



**Universidade de Brasília - UnB
Instituto de Geociências - IG
Programa de Pós-Graduação em Geologia**

***FOOTPRINT SUPERFICIAL DO DEPÓSITO AURÍFERO PEQUIZÃO,
GREENSTONE BELT CRIXÁS, GOIÁS, BRASIL***

ANDERSON MATIAS DOS SANTOS

DISSERTAÇÃO DE MESTRADO

Nº 484

Orientadora: Prof. Dra. Adalene Moreira Silva

**Brasília
2021**

ANDERSON MATIAS DOS SANTOS

***FOOTPRINT SUPERFICIAL DO DEPÓSITO AURÍFERO PEQUIZÃO,
GREENSTONE BELT CRIXÁS, GOIÁS, BRASIL***

Dissertação de mestrado apresentada junto ao curso de Pós-Graduação em Geologia (Área de concentração em Prospecção e Geologia Econômica), Instituto de Geociências da Universidade de Brasília, como requisito para obtenção do título de Mestre em Geologia.

Orientadora: Prof. Dra. Adalene Moreira Silva

Banca examinadora:

Prof. Adalene Moreira Silva (Presidente)
Prof. José Carlos Sícoli Seoane
(Universidade Federal do Rio de Janeiro)
Prof. Roberta Mary Vidotti
(Universidade de Brasília)

Brasília
2021

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

SS237f Santos, Anderson Matias dos
Footprint Superficial Do Depósito Aurífero Pequizão,
Greenstone Belt Crixás, Goiás, Brasil / Anderson Matias dos
Santos; orientador Adalene Moreira Silva. -- Brasília, 2021.
133 p.

Dissertação (Mestrado - Mestrado em Geologia) --
Universidade de Brasília, 2021.

1. Greenstone belt Crixás. 2. Depósito de ouro Pequizão.
3. Geoquímica de Solo. 4. Espectroscopia de Reflectância. 5.
Análise multivariada na exploração mineral. I. Silva,
Adalene Moreira, orient. II. Título.

"May the odds be ever in your favor."
— *Suzanne Collins*

AGRADECIMENTOS

A elaboração desta dissertação contou com a colaboração e amparo de muitos, para isso gostaria de agradecer imensamente:

- Aos meus pais e irmãs por todo o apoio e atenção durante todo meu caminho de estudos até aqui;
- A todos os meus professores e minhas orientadoras Adalene e Catarina que desde o início me apoiaram e estiveram colaborando para o meu desenvolvimento tanto acadêmico quanto pessoal;
- Ao geólogo e mestre Rodrigo da AngloGold Ashanti pelo apoio desde o início, concepção inicial do projeto, bem como pela coleta dos dados de geoquímica de solo e disponibilidade em todas as fases de desenvolvimento desta pesquisa;
- À Mineração Serra Grande pelo fomento através do Convênio de Cooperação e Apoio Técnico Científico entre a UnB (nº 23106.006453/2016-32 - Instituto De Geociências / UnB - Código 9961) sob a coordenação da Prof. Adalene.
- Ao Professor Diego Ducart do Departamento de Geologia e Recursos Naturais do IG-UNICAMP pelo apoio no processamento de dados espectrorradiométricos;
- Aos laboratórios da UnB e todos que me ajudaram nas análises, em especial a doutoranda Gabriella Fazzio e os técnicos do laboratório de DRX;
- Agradeço a doutoranda Luana M. Campos pela interação no processamento de dados de espectrometria de reflectância e a todos os membros do grupo de pesquisa - Geodinâmica de Terrenos Pré-Cambrianos e Sistemas Minerais Associados (SisMineral);
- Pela oportunidade de acesso a universidade pública e de qualidade a todos. E este trabalho contou com o apoio recebido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, e pela bolsa de pesquisa de mestrado nº 88882.347173/2019-01;
- A todos os colegas e amigos cientistas que conheci e tive a oportunidade de trocar conhecimentos geológicos e de vida durante todos esses anos em Brasília.
- E a Universidade de Brasília e todos os funcionários, técnicos e corpo docente.

RESUMO

Esta dissertação apresenta os resultados da caracterização do *footprint* superficial da mineralização aurífera do depósito Pequizão, localizado no *Greenstone Belt* Crixás, utilizando análise multivariada de dados geoquímicos e de espectrorradiometria de reflectância. Para tanto, foi efetuada uma coleta de 939 amostras em uma malha orientada de solo de aproximadamente 0,35 km² com perfis transversais de direção E-W. Os perfis possuem espaçamento de 25 m tanto entre os perfis quanto entre as estações de amostragem. Em cada estação foram coletadas três amostras que correspondem a três profundidades a saber: a partir de 45 cm do nível do solo coletou-se o nível A, e repetindo a mesma distância para os subsequentes níveis B e C, totalizando um perfil de 135 cm de profundidade. Os dados geoquímicos obtidos foram processados utilizando técnicas de estatística multivariada como análise de componentes principais e análise de fatores, assim como algoritmos de aprendizado em conjunto (*ensemble learning*) e agrupamento hierárquico (*clustering*). Foi utilizada análise geoquímica univariada como abordagem inicial, identificando que elementos *pathfinders* de Au aumentam a concentração com o aumento de profundidade. Com a análise de *clustering* conseguiu-se identificar que Ag, Hg, Sb, Te, e Tl estão associados com o Au, Com a técnica de *ensemble learning* pode-se mapear que o grupamento Zn, Al, Ga, e Pb também pode prever a ocorrência de amostras mineralizados a ouro na base de dados. Resultados obtidos a partir de dados de espectroradiométricos indicam abundância de caolinita, gibsita, anfibólio e clorita, com predomínio de mica branca no filito carbonoso e dolomito. A mica branca é um importante mineral para análise de alteração mineral por ser altamente resistente ao intemperismo e ocorre como muscovita nas amostras de solo analisadas, porém ocorre como muscovita fengítica (mais pobre em Al) nas amostras mineralizadas. A análise de dados de espectroscopia de reflectância mostra resultados robustos para detectar e mapear minerais indicativos de alteração hidrotermal como muscovita e a gibsita e que podem ser classificados como minerais-índice das estruturas mineralizadas Palmeiras e Estrutura IV, respectivamente. O *footprint mineralógico* definido em furos de sondagens por outros estudos também não foi preservado durante o intemperismo. A análise de componentes principais permitiu a identificação no PC1 dos elementos como a associação de Fe-Al-Cr-Ni-Cu, em contraste com K-Na-Ba-Tl. PC2 conseguiu definir os elementos de maior importância na predição de ouro, onde o PC3 destaca os elementos As-Ag-Te-Ca-Mg e gibsita associados diretamente à mineralização. Na análise de fatores, o Fator 1 prediz as unidades das rochas parentais, enquanto o Fator 3 caracterizou e delimitou as zonas da assinatura da mineralização como representativos da alteração hidrotermal em superfície. Comparando amostras de Pequizão com solo distal e não mineralizado, podemos inferir enriquecimento Sb, V, Na, Ba, e W e depleção de Zn, Ga e Pb, e que pode ser associado ao Fator 2. Com base nesses últimos elementos, é possível dizer que a predição de conteúdo de Au pelo *ensemble learning* se deve à ausência de ouro na presença de altas concentrações de Zn, Ga e Pb no Pequizão. De acordo com os resultados, sugere-se a aplicação de estudos pilotos para ampliar a aplicação na exploração em *brownfield* com base em métodos de solo, mais baratos, rápidos, dependentes de menor infraestrutura e de maior abrangência — até em locais de mais difícil acesso — do que sondagens. Para isso, é preciso uma expansão do conhecimento pedológico brasileiro e na relação entre espessura, clima e vegetação com a mudança e preservação de *footprints* de depósitos, além de mais estudos detalhados de espectrometria que possui rapidez na identificação e delimitação da expressão da alteração hidrotermal em superfície.

PALAVRAS CHAVE: Greenstone belt Crixás; Depósito de ouro Pequizão; Geoquímica de Solo; Espectroscopia de Reflectância; Análise multivariada na exploração mineral

ABSTRACT

This dissertation presents the results of the characterization of the surface footprint of the auriferous mineralization of the Pequizão deposit, located in the Crixás Greenstone Belt, using multivariate analysis of geochemical data and reflectance spectroscopy. Thus, there is a need to understand the extent, behavior, and preservation of the footprint after weathering. Therefore, a collection of 939 samples was carried out on an oriented soil grid of approximately 0.35 km² with cross-sectional profiles of E-W direction. The profiles are 25 m apart between each line and also between the sampling stations. At each station, three samples corresponding to three depths were collected: starting at 45 cm from ground level, level A was collected, repeating the same distance for subsequent levels B and C, resulting in a 135 cm-deep profile. The obtained geochemical data were processed using multivariate statistical techniques such as principal component analysis and factor analysis, as well as ensemble learning and clustering. Univariate geochemical analysis was used as an initial approach, identifying that Au pathfinder elements increase in concentration with depth. Clustering analysis identified that Ag, Hg, Sb, Te, and Tl are clustered with Au. With the ensemble learning technique, it can be mapped that the Zn, Al, Ga, and Pb clustering can also predict the occurrence of gold mineralized samples in the database. Results obtained from spectroradiometric data indicate an abundance of kaolinite, gibbsite, amphibole, and chlorite, with the predominance of white mica in the carbonaceous phyllite and dolomite. White mica is an important mineral for mineral alteration analysis because it is highly resistant to weathering and occurs as muscovite in the soil samples analyzed, but as phengitic muscovite (Al-poor) in the mineralized samples. Reflectance spectroscopy analysis proved to be a powerful technique in detecting and mapping minerals indicative of hydrothermal alteration, identifying gibbsite and muscovite and can be classified as index minerals of the Palmeiras and Structure IV mineralized structures, respectively. The mineralogical footprint defined in drill holes by other studies was also not preserved in the regolith. Principal component analysis allowed the identification in PC1 of the elements as the Fe-Al-Cr-Ni-Cu association, in contrast to K-Na-Ba-Tl. PC2 was able to define the elements of greatest importance in predicting gold, where PC3 highlights the elements As-Ag-Te-Ca-Mg and gibbsite associated directly with mineralization. Factor 1 predicted the parent rock units in factor analysis, while Factor 3 characterized and delineated the zones of the mineralization signature as representative of hydrothermal alteration at the surface. Comparing Pequizão samples with distal, unmineralized soil, we can infer Sb, V, Na, Ba, and W enrichment and Zn, Ga, and Pb depletion, thereby interpreting the result of Factor 2. Based on the latter elements, it is possible to say that the prediction of Au content by the learning ensemble is due to the absence of gold in the presence of high Zn, Ga, and Pb concentrations in Pequizão. According to the results, pilot studies are suggested to expand the application in brownfield exploration based on soil methods, which are cheaper, faster, less infrastructure-dependent, and more feasible - even in places of hard access - than drilling. This requires an expansion of Brazilian pedological knowledge and the relationship between thickness, climate, and vegetation with changes and preservation of the footprint from the deposit, as well as more detailed spectrometric studies and database to increase the ease and quickness in identifying and delimitation of the expression of hydrothermal alteration at the surface.

KEYWORDS: Pequizão gold deposit; Reflectance spectroscopy; Multivariate analysis; Crixás greenstone belt; Soil geochemistry.

LISTA DE FIGURAS

Figura 1.1 - Geologia regional simplificada da parte norte da Faixa Brasília com destaque para o Domínio Crixás-Goiás (modificado de Cordeiro e Oliveira, 2017; Borges et al., 2021b). O polígono em vermelho destaca a localização da Faixa Brasília no cenário brasileiro em (b) o domínio Crixás-Goiás.....3

Figura 1.2 – Representação esquemática dos diferentes tipos de assinaturas que um corpo mineralizado possui e seus respectivos limites próximos ao depósito (modificado de Leshner et al., 2017).....4

Figura 1.3 – Mapa simplificado de localização da cidade de Crixás-GO. A AngloGold Ashanti - Unidade Serra Grande, à sul da zona urbana, com toda sua área de produção em imagem RGB via satélite (Imagery – 2019). Rodovias principais em amarelo, escala 1:35.000. O polígono vermelho destaca a área de estudo.....5

Figura 1.4 – Scree-plot das análises de principais componentes em 30 variáveis (elementos e minerais) e o valor da variância encontrados na base de dados geoquímico e espectrométrico do depósito Pequizão.....11

Figura 1.5 – (a) Espectrorradiômetro FieldSpec3 HiRes® (ASD). (b) Preparação das amostras de solo para coleta de dados espectrais com uma quantidade representativa de cada. (c) Sacos de amostras de PEQSO0016B até PEQSO0019B, evidencia diferenças de coloração em amostras próximas e de mesmo nível.....16

Figura 1.6 – Stack plot com amostras de exemplo do espectro característico de alguns minerais individuais. No espectro do SWIR mostrando a principal região do comprimento de onda produzido pelas moléculas dos minerais nas amostras (modificado de Velasco et al., 2005).....18

Fig. 2.1. Simplified maps showing the location of the Crixás-Goiás Domain, Crixás Greenstone Belt, and the study area of the Pequizão deposit. (a) Simplified map of the main tectono-geological regions of Brazil, highlighting the Crixás-Goiás Domain and its greenstone belts, in gray. The red square is the location of (b) Regional map of the Crixás Greenstone Belt showing other active gold deposits (white stars) and emphasizing the Pequizão deposit (after Borges et al., 2021a); (c) Simplified tectonostratigraphic column of the Crixás Greenstone Belt, the star symbol represents the formation with a high gold content (after Jost et al., 2019); (d) Geological map (1:25,000) of the Pequizão deposit, elaborated by AngloGold Ashanti - Serra Grande Unit. Circles are the surface location of the underlying soil samples. Lines with triangles are thrust faults: Palmeiras (left) and Structure IV (right).....26

Fig. 2.2. Violin plots and sample distribution for selected elements separated into the metasedimentary unit (green) and metabasic unit (purple) with the levels A, B and C. Concentration of all elements are in Supplementary Data 2.....40

Fig. 2.3. Correlation matrix of selected elements from level A (313 samples) with significant correlations in this study. The complete correlation with all elements from B and C are in the Supplementary Data. Significance of 0.90.....42

- Fig. 2.4.** Dendrogram of soil sample using Euclidean-complete linkage, evidencing 4 clusters called C1 (in pink), C2 (in yellow), C3 (in green) and C4 (in blue) (pruning: 10 levels).....43
- Fig. 2.5.** (a) Predicted \times Observed values according to the used model. The thin black line is a 1:1 line, and the magenta line is the fitted linear regression with the gray area representing a 95% confidence interval. (b) Variable importance plot, with the numbers indicating the importance value of the respective transformed element.....45
- Fig. 2.6.** Stacked reflectance spectra of the samples containing iron oxides (860-940 nm), phyllosilicates and clay-minerals of the Pequizão deposit, including kaolinite, chlorite, amphibole, gibbsite and white-mica. All samples contain some iron oxides. The main absorption features are highlighted.....47
- Fig. 2.7.** Histograms of the minerals from the reflectance spectral results and showing Au presence; (a) White-mica histogram with 174 samples; (b) Fe-oxides histogram with 487 samples; (c) Amphibole histogram with 640 samples; (d) Chlorite histogram with 779 samples. Abbreviations (after Warr, 2021): act = actinolite; goe = goethite; hbl = hornblende; hm = hematite; ms = muscovite; phe* = tending to phengite.....52
- Fig. 2.8.** PC and FA loadings of centered log-ratio-transformed $n = 938$ data; (a) The biplot projection of PC1 versus PC2 shows loadings of the elements. Each dot corresponds to an individual sample, in factor-plane 1x2; (b) The biplot projection of PC1 versus PC3 and the factor-plane with mineralized samples ($Au > 0.2$ ppm) in red and barren samples ($Au < 0.2$ ppm) in blue.....55
- Fig. 2.9.** Map of the interpolated scores from factor 1 to 3 in the soil samples from the B level; (a) Factor 1 map, and the black dots are samples location; (b) Map of the four pedochemical domains based on the factor 1 interpolation; (c) Factor 2 map and thrust faults (left: Palmeiras, right: Structure IV); (d) Geological map and isolines from factor 2, highlighting only positive scores; (e) factor 3 map with mineralized soil samples in black dots ($Au > 0.2$ ppm); (f) Pequizão open pit with the highest factor 3 scores on satellite image also showing high-grade samples in white dots.....57
- Fig. 2.10.** Gold ranges in geochemical analysis from soil samples of level C from Pequizão deposit and the box-plot in log-ratio resulted in three different anomalies. Third-order anomalies in yellow, second-order anomalies in orange and first-order anomalies in red.59
- Fig. 2.11.** Integrated map with the IRS results and their localization in the geological map, using a scale according to the abundance from: (a) white mica; (b) chlorite; (c) amphibole; (d) gibbsite; (e) kaolinite; (f) hematite/goethite.....63
- Fig. 2.12.** Mineral classification at SWIR spectra. (a) kaolin group index and crystallinity plot for kaolinite crystallinity. (b) Map from the classification of the mineral and highlight the thrust-faults from the deposit.....64
- Fig. 2.13.** Tukey boxplot of enrichment factor in the Pequizão samples from Factor 2 comparing the Negative Factor Scores (NFS) and the Positive Factor Scores (PSF)

relative to the reference sample. The horizontal line at $EF = 1$ represents equivalence between the Pequizão and reference ratios. Elements with different X_{Peq}/X_{Ref} ratios between Pequizão and reference soil samples are labeled in bold, with significantly enriched elements in red, and significantly depleted elements in blue.....67

