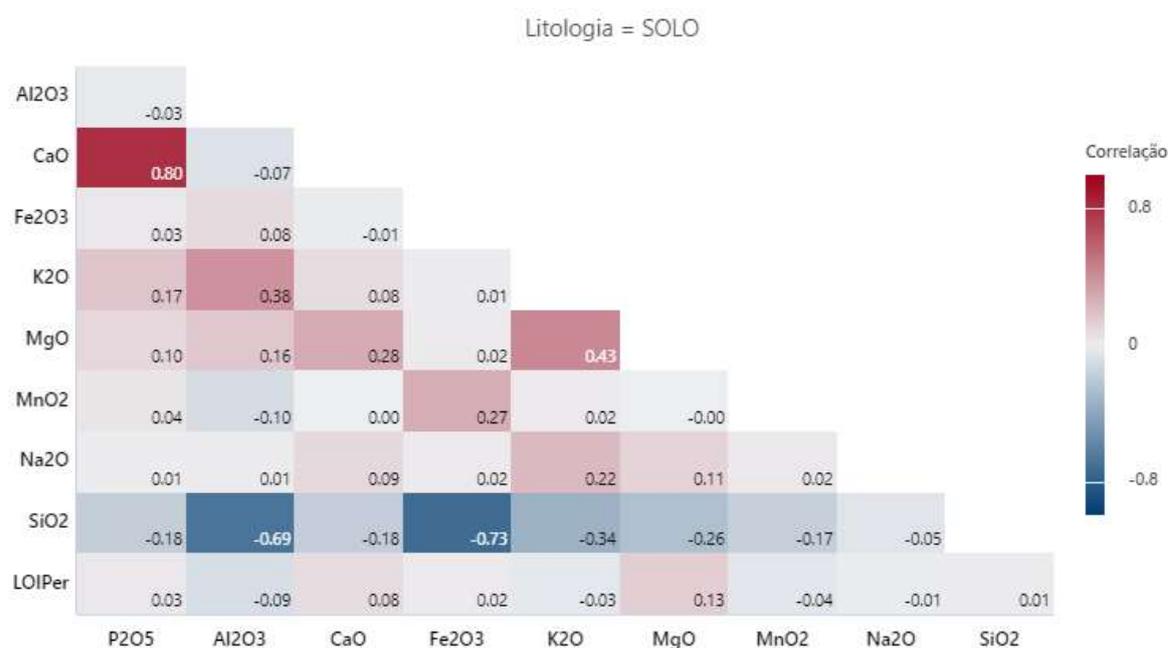


Figura 5.29: Correlograma com os coeficientes de correlação de Pearson entre os óxidos para as amostras descritas como "Solo".

Além disso, o *boxplot* com os teores das análises química mostra uma grande quantidade de *outliers*, que demonstram uma variação química grande. Tal variação deve advir do fato de que a rocha matriz não foi levada em consideração na descrição da litologia, o que acarreta diferentes tipos de solo agrupados em apenas uma litologia (Figura 5.30).

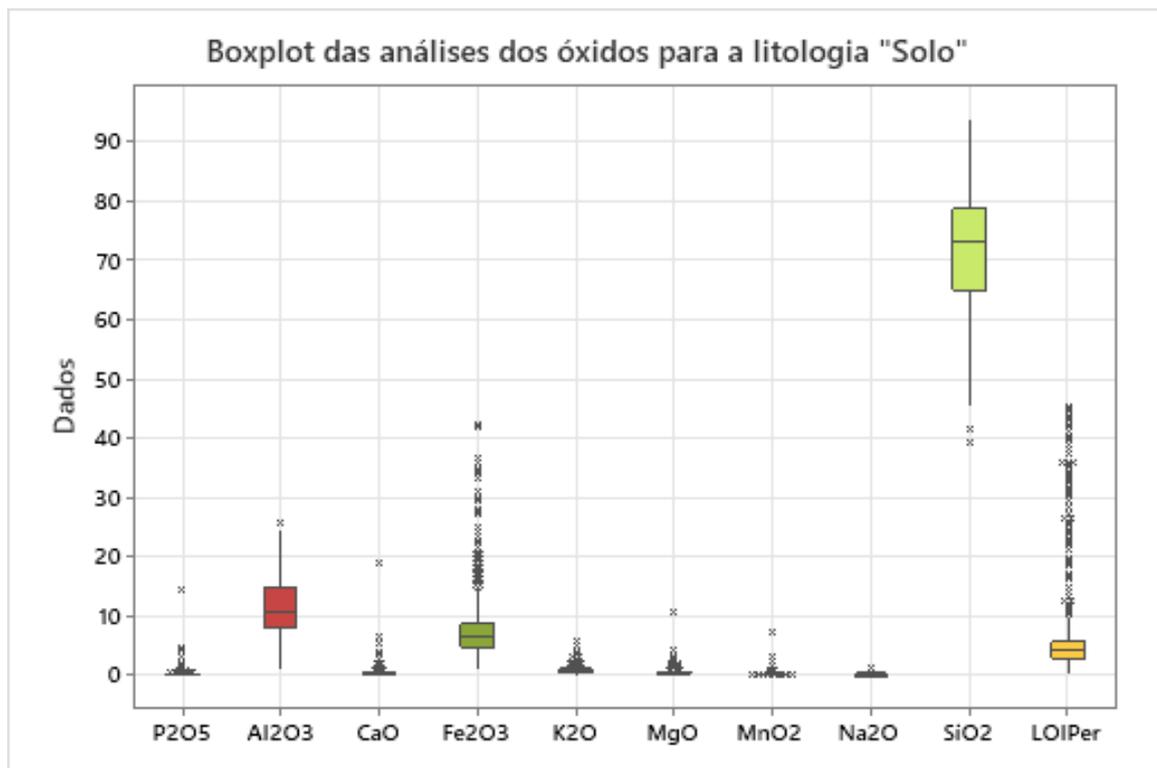


Figura 5.30: Boxplot das análises dos óxidos para as amostras descritas como "Solo".

5.2 Análise de Agrupamentos

A análise de agrupamentos foi feita no *software* Minitab utilizando o algoritmo *k-means*. Como os valores dos óxidos possuem intervalos diferentes para as diferentes litologias, alguns com quantidade muito superior que os outros, foi realizada a padronização das variáveis. A padronização corresponde a subtração da média e a divisão pelo desvio padrão de todas as variáveis.

A fim de se determinar a quantidade de classes a serem utilizadas, realizou-se o teste da silhueta (Fig. 5.31), que indicou a melhor quantidade de classes como sendo igual a duas. Entretanto, como visto no Capítulo 5.1, a utilização de apenas duas classes não cobre a variação litológica encontrada no depósito. É possível notar, no entanto, um pequeno salto no valor do Teste da Silhueta, nos valores de 5 e 10 classes. Dessa maneira, levando em consideração a existência de 10 litologias no banco de dados e esse aumento no valor do Teste da Silhueta, utilizou-se 10 classes para a utilização do algoritmo *k-means* do Minitab.

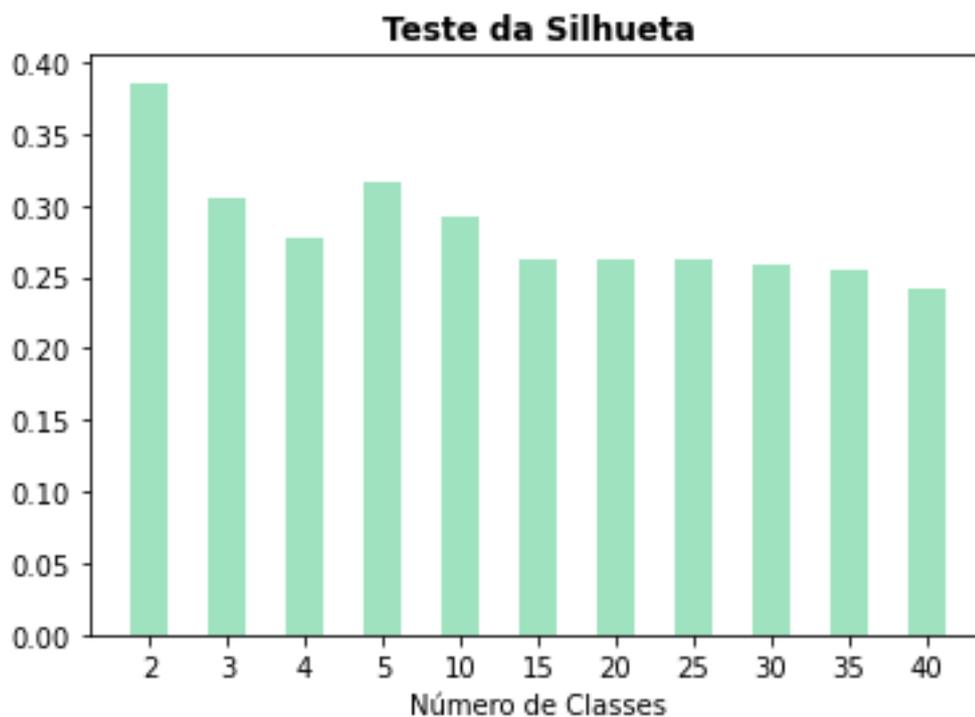


Figura 5.31: Teste da Silhueta para diferentes quantidades de classes.

O resultado do algoritmo *k-means* está disposto na Tabela 5.1 e Figura 5.32. É possível notar que a maioria das litologias ficaram restritas a poucas classes, com uma predominância clara de uma classe, esse é o caso das litologias Arenito, Brecha-Tipo I, Calcário, Dolomito, Granito, Marga, Silexito e Siltito-Brechado. As únicas duas litologias que não apresentaram esse comportamento foram o Solo e o Siltito. No caso do Solo, a sua distribuição em diferentes classes, sem predominância de apenas uma deve-se, provavelmente, ao fato de que a descrição não levou em consideração a rocha matriz, o que acarretou a solos com diferentes composições químicas a serem considerados como sendo iguais. A litologia mais abrangente, como era esperado, é o siltito, que foi encontrado em todas as classes (Figuras 5.32 e 5.33), sendo principalmente definido nas Classes 3,4 e 6, porém com quantidades razoáveis em todas as outras classes, com exceção da Classe 7 e 8, que possuem dominância das litologias Granito e Brecha-Tipo I, respectivamente.

Tabela 5.1: Distribuição das litologias descritas nas diferentes Classes.

Litologia	Classes										Total
	1	2	3	4	5	6	7	8	9	10	
Arenito	2	4	1	44	0	0	0	0	0	0	51
Brecha-Tipo I	0	1	1	7	134	1	1	388	1	0	534
Calcário	0	7	5	5	44	24	0	0	0	263	348
Dolomito	3	11	8	4	43	68	22	1	1	606	767
Granito	0	0	13	15	0	82	304	0	1	2	417
Marga	0	0	1	13	37	44	9	0	0	246	350
Silexito	0	2	3	20	0	0	0	0	2	6	33
Siltito	151	788	1787	2103	748	1058	49	11	354	219	7268
Siltito-Brechado	0	0	0	0	48	0	0	0	0	1	49
Solo	65	77	275	233	0	13	0	1	12	0	676
Total	221	890	2094	2444	1054	1290	385	401	371	1343	10493

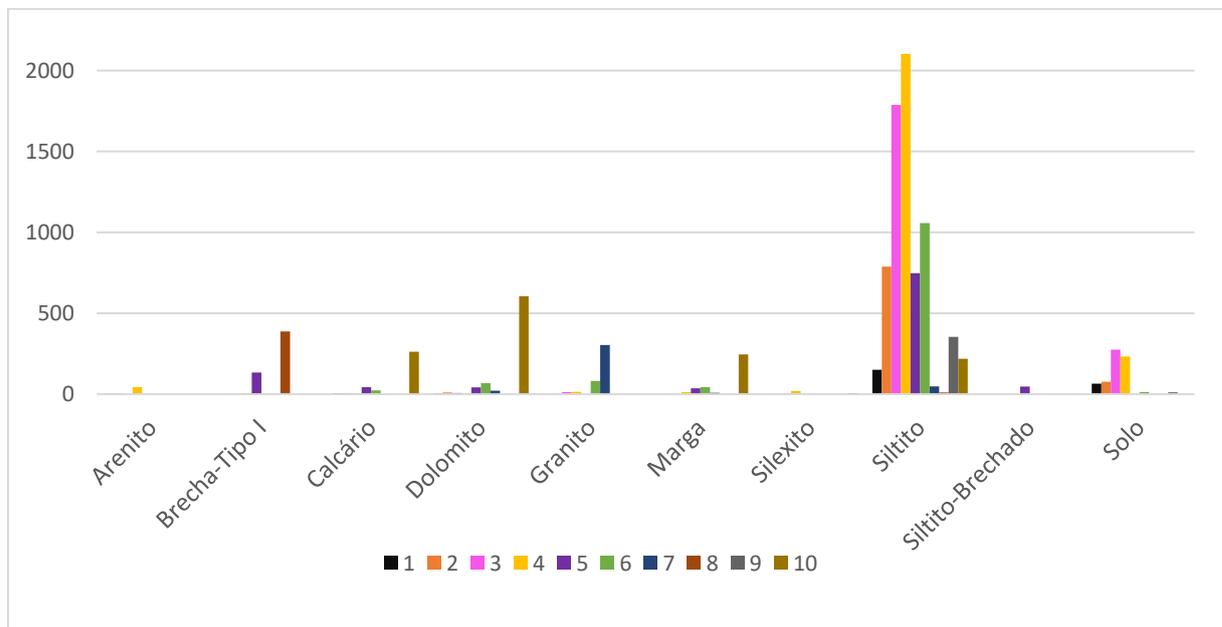


Figura 5.32: Histograma de Frequência das Litologias por Classes.



Figura 5.33: Ocorrência de amostras descritas como Siltito em cada Classe, notar a concentração nas Classes 4, 3 e 6.