Fig. 2.14. (a) Geological map from Crixás greenstone belt highlighting the Pequizão deposit, scale 1:10.000; (b) Cross-section (A-B) from the deposit, according to previous studies in the area using drill cores, modified after Borges et al., 2021b. Abbreviations (after Warr, 2021): Gbs = gibbsite; Kln-wx = Well-crystalline kaolinite; Ms (Phe) = muscovite tending to phengite.....70

LISTA DE TABELAS

Tabela 1 - Estatísticas resumidas dos elementos das amostras de solo do depósito Pequizão.....8

Tabela 2 - Amostras selecionadas para análise de difração de raios-X (DRX) divididas em mineralizadas e estéreis, em cada nível e litologias. Resultados estão no anexo 2.....12

Tabela 3 - Hiperparâmetros e seu valor escolhido após o ajuste ($R^2 = 0,71$ e $RMSE = 0,50$).....15

Table 2.1. Statistical parameters of raw geochemical data of all soil samples from the Pequizão deposit. Min: minimum concentration; Max: maximum concentration; S. D.: Standard Deviation; Q25: 25th percentile of the data set; Median: 50th percentile of the data set and Q75: 75th percentile of the data set.....30

Table 2.2. Hyperparameters and its chosen value after tuning ($R^2 = 0.71$ and $RMSE = 0.50$).....36

Table 2.3. Main absorption bands according to the presence of ions or molecules in minerals (after Pontual et al., 1997).....38

Table 2.4. Parameters from infra-red spectroscopy.....46

Table 2.5. Mineralogical interpretation of infrared reflectance spectroscopy (IRS) and X-ray diffraction (XRD) was conducted on selected samples.....50

Table 2.6. Principal components and the loadings of elements in the first three principal components. Ordered according to the first component. Scores below ± 0.1 were excluded.....53

Table 2.7. Factor analysis results with 3 selected factors and their respective geochemical association. Abbreviations (after Warr, 2021): Kln = kaolinite; IrOx = iron oxides; Amp = amphibole; Wm = white mica; Gbs = gibbsite.....55

SUMÁRIO

AGRADECIMENTOS	I
RESUMO	II
ABSTRACT	III
LISTA DE FIGURAS	IV
LISTA DE TABELAS	VI
CAPÍTULO 1	
1. Introdução	1
1.1 Apresentação.....	1
1.2 Objetivos.....	4
1.3 Localização.....	5
2. Materiais e métodos	5
2.1 Dados Cartográficos.....	6
2.2 Amostragem	6
2.3 Dados geoquímicos.....	7
2.4 Análises geoquímicas.....	9
2.5 Difractometria de raios-X (DRX)	12
2.6 Sistema de combinações <i>Ensemble (Boosting)</i>	13
2.7 Espectrometria de reflectância.....	15
3. Estrutura da dissertação	18
CAPÍTULO 2	
Artigo científico: “New near-mine prospecting approach using multivariate analysis and reflectance spectroscopy to define surface footprint: a case study of the Pequizão Gold Deposit, Crixás Greenstone Belt, Central Brazil”.....	19
1 Introduction	20
2 Geological background	21
2.1 Gold mineralization.....	27
3 Materials and methods	29
3.1 Sampling strategy.....	29
3.2 Chemical analyses of soil samples and methodology.....	30
3.3 Statistical analyses.....	31
3.4 Ensemble Learning Framework.....	34
3.5 Reflectance spectroscopy and XDR analyses.....	37
4 Results	39
4.1 Descriptive statistics.....	39
4.2 Hierarchical cluster analysis.....	42
4.3 Ensemble learning results.....	44
4.4 Reflectance spectroscopy results.....	45
4.5 FA and PCA results.....	52
5 Discussion	58
6 Conclusions	71
Acknowledgments	72
References	72
CAPÍTULO 3	
3. Discussões e Conclusões	84
CAPÍTULO 4	
4. Referências Bibliográficas	88
ANEXOS	

1. Introdução

1.1 Apresentação

Os greenstone belts encerram registros importantes da história evolutiva do planeta no arqueano e paleoproterozóico e possuem vocação metalogenética para hospedar depósitos de ouro e metais base. Eles são constituídos de sequências metavulcanosedimentares com evolução complexa normalmente marcada por vários eventos termotectônicos e magmáticos superpostos (Anhaeusser, 2014; Kusky et al., 2021). Depósitos auríferos relacionados a esses ambientes são importantes alvos exploratórios em cinturões pré-cambrianos no mundo e sua investigação é essencial para o entendimento da gênese desses depósitos (Groves et al., 2020; Goldfarb, 2021).

A região de estudo se situa dentro do denominado Terreno Arqueano-Paleoproterozoico Crixás-Goiás (Jost et al., 2014; Borges et al., 2017; Campos et al., 2017) ou Domínio Crixás-Goiás (Cordeiro and Oliveira, 2017; Borges et al., 2021a,b). Dentro deste trabalho, adotaremos a nomenclatura proposta por Cordeiro e Oliveira (2017) e adotada por Borges et al (2021a,b). O Domínio Crixás-Goiás consiste em uma associação de seis complexos TTG arqueanos (Uvá e Caiçara, ao sul, e Anta, Caiamar, Moquém e Hidrolina, ao norte) e cinco *greenstone belts* com idades arqueanas e paleoproterozóicas (Faina e Serra de Santa Rita, ao sul, e Crixás, Pilar de Goiás e Guarinos, ao norte). Os *greenstone belts* ocorrem como sequências metavulcanosedimentares estreitas situadas entre os complexos TTG. Os cinco greenstones belts são compostos de metakomatiitos e metabasaltos basais sobrepostos por sucessões metassedimentares arqueanas a paleoproterozóicas que hospedam a mineralização aurífera (Jost et al., 2014; Jost et al., 2019) (Figura 1.1).

No *greenstone belt* Crixás, alvo desta pesquisa, a sequência supracrustal possui na base uma unidade metavulcânica arqueana representada por de metakomatiitos (Formação Córrego Alagadinho) e metabasaltos (Formação Rio Vermelho) metamorfizados sob condições de fácies xisto verde e anfibolito inferior (Arndt et al., 1989; Jost & Oliveira, 1991; Fortes et al., 2003) sotopostos a uma sequência de rochas metassedimentares paleoproterozóicas (Tassinari et al., 2006; Jost et al., 2008; Borges et al., 2021a). A seção metassedimentar de Crixás compreende a Formação Ribeirão das Antas, caracterizada por uma sequência de filitos carbonosos intercalados com lentes de dolomitos (Jost et al., 2014). No topo, afloram metagravaucaas da Formação Córrego Geral (Jost et al., 2010, 2019).

A mineralização aurífera ocorre como lentes de sulfeto maciço controlados estruturalmente, veios de quartzo e vários corpos minerais disseminados hospedados em falhas de empurrão de baixo ângulo dentro da sucessão metassedimentar (Jost et al., 2010, 2014). Diversos estudos têm sido desenvolvidos na região enfocando os aspectos metalogenéticos, como a pesquisa pioneira de Jost & Fortes (2001) ou trabalhos como o de Tassinari et al. (2006) e Marques et al. (2013) com foco na datação da mineralização. Recentemente, Campos (2019) apresentou um estudo inovador sobre a assinatura do *footprint* primário das mineralizações de ouro do *Greenstone Belt* Crixás utilizando petrografia, técnicas espectrais para mapeamento mineral, geoquímica e propriedades físicas das rochas, concentrando-se nas estruturas Palmeiras, Estrutura III e Estrutura IV que controlam os principais corpos de minério conhecidos. Com este estudo foi possível caracterizar as zonas de alteração hidrotermal associadas às estruturas mineralizantes do ponto de vista químico, mineralógico e petrofísico.

Por outro lado, dados gamaespectrométricos de alta resolução apontam anomalias que indicam altos valores de radioelementos (K, Th, U) que mapeiam o ambiente mineralizado da região do depósito Pequizão. A mineralização está hospedada nas rochas metassedimentares da formação Ribeirão das Antas e é controlada pela Estrutura IV. Afloram filitos carbonosos e metagrauvas cortadas por veios de quartzo em contato com dolomitos impuros. A mineralização é do tipo disseminada e está orientada preferencialmente na direção N297/10°-25°. A assembleia mineral é composta por arsenopirita + pirita + pirrotita ± calcopirita, bem como ouro livre em venulações de quartzo (Jost & Fortes, 2001).

O estudo do *footprint* primário efetuado por Campos (2019) e a anomalia gamaespectrométrica motivou a realização de um estudo inédito sobre o *footprint* superficial das mineralizações auríferas dentro do *Greenstone Belt* Crixás. Para isso, foi efetuada coleta sistemática de amostras de solo em três profundidades diferentes que recobrem a área do depósito para se verificar a variação com a profundidade. Frequentemente, o horizonte B é o amostrado em trabalhos de prospecção mineral, mas algumas contribuições mais recentes mostram a importância de estudos pilotos locais com amostragem em níveis fixos e que entender a variação espacial do regolito em geral, se *in situ* ou não, representa uma nova fronteira no entendimento do *footprint* de sistemas minerais (i.e., UNCOVER Project, 2012).

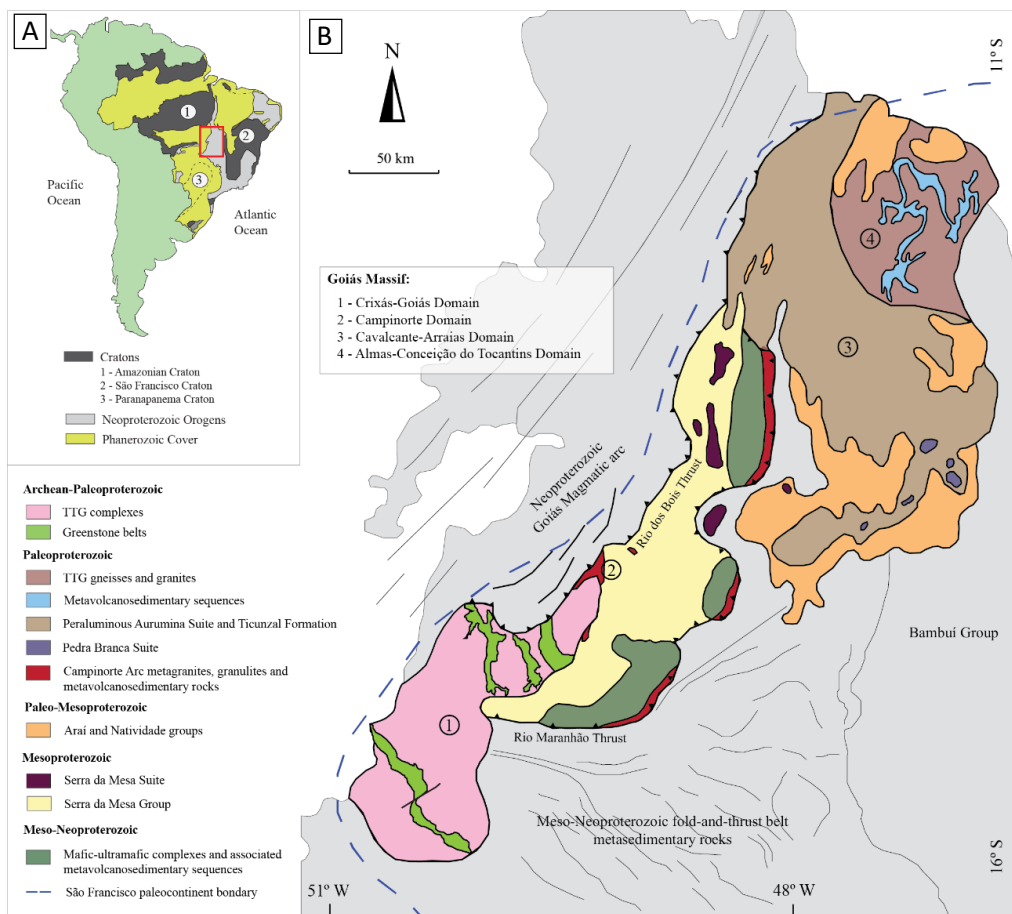


Figura 1.1 - Geologia regional simplificada da parte norte da Faixa Brasília com destaque para o Domínio Crixás-Goiás (modificado de Cordeiro e Oliveira, 2017; Borges et al., 2021b). O polígono em vermelho destaca a localização da Faixa Brasília no cenário brasileiro em (b) o domínio Crixás-Goiás.

As pesquisas prospectivas de ouro vêm evoluindo conforme os anos e aprimorando e integrando cada vez mais métodos distintos que auxiliam a definir vetores e zonas de alteração próximas ao corpo mineralizado (Okada, 2020; Bacha et al., 2021; Gisbert et al., 2021). Podemos definir vetores para identificação e estudo de *footprints* litogeoquímicos produzidos pelo sistema hidrotermal mineralizante ou por remobilizações próximas (Figura 1.2) e ao entorno de depósitos de minério (e.g., Ames et al., 2016; Gisbert et al., 2021; Medina et al., 2021). Possuem potencial para detectar uma possível proximidade com um depósito além de fornecer informações sobre suas principais características, normalmente baseados na litologia, mineralogia, geoquímica e química mineral (Cooke et al., 2017; Hollis et al., 2021).



Figura 1.2 – Representação esquemática dos diferentes tipos de assinaturas que um corpo mineralizado possui e seus respectivos limites próximos ao depósito (modificado de Lesher et al., 2017).

A espectroscopia de refletância permite uma caracterização mineral econômica e rápida, contribuindo com os desafios da exploração mineral e da mineração. Utilizada comumente entre trabalhos que utilizam furos de sondagem por possuir resposta espectral mais nítida (Herrmann et al., 2001; Ducart et al., 2006; Harraden et al., 2013; Prado et al., 2015; Almeida, 2021; Campos et al., 2022) Definir minerais de alterações em zonas mineralizadas utilizando espectroscopia de reflectância vem sendo comum em projetos de prospecção minerais na Austrália (e.g., Tappert et al. 2013; Lampinen et al., 2019; Laukamp et al., 2021), que assim como o Brasil possuem uma cobertura de solo bem desenvolvida. Os espectros de refletância medidos de materiais geológicos, como rochas e solos, contêm assinaturas espectrais ou *footprints* de seus minerais constituintes e outros componentes e propriedades, das quais podem ser derivadas informações, incluindo abundância, química, espécies minerais e cristalinidade.

A meta central desta pesquisa é, portanto, a definição do *footprint* superficial da mineralização de ouro do *Greenstone Belt* Crixás através do mapeamento do regolito que recobre o depósito de Pequizão. Para tanto, efetuou-se uma análise multivariada de dados geoquímicos e de espectrorradiometria de reflectância visando a caracterização das zonas mineralizadas em superfície e delimitação de novos vetores de exploração na região de estudo.

1.2 Objetivos

O objetivo central é caracterizar o *footprint* superficial da mineralização de ouro do *greenstone belt* Crixás através do mapeamento do regolito que recobre o depósito de

Pequição através da análise multivariada de dados geoquímicos de solo e de espectrorradiometria de reflectância.

Os objetivos específicos compreendem:

- Descrição e coleta de dados espectrais em amostras de solo para reconhecer os minerais que podem estar associados as zonas de alteração hidrotermal em superfície;
- Processamento e interpretação de dados espectrorradiométricos das amostras de solo do depósito efetuadas em três níveis (ponto a ponto);
- Análises de difratometria de raios-X (DRX) com intuito de auxiliar a identificação mineralógica e apoiar a interpretação dos dados;
- Análise multivariada de dados de geoquímica de solo que recobrem o depósito, buscando compreender associações e a mobilidade dos elementos no sistema hidrotermal mineralizante nas porções mais superficiais;
- Integrar os resultados obtidos a partir dos dados geoquímicos e espectrais para caracterização do *footprint* superficial e delinear vetores exploratórios.

1.3 Localização

A área de estudo está localizada na região Centro-Oeste do Brasil, município de Crixás, no noroeste do estado de Goiás. As instalações da Unidade Serra Grande, pertencente a empresa AngloGold Ashanti, se inserem na parte sul da cidade, com acesso pela rodovia GO-336, à 330 km de distância de Goiânia e 390 km de Brasília (Figura 1.2).

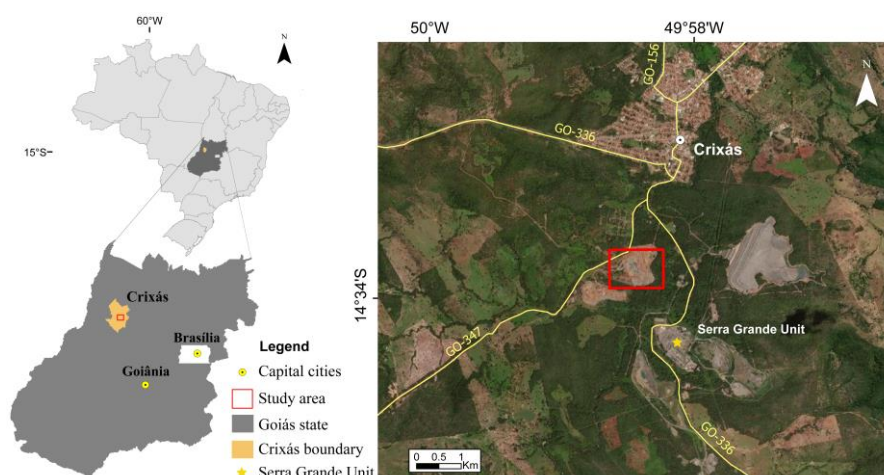


Figura 1.3 – Mapa simplificado de localização da cidade de Crixás-GO. A AngloGold Ashanti - Unidade Serra Grande, à sul da zona urbana, com toda sua área de produção em imagem RGB via satélite (Imagery – 2019). Rodovias principais em amarelo, escala 1:35.000. O polígono vermelho destaca a área de estudo.