A fim de se comparar as descrições litológicas com as classes geradas pelo *k-means*, realizou-se a análise de componentes principais para 3 componentes. A partir das análises de componentes principais, juntamente com o discutido anteriormente, nota-se uma clara sobreposição do siltito em toda a área do gráfico (Fig. 5.34), o que claramente indica a existência de uma grande quantidade de amostras descritas de maneira inadequada e pode justificar as inconsistências encontradas no modelo geológico e os problemas operacionais da lavra.

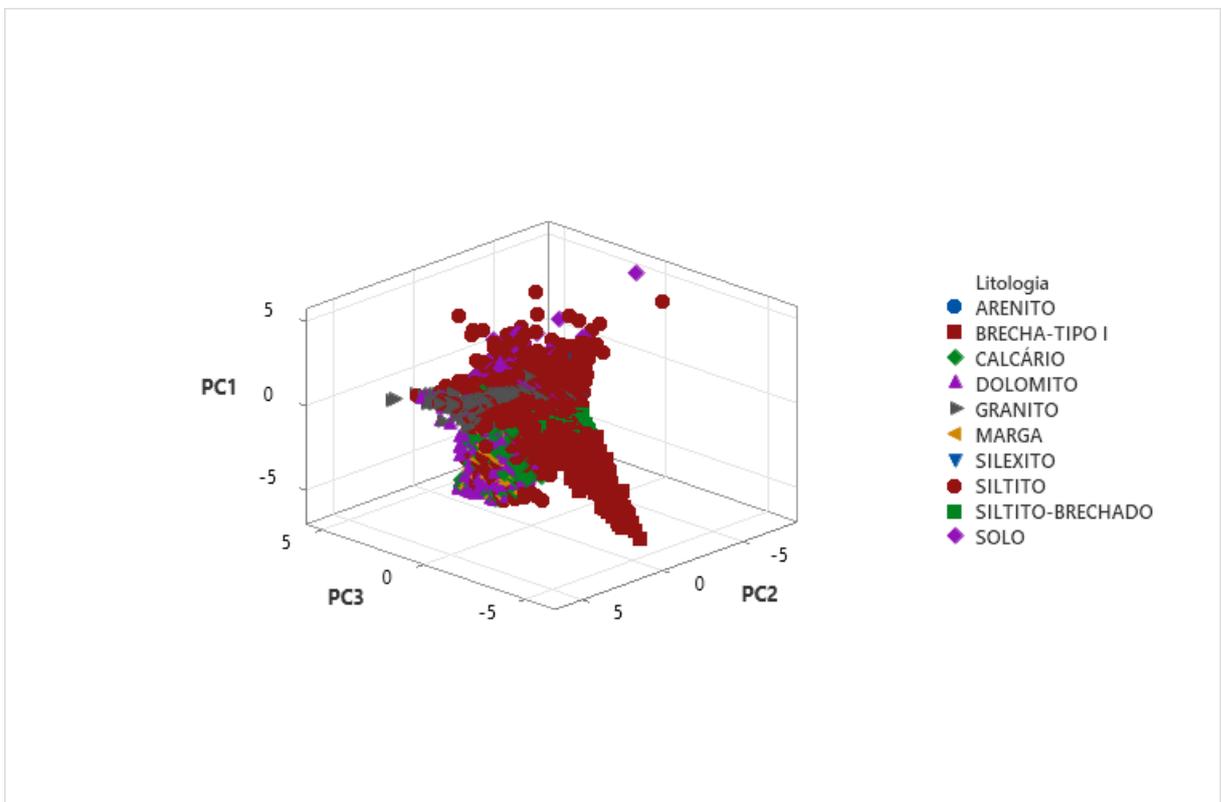


Figura 5.34: Análise de componentes principais para três componentes de acordo com as litologias descritas.

A Figura 5.35 mostra a análise de componentes principais de acordo com as classes geradas pelo *k-means*. É possível perceber uma melhor separação entre os grupos gerados quando comparados com a litologias descritas, entretanto, ainda observa-se uma sobreposição entre os grupos.

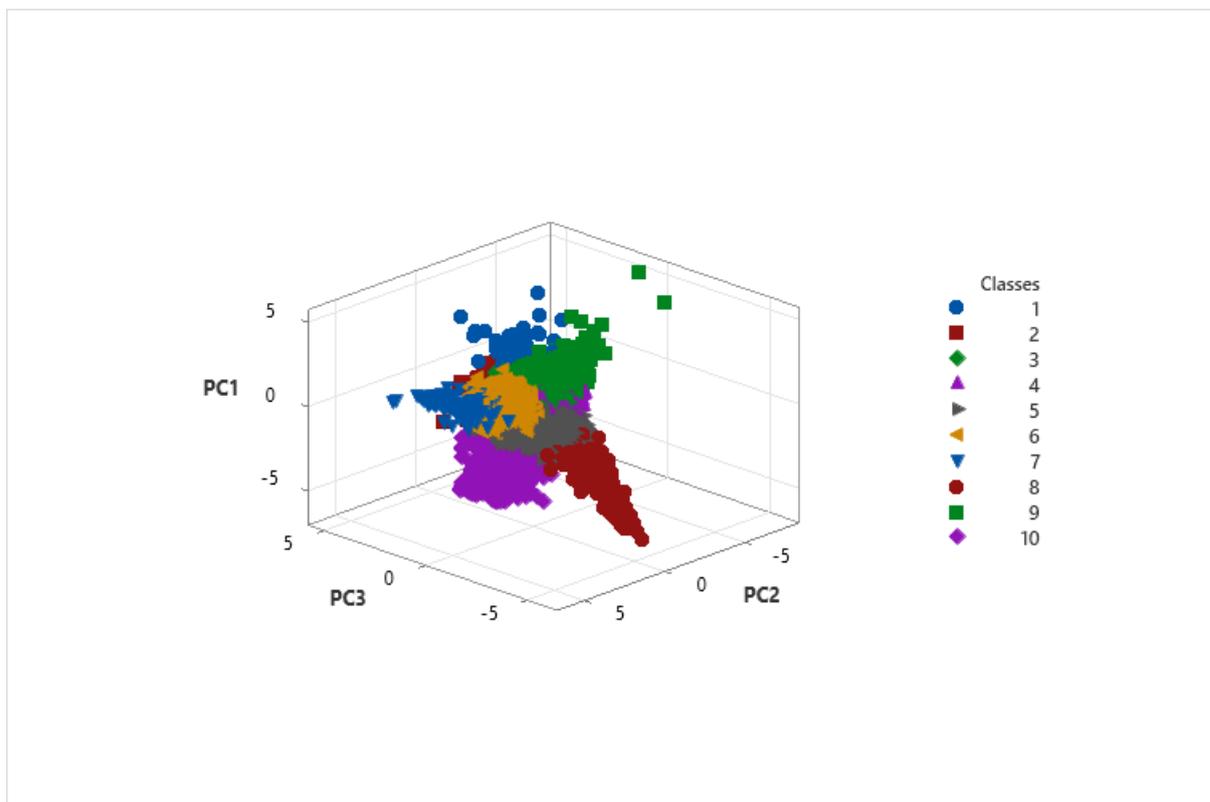


Figura 5.35: Análise de componentes principais para três componentes de acordo com as classes geradas.

De acordo com o resultado da análise de variância multivariadas (Tabela 5.2), para o presente número de classes, a hipótese nula foi rejeitada com pelo menos 99,9% de nível de confiança, indicando que os grupos são distintos entre si.

Tabela 5.2: Resultado da análise de variância multivariada (MANOVA) para as classes geradas.

Teste	Graus de liberdade (GF)	Estatística	F aproximado	Num. GF	Den. GF	p-valor
De Wilks	9	0,00003	2811,856	90	71048	0,000
Lawley-Hotteling	9	29,653	3450,364	90	94250	0,000
De Pillai	9	5,26925	1480,461	90	94338	0,000
Maior raiz de Roy	9	10,71928				

O próximo tópico discutirá a composição química e a distribuição espacial de cada classe. Tal tópico buscará relacionar as classes encontradas com a geologia local.

5.2.1 Classe 1

A Classe 1 foi atribuída a 221 amostras do banco de dados (Tabela 5.1). De acordo com a Figura 5.36 nota-se que o coeficiente da silhueta assume valores tanto positivos quanto negativos, além de razoavelmente baixos. Essa situação indica que a classe possui uma variabilidade grande, sendo que metade dos dados talvez se encaixassem melhor em outra classe (valores negativos para o coeficiente da silhueta).

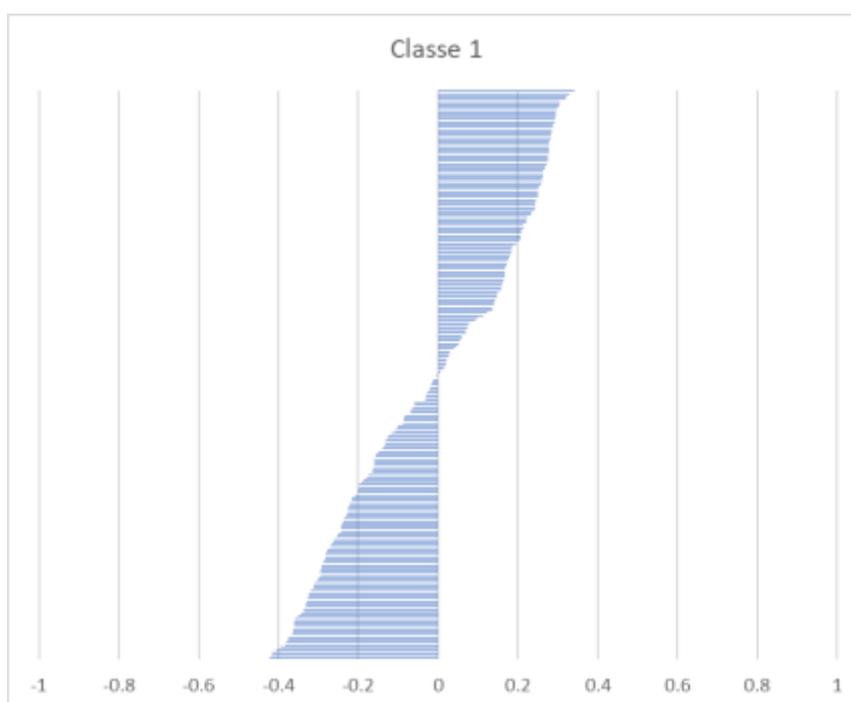


Figura 5.36: Coeficiente da silhueta para as amostras agrupadas na Classe 1.

Essa variabilidade também é observada na composição química da classe (Figura 5.37), com a grande presença de *outliers* no *boxplot*. De maneira geral, nota-se que a Classe 1 é rica em SiO_2 , Fe_2O_3 e Al_2O_3 , além de possuir grande quantidade de minerais hidratados (Alto LOIPer).

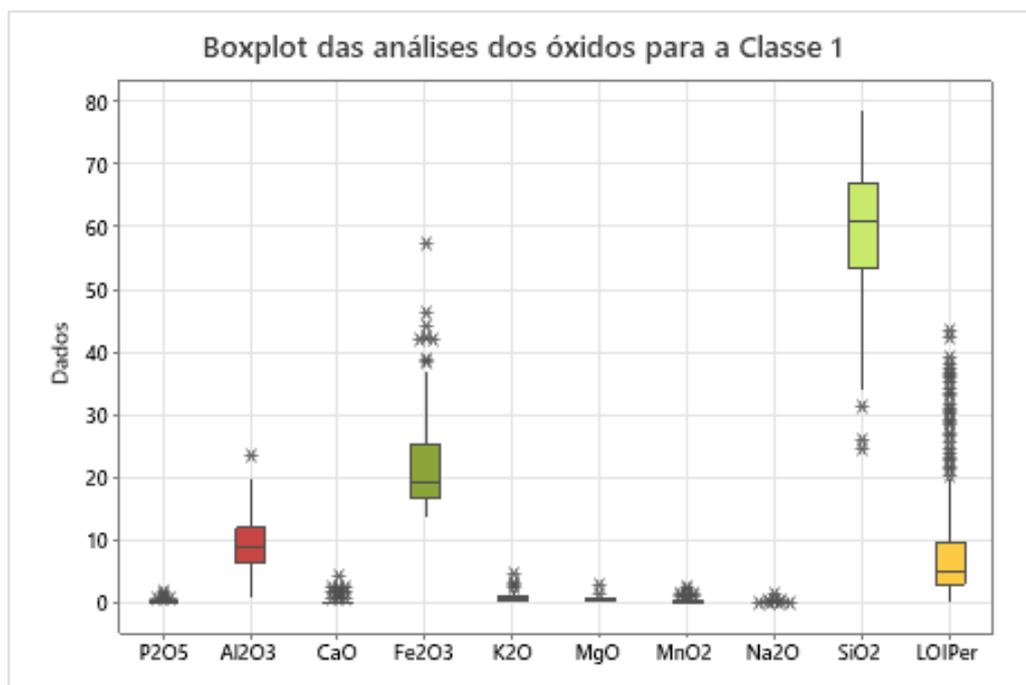


Figura 5.37: *Boxplot* das análises dos óxidos para a Classe 1. Notar a grande quantidade de *outliers* e os valores altos de SiO₂, Al₂O₃, Fe₂O₃ e de LOIPer.

A variação também é observada na distribuição espacial da Classe 1. A grande maioria dela ocorre na porção norte do depósito e relaciona-se com amostras da superfície, porém, uma parcela das observações ocorre em profundidades médias (Figura 5.38).

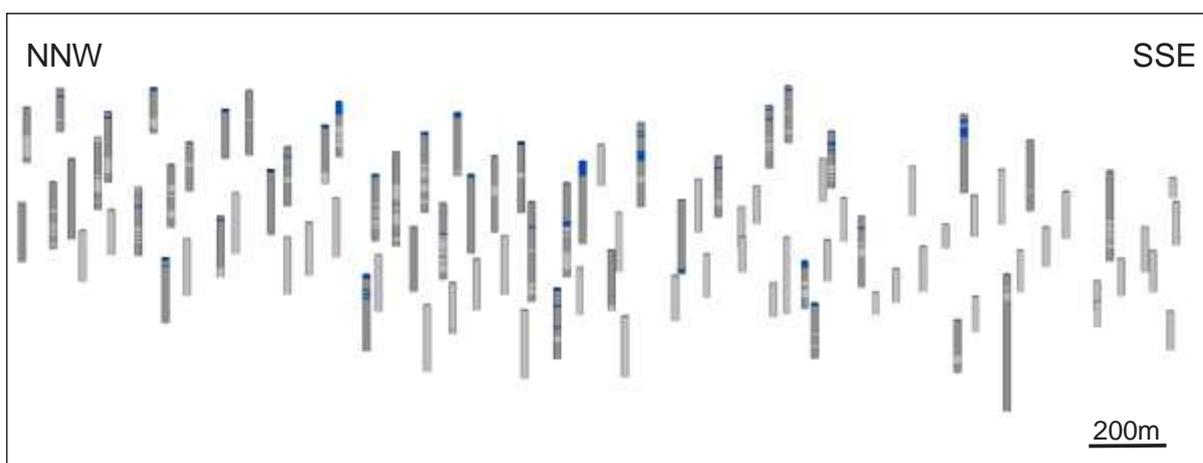


Figura 5.38: Disposição típica da Classe 1 (em azul) na porção norte do depósito. Notar a concentração na superfície e a ocorrência em menor quantidade em subsuperfície.

Diante do exposto, provavelmente, uma parte da Classe 1 deve estar relacionada a um perfil intempérico, uma vez que a grande quantidade de Al₂O₃, Fe₂O₃ e LOIPer devem estar relacionadas com óxidos e hidróxidos desses elementos, enquanto outra parte deve estar relacionada a pequenos intervalos correspondentes aos *outliers* do *boxplot* (Fig. 5.37), possivelmente intervalos ricos em argilominerais.

5.2.2 Classe 2

A Classe 2 foi atribuída a 890 amostras no banco de dados (Tabela 5.1). O coeficiente da silhueta para a Classe 2 (Figura 5.39) apresenta apenas alguns resultados negativos, sendo que a maioria dos valores se encontra no intervalo entre 0,3 e 0,58, o que indica uma boa consolidação do agrupamento.

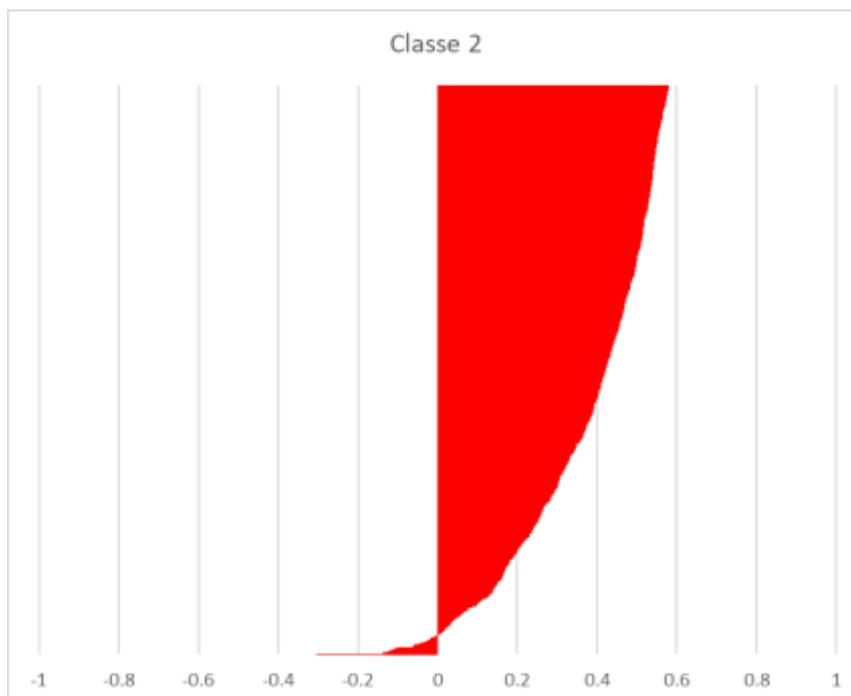


Figura 5.39: Coeficiente da silhueta para as amostras agrupadas na Classe 2.

Quimicamente, possui certa semelhança com a Classe 1, com valores menores de Fe_2O_3 e valores maiores de LOIPer (Figura 5.40). Além disso, o comportamento espacial também é parecido com o da Classe 1, entretanto, possui maior espessura e ocorre em profundidades um pouco maiores (Figura 5.41).

De acordo com o comportamento espacial e a composição química da Classe 2, ela provavelmente corresponde a intervalos de siltito estéril (<1,5% de P_2O_5). Além disso, as amostras com coeficientes de silhueta negativos devem corresponder a amostras em contatos litológicos, uma vez que os avanços na sondagem de circulação reversa são constantes.

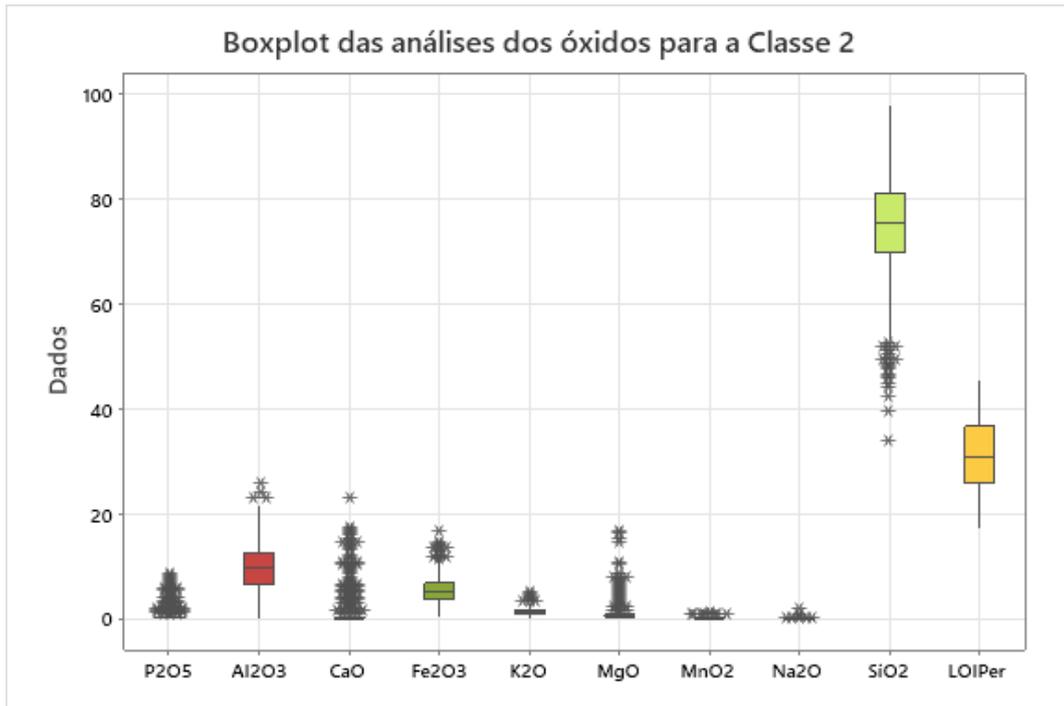


Figura 5.40: Boxplot das análises dos óxidos para a Classe 2.

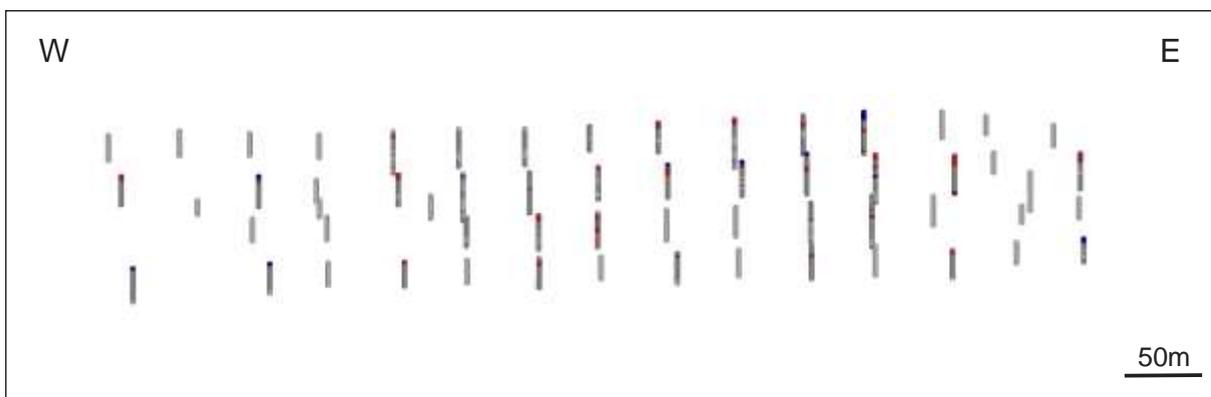


Figura 5.41: Disposição típica da Classe 2 (em vermelho), notar relação espacial com a Classe 1 em alguns pontos (em azul).

5.2.3 Classe 3

A Classe 3 foi atribuída a 2094 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta para as amostras mostra uma quantidade significativa de amostras com valores negativos (Fig. 5.42). Tais valores podem estar relacionados, de maneira análoga à Classe 2, a

contatos litológicos. Essa hipótese é reforçada pela quantidade de amostras atribuídas a essa classe, o que acarreta em uma maior distribuição espacial.

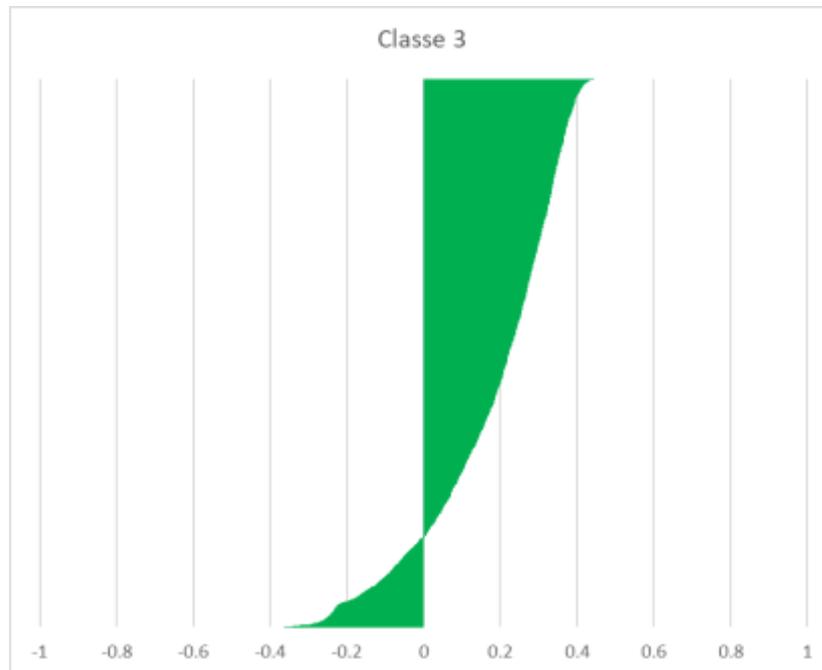


Figura 5.42: Coeficiente da silhueta para as amostras agrupadas na Classe 3.

Quimicamente, a Classe 3 se destaca pelos altos valores de SiO_2 e valores significativos de Al_2O_3 e expressivos de Fe_2O_3 (Figura 5.43). Apesar do grande número de amostras relacionadas à essa classe, o teor de P_2O_5 ainda é baixo e somente os *outliers* possuem valores de P_2O_5 maiores que 1,5%. Tais teores anômalos devem estar relacionados com contatos entre essa Classe e classes mais ricas em P_2O_5 .

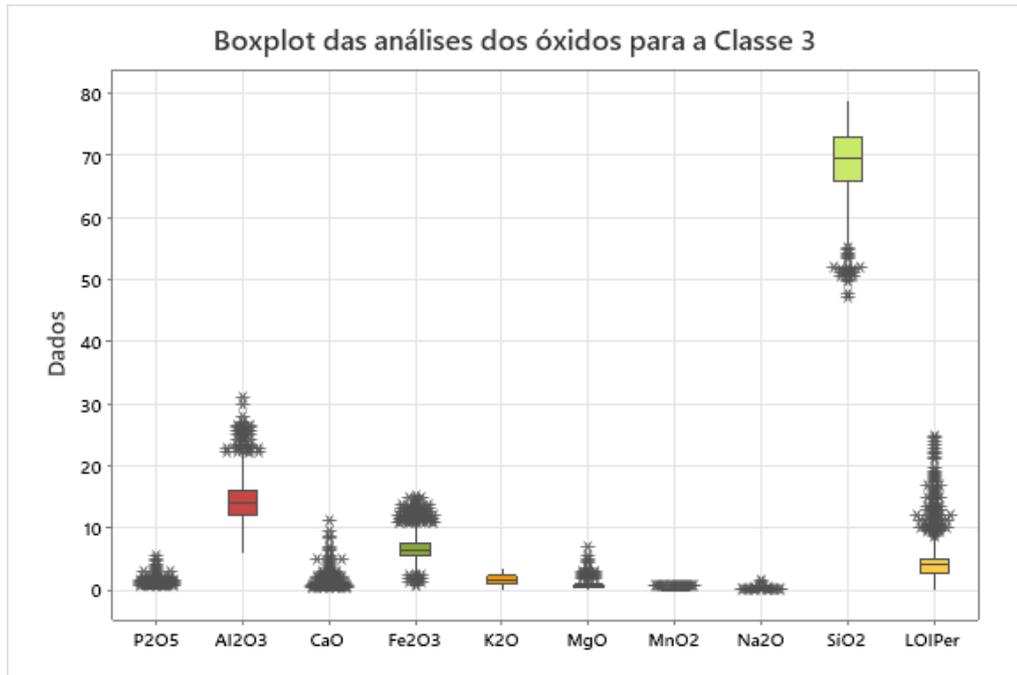


Figura 5.43: Boxplot das análises dos óxidos para a Classe 3.