2. Materiais e métodos

Os materiais utilizados compreendem dados cartográficos e geoquímicos de solo de alta resolução espacial. Adicionalmente, foram coletados dados espectrais em solo. O detalhamento dos materiais e do conjunto de métodos são apresentados abaixo.

2.1 Dados Cartográficos

As bases cartográficas utilizadas dentro da pesquisa foram elaboradas pela AngloGold Ashanti - Unidade Serra Grande (USG), no início de 2019 e anteriormente a abertura da cava do Pequizão. O mapa geológico na escala de 1:25.000 foi empregado dentro do trabalho e manteve-se a descrição dos litotipos e abreviações utilizadas pela empresa. Ou seja, as rochas metabásicas são descritas como quartzo-clorita-carbonato-sericita xisto (CBCX), anfibólio xisto (AMP), metabasalto (MBA), e a porção de rochas metassedimentares inclui o filito carbonoso (GXN) e o dolomito (DOL). A região foi selecionada como alvo pois: (a) ocorre a *up-plunge* do corpo; (b) a projeção desse *plunge* corrobora com anomalias de potássio impresso na gamaespectrometria realizada pela empresa.

2.2 Amostragem

As amostras de solo foram cedidas pela Unidade Serra Grande dentro do Convênio de Cooperação e Apoio Técnico Científico com a Universidade de Brasília (nº 23106.006453/2016-32 - (Instituto De Geociências) / UnB - Código 9961) sob a coordenação da Professora Adalene Moreira Silva. A coleta foi realizada durante o final do segundo semestre de 2018 e início de 2019. A malha orientada de solo possui uma área de aproximadamente 0,35 km² (500 por 700 m) com as estações e as linhas de amostragem espaçadas em 25 m. Atentou-se coletar, em superfície, amostras próximas ao corpo mineralizado.

Em cada local amostrado foram coletados 3 níveis de profundidades diferentes. Em cada estação de amostragem foram coletadas 3 amostras que correspondem às 3 profundidades, onde a partir de 45 cm, coletou-se no nível A, depois mais 45 cm de profundidade para a coleta no nível B, e mais 45 cm para a coleta das amostras nível C,

totalizando um perfil de 1,35 m de profundidade e com 0,8 a 1 quilograma cada amostra. Nesse contexto, há 313 pontos de amostragem distribuídos em 11 linhas de solo, totalizando assim 938 amostras de solo, todas armazenadas no Instituto de Geociências da Universidade de Brasília. Em alguns resultados e para organização do banco de dados, resolvemos manter a nomenclatura proposta pela empresa, onde PEQ é a abreviação do Corpo Pequizaão, SO de solo, seguido do número da amostra e seu nível correspondente (A, B ou C), por exemplo, PEQSO0123A. Segundo a classificação do IBGE (2015) o solo na região do distrito aurífero de Crixás possui duas características pedogenéticas principais: (1) PVD1: argissolo vermelho distrófico de textura média/argilosa com horizonte A moderado e relevo ondulado; (2) PVAd25: argissolo vermelho-amarelo distrófico Tb (argila de atividade baixa) de textura médio/argiloso com horizonte A moderado de fase pedregosa I (com calhaus no perfil até profundidade maior que 40 cm) também com relevo ondulado.

2.3 Dados geoquímicos

As dosagens geoquímicas das 938 amostras de solo foram obtidas na ALS Laboratory para análise de concentrações de elementos maiores (%), menores e traços (ppm) por ICP-MS (espectrometria de massa por plasma acoplado indutivamente) e fornecidas para este estudo. Os lotes de amostras foram preparados com os originais e amostras em branco para controle de qualidade adicionadas na proporção de 20:1. A primeira amostra de cada lote é uma amostra branca que permite avaliar se houve contaminação no manuseio e/ou durante o preparo do material, e a qualidade dos reagentes. O solo foi seco a $<60^{\circ}\text{C}$ para evitar a perda de mercúrio e peneirado até 80 mesh. A análise do ouro foi feita por Fire Assay e espectrofotometria de absorção atômica (*atomic absorption spectrophotometry* – AAS). As amostras foram digeridas com quatro ácidos e posteriormente analisadas por ICP-MS. Foram usados neste trabalho 35 elementos: Ag, Al, As, Au, Ba, Bi, Ca, Co, Cr, Cu, Fe, Ga, Hf, Hg, K, Li, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Rb, Sb, Sr, Te, Th, Ti, Tl, V, W, Zn e Zr. Outros 14 elementos também foram analisados, mas censurados - Be, Cd, Ce, Cs, Ge, In, La, Re, S, Sc, Se, Sn, Ta, U e Y – pois mostram concentrações fora do limite de detecção e, portanto, não confiável (Tabela 1). O processamento dos dados geoquímicos do solo tem como foco a análise e definição de anomalias utilizando os softwares Geosoft Oasis Montaj e ArcGIS - ArcMap. O primeiro processamento de interpolação dos dados geoquímicos de solo da

área de estudo tem como foco a análise e definição de anomalias unielementares utilizando o software Geosoft Oasis Montaj, onde todos os mapas elementares selecionados estão no Anexo 1.

Tabela 1 – Estatísticas resumidas dos elementos das amostras de solo do depósito Pequizeão.

Element (ppm)	Variation coeff. (%)	Mean	Minimum	1st Quartile	Median	3rd Quartile	2nd Order Threshold (Q3+1,5[Q3-Q1])	1st Order Threshold (Q3+3[Q3-Q1])	Maximum
Ti	39.87	8274	1030	5520	8985	10700	18470	26240	16400
Al	21.17	84707	36600	73900	89000	97200	132150	167100	125000
Fe	41.91	94960	12100	51900	110250	125500	235900	346300	177000
Mg	145	1739	400	800	1200	1700	3050	4400	40800
Ca	121	823	50	300	500	1000	2050	3100	6500
Na	248	440	50	50	100	500	1175	1850	9200
K	150	4798	50	200	300	10100	24950	39800	31900
Ag	278	0.04	0.01	0.02	0.03	0.04	0.07	0.10	3.05
As	201	33.60	2.600	7.0	8.900	26.40	55.50	84.60	531
Au	438	0.132	0.005	0.016	0.021	0.039	0.074	0.1080	7.470
Ba	152	590	10	10	30	1260	3135	5010	4880
Bi	98.08	0.1126	0.005	0.04	0.06	0.15	0.3150	0.48	0.94
Co	106	43.16	1.100	20.40	36.05	53.60	103	153	706
Cr	61.03	523	43	219	542	735	1509	2283	1780
Cu	31.76	118	19.10	94.10	127	145	220	296	230
Ga	19.18	19.77	9.290	17.55	20.20	22.20	29.17	36.15	32.60
Hf	105	1.287	0.200	0.300	0.500	2.30	5.3	8.300	6.80
Hg	42.34	0.024	0.003	0.017	0.022	0.029	0.047	0.065	0.074
Li	37.59	14.63	4.70	11	13.35	16.70	25.25	33.80	57.70
Mn	63.43	819	54	528	789	1020	1758	2496	5710
Mo	65.95	0.53	0.06	0.31	0.41	0.62	1.085	1.550	3.460
Nb	32.83	5.507	0.90	4.30	5.30	6.50	9.8	13.10	15.40
Ni	59.77	140	11.80	79.70	144	178	325	473	751
P	42.66	313	40	240	280	360	540	720	1790
Pb	117	11.24	1.100	3.200	4.400	17.10	37.95	58.80	142
Rb	140	18.30	0.100	0.900	1.700	40.40	99.65	159.00	104
Sb	42.06	0.7219	0.080	0.500	0.740	0.890	1.475	2.06	1.920
Sr	165	12.03	0.70	1.90	3.70	16.70	38.90	61.10	143
Te	124	0.08	0.03	0.03	0.05	0.09	0.1875	0.285	0.960
Th	122	2.78	0.26	0.70	0.98	4.98	11.40	17.82	43.90
Tl	134	0.15	0.01	0.02	0.03	0.30	0.720	1.140	1.030
V	43.30	260	36	154	299	340	619	898	522
W	71.82	2.747	0.30	1.50	2.40	3.40	6.250	9.100	23.90
Y	38.48	17.33	5.600	12.40	16.45	20.60	32.90	45.20	66.20
Zn	46.58	59.62	22	42	52	68	107	146	265
Zr	119	42.90	2.800	7.900	11.90	84.70	200	315	248

2.4 Análises geoquímicas

De acordo com Carranza (2009), uma vez que os dados geoquímicos brutos são geralmente distorcidos, ou seja, não obedecem uma distribuição gaussiana ou normal. O pré-processamento de dados geoquímicos mostram que as concentrações dos elementos mostram usualmente uma distribuição de densidade empírica assimétrica e deve ser transformada antes de fazer qualquer análise adicional do processamento dos dados. A transformação logarítmica é um método útil para melhorar a simetria da distribuição empírica da densidade das concentrações uni-elementares, porém, após a sua aplicação várias variáveis ainda apresentam uma relação de dependência, por isso, apenas a transformação logarítma não consegue resultados efetivos para geologia. A Análise de Dados de Composição (*Compositional Data Analysis - CoDA*) é uma abordagem baseada em transformações de razão logarítmica, usada para evitar o efeito de "fechamento" e não restringir os dados (Aitchison, 1986; Filzmoser et al., 2018). Entre as diferentes transformações de razão logarítmica e devido às suas propriedades isométricas e isomórficas, dentro deste estudo optou-se pela transformação logarítmica-centrada (*centered-log transformation - clr*) (Aitchison, 1986). Segundo Filzmoser et al. (2018), a transformação logarítmica-centrada (clr) é a única transformação que projeta o espaço amostral dos dados para o espaço real sem perder suas dimensões. Além disso, o clr preserva a distância e a estrutura, uma condição que se ajusta bem ao estudo geológico deste trabalho.

Diversas técnicas de agrupamento vêm sendo utilizadas. Onde os mais comuns são: “ligação simples” (ou *nearest neighbor*); “ligação completa” (ou *farthest neighbor*); e “variância mínima” (*minimum variance clustering* ou *Ward’s method of sum-of-squares method*). O agrupamento, ou *cluster*, por ligação completa foi escolhido, e funciona de maneira similar ao que o agrupamento de ligação simples, com uma exceção importante: não são consideradas as menores distâncias, mas as maiores. Então, o critério para a fusão de clusters é:

$$d(C_i \cup C_j, C_k) = \max\{d(C_i, C_k), d(C_j, C_k)\} \quad Eq. 1$$

O algoritmo de ligação completo tende a produzir um dendrograma balanceado. A distância máxima das observações de um *cluster* para qualquer outra observação de outro *cluster* é considerada, onde o mínimo ao longo dessas distâncias é selecionado para

o critério de ligação completa (Filzmoser et al., 2018). No dendrograma estão dispostas linhas ligadas segundo os níveis de similaridade que agrupam pares de variáveis. Esta forma de agrupamento hierárquico é uma simplificação em duas dimensões de uma relação n-dimensional e pode ocorrer algumas distorções quanto à similaridade entre as variáveis.

Os dados de análise geoquímica são conhecidos por terem uma alta dimensionalidade, ou seja, grande número de variáveis (elementos/óxidos). Quando existem muitas dimensões há problemas ao tentar entender a variabilidade dos dados e observar padrões, pois diagramas bidimensionais (2D) ou tridimensionais (3D) podem apenas demonstrar subconjuntos limitados dos dados, portanto, apenas tendências e estruturas desses subconjuntos (Jolliffe e Cadima, 2016; Varmuza e Filzmoser, 2009). Portanto, a fim de reduzir a dimensionalidade do conjunto de dados, usamos a análise de componentes principais (APC).

A ACP é um método estatístico não supervisionado que usa transformações lineares (por exemplo, centralização e rotação) para a matriz de dados original com um grande número de variáveis em um novo conjunto com menor número de variáveis, chamados de componentes principais, que não são correlacionados e ordenados de forma que as primeiras componentes principais (CP) retenham o máximo possível de informações das variáveis originais (Jolliffe, 2002). Para cada CP há um "autovetor" (*eigenvector*) de 1 unidade de comprimento que é composto de proporções de cada uma das variáveis originais e denominada de pontos de carregamento ou loading score. Para cada autovetor tem um "autovalor" (*eigenvalue*) que representa a soma das distâncias quadradas dos dados que apontam para a origem após a centralização dos dados. Para cada valor de autovetor é possível classificar os autovetores em ordem decrescente para que aqueles com altos autovalores representem uma maior variabilidade dos dados (Abdi e Williams, 2010). Resumindo, é o cálculo dos autovalores e correspondentes autovetores de uma matriz de coeficientes de correlação entre variáveis ou variâncias-covariâncias, que são idênticas pós transformação logarítmica. Consiste em uma transformação linear de 'x' variáveis em novas variáveis, onde a primeira variável se torna responsável pela maior variação dentro do conjunto de dados, assim por diante.

Podemos representar graficamente a porcentagem que cada CP tem na variação total em um "scree-plot" para determinar o número mínimo de principais componentes que representa a variância das respectivas componentes. O "scree test" (Cattell, 1966)

consiste em analisar a representação gráfica dos autovalores do conjunto de fatores gerados em busca de quebras na curva e considerar apenas o número de fatores antes da "quebra", desconsiderando os fatores posicionados em a parte mais plana da curva (Figura 1.3). Uma vantagem do gráfico de ACP para os 2D e 3D, como por exemplos gráficos de sílica por óxidos, é que as informações de todas as variáveis são integradas e, como os CPs são ortogonais entre si, também podem ser facilmente visualizados em um gráfico 2D ou 3D (Jolliffe e Cadima, 2016). Uma vez que o ACP tende a maximizar a variabilidade, os pontos de dados (amostras) que são semelhantes tendem a se agrupar, enquanto se separam dos pontos com uma variância mais alta, tornando o ACP uma técnica comum para *clustering* (e.g., Xue et al., 2011). Utilizamos o software IBM SPSS® Statística v.26 para calcular e plotar a análise de componentes principais e seus parâmetros estatísticos relacionados.

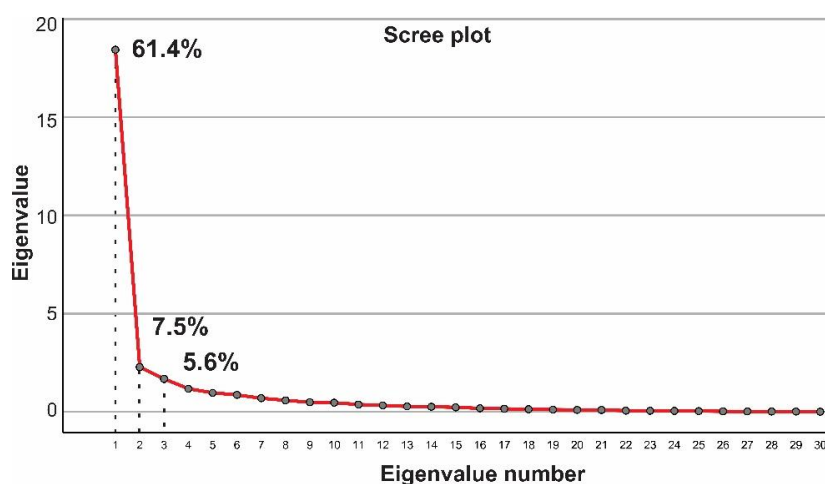


Figura 1.4 – Scree-plot das análises de principais componentes em 30 variáveis (elementos e minerais) e o valor da variância encontrados na base de dados geoquímico e espectrométrico do depósito Pequizão.

A análise fatorial (AF) é uma técnica estatística usada na análise multivariada de dados e pode identificar fatores de comportamento comuns entre as variáveis em termos de sua correlação (Reimann et al., 2002). Na análise fatorial as relações existentes dentro de um conjunto de “z” variáveis seja o reflexo das correlações de cada uma dessas variáveis com “n” fatores, não correlacionáveis entre si. Assim, em estudos geoquímicos é possível encontrar associações químicas elementares em representantes de diferentes litotipos parentais. Fatores são entidades estatísticas e podem ser visualizados graficamente como um eixo de classificação ao longo do qual as variáveis podem ser plotadas. A intensidade da correlação entre variáveis e um fator é dada por sua coordenada

ao eixo. As projeções traçadas de cada amostra de solo foram baseadas na análise fatorial. O método das componentes principais foi escolhido como método de extração dos fatores de cada elemento químico neste trabalho, pois foi o único método que apresentou resultados muito consistentes e coesos com os dados geológicos.

É uma abordagem que está intimamente relacionada ao ACP, de acordo com Rollinson e Pease (2021), e também busca minimizar a variação nos dados multivariados para o mínimo de "fatores" possível. Ele difere da ACP porque o número de eixos não é igual ao número de variáveis (como na ACP). Em vez disso, AF define um pequeno número de "fatores" que explicam a maior proporção dos dados. AF é realizada usando a matriz de correlação padronizada, ponderando todas as variáveis igualmente, e convertendo vetores de componentes principais em vários "fatores". É amplamente utilizado na sedimentologia (Hofer et al., 2013), nas ciências ambientais (Rezaei et al., 2017) e geologia econômica (Zhao et al., 2017).