Espacialmente, ocorre desde a superfície até a profundidades mais elevadas, sendo amplamente distribuído (Figura 5.44). Dessa maneira, deve estar relacionado à um siltito arenoso ou a um arenito fino.

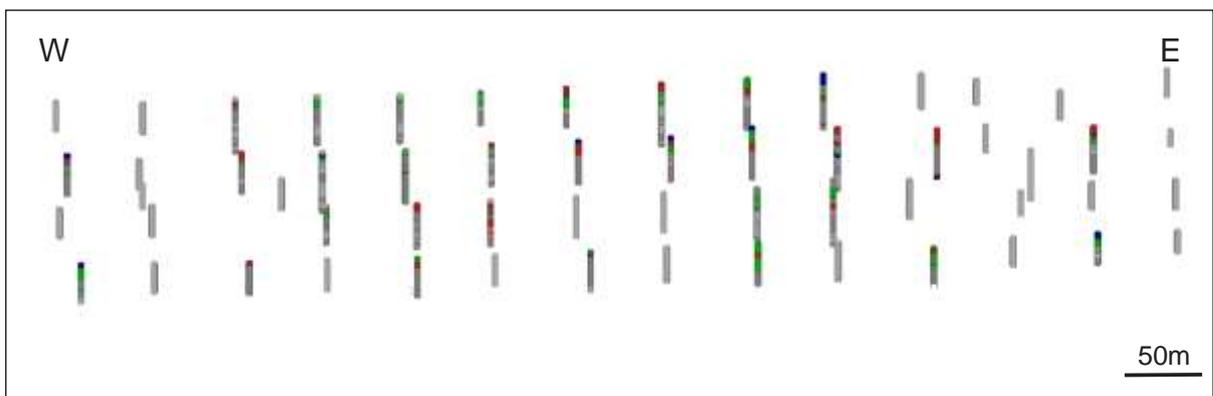


Figura 5.44: Disposição típica da Classe 3 (em verde). Classe 1 em azul e Classe 2 em vermelho.

5.2.4 Classe 4

É a Classe que mais foi atribuída as amostras do banco de dados, com 2444 amostras atribuídas a ela (Tabela 5.1). O coeficiente da silhueta para as amostras mostra uma silhueta muito

parecida com a Classe 2, com somente poucas amostras mostrando valores negativos e a grande maioria mostrando valores entre 0,3 e 0,58 (Figura 5.45). Quimicamente, a Classe 4 é bem semelhante a Classe 3, entretanto, apresenta valores mais altos de SiO_2 e valores menores de Al_2O_3 , Fe_2O_3 e K_2O (Figura 5.46).

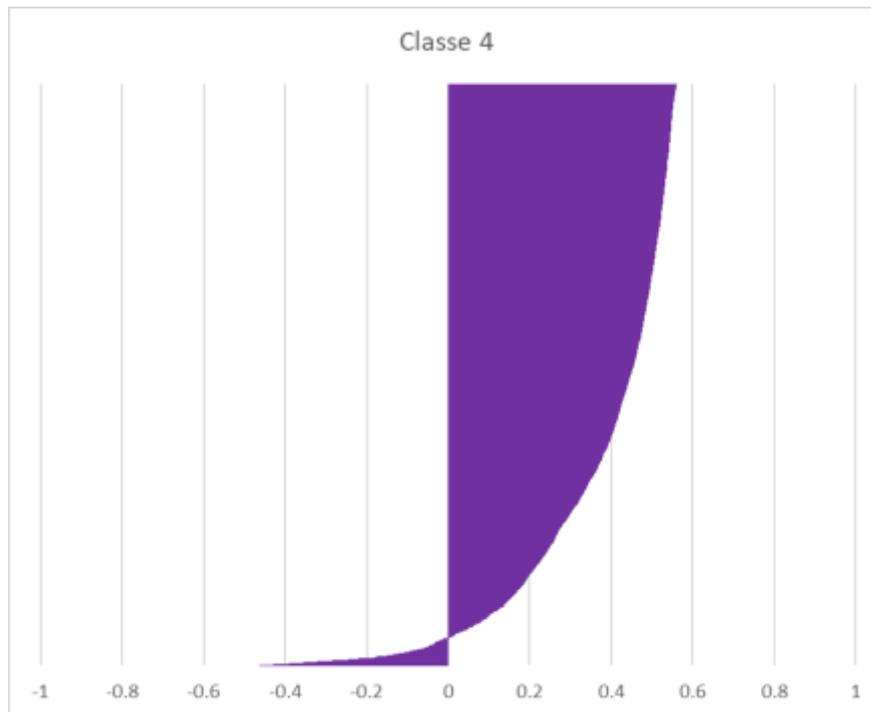


Figura 5.45: Coeficiente da silhueta para as amostras agrupadas na Classe 4.

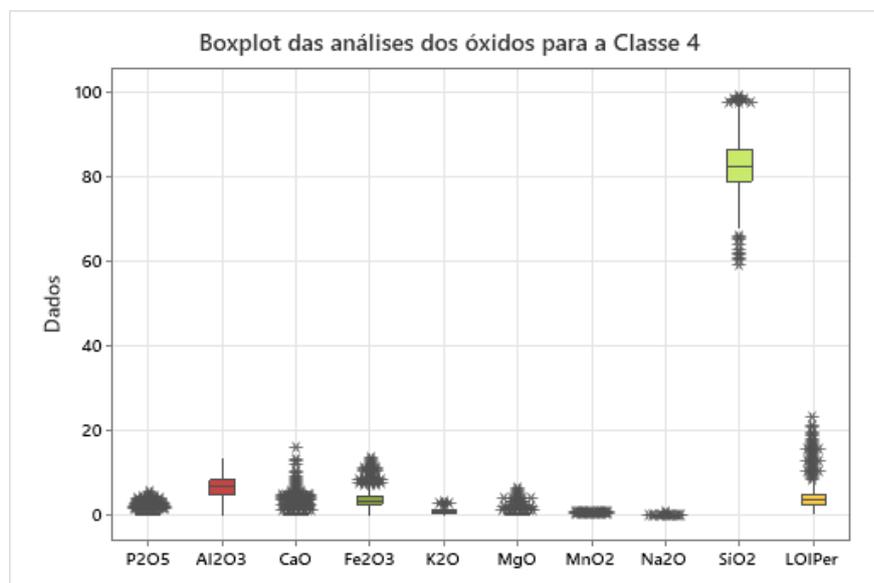


Figura 5.46: Boxplot das análises de óxidos para a Classe 4.

Espacialmente, de maneira similar à Classe 3, a Classe 4 ocorre de profundidades superficiais até profundidades maiores (Fig. 5.47). Ocorre associada principalmente às Classes 2 e 3. Corresponde, provavelmente, de maneira análoga a Classe 3, a um siltito arenoso ou arenito fino, com altas quantidades de quartzo. Os teores de P_2O_5 maiores que 1,5% também são apenas *outliers*.

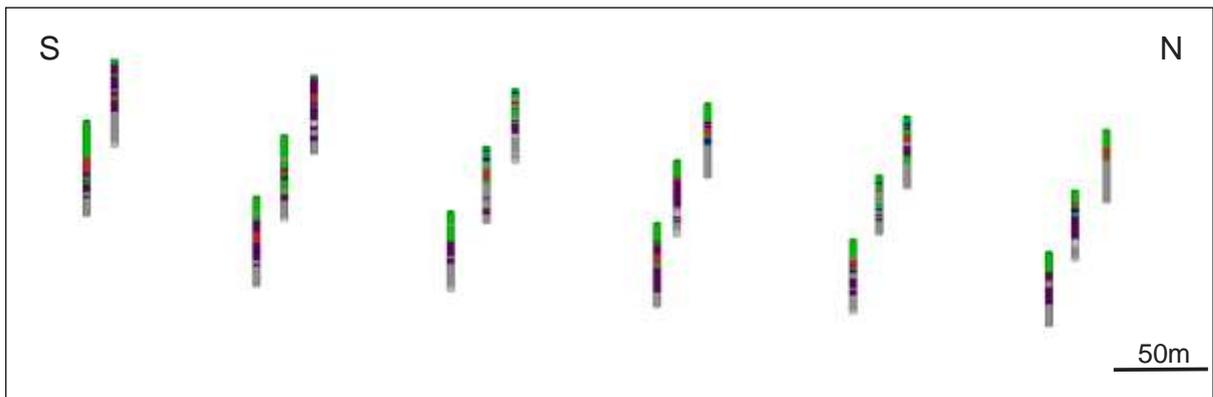


Figura 5.47: Disposição da Classe 4 (em roxo). Classe 1 em azul, Classe 2 em vermelho e Classe 3 em verde.

5.2.5 Classe 5

A Classe 5 foi atribuída a 1054 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta para as amostras mostra um comportamento similar ao da Classe 1, com aproximadamente metade das amostras apresentando valores negativos (Figura 5.48).

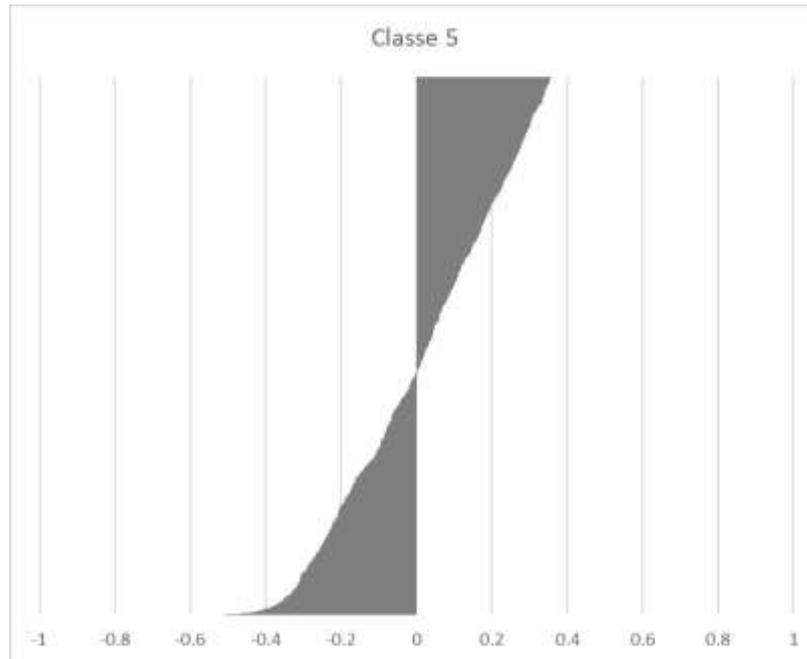


Figura 5.48: Coeficiente da silhueta para as amostras da Classe 5.

Quimicamente, a Classe 5 apresenta teores elevados de SiO_2 , P_2O_5 , Al_2O_3 e CaO . Dessa maneira, pode ser considerada como minério, sendo que o primeiro quartil do *boxplot* para os teores de P_2O_5 possui valor de 4.36% (Figura 5.49).

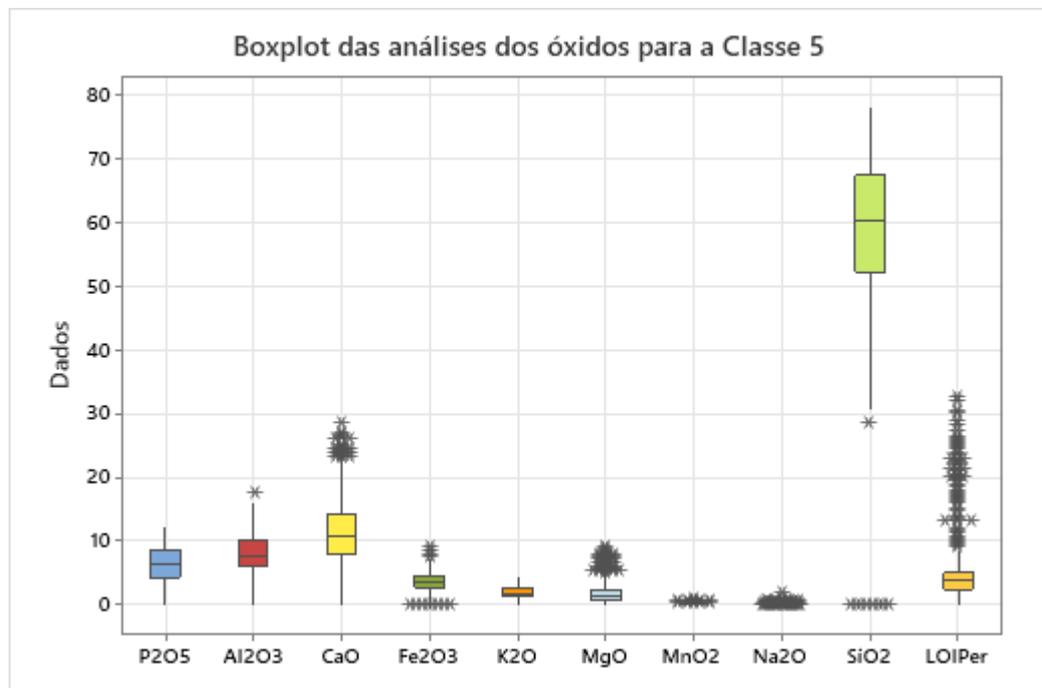


Figura 5.49: Boxplot das análises dos óxidos para a Classe 5.

É interessante notar, entretanto, que de acordo com o correlograma da Figura 5.50 o coeficiente de correlação de Pearson entre CaO e P₂O₅ é igual a 0,01, muito distante do valor 1,00 esperado para um minério rico em apatita.

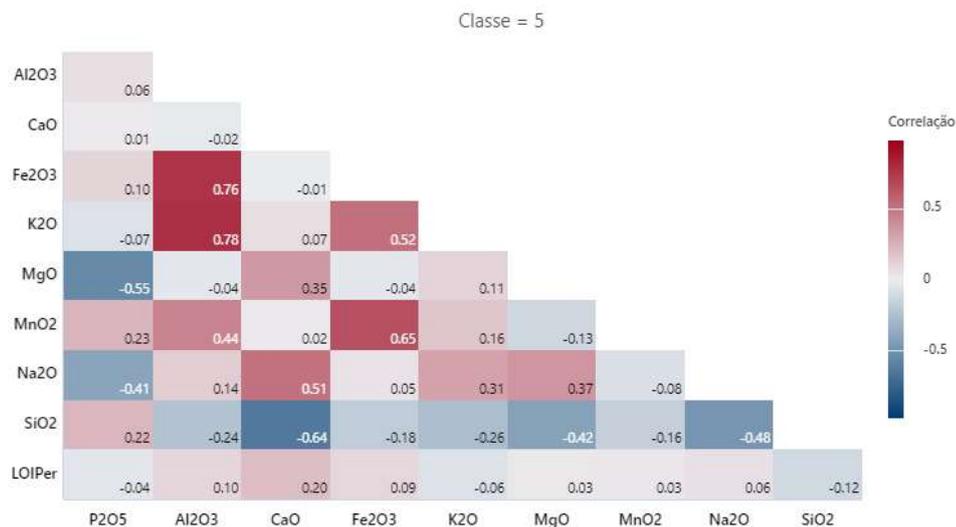


Figura 5.50: Correlograma dos óxidos da Classe 5.

Dessa maneira, o coeficiente da silhueta, juntamente com o correlograma mostram a existência de uma distribuição bimodal das amostras, que fica evidente no diagrama de dispersão entre P₂O₅ e CaO da Figura 5.51. Nessa figura, é possível notar 2 tendências, uma compatível com um minério rico em apatita (Teores diretamente proporcionais de CaO e P₂O₅) e outra compatível com uma rocha rica em calcita e pobre em apatita (Teores altos de CaO e baixos de P₂O₅).

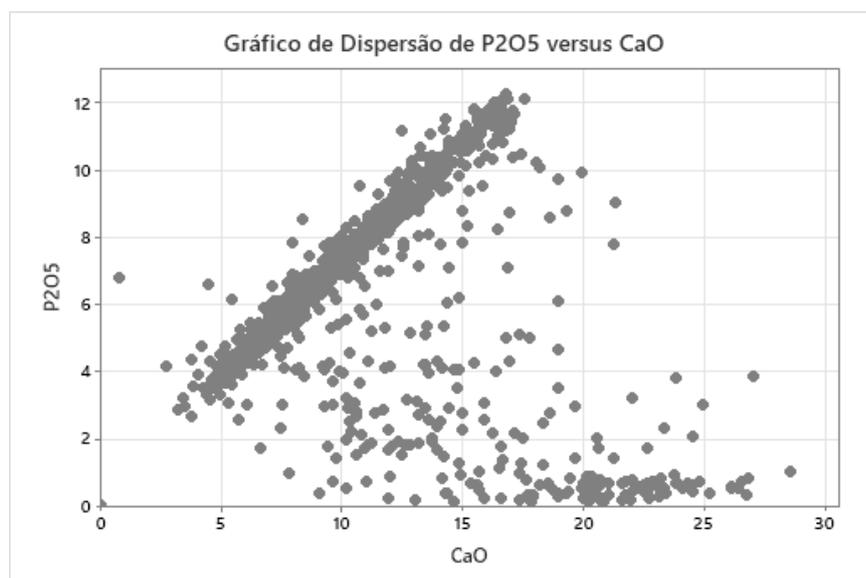


Figura 5.51: Gráfico de Dispersão de P₂O₅ e CaO para a Classe 5.

A Figura 5.52 mostra que a Classe 5 ocorre na parte inferior do depósito, exatamente no contato dos siltitos da Fm. Sete Lagoas com os carbonatos da mesma formação. Dessa maneira, o que temos na Figura 5.2.5.4 são diferentes níveis de contaminação, passando de um siltito mineralizado puro (Teores diretamente proporcionais de CaO e P₂O₅) até um carbonato puro (Teores altos de CaO e baixos de P₂O₅). Isso ocorre devido à natureza da sondagem de circulação reversa. Além disso, pode-se ter uma margas com diferentes níveis de contribuição terrígena: Pontos com razão CaO/P₂O₅ mais baixas indicam pontos de maior contribuição terrígena, e pontos com razão maiores de CaO/P₂O₅, indicam zonas de maior precipitação química.



Figura 5.52: Localização das amostras atribuídas à Classe 5 (Em amarelo). Notar a associação com as amostras da Classe 10 (Em azul claro), relacionada aos carbonatos.

5.2.6 Classe 6

A Classe 6 foi atribuída a 1290 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta das amostras mostra um comportamento semelhante ao da Classe 3, com uma porção significativa das amostras com coeficiente negativo de silhueta e a maioria com valores no intervalo de 0,2 a 0,4 (Figura 5.53).

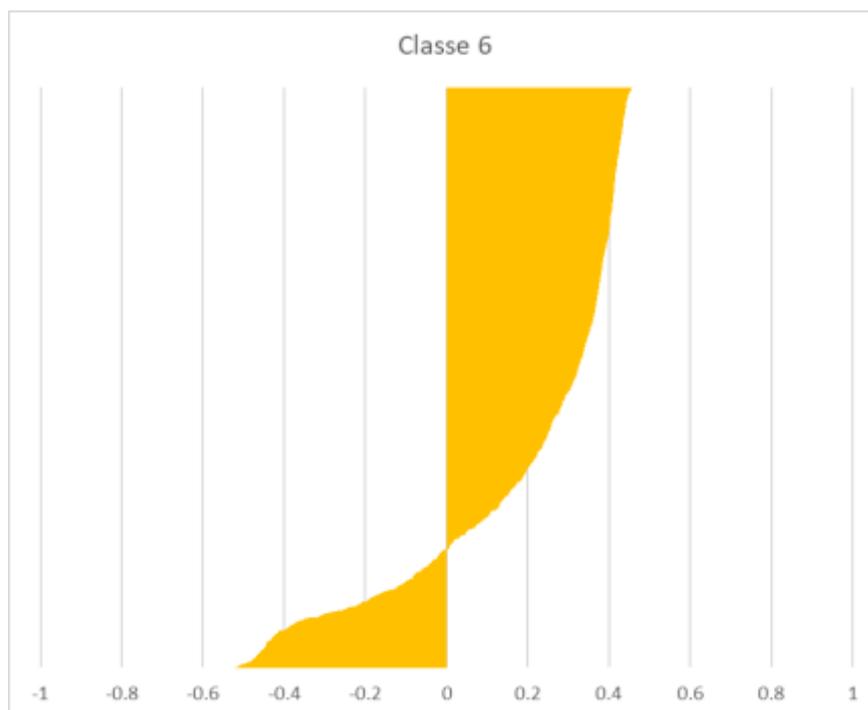


Figura 5.53: Coeficiente da silhueta para as amostras atribuídas a Classe 6.

Quimicamente, a Classe 6 possui certa semelhança com a Classe 5, porém com teores de P_2O_5 e CaO mais baixos e teores mais altos de Al_2O_3 , Fe_2O_3 e K_2O (Figura 5.54). Nota-se, também, de maneira análoga, mas não tão acentuada, uma distribuição bimodal no gráfico de dispersão entre P_2O_5 e CaO (Figura 5.55). O que também indica a presença de amostras correspondentes a rochas carbonáticas que foram associadas a Classe 6. Apesar de um teor mais elevado, a grande maioria das amostras não atinge o teor de P_2O_5 igual a 1,5% ou estão nesse limite (Terceiro quartil 1,62% P_2O_5)

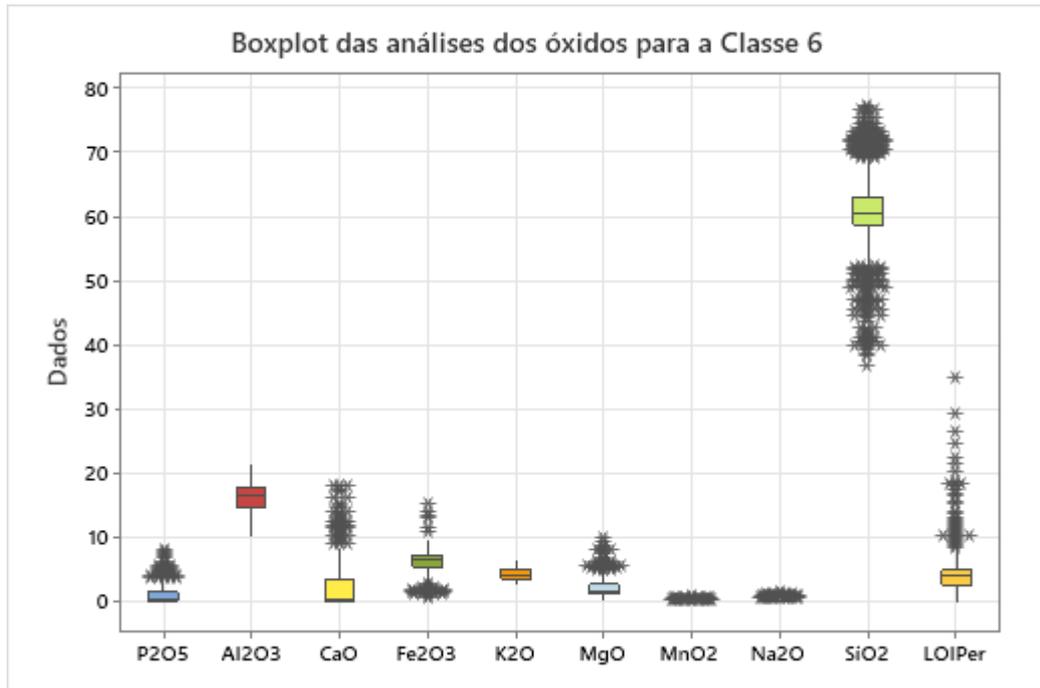


Figura 5.54: Boxplot das análises dos óxidos para a Classe 6.

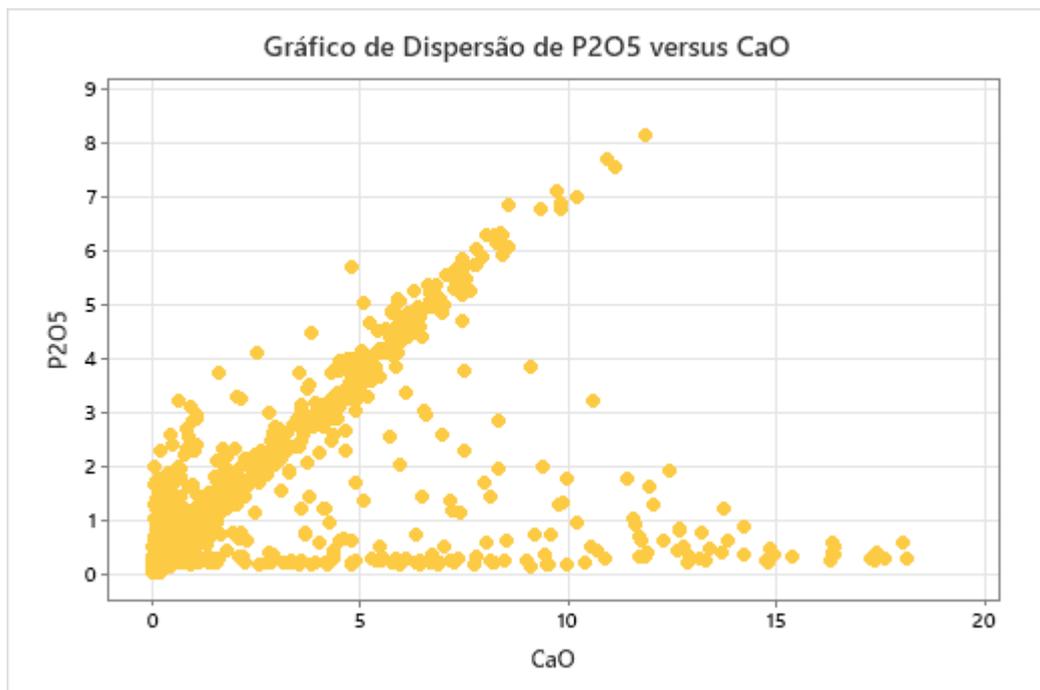


Figura 5.55: Gráfico de Dispersão de P2O5 e CaO para a Classe 6.

Especialmente, a Classe 6 também possui um comportamento similar a Classe 5, ocorrendo em alternância com as outras Classes associadas ao siltito, porém restritas mais a norte do depósito (Figura 5.56).

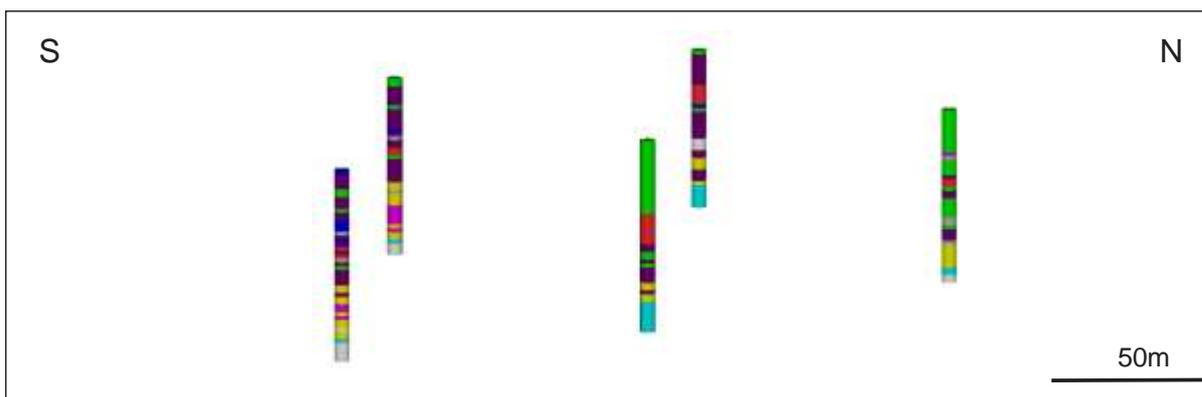


Figura 5.56:Localização das amostras atribuídas à Classe 6 (Em rosa). Classe 1 em azul, Classe 2 em vermelho, Classe 3 em verde, Classe 4 em roxo, Classe 5 em amarelo.