2.5 Difractometria de raios-X (DRX)

Com o objetivo de comparar e verificar a qualidade da análise espectral, 12 amostras (Tabela 2) representativas da área foram analisadas por difração de raios-X (XRD) utilizando, na Universidade de Brasília, o instrumento RIGAKU - D / MAX - 2A / C com radiação de Cu ($k = 1,5406 \text{ \AA}$), gerador de tensão e corrente definido para 35 kV e 15 mA.

Tabela 2 - Amostras selecionadas para análise de difração de raios-X (DRX) divididas em mineralizadas e estéreis, em cada nível e litologias. Todos resultados são apresentados no anexo 2.

	A	B	C
i) Mineralizadas	PEQSO036	PEQSO09B	PEQSO289C
(Au: 0.2 – 7 ppm)	Filito carbonoso	Filito carbonoso	Filito carbonoso
	PEQSO050	PEQSO90B	PEQSO125C
	Anfibólio xisto	Qz-chl-cb-ser Xisto	Anfibólio xisto
ii) Estéreis	PEQSO011	PEQSO132B	PEQSO028C
(Au: < 0.2 ppm)	Filito carbonoso	Filito carbonoso	Filito carbonoso
	PEQSO111	PEQSO229B	PEQSO280C
	Qz-chl-cb-ser Xisto	Metabasalto	Metabasalto

Abreviações (Warr, 2021): cb = carbonato; chl = clorita; ser = sericita; qz = quartzo.

2.6 Sistema de combinações *Ensemble (Boosting)*

Ensemble learning, ou aprendizado em conjunto, é um termo para métodos que utilizam e combinam modelo base (algoritmos de aprendizagem de máquina) para que cada base de aprendizado compense os erros uns dos outros, resultando em um conjunto (*ensembles*) com melhor desempenho geral (Opitz & Maclin, 1999; Rokach, 2010). Esses algoritmos base podem ser uma infinidade de algoritmos como, por exemplo, *árvores de decisão* e *redes neurais*. Dentro da análise, podem ser utilizados dados de mesma natureza, resultando em conjuntos homogêneos, ou de diferentes tipos, formando conjuntos heterogêneos (Zhou, 2012). Como o conjunto é uma coleção de algoritmos bases pequenos e simples, eles também podem ser chamados de modelos simples (*weak learners*). Um dos principais métodos de *ensemble* é o *boosting*, um método que dá mais importância aos dados classificados incorretamente ao seccionar o conjunto de dados para modelo simples subsequentes, com o objetivo de aumentar a precisão da previsão do modelo (Schapire, 1990, 2003), enquanto o *bootstrap aggregating (bagging)* é um conjunto com particionamento aleatório de dados que visa reduzir a variância do modelo (Breiman, 1996). Uma árvore de decisão (Loh, 2014) é um dos tipos mais comuns de *weak learner* e a base de métodos predominantes e poderosos, como *Random Forest* (Breiman, 2001), *Adaptive Boosting (AdaBoost)* (Freund & Schapire, 1997), *Gradient Boosting Machine (GBM)* (Friedman, 2001), *Extreme Gradient Boosting (XGBoost)* (Chen & Guestrin, 2016) e *Light Gradient Boosting Machine (LightGBM)* (Ke et al., 2017).

Como estado da arte, o XGBoost foi empregado para realizar uma modelagem de regressão no conjunto de dados geoquímicos desta pesquisa com o objetivo de prever o teor de Au nas amostras e, em seguida, extrair as informações de quais características (os elementos geoquímicos) foram usadas e as suas devidas importâncias dadas para o modelo, visando assim, compreender melhor o padrão geral associado ao conteúdo de ouro. O XGBoost foi implementado em R (versão 4.1.0; R Core Team, 2021) usando o *framework* do pacote *tidymodels* (v. 0.1.3; Kuhn e Wickham, 2020) e o pacote XGboost (v. 1.4.1.1; Chen et al., 2021). O conjunto de dados geoquímicos usado foram os CLR-transformados, que foi posteriormente dividido em conjuntos (ou *sets*) de treinamento e teste (nas proporções 0,75:0,25) e "estratificados" usando a variável Au, ou seja, separar em populações relativamente homogêneas de acordo com certa variável para garantir uma distribuição uniforme de valores altos e baixos entre estes conjuntos. Para evitar

overfitting, utilizamos a validação cruzada de 10 vezes no conjunto de treinamento, "estratificada" com Au e repetindo a validação cruzada 5 vezes. Embora a repetição consuma muito tempo computacional, ela pode beneficiar o modelo, pois terá mais conjuntos para o treinamento. A fórmula usada foi Au como resultado e os outros elementos — Al, Ca, Fe, K, Mg, Na, Ti, Ag, As, Ba, Co, Cr, Cu, Ga, Hg, Mn, Ni, P, Pb, Sb, Te, Tl e Zn — como preditores.

Como o XGBoost é um método com hiperparâmetros (ou seja, um parâmetro cuja configuração é externa ao modelo e deve ser inserida pelo usuário), podemos otimizar os valores de entrada usando os dados existentes para identificar as configurações do parâmetro do modelo que geram desempenho preditivo mais realista, um processo denominado "ajuste" (ou *tuning*). O algoritmo foi ajustado por meio do pacote *finetune* (v. 0.1.0; Kuhn, 2021) usando *simulated annealing* (Bohachevsky et al., 1986; Kirkpatrick et al., 1983), um modelo de pesquisa global que emprega uma rotina de otimização baseada em gradiente que pode reavaliar soluções anteriores para encontrar a melhor rota de ajuste de hiperparâmetros, sendo utilizado 50 iterações em todo o conjunto validado (*cross-validated set*). Os ajustes estimados foram avaliados de acordo com as métricas do coeficiente de determinação R^2 e da raiz do erro quadrático médio (RMSE). O R^2 quantifica a determinação de uma variável dependente (variável de resultado) pela variável independente (variável preditora) como proporção da variância, demonstrando ser uma métrica preferida para análise de regressão (Chicco et al., 2021). Outra forma de avaliar um modelo é considerar a distância da linha de regressão aos pontos de dados (ou seja, o erro), como o erro quadrático médio (MSE), com RSME sendo usado para padronizar as unidades de medidas. A faixa de R^2 vai de 1, um ajuste perfeito, a 0, um ajuste trivial e "aleatório", enquanto o RMSE vai de 0, um ajuste perfeito, a valores de ajuste positivos infinitamente crescentes indicando um modelo ruim. O conjunto de parâmetros ajustados mais bem avaliado é escolhido com base no valor mais alto de R^2 e no valor mais baixo de RMSE. Os hiperparâmetros ajustados do XGBoost estão listados na tabela 3, com os parâmetros selecionados para o modelo.

Em seguida, o modelo base é usado para modelar o conjunto de treinamento e, em seguida, ajustado ao conjunto de teste, onde o modelo é testado por um modelo de regressão linear do conteúdo de Au verdadeiro *versus* a estimativa de Au prevista e avaliada com R^2 . Os valores de importância do recurso para o modelo são extraídos usando o pacote *vip* (v. 0.3.2; Greenwell e Boehmke, 2020) para entender quais variáveis foram consideradas mais importantes para o modelo prever Au.

Tabela 3 - Hiperparâmetros e seu valor escolhido após o ajuste ($R^2 = 0,71$ e $RMSE = 0,50$).

Hiperparâmetros	Descrição	Valores de intervalo	Valores selecionados
trees	Número de árvores de decisão utilizadas no conjunto.	500 – 1000	973
min_child_weight	Soma mínima dos pesos de observação necessários para dividir um nó da árvore de decisão. Quanto maior, mais conservador é o modelo. Varia de 0 ao infinito positivo.	2 – 40	40
max_depth	Tamanho máximo de uma árvore (ou seja, número de divisões). Valores altos tornam os modelos mais complexos sujeitos a <i>overfitting</i> . Varia de 0 a infinito positivo.	3 – 12	7
eta	Taxa de aprendizagem. Determina a rapidez com que o modelo se ajusta ao erro residual utilizando modelos fracos adicionais. Reduz os pesos dos recursos para evitar ajustes excessivos. Varia de 0 a 1.	10^{-5} – 10^{-1}	6.998×10^{-2}
gamma	Redução de perda mínima (erro) necessária para criar uma nova divisão de árvore. Quanto maior, mais conservador. Varia de 0 a infinito positivo.	10^{-10} – 10^{-1}	1.89×10^{-4}
early stopping	Número de iterações sem melhorias antes que o algoritmo pare.	10 – 50	40

2.7 Espectrometria de reflectância

A espectrorradiometria de reflectância é uma técnica efetiva, rápida e não destrutiva para a identificação mineralógica de depósitos minerais, utilizada nas diferentes etapas da exploração mineral (Gladwell et al., 1983; Cudahy and Ramanaidou, 1997; Ducart et al., 2006). A análise espectral dos depósitos permite o melhor entendimento das zonas mineralizadas, auxiliando no entendimento da distribuição espacial do minério.

Para esta dissertação, foram coletados dados espectrais dos níveis A, B e C utilizando o Espectrorradiômetro FieldSpec3 HiRes® (ASD) dentro do Laboratório de Geofísica Aplicada (IG-UnB) (Figura 1.4).

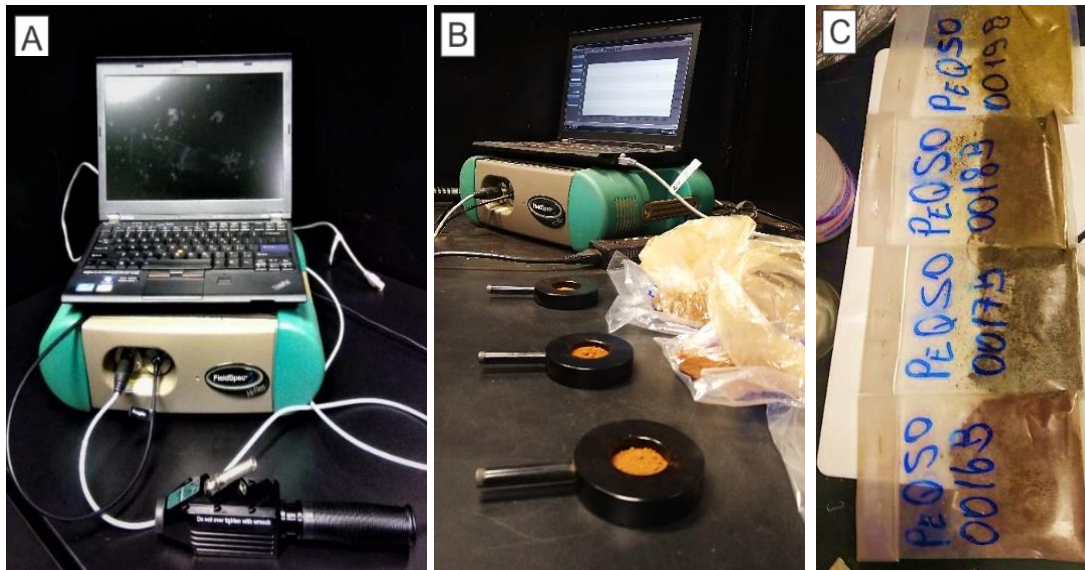


Figura 1.5 – (a) Espectrorradiômetro FieldSpec3 HiRes® (ASD). (b) Preparação das amostras de solo para coleta de dados espectrais com uma quantidade representativa de cada. (c) Sacos de amostras de PEQSO0016B até PEQSO0019B, evidencia diferenças de coloração em amostras próximas e de mesmo nível.

Para medir a distribuição espectral de potência (DEP) de radiação emitida pelas amostras de solo foi utilizado um espectrômetro de reflectância pontual e uma sonda de contato para a coleta dos dados. A sonda de contato tem iluminação interna, e a coleta do dado é feita com a sonda em contato direto com a amostra. O dispositivo coleta os espectros nas regiões do visível ao infravermelho próximo (*visible and near-infrared*, VNIR: 350-1000 nm e, *shortwave-infrared*, SWIR: 1000-2500 nm).

O intervalo de coleta do espectrômetro é de 1,4 nm no VNIR e 2,0 nm do SWIR, com precisão de comprimento de onda de $\pm 0,5$ nm. A resolução espectral é de 3 nm no VNIR e 8 nm no SWIR (www.asdi.com). Os minerais que interagem com a luz absorvem preferencialmente certos comprimentos de onda e transmitem em outros comprimentos de onda especificamente, onde a profundidade, a forma e a posição dessas absorções são controladas por sua estrutura e composição química (Van der Meer, 2004) (Figura 1.5). Depois de analisar as principais características dos dados espectrais definimos parâmetros espectrais como abundância, cristalinidade e também composição.

Os parâmetros utilizados para identificação mineralógica seguem a metodologia proposta por padrões das bibliotecas espectrais de Stonntag et al. (2012) e Haest et al. (2012): 900D (óxidos de Fe), 900Wv1 (diferenciar a hematita de goethita), 2200Dk (caulinita), 2200SL (cristalinidade da caulinita), 2200D (abundância de mica branca),

2200Wvl (composição da mica branca), 2260D (gibbsite), 2345D (clorita), 2345Wvl (composição das cloritas), 2390D (anfíbólios), 2390Wvl (composição dos anfíbólios).

Segundo Pontual et al. (2008b), usar dados espectrais de solo envolve considerações similares para programas geoquímicos envolvendo amostras de solo. A consideração principal em usar as amostras de solo é que essas precisam ser representativas do assoalho rochoso subjacente e, recomenda-se, preferencialmente, coletá-las sempre num mesmo horizonte ou nível, além de observar se é material residual (i.e., não transportado). São potencialmente úteis em particular para análises espectrais em projetos em que possuem expressiva quantidade de solo (com no mínimo 100 amostras) e dado geoquímico, e em locais onde o perfil de solo sobre as rochas seja raso e pouco desenvolvido. As vantagens são: (a) o mapeamento litológico e alterações; (b) podem ser úteis em interpretar associações geoquímicas dentro do banco de dados; (c) alto custo-benefício e método rápido de produzir mapas de minerais e alterações. Por outro lado, as desvantagens são: (a) o dado espectral pode apresentar ruídos e assinaturas de minerais mais tênues que amostras de outros tipos (rocha ou fragmentos); (b) material biogênico pode dificultar a análise; (c) a cristalinidade dos minerais e possivelmente a composição podem mudar em comparação com a rocha-mãe, mesmo em solo espesso.

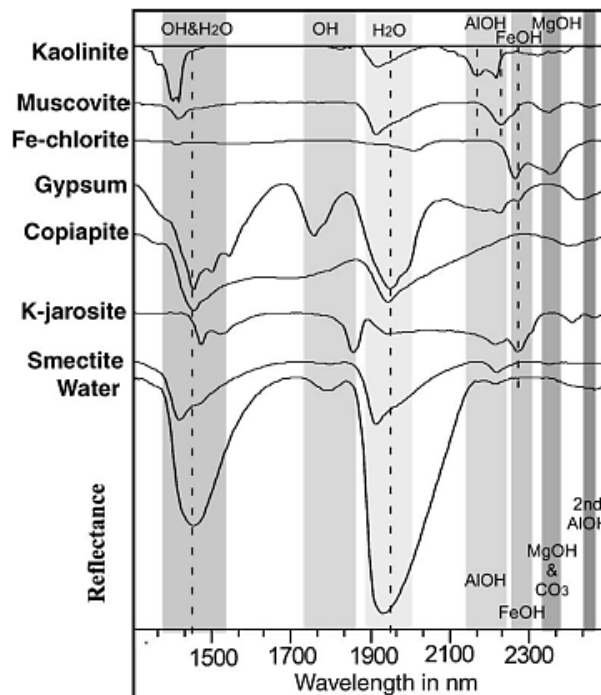


Figura 1.6 – Stack plot com amostras de exemplo do espectro característico de alguns minerais individuais. No espectro do SWIR mostrando a principal região do comprimento de onda produzido pelas moléculas dos minerais nas amostras (modificado de Velasco et al., 2005).

O processamento e tratamento dos dados foi efetuado utilizando-se o software *The Spectral Geologist* (TSG Core™ v.8.0.7, CSIRO Earth Science and Resource Engineering - CESRE - Sydney, Austrália). Os espectros foram comparados com os espectros de bibliotecas espectrais de referência do GMEX (Pontual et al., 2008a, 2008b) e USGS (Clark et al., 2007). A identificação do mineral foi feita com base nas características de absorção diagnóstica nos espectros, como largura, profundidade e posição do comprimento de onda. O TSG fornece ferramentas para visualização e quantificação da mineralogia por parâmetros pré-estabelecidos.

3. Estrutura da dissertação

A dissertação de mestrado foi elaborada em forma de artigo intitulado “*New brownfield prospecting approach using multivariate analysis and reflectance spectroscopy to define surface footprint: a case study of the Pequizão Gold Deposit, Crixás Greenstone Belt, Central Brazil*” a ser submetido para uma revista indexada após a defesa. Seccionada em três partes, a primeira resume toda introdução e detalha a metodologia utilizada durante o projeto. A segunda parte consiste no artigo completo com todos os resultados e discussões com o objetivo de definir as assinaturas mineralógicas e geoquímicas do depósito aurífero. A última parte conclui a dissertação com uma discussão e considerações finais para tentar otimizar a prospecção de ouro em terrenos férteis. O trabalho foi desenvolvido no Instituto de Geociências da Universidade de Brasília, sob a orientação da Professora Dra. Adalene Moreira Silva, como parte do Convênio de Cooperação e Apoio Técnico Científico entre a UnB e a Mineração Serra Grande (nº 23106.006453/2016-32 - (Instituto De Geociências) / UnB - Código 9961) sob a coordenação da Professora Adalene Moreira Silva (“Crixás Project”) e dentro do Grupo de Pesquisa CNPq - Geodinâmica de Terrenos Pré-Cambrianos e Sistemas Minerais Associados (SisMineral) e liderados pelas Professoras Adalene Moreira Silva e Catarina L.B. Toledo.

New near-mine prospecting approach using multivariate analysis and reflectance spectroscopy to define surface footprint: a case study of the Pequizão Gold Deposit, Crixás Greenstone Belt, Central Brazil

Anderson Matias dos Santos ^{a,*}, Adalene Moreira Silva ^a, Catarina Lauboré Bemfica Toledo ^a, Rodrigo dos Reis Salles ^b, Diego Fernando Ducart ^c, Luana Machado Campos ^a, Caio Cesar Aguiar Borges ^a, Marcelo Pereira Campos ^d, Bruno Araújo dos Santos ^d

a. Instituto de Geociências, Universidade de Brasília, 70910-900, Brasília, DF, Brazil

b. AngloGold Ashanti Ltd., Brownfields exploration – CTOP, 34000-000, Nova Lima, MG, Brazil

c. Instituto de Geociências, Universidade Estadual de Campinas, 13083-855, Campinas, SP, Brazil

d. AngloGold Ashanti Ltd., Rodovia GO-336, 97000 Km 97, 76510-000, Crixás, GO, Brazil

* Corresponding author. E-mail address: and.matias.santos@gmail.com (A. M. Santos).

ABSTRACT

The Pequizão gold deposit, located in the Crixás greenstone belt, is structurally-controlled mineralization hosted mainly in carbonaceous phyllite with pervasive hydrothermalized zones and large amounts of disseminated sulfides and gold. Usually, regoliths of Au deposits are either unstudied or understudied in exploration surveys because of lower Au content and the difficulty and complexity of soil profiles. However, such investigation can be beneficial in brownfield exploration to assess larger areas than with drilling. In this work, results of clustering and machine learning of soil geochemistry, and reflectance spectroscopy integrated with a multivariate approach to determine soil footprints and to target new deposits in the surface of the Crixás greenstone belt. Reflectance spectroscopy was applied in 939 soil samples and is a valuable tool in mineral exploration for fast investigation of mineral assemblage from the target. It was calibrated with X-ray diffraction, geochemical, and multivariate approaches providing consistent vectors toward mineralization, with Pequizão soil samples developing gibbsite and phengitic-white mica as minerals related to mineralized samples. Multivariate analysis reveals that the deposit has a typical orogenic gold deposit chemical signature. Principal components analysis and factor analysis firstly defined samples derived from carbonaceous phyllite and dolomite as the main ore hosts. The chemical aspects of hydrothermal alteration are As, Ag, Te, Ca, and Mg, with enrichment of Sb, V, Na, Ba, and W and depletion of Zn, Ga, and Pb that, according to ensemble learning, they have significant importance in the detection of gold. The understanding of surface footprints of known deposits can be used as an exploratory guide for finding new soil geochemical halos related to mineralization zones. The research revealed new associations of minerals and chemical elements that can be determined as exploration vectors. It was possible to further clarify the knowledge about the footprint of mineralization in soil. The combination of the soil-related methods applied in this study together with a broad-coverage spectral approach in soil and drill cores can enhance success in prospecting in brownfield areas and expand to greenfields.

Keywords: Gold; Soil sampling; Reflectance spectroscopy; Exploration vectoring; Multivariate geochemical analysis.

1. Introduction

Prospective gold research has been evolving over the years and increasingly improving, integrating distinct methods that help to define alteration vectors and zones near ore bodies (Okada, 2020; Bacha et al., 2021; Gisbert et al., 2021). These vectors can be defined for the identification and study of footprints produced by mineralizing hydrothermal systems or by remobilizations near and around ore deposits (e.g., Ames et al., 2016; Gisbert et al., 2021; Medina et al., 2021). They have the potential to detect proximity of a mineralization in addition to providing information about its main characteristics, usually based on rock type, mineral and chemical compositions (Cooke et al., 2017; Hollis et al., 2021). A recent effort has been made to study the footprints of the orogenic gold mineral system in Brazil (Campos, 2019; Almeida, 2021; Bacha et al., 2021), but there is still a lack of studies detailing the exploration vectors based on multi-elemental soil sampling in mineralized fields. This stems from the fact that in Brazil, generally, the main pieces of information about an ore body are not easily available for research, especially above a deposit before mining operations. In comparison, the definition of hydrothermal- and weathering-related alteration minerals of mineralized zones has been a relatively common practice in mineral exploration projects in Australia (e.g., Anand et al., 2007; Tappert et al. 2013; Lampinen et al., 2019; Laukamp et al., 2021).

The auriferous region of the Crixás greenstone belt occurs within the Crixás-Goiás Domain of the Goiás Massif in Brazil, with thrust faults controlling the mineralization. Of those faults, the most important are Structure III and IV, which have the highest ore potential (Jost et al., 2010, 2014), and the main ore bodies are the Palmeiras, Ingá, Baru, IV, V, XI, Sucupira, Forquilha, Pequizão, and Urucum deposits. Campos (2019) presented an innovative study on the primary footprint signature of gold mineralizations in the Crixás greenstone belt using petrography, and spectral techniques for mineral mapping in the mineralized thrust-faults in the area and successfully determined the ore footprints.

Contributing to solving the challenges of mineral exploration, reflectance spectroscopy is less destructive, more cost-effective method for a rapid mineral characterization in comparison to petrography, X-ray diffraction, Raman spectroscopy and electron microprobe analyzer, while still being a consistent and repeatable method and not requiring dedicated sample preparation and laboratory. In the last decade, field-based spectrometers studies exclusively of drill cores and fresh rocks analyzed spectral data to define hydrothermal footprints and to characterize the mineralization in ore deposits (e.g., [Prado et al., 2016](#); [Pereira, 2017](#); [Lampinen et al., 2019](#); [Simpson & Christie, 2019](#); [Medina et al., 2021](#)). Henceforth, the usage of field-based spectrometers in soil samples can be successfully employed for a quick and effective surface mineral exploration ([Velacos et al., 2005](#)), as attested by the common usage for general soil properties studies or agricultural sciences (e.g., [Brown et al., 2006](#); [Christy, 2008](#)).

Widely used for the statistical processing for mineral exploration, multivariate data analysis, such as clustering, principal component and factor analysis, is an efficient method for the visualization and exploration of geochemical data, and it is commonly used to reduce the high dimensionality of variables, i.e. analyzed elements ([Reimann et al., 2008](#); [Carranza, 2010](#); [Zuo, 2011](#); [Sadeghi et al., 2015](#)). Geochemical multivariate analysis can be used to interpret bedrocks, soil types, and even to define metallogenic signatures and distributions (e.g., [Lin et al., 2014](#); [Sadeghi et al., 2015](#)). Just as the distribution of the clustering results should be assessed against the known properties of the research area, for example, based on the knowledge of the regional geology, the factor analysis applied to geochemical data sets will result in subsets showing distinct geochemical behaviors ([Templ et al., 2008](#)).

This study aims to understand and outline the spatial area, mineralogical and chemical composition, and influence of weathering on the hydrothermal overprint of the Pequizão gold deposit, in the Crixás greenstone belt (Central Brazil). We used geochemical and reflectance spectroscopy data of a high-resolution soil sampling, enabling the characterization of the mineral and chemical assinature of the mineralization footprint based on surface hydrothermal alteration zones and delineation of new exploration vectors in the region. This is

a pilot study within the Crixás goldfield, delineating new exploration vectors of brownfield areas based on a near-mine case study, while simultaneously displaying the importance and ease of detailed soil studies for both exploration and exploitation of an ore body.

2. Geological background

The Tocantins Province was formed during the collision of three cratons — Amazonian, São Francisco-Congo, and Paranapanema — and is defined as a Neoproterozoic orogenic system known as Brasiliano-Pan African that resulted in the amalgamation of the western Gondwana supercontinent. This province is located in central Brazil and is divided into the Araguaia, Paraguay and Brasília fold belts (Almeida et al., 1981; Pimentel et al., 2000; Fuck et al., 2014). The Brasília Belt extends for ca. 1000 km in the N–S direction and comprises: (i) a craton-verging fold-and-thrust belt in the east, (ii) folded metasedimentary sequences, exposed crustal block, and granite-gneiss terrains in the north, (iii) the Anápolis-Itauçu granulites and the metasedimentary rocks of the Araxá Group in the west, and (iv) the orthogneisses of the Goiás Magmatic Arc (930–800 Ma) and the Archean-Paleoproterozoic granite-greenstone terrain of the Goiás Massif in the south-west (Fuck et al., 2014). Cordeiro and Oliveira (2017) sectioned the Goiás Massif into four tectonic domains: Campinorte, Cavalcante-Arraias, Almas-Conceição do Tocantins, and Crixás-Goiás domains. It is postulated that the domains of the Goiás Massif were amalgamated in the Paleoproterozoic during successive orogenic cycles and accretion of continental and intraoceanic arcs, forming the western margin of the São Francisco-Congo paleocontinent (Martins-Ferreira et al., 2020; Borges et al., 2021a, 2021b).

The Crixás-Goiás Domain is located in the central area of the Brasília Belt (Fig. 2.1a) and it is an Archean-Paleoproterozoic block composed of tonalite-trondhjemite-granodiorite rocks (TTG) complexes enveloped by greenstone belts (Jost et al., 2010). Previously interpreted as an allochthonous microcontinent amalgamated to the western border of the Brasília Belt in the late Neoproterozoic

(Pimentel et al., 2000, 2011), recent works about the Goiás Massif regional geology concluded that this region, and the other Paleoproterozoic domains of the massif, represented a continuous landmass of the São Francisco paleocontinent that was amalgamated in the late Paleoproterozoic (Cordeiro and Oliveira, 2017). The Crixás-Goiás Domain was built up from six blocks of TTG complexes: the northern Hidrolina, Moquém, Caiamar and Anta, and the southern Uvá and Caiçara complexes. Queiroz et al. (2008) established two stages of granitic plutonism, represented firstly by earlier tonalitic to granodioritic with minor granitic gneisses (ca. 2.84–2.78 Ga) and secondly by later granodioritic to granitic gneisses (ca. 2.71–2.70 Ga). The Hidrolina Complex comprises mostly granodiorite and minor tonalitic gneisses. The Anta Complex consists of an association of granodiorite, tonalities, and minor granites grouped in the Chapada Granite Suite. The Caiamar Complex is subdivided, from south to north, into granodiorite gneisses, tonalities, and tonalitic gneisses. The Moquém Complex consists of granitic, granodioritic and tonalitic gneisses (Jost et al., 1994a; Jost et al., 1994b; Lacerda Filho et al. 1999). The Uvá Complex, in the southernmost part of the Crixás-Goiás Domain, is composed of two major domains of orthogneisses: The Rio do Índio Gneisse — tonalitic, dioritic and granodioritic, and the Rio Vermelho Gneisse — with intrusions of monzogranite and tonalities (Jost et al., 2005, 2013). At last, the Caiçara Complex was divided by Beghelli Junior (2012) into four distinct portions of orthogneisses, granodiorites, monzogranites, and charnockites. In addition to the TTGs, the Crixás-Goiás Domain includes five greenstone belts, the southwestern Faina and Serra de Santa Rita belts, and the northern Guarinos, Pilar de Goiás, and Crixás Greenstone belts.

The Crixás Greenstone Belt is an N–S elongated belt about 6 km wide and 45 km long (Fig. 2.1b), stratigraphically composed of — from its base to the top — the Córrego Alagadinho, Rio Vermelho, Ribeirão das Antas, and Córrego Geral formations (Fig. 2.1c) (Saboia et al., 1981; Jost and Oliveira, 1991; Jost et al., 2019; Borges et al., 2021a), and limited to the west by the Anta Complex, to the east and southeast by the Caiamar Complex, and to the north by the Neoproterozoic metasedimentary sequence of the Santa Terezinha (Mara Rosa

Arc). The Córrego Alagadinho Formation has an average thickness of 500–900 m and comprises mainly metakomatiites roughly dated at ca. 2.82 Ga (whole-rock Sm-Nd data) and 2.72 Ga (whole-rock Pb-Pb data) (Arndt et al., 1989), and ca. 3.0 Ga (whole-rock Sm-Nd; Fortes et al., 2003). The typical internal structure of the individual flows is given by cumulatic spinifex-textured layers, and the protoliths of the Al-depleted metakomatiites (Arndt et al., 1989; Kuyumjian and Jost, 2006; Borges et al., 2021b) were metamorphized into rocks composed mostly of Mg-rich chlorite, serpentine, tremolite, chromite, magnetite, and local carbonate (Jost et al., 2019; Borges et al., 2021b). The Rio Vermelho Formation comprises a ~300 m-thick sequence of tholeiitic metabasalts and mafic schists. Locally, there are preserved pillow structures, amygdales and vesicles, and intercalated of metachert, banded iron formation (BIF), and manganese layers. The metabasalts are composed of Mg-hornblende, plagioclase, carbonate, and trace amounts of chlorite, actinolite, epidote, and ilmenite. The mafic schists in this formation are basically highly-foliated rocks composed of amphiboles, albite, oligoclase, titanite, chlorite, quartz, and some pyrite and magnetite (Sabóia and Teixeira, 1980; Jost et al., 2019).

Metasedimentary rocks occur in the central and northwestern sectors and are grouped in the Ribeirão das Antas and Córrego Geral formations and host most of the known gold ore bodies in the Crixás belt. The Ribeirão das Antas Formation (ca. 500 m thick) consists of carbonaceous phyllite with quartz, non-crystalline carbonaceous matter, muscovite, biotite, trace amounts of plagioclase, chlorite, garnet, and disseminated pyrite/pyrrhotite, and (Fe-)dolomite lenses (6 m-15 m thick) with rare oolitic textures preserved. The rhythmic layering of black shales across a transition interval indicates deposition in a shallow water environment (Theodoro and Jost, 1996; Jost et al., 2010). The phyllite is dark-gray or black, laminated, well foliated, very fine-grained, and the lepidoblastic texture is well defined by the orientation in one or more directions of muscovite and biotite crystals, usually with crenulated foliation. The groundmass is composed of opaque and non-crystalline organic matter, indicating metamorphism under low to medium greenschist-facies conditions, as opposed to the metabasalts containing hornblende and plagioclase that underwent

metamorphism with a peak at amphibolite-facies condition (Borges et al., 2021b). This succession overrides the metavolcanic rocks above a nonconformity obliterated by thrust faults and an overturning of the stratigraphy (Jost et al., 2014). The Córrego Geral Formation rocks are metagraywackes composed of coarse-grained clasts of quartz albite and rare rock fragments embedded in a fine-grained matrix formed of muscovite, biotite, chlorite, quartz, and oligoclase/andesine (Jost et al., 2019; Borges et al., 2021a; Ferreira et al., 2021).

The structural configuration of the set units shows a complex framework developed after multiphase deformational events forming thin-skinned thrust faulting, tight-isoclinal folding, and duplex structures (Jost et al., 2019). Low-angle thrust faults and/or strike-slip faults, and ductile-brittle shear zones control the deposits hosted in the metasedimentary units. Usually, these structures are located parallel to the lithological contacts (Jost et al., 2014). Gold mineralization and the geometry of the ore shoots are associated with thrust faults and shear zones. The Au-related hydrothermalism postdates the peak metamorphism since it overprints the greenschist facies metamorphic assemblages (Jost et al., 2001; Bogossian et al., 2021).

The geomorphological aspects in the Crixás region are conditioned mainly by the lithology. The elevations in the mafic-ultramafic rocks vary between 500-600 m, while in the metasedimentary rocks it varies around 400 m with some higher elevations being supported by quartz-rich rocks. Soils originated from the metamafic-ultramafic rocks is reddish to yellow, with a clayey to silty aspect, often magnetic, while the soil derived from the metasedimentary rocks varies from quartz to clayey, with a yellowish to reddish color, and sometimes dark gray (possibly derived from carbonaceous phyllites). The climate is tropical semi-humid with annual rainfall between 1,750 and 1,850 mm (Sobiesiak, 2011).