5.2.7 Classe 7

A Classe 7 foi atribuída a 385 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta para as amostras dessa Classe mostra um comportamento parecido com o da Classe 6, com parte significativa dos dados apresentando coeficiente da silhueta negativo (Figura 5.57).

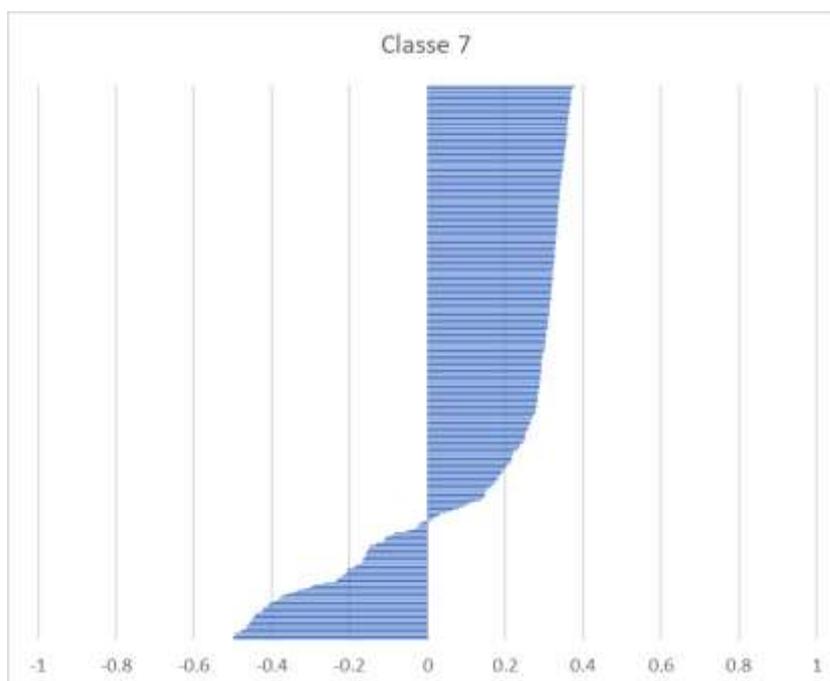


Figura 5.57: Coeficiente da silhueta para amostras atribuídas a Classe 7.

Quimicamente, a Classe 7 é a mais distinta das Classes, com os maiores teores de Na_2O de todas as Classes (Figura 5.58). De acordo com a Tabela 5.1, a maioria das amostras que foram atribuídas à essa Classe foram descritas como “Granito” (304 das 385 amostras), dessa forma,

o *boxplot* com para as análises dos óxidos ficou similar ao da Figura 5.15, mostrando também uma grande quantidade de *outliers*.

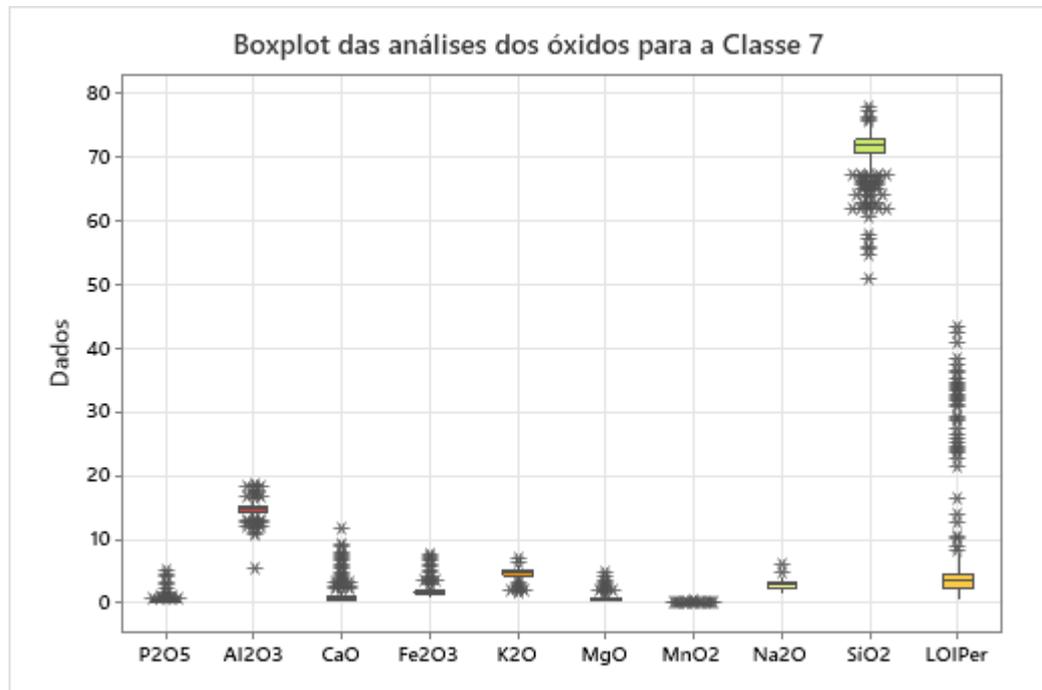


Figura 5.58: Boxplot das análises dos óxidos para a Classe 7.

Especialmente, as amostras atribuídas à Classe 7 estão em consonância com a ocorrência dos granitos na região (Suíte Aurumina), sendo restrita a base do depósito ao norte e ao sul (Figura 5.59).

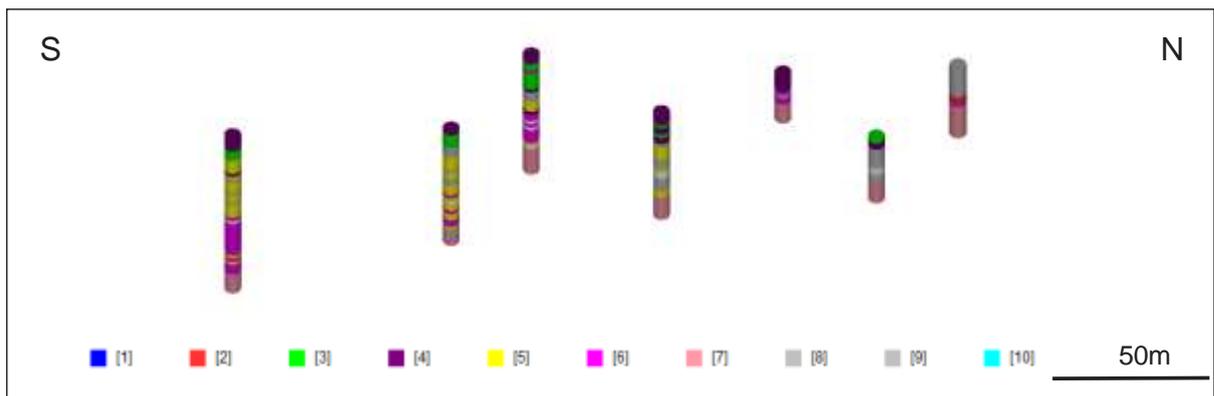


Figura 5.59: Localização da Classe 7 no setor norte do depósito.

5.2.8 Classe 8

A Classe 8 foi atribuída a um total de 401 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta para as amostras da Classe 8 mostra um comportamento similar 2 e 4, com poucas amostras apresentando um valor negativo e a maioria mostrando valores entre 0,3 e 0,59 (Figura 5.60).

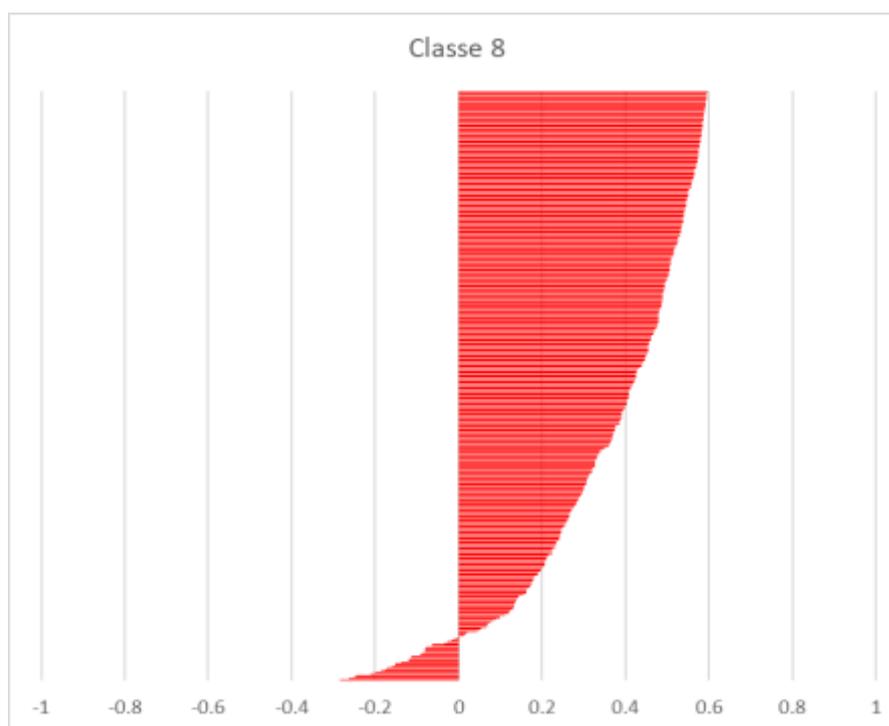


Figura 5.60: Coeficiente da silhueta para as amostras atribuídas a Classe 8.

Quimicamente, a Classe 8 é a que apresenta maiores teores de P_2O_5 , que variam de 11,85% a 30,70% (Figura 5.61), apresentando um coeficiente de correlação de Pearson de 0,98 entre CaO e P_2O_5 (Figura 5.62). Isso indica um minério rico em apatita, além de mostrar uma baixa contaminação com carbonatos. O coeficiente negativo de silhueta para algumas amostras deve indicar amostras mistas obtidas em contatos litológicos por conta dos avanços constantes da sondagem de circulação reversa.

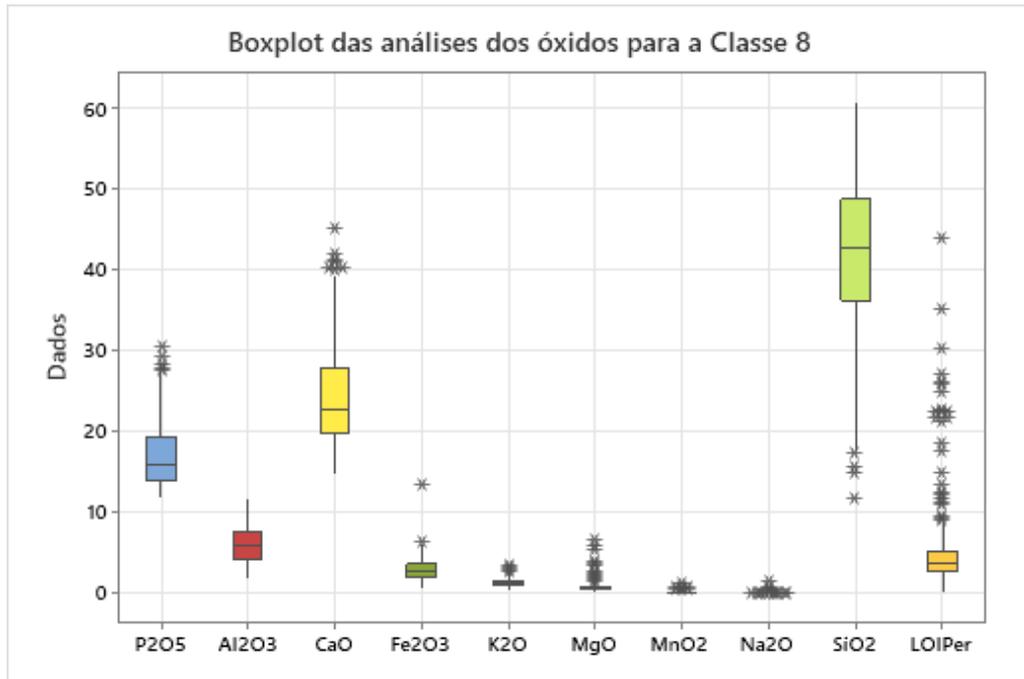


Figura 5.61: Boxplot das análises dos óxidos para a Classe 8.

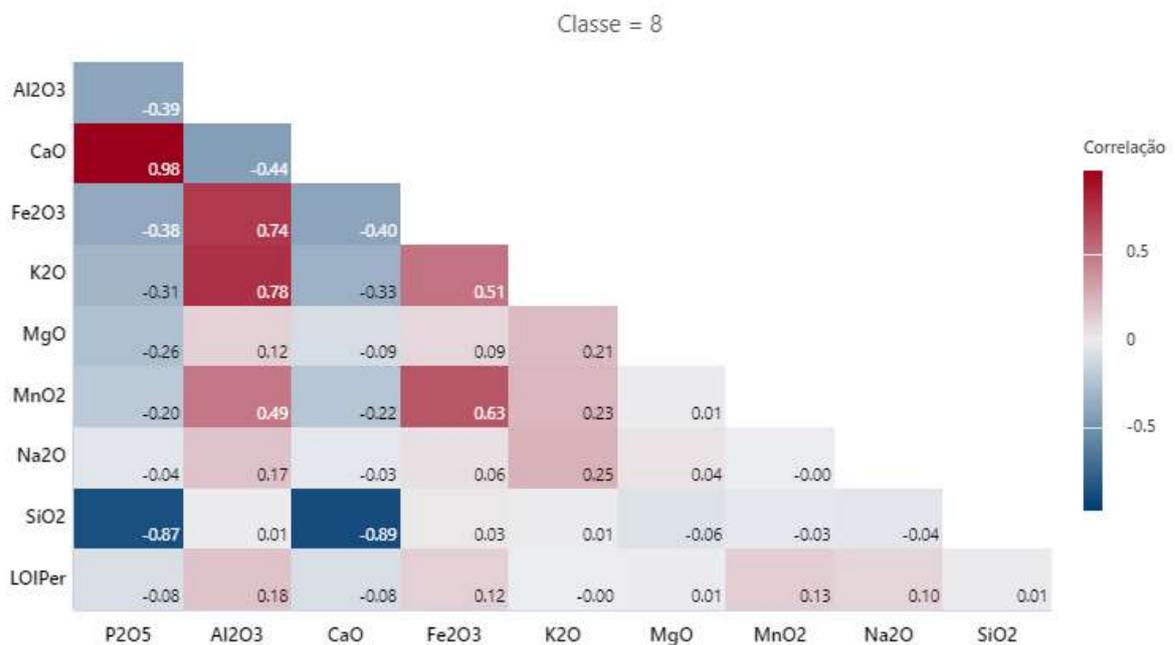


Figura 5.62: Correlograma entre os valores de óxidos para a Classe 8.