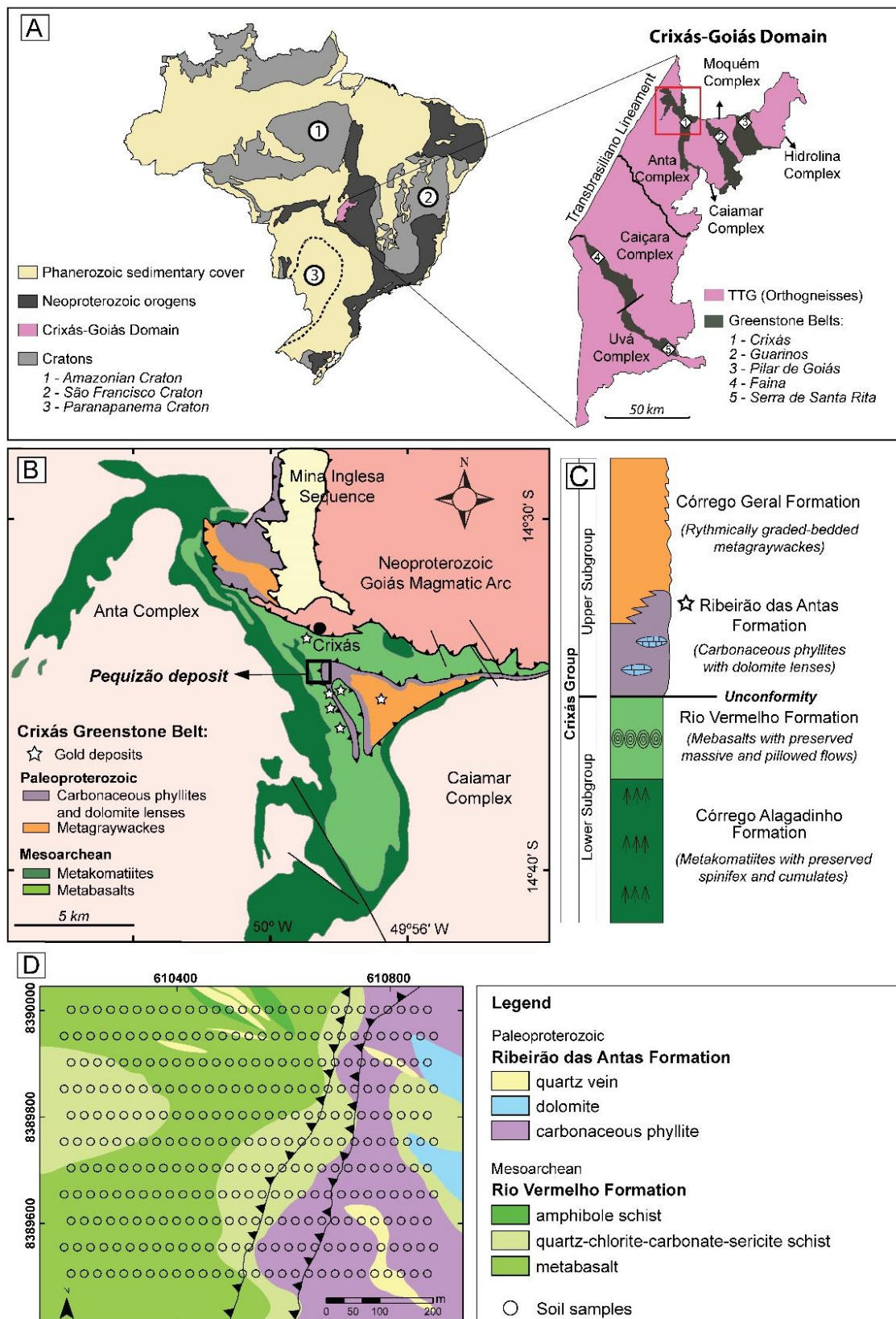


Fig. 2.1. Simplified maps showing the location of the Crixás-Goiás Domain, Crixás Greenstone Belt, and the study area of the Pequizão deposit. (a) Simplified map of the main tectono-geological regions of Brazil, highlighting the Crixás-Goiás Domain and its greenstone belts, in gray. The red square is the location of (b) Regional map of the Crixás Greenstone Belt showing other active

gold deposits (white stars) and emphasizing the Pequizão deposit (after [Borges et al., 2021a](#)); (c) Simplified tectonostratigraphic column of the Crixás Greenstone Belt, the star symbol represents the formation with a high gold content (after [Jost et al., 2019](#)); (d) Geological map (1: 25,000) of the Pequizão deposit, elaborated by AngloGold Ashanti - Serra Grande Unit. Circles are the surface location of the underlying soil samples. Lines with triangles are thrust faults: Palmeiras (left) and Structure IV (right).

2.1. Gold mineralization

[Jost and Fortes \(2001\)](#) characterized the Mina III ore body (~65 tons Au), as an example to describe the mineralization style in the Crixás goldfield (<3.7 Moz Au). The ore body was classified as an auriferous massive sulfide deposit with preeminent arsenopyrite and pyrrhotite. This deposit comprises three different styles: (i) stacked massive sulfide near BIFs, (ii) auriferous quartz veins, and (iii) disseminated sulfide minerals in carbonaceous phyllites. In addition to the Mina III ore body, there are other smaller occurrences that are also structurally controlled by regional shear zones (thrust/strike-slip faults).

The ore bodies can be divided according to the main controlling structure: Palmeiras Structure (Palmeiras, Palmeiras Norte and Baru bodies), Upper Zone (several sub-bodies), Structure III (Mine III, Mina Nova, Ingá, and Body XI), and Structure IV (Forquilha, Body IV, Body V, Sucupira, and Pequizão bodies). They differ from each other in relation to Au content, size, host rocks, ore minerals, and hydrothermal alteration assemblages. The hydrothermal alteration in the ore zone is, generally, carbonatization and sulfidation within strike-slip faults, including potassic, propylitic, and sericitic alteration zones ([Jost and Fortes, 2001](#); [Petersen Junior, 2003](#); [Jost et al., 2010, 2014](#)).

The Structure IV thrust fault has lower quartz veining occurrence compared with nearby faults and its mineralization is disseminated arsenopyrite, pyrrhotite, and chalcopyrite, with gold associated with silicification. The mineralization occurs in the carbonaceous phyllites of the Ribeirão das Antas Formation and consists of gold-rich portions 1.5 m thick and 200 m long on

average, extending along the dip for depths greater than 1 km (Fortes, 1996; Petersen Junior, 2003).

The deposition of the protoliths of the Crixás metasedimentary rocks during the Rhyacian lasted until it underwent progressive compression, detachment and the first gold mineralization. This first mineralization was dated at 2.165 Ga by Tassinari et al. (2006) from U-Pb data of zircon grains from a silicified metagraywacke interpreted to be of hydrothermal origin, while a Re-Os dating method was used by Marques et al. (2013) and Borges et al. (2021b) to date arsenopyrite from sulfide lenses at 2.12 Ga and 2.13 Ga, respectively, interpreted as the approximate timing of the gold mineralization.

Pequizão ore body is strongly controlled by Structure IV, with an intense brittle deformation, and the percolation of hydrothermal fluids generated distinct hydrothermal zones. The underground ore is represented by two main rock types from the Ribeirão das Antas Formation: carbonaceous phyllite and metagraywacke rocks ca. 60–90 m thick. In the mineralized rocks, there is an abundance of disseminated gold-rich arsenopyrite (marking the fault zone) and a minor amount of chalcopyrite and pyrrhotite, but also with gold associated with quartz-rich carbonate veins (preferential orientation is N297/10°-25°). The capping barren rock is quartz-chlorite-sericite-garnet schist, and at the base, there is an alternation of impure dolomite with carbonaceous phyllite and quartz/carbonate veins. Two stages of different hydrothermal alterations occur in the Pequizão: (i) an early stage of silicification, potassic and phyllic alterations with transitional contacts, and (ii) a late stage of chloritic and carbonatic alterations superimposing the earlier stage (Sobiesiak, 2011).

3 Materials and methods

3.1. Sampling strategy

Pequizão soil samples were acquired by the AngloGold Ashanti company - Serra Grande Unit in cooperation with the University of Brasília to obtain information on the surface footprint of gold mineralization. A pilot study was

carried out on the Pequizão deposit in 2019, before the opening of the open pit. The soil geochemistry survey included coverage of 11 EW, 700 m long profiles spaced 500 m, with a sampling spacing of 25 meters (Fig. 2.1d). About 0.8 to 1 kg were gathered at three constant intervals along the traverse lines. The sampling strategy for the study addressed three different soil levels named A, B, and C, which are 45, 90, and 135 cm below the surface, respectively, standardizing the depth of each sample, thus disregarding true pedological horizons. The sampling goal was to collect the 3 samples in every 313 points near the Au-enriched ore body and characterize their signatures on the surface. In order to simplify, we are using the parental rock names for the soil samples.

The geological map and all the cartographic information used were designed by the AngloGold Ashanti - Serra Grande Unit (SGU) at a scale of 1:25,000 scale. The rock names and their acronyms are the same used by the company. Metabasic rocks comprises quartz-chlorite-carbonate-sericite schist (CBCX), amphibole schist (AMP), and metabasalt (MBA); lithotypes related to the metasedimentary rocks are the carbonaceous phyllite (GXN) and dolomite (DOL).

3.2 Chemical analyses of soil samples and methodology

The 938 soil samples were analyzed at the ALS Laboratory for concentrations of major (wt.%), minor and trace elements (ppm) using ICP-MS (inductively coupled plasma mass spectrometry). The sample batches were prepared with the originals and blank samples for quality control added at a 20:1 ratio. The first sample of each batch is a white sample that allows evaluating if there is contamination in the handling and/or during the preparation of the material, and the quality of the reagents. The soil was dried out at < 60° C to avoid the loss of the mercury and sieved to 80 mesh. Gold analysis was done by Fire Assay and atomic absorption spectrophotometry (AAS). The samples were digested using four acids and later analyzed by ICP-MS. 35 elements were used in this work: Ag, Al, As, Au, Ba, Bi, Ca, Co, Cr, Cu, Fe, Ga, Hf, Hg, K, Li, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Rb, Sb, Sr, Te, Th, Ti, Tl, V, W, Zn, and Zr. Another 15

elements were also analyzed but excluded — Be, Cd, Ce, Cs, Ge, In, La, Re, S, Sc, Se, Sn, Ta, U, and Y — because they have concentrations below the limit of detection and are, thus, unreliable. The geochemical data processing relied on analyzing and defining main anomalies on the software Geosoft Oasis Montaj and ArcGIS. The statistical parameters can be found in [Table 2.1](#).

Table 2.1. Statistical parameters of raw geochemical data of all soil samples from the Pequizão deposit. Min: minimum concentration; Max: maximum concentration; S. D.: Standard Deviation; Q25: 25th percentile of the data set; Median: 50th percentile of the data set and Q75: 75th percentile of the data set.

	Ti	Al	Fe	Mg	Mn	Ca	Na	K	Ag	As	Au	Ba	Bi	Co	Cr	Cu	Ga
Max.	1.64	12.5	17.7	4.08	0.57	0.65	0.92	3.19	3.05	531	7.47	4880	0.94	706	1780	230	32.6
Min.	0.1	3.66	1.21	0.04	0.01	0.01	0.01	0.01	0.01	2.6	0.01	10	0.01	1.1	43	19.1	9.3
Median	0.89	8.9	11	0.12	0.08	0.05	0.01	0.03	0.03	8.9	0.02	30	0.06	36.1	541.5	127	20.2
Mean	0.83	8.47	9	0.17	0.08	0.08	0.04	0.48	0.04	33.6	0.13	590	0.11	43.2	522.9	118.4	19.8
S. D.	0.33	1.79	3.98	0.25	0.05	0.1	0.11	0.72	0.11	67.6	0.58	899	0.11	45.7	319.1	37.6	3.8
Q25	1.07	9.72	12.6	0.17	0.1	0.1	0.05	1.01	0.04	26.4	0.04	1263	0.15	53.8	735	144.5	22.2
Q75	0.55	7.39	5.19	0.08	0.05	0.03	0.01	0.02	0.02	7	0.02	10	0.04	20.4	218.5	94.1	17.6

Ti, Al, Fe, Mg, Mn, Ca, Na and K in %; others in ppm.

	Hg	Li	Mo	Nb	Ni	P	Pb	Rb	Sb	Sr	Te	Th	Tl	V	W	Zn	Zr
Max.	0.074	57.7	3.46	15.4	751	1790	141.5	103.5	1.92	143	0.96	43.9	1.03	522	23.9	265	248
Min.	0.003	4.7	0.06	0.90	11.8	40	1.1	0.1	0.08	0.7	0.03	0.26	0.01	36	0.3	22	2.8
Median	0.022	13.4	0.41	5.30	143.5	280	4.4	1.7	0.74	3.7	0.05	0.98	0.03	299	2.4	52	11.9
Mean	0.024	14.6	0.53	5.51	139.7	313.3	11.2	18.3	0.72	12	0.08	2.77	0.15	260	2.75	59.6	42.9
S. D.	0.01	5.5	0.35	1.81	83.5	133.6	13.1	25.6	0.30	19.9	0.10	3.39	0.21	112	1.97	27.8	50.8
Q25	0.029	16.7	0.62	6.53	178.4	360	17.1	40.4	0.89	16.7	0.09	4.98	0.30	340	1.5	68	84.8
Q75	0.017	11	0.31	4.30	79.7	240	3.2	0.9	0.50	1.9	0.03	0.70	0.02	154	3.4	42	7.9

Cont. Table 2.1.

3.3 Statistical analyses

After organizing the geochemical database and to improve the procedure of geochemical anomaly identification using the traditional threshold-based techniques, the variables were classified to basic univariate statistical processing, with the design of histograms, normal probability plot and box plot, in addition, statistical summaries of each element, showing relevant statistical parameters of the variables of each level collected in the field. To provide an excellent means of identifying or testing natural geological laws all data were log-transformed to univariate analyses. It is common that compositional data are log-transformed prior to processing to remove the closure problem (Filzmozer et al., 2018). In this case, the log-transformation showed to be more consistent and fully satisfactory to handle and interpret the univariate graphs.

The symbology in the geochemical maps has been categorized into quartiles presented in the box and violin plots for each element based on the methodology used by Grunsky (2010). Regarding the symbols, the smaller circle represents the lower outliers, growing gradually upward: the minor circle represents the interquartile range - AIQ - and also the estimate of the background), and the middle circle is the range between the third quartile (Q3) and the upper whisker. Outliers can be further divided into: third-order anomalies from the upper whisker up to $Q3 + 3 \cdot AIQ$, the second-order anomaly is higher than the third-order but lower than $Q3 + 4.5 \cdot AIQ$, and the first-order anomalies are those greater than $Q3 + 4.5 \cdot AIQ$ (Reimann et al., 2005).

To perform the multivariate analysis, the data were transformed using CoDA-Pack (Comas-Cufi et al., 2011) by centered log-ratio (CLR) algorithm (Aitchison, 1986), due to the problems of the constant sum constraints and spurious correlation of compositional data (Pearson, 1897; Chayes, 1960), such as geochemical analysis that are reported as a part of a whole, as a percentage or in parts per million (ppm). The Compositional Data Analysis (CoDA) is based on log-ratio transformations used to overcome the “closure” effect and unconstraint the data. However, there was no removal of samples considered outliers (a procedure that avoids the influence of anomalous samples), in order

to highlight the anomalous zones. The centered log-ratio transformation does not depend on the results of one single other variable but uses the average of all variables, calculated as the log of the individual measurement divided by the geometric mean of the element across the whole dataset (Templ et al., 2008; Filzmoser et al., 2018). The use of the CLR allows some factors to present information about two geochemical associations, positive and negative loadings with the factor.

$$clr(x) = \left[\ln \frac{x_1}{g_m(x)} ; \dots ; \ln \frac{x_D}{g_m(x)} \right]$$

where $g_m(x)$ is the geometric mean of the composition

There are a large number of techniques available that address problems in multidimensional space related to different populations, with the most used methods being cluster analysis and principal component analysis since geochemical data have a lot of analyzed variables, so, the multidimensional statistical techniques become a fundamental tool.

The hierarchical clustering method used starts with single object clusters and aggregate observations to enlarge them stepwise. Hierarchical cluster analysis (HCA) is a graphical statistical method of data classification into a hierarchy of clusters. It can separate the samples into groups based on similarities in their measured attributes (Lin et al., 2014). For this study, hierarchical cluster analysis was applied to geochemical samples from all soil samples in order to investigate the mineralization effects. Templ et al. (2008) explain that the horizontal lines of the dendrogram indicate the linkage of two clusters and the vertical axis the similarity as a measure of distance. Specifically, the integration of Euclidean distance as the distance measurement and complete method as the linkage rule was found to be structured and the distinct hierarchical groups within the geochemical data set compared to other methods. The complete linkage is the distance between two groups defined as being the distance between the most distant remaining of the two groups (maximum distance) and this method eliminates problems that the simple linkage method

has. Dendrograms were generated on Orange: Data Mining Toolbox in Python (version 3.29.3) ([Demsar et al., 2013](#)).

Factor analysis (FA) is a statistical method used in multivariate data analysis that can identify factors among variables in terms of their correlation ([Reimann et al., 2002](#)). Thus, it is possible to find elementary chemical associations representative of different parental lithotypes. Factors are statistical entities and they can be visualized graphically as a classification axis along which variables can be plotted. The intensity of the correlation between variables and a factor is given by its coordinate to the axis. The projection plotted from each soil sample was based on factor analysis. The principal components method was chosen as the factor extraction method from each chemical element in this work because it was the only method that presented results that were very consistent and cohesive with the geological data.

Geochemical analysis data is well known for having a high dimensionality, i.e., the high number of variables as elements (or oxides). The high dimensionality poses problems when trying to understand data variability and searching for patterns as two- (2D) or three-dimensional (3D) diagrams can identify limited subsets of the data and thus only trends and structures of these subsets (e.g., [Jolliffe and Cadima, 2016](#); [Varmuza and Filzmoser, 2009](#)). So, in order to reduce the dimensionality of the data set we used Principal Component Analysis.

The Principal Component Analysis (or PCA) is an unsupervised statistical method that uses linear transforms (e.g., centering and rotation) to the original data matrix with a large number of variables into a new set with a smaller number of variables, called principal components (PCs), that are not correlated and ordered so that the first few PC retain the maximum possible information of the original variables ([Jolliffe, 2002](#)). Each PC is a unit long “eigenvector” that is composed of proportions of each of the original variables (this proportion is called “loading score”), and each eigenvector has an “eigenvalue” that represents the sum of the squared distances of the data points to the origin after the data has been centered, this gives a value for the eigenvector and it is possible to sort eigenvalues in descending order so that eigenvectors that have higher eigenvalues accounts for a higher variability of the data ([Abdi and Williams,](#)

2010). We can plot the percentage that each PC has on the total variation in a “scree-plot” to determine the minimum number of PCs that represents the largest variation in the data. The “scree test” (Cattell, 1966) is widely used and consists of analyzing the graphical representation of the eigenvalues from the factors set generated looking for breaks in the curve and considering only the number of factors before the break, disregarding the factors positioned in the flattest portion of the curve. An advantage of PCA to 2D and 3D plots such as plots of oxides or element-ratios is that information is integrated from all variables, and because PCs are orthogonal to each other they can also be easily visualized in a 2D or 3D plot (Jolliffe and Cadima, 2016). Since PCA tends to maximize variability, data points (samples) that are similar will tend to cluster with each other, while separating from points with higher variance, making PCA a common technique for clustering (e.g., Xue et al., 2011). We used the software IBM SPSS® Statistica v.26 to calculate and plot the PCA and its related statistical parameters.

3.4 Ensemble Learning Framework

Ensemble learning is a term for methods that utilize and combine base learners (machine learning algorithms) so that each base learner will compensate for the errors of previous learners and thus resulting in an ensemble with better overall performance (Opitz & Maclin, 1999; Rokach, 2010). These base learners can be a multitude of algorithms (e.g., decision trees and neural networks) and they can be all the same type, resulting in homogenous ensembles, or from different types to form heterogeneous ensembles (Zhou, 2012). Since the ensemble is a collection of small and simple base learners, they can also be called weak learners. One of the main ensemble methods is boosting, a method that gives more importance to misclassified data when partitioning the data set to subsequent weak learners, with a goal to increase the prediction accuracy of the model (Schapire, 1990, 2003), while bootstrap aggregating (bagging) is an ensemble with random data partitioning that aims to reduce model variance (Breiman, 1996). A Decision Tree (Loh, 2014, and references therein) is the most common type of weak learner and the basis of prevailing and powerful methods, such as Random Forest (Breiman, 2001), Adaptive Boosting (AdaBoost; Freund & Schapire, 1997), Gradient Boosting Machine (GBM; Friedman, 2001), Extreme

Gradient Boosting (XGBoost; [Chen & Guestrin, 2016](#)), and Light Gradient Boosting Machine (LightGBM; [Ke et al., 2017](#)).

As a state-of-the-art ensemble method, XGBoost was employed to perform a regression modelling into the geochemical data set aiming to predict the Au content in the samples, and then to extract the information of what features (geochemical elements) were used and their importance for the learner, in order to better grasp the general pattern associated with Au content. The XGBoost was implemented in R (version 4.1.0; [R Core Team, 2021](#)) using the tidymodels framework (v. 0.1.3; [Kuhn and Wickham, 2020](#)) and the xgboost package (v. 1.4.1.1; [Chen et al., 2021](#)) as an engine for the model. The geochemical data set used was the CLR-transformed data set, which was further split into training and testing sets (0.75:0.25) and stratified using the Au variable to ensure an even distribution of low and high values between the sets. To prevent overfitting, we utilized 10-fold cross-validation in the training set, stratified with Au, and repeated the cross-validation 5 times. While the repetition is time-consuming, it can vastly benefit the learner as it'll have more sets to train. The formula used was Au as the outcome and the other elements — Al, Ca, Fe, K, Mg, Na, Ti, Ag, As, Ba, Co, Cr, Cu, Ga, Hg, Mn, Ni, P, Pb, Sb, Te, Tl, and Zn — as predictors.

XGBoost is a method with hyperparameters (i.e., a parameter whose configuration is exterior to the model and must be input by the user) and we can optimize the input values by using the existing data to identify settings for the parameter of the model that yield the best and most realistic predictive performance, so-called tuning. The learner was tuned via finetune package (v. 0.1.0; [Kuhn, 2021](#)) using simulated annealing ([Bohachevsky et al., 1986](#); [Kirkpatrick et al., 1983](#)), a global search model that employs a gradient-based optimization routine that can reassess previous solutions to find the best “evaluation path”, with 50 iterations across the cross-validated set. The tuned estimates were evaluated with the coefficient of determination R^2 and root mean squared error (RMSE) metrics. R^2 quantifies the determination of a dependent variable (outcome variable) by the independent variable (predictor variable) as the proportion of variance, shown to be a preferred metric for regression analysis ([Chicco et al., 2021](#)). Another way to evaluate a model is to consider the distance of the regression line to the data points (i.e., the error) such as the mean squared

error (MSE), with RSME being used to standardize the units of measures. The range of R^2 goes from 1, a perfect fit, to 0, a trivial and “random” fit, while RMSE goes from 0, a perfect fit, to infinitely-growing positive values of increasingly worse fit. The best-evaluated set of tuned parameters is chosen on the basis of the lower value of R^2 and the highest value of RMSE.

The tuned xgboost hyperparameters are listed in [Table 2.2](#) with the selected parameters for the learner. Then, the learner is used to model the training set and then fitted to the testing set, where the model is tested by a linear regression model of the true Au content versus the predicted Au estimate and evaluated with R^2 . The feature importance values for the model are extracted using the VIP package (v. 0.3.2; [Greenwell and Boehmke, 2020](#)) in order to understand which variables were considered more important for the learner to predict Au.

Table 2.2. Hyperparameters and its chosen value after tuning ($R^2 = 0.71$ and $RMSE = 0.50$).

Hyperparameters	Description	Value range	Chosen Value
trees	The number of decision trees employed in the ensemble.	500 – 1000	973
min_child_weight	Minimum sum of observation weights needed to split a “children” node. The larger, the more conservative. Ranges from 0 to positive infinity.	2 – 40	40
max_depth	Maximum size of a tree (i.e., number of splits). High values make more complex models prone to overfit. Ranges from 0 to positive infinity.	3 – 12	7
eta	Learning rate. Determines how fast the model fits the residual error utilizing additional weak learners. Shrinks feature weights to prevent overfitting. Ranges from 0 to 1.	$10^{-5} - 10^{-1}$	6.998 $\times 10^{-2}$
gamma	Minimum loss (error) reduction required to create a new tree-split. The larger, the more conservative. Ranges from 0 to positive infinity.	$10^{-10} - 10^{-1}$	1.89 $\times 10^{-4}$

early stopping	Number of iterations without improvement before the algorithm stops.	10 – 50	40
----------------	--	---------	----

3.5 Reflectance spectroscopy and XRD analyses

Reflectance spectroradiometry is an effective, fast, and non-destructive technique for mineralogical identification of mineral deposits, used in different stages of mineral exploration ([Gladwell et al., 1983](#); [Cudahy and Ramanaidou, 1997](#)). The reflectance spectroscopy analysis of the deposits allows a better understanding of the alteration zones, helping to grasp the spatial distribution of the ore. Minerals interacting with light preferably absorb certain wavelengths and also transmit at other wavelengths specifically, where the depth, shape, and position of these absorptions are controlled by their structure and chemical composition ([Van der Meer, 2004](#)) ([Table 2.3](#)).

Point-based spectral measurements were performed with an Analytical Spectral Device (ASD) FieldSpec 3 Hi-Res spectrometer, at the University of Brasília. Four measures were taken in the same amount of soil samples, resulting in 3756 spectra. The spectra are obtained from a contact probe composed of a halogen lamp (1 cm) and data acquiring is done with direct contact on the sample. The spectrometer operates in the visible-near infrared (VNIR) in the 350 to 1000 nm wavelength range and in the short-wave infrared (SWIR) regions in 1000 to 2500 nm range. The sample spacing and spectral resolutions are 1.4 and 3 nm for VNIR, and 2 and 8 nm for SWIR, respectively (www.asdi.com).

The analysis and processing of spectral data were carried out into The Spectral Geologist (TSG Core™ v.8.0.7, CSIRO Earth Science and Resource Engineering - CESRE - Division, Sydney, Australia) software. The spectra were compared with spectra of reference spectral libraries from GMEX ([Pontual et al., 2008a, 2008b](#)) and USGS ([Kokaly et al., 2017](#)). The mineral identification was taken on the basis of diagnostic absorption features in the spectra, such as width, depth, and wavelength position. In our 938 samples, some were excluded (~5) as they contained biogenic materials (i.e., dry vegetation) and could mask

minerals. Four measures were taken in each sample and the fourth was chosen for the analysis. After analyzing key features from the spectra data in the software and other spectral libraries, following baseline literature from [Haest et al. \(2012\)](#) and [Sonntag et al. \(2012\)](#), for extraction of semi-quantitative data we defined the spectral parameters of abundance, crystallinity, and composition. For the purpose of comparing and verifying the quality of the spectral analysis, 12 collected samples were analyzed with bulk sample by X-ray diffraction (XRD) using a RIGAKU - D/MAX - 2A/C instrument with Cu radiation ($k = 1.5406 \text{ \AA}$), voltage and current generator set to 35 kV and 15 mA.

Table 2.3. Main absorption bands according to the presence of ions or molecules in minerals (after [Clark et al., 2007](#)).

Ions and molecules	Main absorption bands (μm)
OH and water	1.400
NH ₄	1.560 (also in 2.020, 2.120)
Water	1.900
Al-OH	2.160-2.220
Fe-OH	2.230-2.295
Mg-OH	2.300-2.360
CO ₃ ²⁻	2.300-2.350 (also in 1.870, 1.990 and 2.155)
Mg-OH	2.370-2.400

4 Results

4.1 Descriptive statistics

The data were firstly characterized by summary statistics such as the measure of central tendency, percentiles, variance, and standard deviation ([Complete table in Supplementary Data](#)). The probability plots were used to determine the background and threshold values for anomalies, discriminating the different populations when possible. For an effective visualization of the elements contained in the soil, concentrations were plotted as violin plots in groups according to their parental rocks and sampling level (A, B, and C). To simplify the nomenclature from the rocks, we designed the metabasic rocks unit (Rio Vermelho Formation) that is composed of metabasalt, quartz-chlorite-carbonate-sericite schist, and amphibole schist, and the metasedimentary rocks unit (Ribeirão das Antas Formation) composed of carbonaceous phyllite and dolomite. The violin plot shows the distribution (kernel density) of the data and outliers ([Fig. 2.2](#)).

The major elements reveal the influence of the parental rock in the subsurface, such as high values of Al, K, Mg, and Na, expected in sedimentary rock-derived soil in comparison to the low concentration of mafic-associated elements such as Cr, V, Cu, Ni, and Co. Note that gold content is higher in metasedimentary rocks. Overall, there is some arsenic and gold in the metabasic unit, so mineralization is not restricted to a single rock type.

Due to certain geochemical similarities between A, B, and C levels in the same rock unit, it is possible to outline the general chemical compositions of the sampling level. Samples of the A level of the metasedimentary unit are characterized by: Al₂O₃ = 4.1–10.2 wt%, FeOt = 1.2–10.6 wt%, MgO = 0.06–0.6 wt%, CaO = 0.01–0.61 wt%, Na₂O = 0.01–0.9 wt%, K₂O = 0.04–3.1 wt%, TiO₂ = 0.1–1.3 wt%, P₂O₅ = 60–1790 ppm, Cr = 43–576 ppm, and Ni = 12–195 ppm. From the metabasic unit of the same level, the samples are characterized by: Al₂O₃ = 5.6–12.3 wt%, FeOt = 3.6–16.6 wt%, MgO = 0.04–2.6 wt%, CaO = 0.01–

0.45 wt%, Na₂O = 0.005–0.13 wt%, K₂O = 0.005–1.95 wt%, TiO₂ = 0.5–1.5 wt%, P₂O₅ = 130–770 ppm, Cr = 124–1750 ppm, and Ni = 45.4–706 ppm.

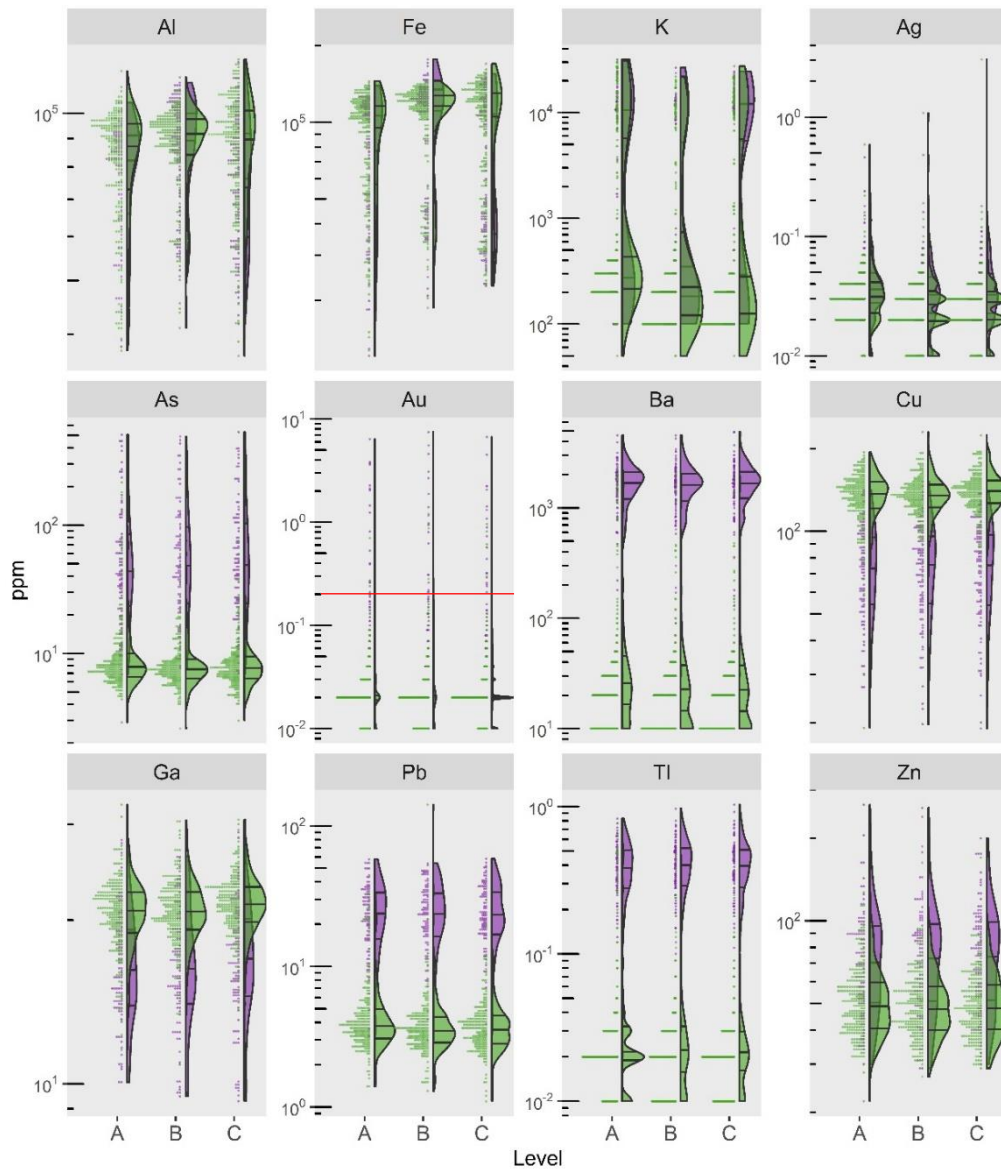


Fig. 2.2. Violin plots and sample distribution for selected elements separated into the metasedimentary unit (green), related to Ribeirão da Antas Formation, and metabasic unit (purple), related to the Rio Vermelho Formation, with the levels A, B, and C. The red line in Au represents 0.2 ppm. The concentration of all elements is in [Supplementary Data 2](#).

For an initial approach to the multivariate analysis of the dataset, we choose to first apply the correlation matrix to help the interpretation of forthcoming results from the other multivariate methods in this work. In the correlation matrix

of centered log-ratio (clr) transformed data for the samples from level A (Fig. 2.3), we identify higher correlation indexes classified as very strong ($r > 0.82$) observed in associations of Fe-Al-Ti, Ni-Cr-Fe, K-Ba, and classified as strong ($r > 0.65$) the Al-(Cr, Ni), Ti-Ni-Cu-Cr, Na-(K, Ba), Fe-P and Mn-(Fe, P, Ti, Ni, Cr) associations. Strongly negative correlations occur for K-(Al, Fe) and Fe-(K, Na). A slight variation of elements occurs on those associations when comparing them between all the soil levels.

We also examined the correlation matrix using raw data (original data without log-ratio transformation) and there is a similarity of clr and raw results, but strong (> 0.8) relationships in raw data may become non-significant in clr. For example, the very strong association of Fe with Al, Cu, Cr, Ni, and others like Ga-Al, Mn-Co, and Na-K-Ba. Additionally, gold shows a fairly strong (> 0.6) correlation with As in raw data, and moderate (~ 0.5) with K, Ba, Na, while these elements have a strong correlation with As. The overall correlation from those elements only decreases at level C.

There is some similarity of elements at the three levels, however, according to the sample depth (from A to C), we can see a gradual content increase of Al_2O_3 , Ag, FeO, MgO, and Ga. The opposite happens to P_2O_5 and Hg as both elements have higher values in level A and lower values in B and C. In level B, only two elements seem to have a contrast, Au and Pb, with higher retention and greater content.