Especialmente, a Classe 8 se encontra na porção central do depósito e está de acordo com a ocorrência da Brecha Tipo-I, o minério de mais alto teor do depósito (Figura 5.63). Sendo que a maioria das amostras enquadradas nessa Classe foram descritas dessa maneira.

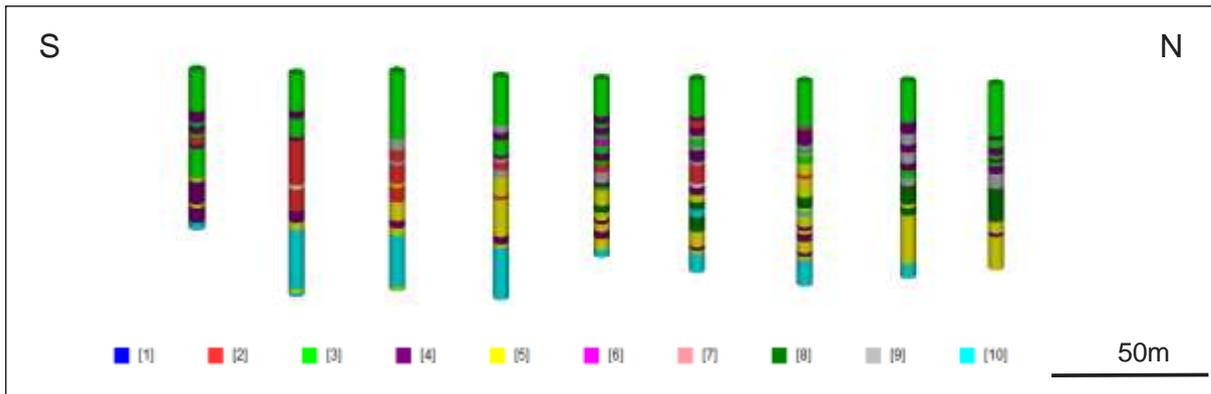


Figura 5.63: Localização de algumas amostras atribuídas a Classe 8 na porção central do depósito.

5.2.9 Classe 9

A Classe 9 foi atribuída a 371 amostras do banco de dados (Tabela 5.1). O coeficiente da silhueta para as amostras dessa classe apresenta comportamento similar aos das Classes 2, 4 e 8, com poucas amostras apresentando valores negativos e a maioria mostrando valores entre 0,2 e 0,5 (Figura 5.64)

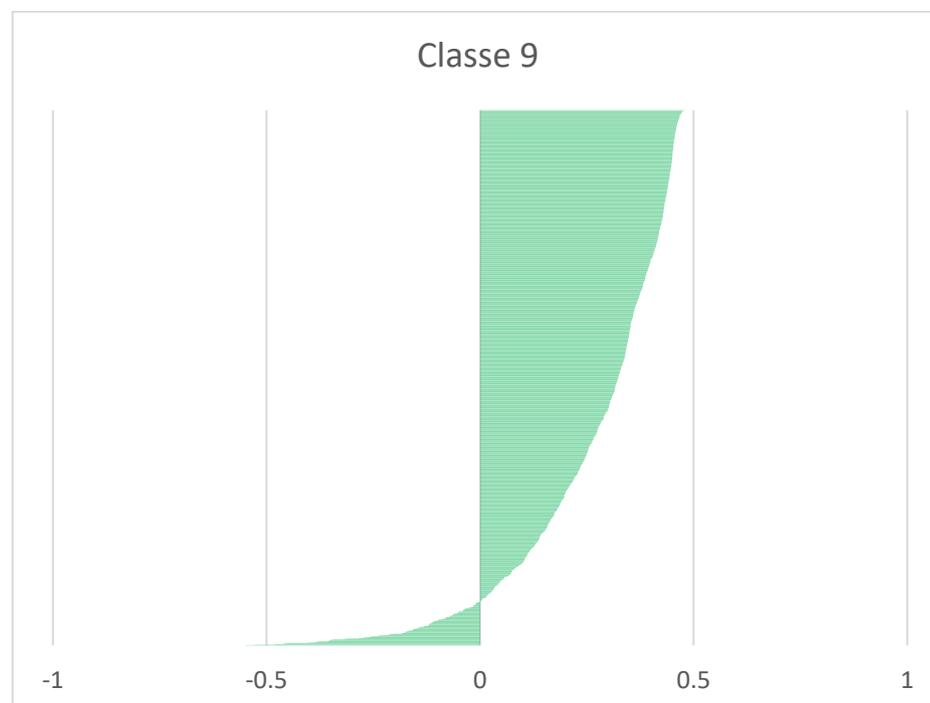


Figura 5.64: Coeficiente da silhueta para as amostras atribuídas a Classe 9.

Quimicamente, a Classe 9 apresenta semelhanças com as Classes 2, 3, 4 e 6 (Figura 5.65), com baixos teores de P_2O_5 , altos teores de SiO_2 e teores medianos de Al_2O_3 e Fe_2O_3 . Dessa forma, também deve ser correlacionada a uma variação de siltito arenoso com baixo teor.

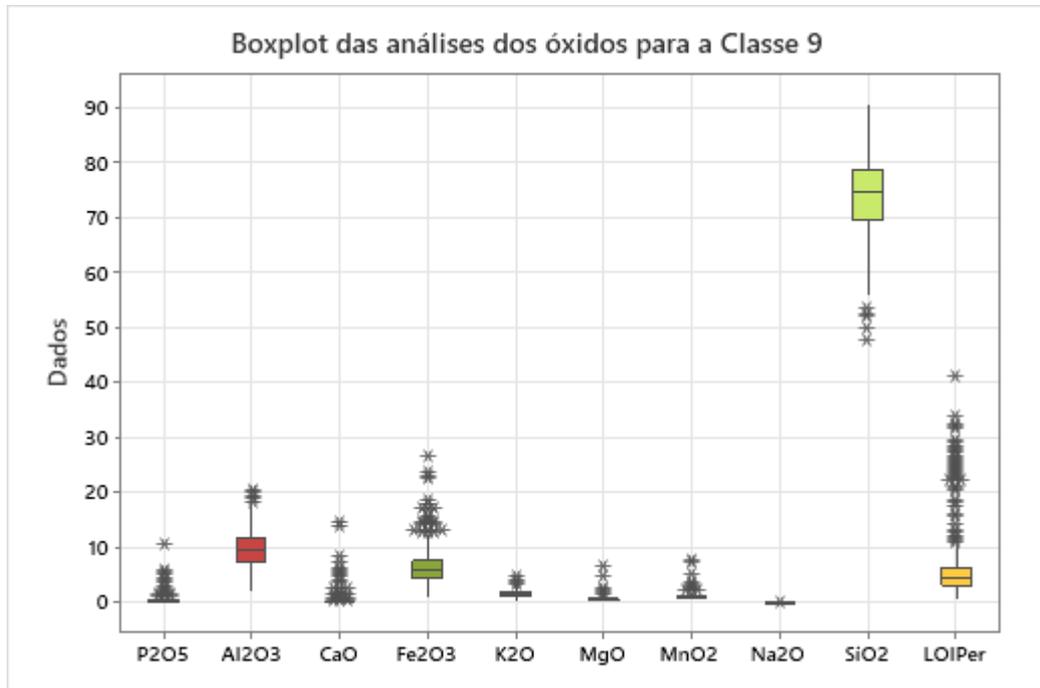


Figura 5.65: Boxplot das análises dos óxidos para a Classe 9.

Especialmente, as amostras atribuídas a essa Classe também ocorrem de maneira alternada com as Classes 2, 3, 4, 5 e 8, sustentando a hipótese de que se trata de uma variação do siltito (Figura 5.66).

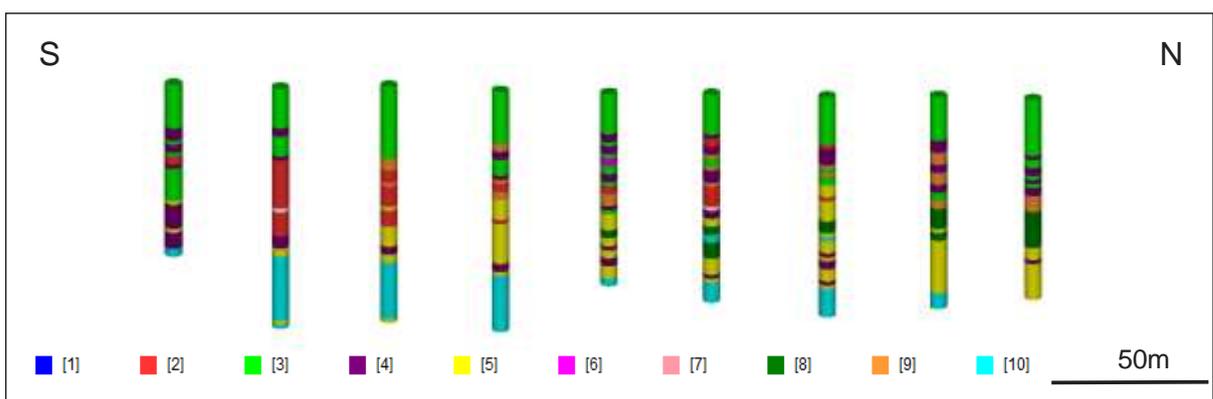


Figura 5.66: Localização de algumas das amostras atribuídas a Classe 9 no centro do depósito.

5.2.10 Classe 10

A Classe 10 foi atribuída a 1343 amostras do banco de dados. O coeficiente da silhueta para as amostras dessa Classe segue uma distribuição similar às das Classes 2, 4, 8 e 9 (Figura 5.67), com poucas amostras apresentando valores negativos e maioria apresentando valores entre 0,2 e 0,4.

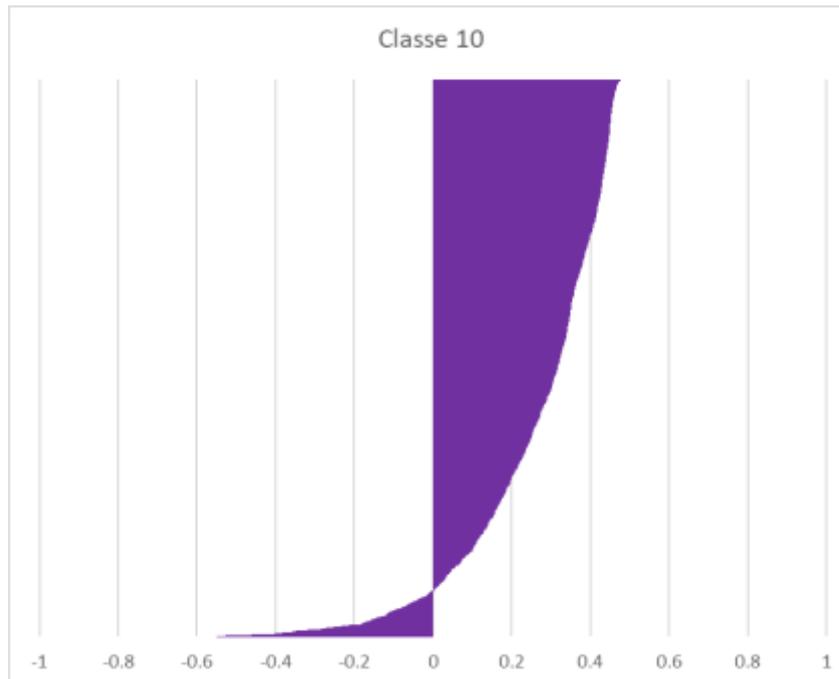


Figura 5.67: Coeficiente da silhueta para amostras atribuídas a Classe 10.

Quimicamente, as amostras atribuídas a Classe 10 são bem distintas das outras Classes, apresentando elevados teores de CaO e MgO, com teores relativamente baixos de SiO₂ (Figura 5.68). Assim sendo, é possível atribuir essa Classe as rochas carbonáticas que ocorrem no depósito. É interessante notar que a grande maioria das amostras atribuídas a essa Classe foram descritas como rochas carbonáticas (Marga, Dolomito e Calcário) (Tabela 5.1) e que apesar dessa classificação visual quanto a textura e composição mineralógica das amostras, essas acabaram agrupadas em apenas uma Classe.

Especialmente, as amostras dessa Classe ocorrem espalhadas em todo o depósito, ocorrendo abaixo das Classes correlacionadas às outras rochas, com exceção da Classe 7, que ocorre ainda mais abaixo. Tal constatação está de acordo com a geologia local, que coloca as rochas carbonáticas como base do depósito da Mina Domingos (Figura 5.69).

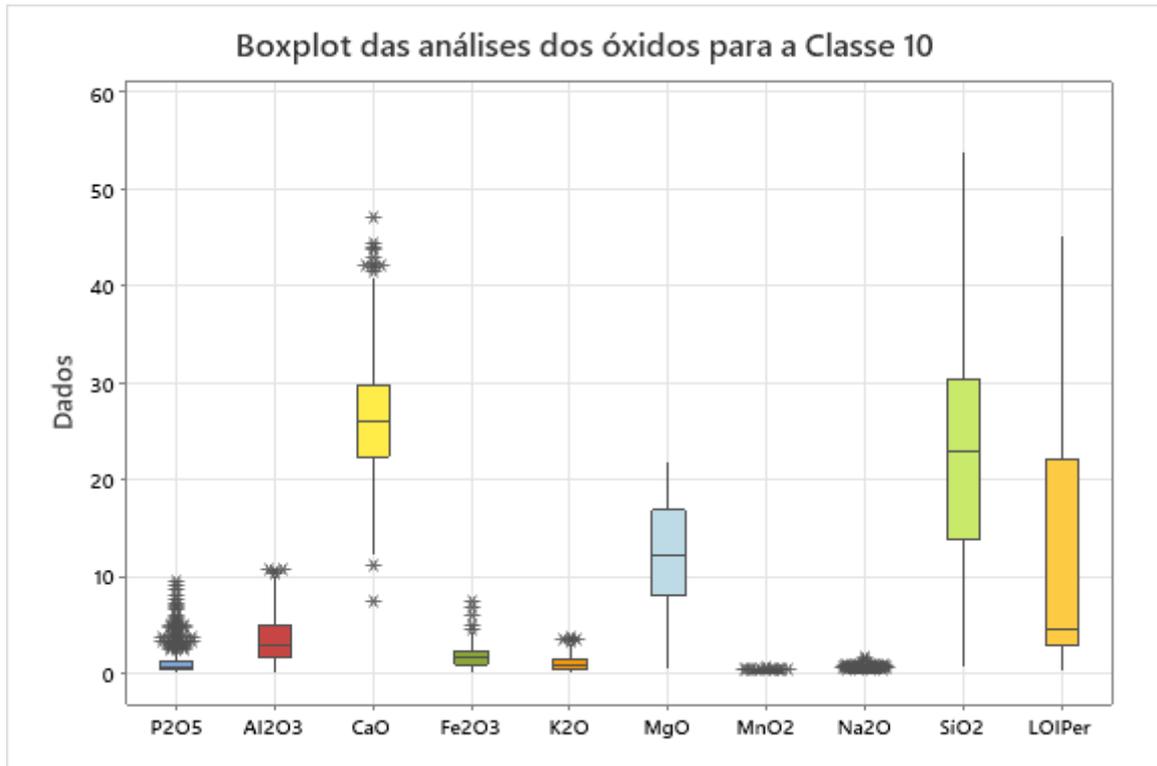


Figura 5.68: Boxplot das análises dos óxidos para a Classe 10

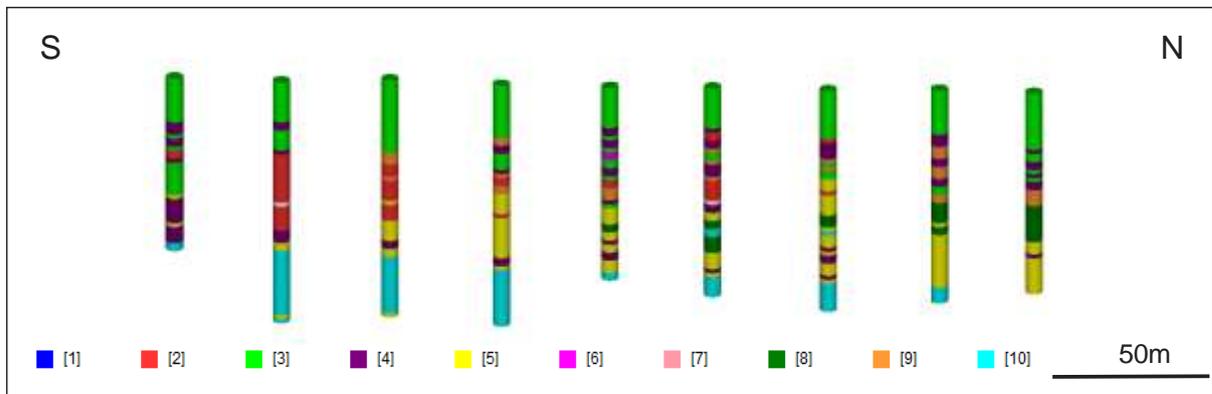


Figura 5.69: Localização das amostras atribuídas a Classe 10 no centro do depósito.

5.2.11 Visão Geral

À primeira vista, as classes geradas podem levar a um entendimento de um modelo geológico lenticular, ou de um modelo de camadas muito finas e, portanto, pouco prático (Figura 5.70). Entretanto, se considerarmos as Classes 2, 3, 4, 6 e 9 como sendo variações de siltito com pouco teor (< 1,5% de P_2O_5) podemos perceber que o modelo pode ser dividido em 3 camadas

distintas: Silito Estéril (Classes 2, 3, 4, 6 e 9), Silito Mineralizado (Classe 5) e Silito de Alto Teor (Classe 8) (Figura 5.70). É possível notar que a camada mineralizada não ocorre desde a superfície, como entendido de acordo com a Figura 4.2. Além disso, é possível notar que as Classes geradas possuem uma maior aderência com o teor de P_2O_5 quando comparado com as Figuras 4.2 e 4.3 (Figura 5.71).

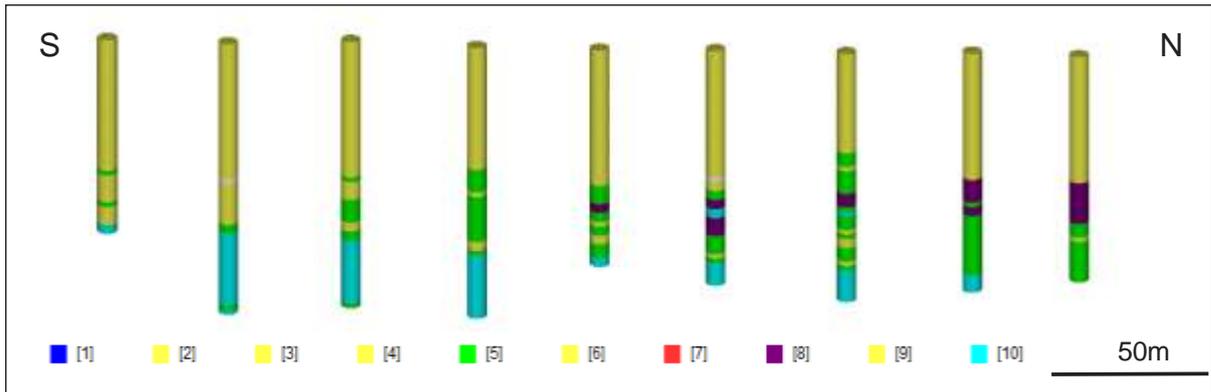


Figura 5.70: Seção mostrando a organização das Classes a partir do entendimento de três classes distintas.

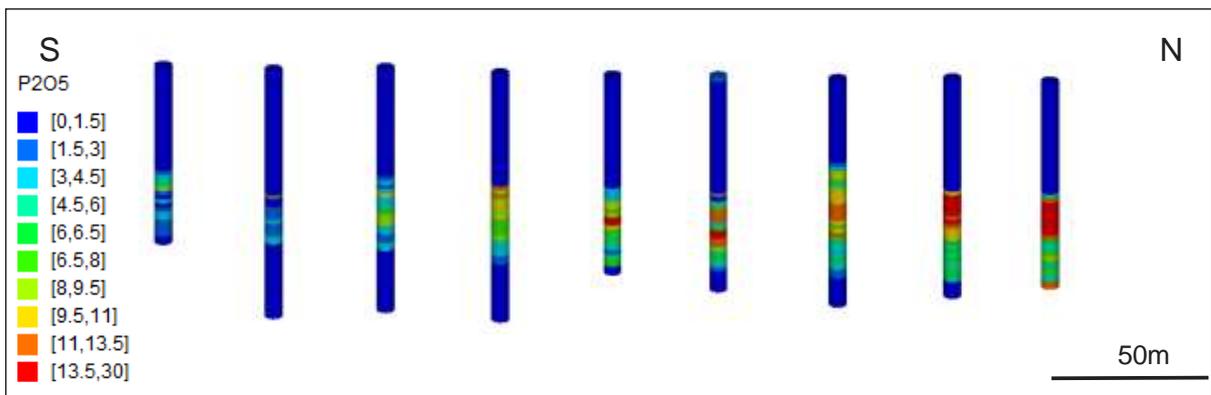


Figura 5.71: Teor de P_2O_5 em diferentes furos de sondagem. Notar a forte correlação das Classes 5 e 8 com os teores de P_2O_5 .

6 CONCLUSÕES

A utilização de técnicas de agrupamento mostra-se cada vez mais presente no dia a dia, abrangendo diferentes áreas do conhecimento. No presente estudo, sua utilização buscou diminuir a influência da subjetividade humana na descrição de amostras geológicas a partir da criação de classes utilizando o algoritmo *k-means*.

É possível perceber que mesmo separando as rochas carbonáticas da região em três distintas litologias, a partir da análise visual de suas texturas e composição mineralógica, com a aplicação do algoritmo, as três litologias foram classificadas em apenas uma classe. Além disso, nota-se que litologias descritas poucas vezes como o Silexito e o Arenito, acabaram por serem aglutinadas dentro de Classes correspondentes ao Siltito.

Também, a falta de distinção entre as diferentes rochas matrizes na descrição do Solo, levou-o a ser aglutinado em diferentes Classes, principalmente as do siltito, com exceção da Classe 1, que por ocorrer majoritariamente na superfície foi considerada como sendo resultante de um perfil de intemperismo.

Como visto no Capítulo 5.2.11, apesar de inicialmente as Classes mostrarem uma tendência lenticular do depósito, ao levar em consideração que várias Classes devem ser apenas variações do Siltito (presença de mais ou menos argila, mais ou menos quartzo, ausência ou presença de apatita), o modelo se torna consonante com o observado localmente e as camadas mineralizadas se tornam bem distintas.

A presença de amostras com valores de coeficiente de silhueta negativos em todas as Classes pode indicar diferentes graus de contaminação por conta da natureza das amostras obtidas a partir da sondagem de circulação reversa. Além disso, tal fator pode ser uma das explicações para a grande quantidade de *outliers* encontrados em todas as Classes.

No mais, a fim de se delimitar com mais precisão e tentar definir no modelo a variação observada localmente nos siltitos, sugere-se a determinação da assembleia mineralógica de cada tipo de siltito, possibilitando a comparação com as análises químicas e sua presença no modelo, além da análise de elementos traços.

REFERÊNCIAS

- DARDENNE, M.A. *et al.* 1986. Proterozoic and Cambrian phosphorites – regional review: Brazil. In: COOK, P, J., SHERGOLD, J, H. **Phosphate deposits of the world. Proterozoic and Cambrian Phosphorites**. London, Cambridge University, p. 116-131
- DINDARLOO, S. R.; SIAMI-IRDEMOOSA, E. **Data mining in mining engineering: results of classification and clustering of shovels failures data**. *International Journal of Mining, Reclamation and Environment*. Taylor & Francis, v. 31, p. 105-118, 2016.
- DRUMMOND, J. R. *et al.* 2015. Neoproterozoic peritidal phosphorite from the Sete Lagoas Formation (Brazil) and the Precambrian phosphorus cycle. *Sedimentology*, v. 62, n. 7, p. 1978-2008.
- FOUEDJIO, F. **A hierarchical clustering method for multivariate geostatistical data**. *Spatial Statistics*. Elsevier, v. 18b, p. 333-351, 2016.
- FUCK, R. A., PIMENTEL, M. M., DEL REY SILVA, L. J. H. 1994. **Compartimentação tectônica da porção oriental da Província Tocantins**. In: SBG, Congresso Brasileiro de Geologia. Cambiriú. Boletim de Resultados Expandido. p. 215-217
- GREENLAND, S. *et al.* 2016. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, v. 31, n. 4, p. 337-350.
- KOCH, I. Analysis of multivariate and high-dimensional data. New York: Cambridge University Press, 2013. ISBN: 978-05-2188-793-9
- HAIR JR., J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. Análise Multivariada de Dados. 6ª ed. Porto Alegre: Bookman, 2009. ISBN: 978-85-7780-534-1
- HENNIG, C. *et al.* **Handbook of cluster analysis**. CRC Press, 2015.
- JOHNSON, R. A.; WICHERN, D. W. Applied Multivariate Statistical Analysis. 6ª ed. Pearson, 2018. 808 p. ISBN-13: 9780131877153.
- MACQUEEN, J. **Some methods for classification and analysis of multivariate observations**. In: FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Oakland, CA, USA, 1967. v. 1, n. 14, p. 281–297.
- MICHAUX, S.; O’CONNOR, L. **How to Set Up and Develop a Geometallurgical Program**. In: GTK Open File Work Report. GTK, 2020.
- MONFORT, M; CUESTAS, J. C.; ORDÓÑEZ, J. **Real convergence in Europe: A cluster analysis**. *Economic Modelling*. Elsevier, v. 33, p. 689-694, 2013.
- MOREIRA, G. C. Análise de agrupamento aplicada à definição de domínios de estimativa para a modelagem de recursos minerais. 2020. 101 f. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul. Porto Alegre, 2020.
- NAZARETH, A. F. D. V. Análise de agrupamentos aplicada à setorização geotécnica de minas a céu aberto. 2020. 170 f. Dissertação (Mestrado em Engenharia Mineral) - Escola de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2020.
- PRASSAS, E. S.; ROESS, R. P.; McSHANE, W. R. **Stochastic Modeling of Shared-Lane Left-Turn Process and Its Effect on Saturation Flow**. *Transportation Research Record*. SAGE Journals, v. 1555, n. 1, p. 105-113, 1996.
- PUNJ, G.; STEWART, D. **Cluster analysis in marketing research: Review and suggestions for application**. *Journal of Marketing Research*. v. 20, n. 2, p. 134-148, 1983.
- ROUSSEEUW, P. J. **Silhouettes: a graphical aid to the interpretation and validation of cluster analysis**. *Journal of Computational and Applied Mathematics*, Elsevier, v. 20, p.53–65, 1987.
- SIAL, A.N. *et al.* 2009. The São Francisco Palaeocontinent. *Developments in Precambrian Geology*, 16: 31-69.