In the binary plot graphs in Fig. 2.3, we can see gaps occurring in values of Fe, K, and Na separating metasedimentary and metabasic units. Elements with pairs that present a strong inflection, i.e., Ba, Cr, and Cu, also help to identify the parental rock units, for example, in the Cr-K graph the flat high levels of Cr represent metabasic unit while the steep low-Cr and high-K represent the metasedimentary unit. Besides the clusters provided by Fe, K, and Na, the pair Al-Ca also provides a good separation of both the units.

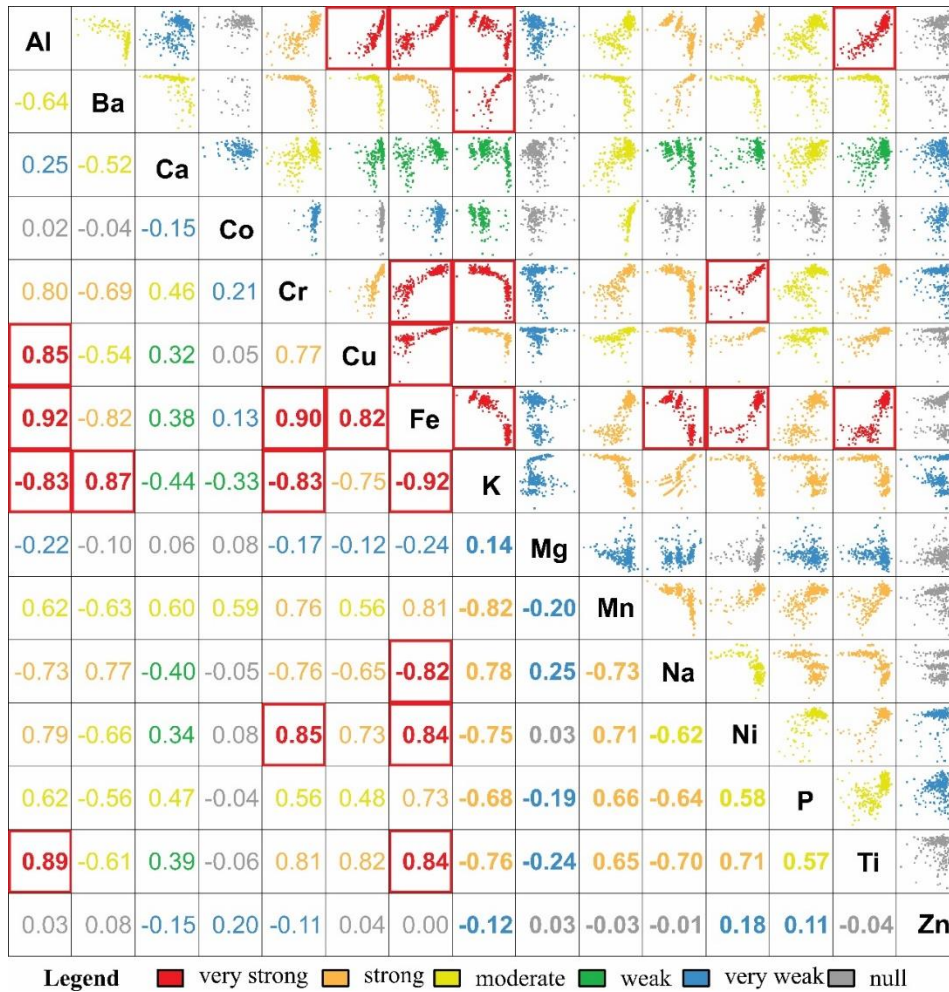


Fig. 2.3. Correlation matrix of selected elements from level A (313 samples) with significant correlations in this study. The complete correlation with all elements from B and C are in [Supplementary Data 2](#). Significance of 0.90.

4.2 Hierarchical cluster analysis

Hierarchical cluster analysis (HCA) was used to identify clusters in the dataset, and it helps distinguish different geochemical relations in the soil samples. The main method used was on the basis of Euclidean distance and complete linkage; other methods of linkage-distance were tested but they do not show geological correlation/significance. This technique was used to classify soil samples in terms of parental-rock and show element associations.

The dendrogram (Fig. 2.4) with 24 variables was clustered into 4 different groups named C1, C2, C3 and C4. The first cluster (C1) is defined by Au, Ag, Hg, Sb, Te and Tl, the second cluster (C2) is defined by Al_2O_3 , FeO_t and TiO_2 , the third cluster (C3) is represented by As, Co, Cu, Ga, Ni, Pb and Zn, at last, the fourth cluster (C4) is defined by Ba, CaO, Cr, K_2O , MgO, MnO, Na_2O , P_2O_5 . In general, HCA shows an association where the C2 cluster (Al-Fe-Ti) have a high influence from the mafic parental rock. The C1 cluster demonstrates association linked to gold-bearing elements, highlighting silver and mercury, for example. For C3 and C4 clusters, just the hierarchy method was not able to properly point out geological signature, however, the C3 cluster shows some clusters in that is related to metabasic rocks and the C4 to the metasedimentary rocks. For this reason, to claim and clarify these results, we used ensemble learning to understand the importance of the elements.

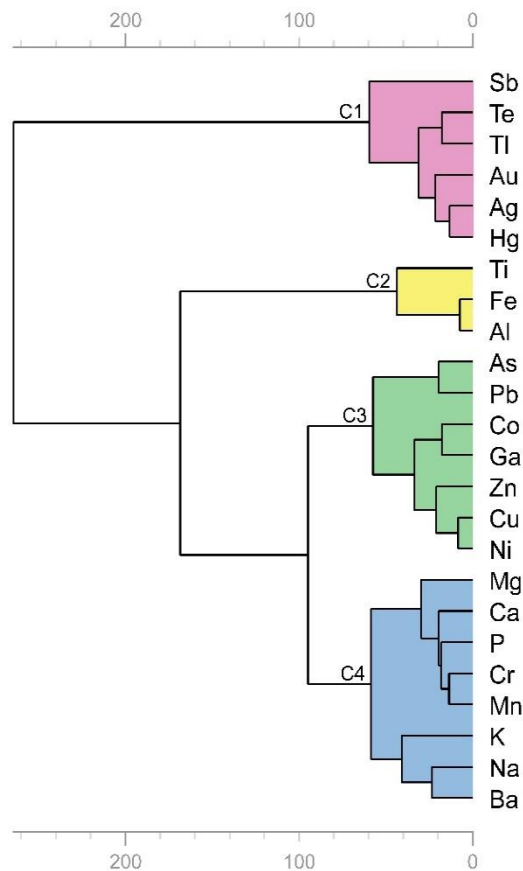


Fig. 2.4. Dendrogram of soil sample using Euclidean-complete linkage, evidencing 4 clusters called C1 (in pink), C2 (in yellow), C3 (in green) and C4 (in blue) (pruning: 10 levels).

4.3 Ensemble learning results

The ensemble learning was generated using the clr-transformed set with XGBoost to predict the Au (clr) content using the other transformed elements. The model was trained and evaluated with $R^2 = 0.711$ and $RMSE = 0.50$, indicating that the training error was relatively low. When the model was applied to the testing set the metrics returned values of $R^2 = 0.78$ and $RMSE = 0.45$, indicating that the generalization error was also relatively low and demonstrating that the model is not overfitted.

By comparing the predicted values versus the observed (true) values of Au (Fig. 2.5a), we can see that the values follow the trend of 1:1 (black line; predicted value is equal to the observed value), as demonstrated by the linear model (magenta line). Even with a low number of samples with high Au (clr) values (which can also be interpreted as higher concentrations in ppm), the predicted values have residuals (the distance from the predicted value to the fitted line) similar to those with lower Au (clr). According to the variable importance plot (Fig. 2.5b), the clr-transformed elements Zn, Al, Ga have considerable importance for the learner to predict Au (clr) values, while Pb, Mg, and P have moderate importance being equal or higher than the mean importance of 0.043. Due to the stochastic (random) nature of a machine learning algorithm such as XGBoost, the importance values may change slightly, which can cause elements (clr) with similar values to change order, but the three larger importance values may not see any change in the order or magnitude of importance.

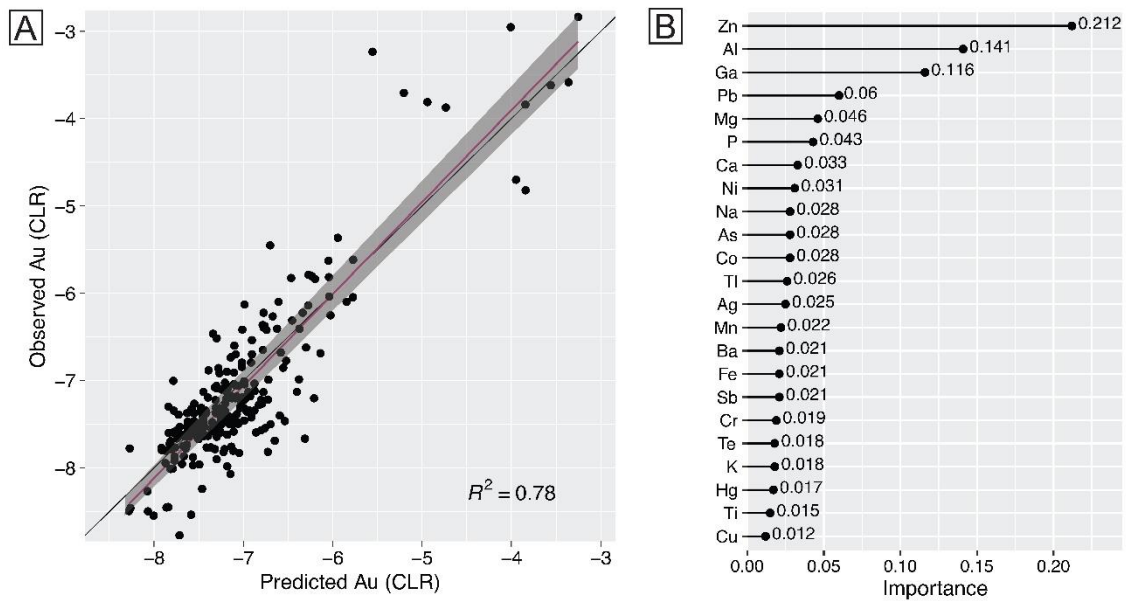


Fig. 2.5. (a) Predicted \times Observed values according to the used model. The thin black line is a 1:1 line, and the magenta line is the fitted linear regression with the gray area representing a 95% confidence interval. (b) Variable importance plot, with the numbers indicating the importance value of the respective transformed element.

4.4 Reflectance spectroscopy results

Spectral techniques in mineral exploration have become increasingly frequent in countries with high prospective demand, as they are useful tools for mineral mapping and, therefore, help to delineate alteration zones associated with mineralization, especially when integrated with geochemical data. Soil samples are rarely employed in spectral projects. However, when they are used, the soil samples must not contain transported material and have to be the underlying bedrock residual material to be representative of the interest area. Soil samples become potentially useful in projects with a large number of samples and geochemical data. As such, it is highly advantageous for alteration mapping, interpreting geochemical associations, and a fast method of producing mineral maps. Although, there are some disadvantages to this process, such as some noisy spectral data and weaker mineral signatures than rock samples. Six groups of minerals were identified: kaolinite, white mica, chlorite, amphibole, aluminium hydroxide, and iron oxides. All the processing, parameters and algorithm

employed in the spectral software were summarized in [Table 2.4](#) for the abundance, composition and crystallinity.

Table 2.4. Parameters from reflectance spectroscopy.

Product name	Function/Mineral	Parameter	Base TSG algorithm
Abundance of Fe-oxides	Identify and quantify abundance of Fe-oxides	900D	PROFILE TSG tool: Depth of the continuum removed spectrum from 762 to 1252 nm, with a radius ranging \pm 245 nm.
Differentiate hematite-goethite	Composition from hematite and goethite	900Wvl	PROFILE TSG tool: Wavelength of the continuum removed spectrum from 762 to 1252 nm, with a radius ranging \pm 245 nm. Masked with: 900D > 0.0329
Kaolinite abundance index	Identify kaolin group minerals (kaolinite)	2200Dk	PFIT TSG tool: Normalized depth of a fitted 4th order polynomial between 2120 and 2245 nm. Masked with: 2160D ((R2138+R2190) / (R2156+ R2179)) > 1.005 (Sonntag et al., 2012)
Kaolinite crystallinity index	Crystallinity of kaolin group minerals ranging from well-ordered kaolinite to poorly-ordered kaolinite; > 1.015 = well-crystalline (Sonntag et al., 2012)	2200SL	[(R2138 + R2173) / R2156] / [(R2156 + R2190) / R2173] 2160D2190 > 1.005. Masked with: 2200D > 0.005 (Sonntag et al., 2012)
White mica abundance index	Identify and quantify white micas	2200D	PFIT TSG tool: Normalized depth of a fitted 4th order polynomial in 2200 nm. Masked with: 2160D < 1.005 and 2350D > 0.00035 and Kaolinite Abundance index < 0.0461 (Sonntag et al., 2012)
White mica composition index	Composition of white mica group minerals ranging from paragonite, muscovite to phengite.	2200Wvl	PFIT TSG tool: Wavelength of absorption minimum calculated using the 1st derivative of a fitted 4th order polynomial between 2120 and 2245 nm. Masked with: Kaolinite Abundance index < 0.0461 (Sonntag et al., 2012)
Gibbsite abundance index	Identify and quantify gibbsite	2260D	PROFILE TSG tool: Depth of the continuum removed spectrum from 2229 to 2290 nm, with a radius ranging \pm 30 nm. Masked with: 2260D > 0.00492
Chlorite abundance index	Identify and quantify chlorite	2345D	PROFILE TSG tool: Depth of the continuum removed spectrum from 2330 to 2360 nm, with a radius ranging \pm 15 nm. Masked with: White mica abundance index < 0.00461

Chlorite composition index	Composition of chlorite group minerals ranging from Mg- to Fe-rich	2345Wvl	PROFILE TSG tool: Wavelength of the continuum removed spectrum from 2330 to 2360 nm, with a radius ranging ± 15 nm. Masked with: White mica abundance index < 0.00462
Amphibole abundance index	Identify and quantify amphibole	2390D	PROFILE TSG tool: Depth of the continuum removed spectrum from 2375 to 2405 nm, with a radius ranging ± 15 nm. Masked with: 2320D > 0.003 and 2390D > 0.005
Amphibole composition index	Composition of amphibole group minerals ranging from actinolite to hornblende	2390Wvl	PROFILE TSG tool: Wavelength of the continuum removed spectrum from 2375 to 2405 nm, with a radius ranging ± 15 nm. Masked with: 2320D > 0.003 and 2390D > 0.005

Samples with the highest abundance and representative from each mineral resulted from spectral analysis and their wavelengths are shown in [Figure 2.6](#).

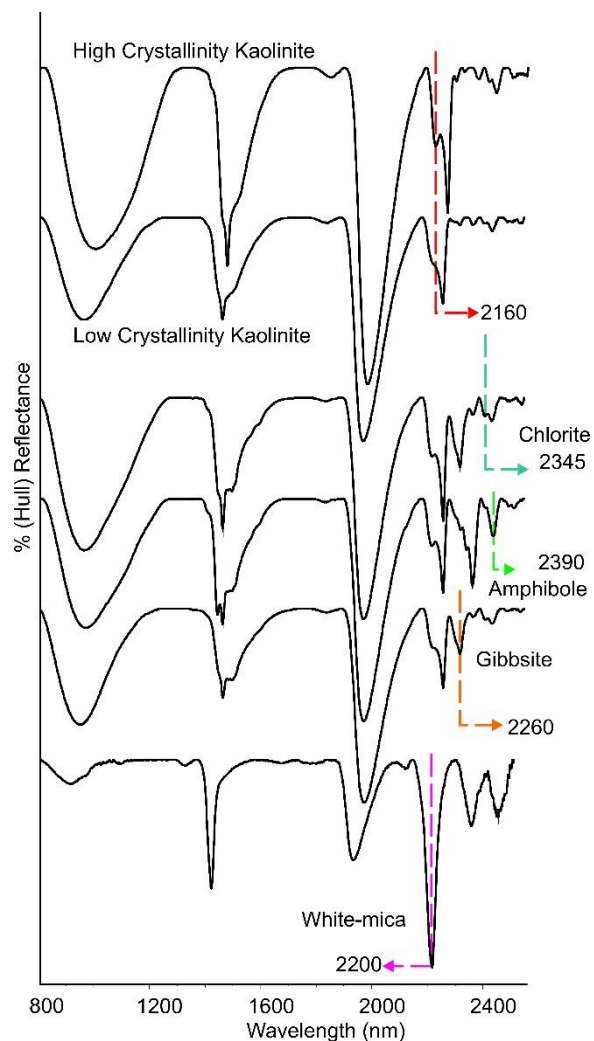


Fig. 2.6. Stacked reflectance spectra of the samples containing iron oxides (860-940 nm), phyllosilicates, and clay-minerals of the Pequizão deposit, including kaolinite, chlorite, amphibole, gibbsite, and white-mica. All samples contain some iron oxides. The main absorption features are highlighted.

Kaolinite has a strong absorption feature at 2200 nm produced by Al-OH. The abundance of kaolinite was obtained from the calculation of the depth of the absorption feature at 2200 nm using a fourth-degree polynomial fitted to the continuum removed spectrum between 2120 and 2245 nm, with the focus ranging between 2160 and 2210 nm (2200Dk, [Table 2.4](#)). It is abundant in the whole set, with ca. 72% of the samples having kaolinite in some proportion. To differentiate the kaolin group from aluminous phyllosilicates, according to [Sonntag et al. \(2012\)](#), we applied a ratio of 4 bands to test the absorption intensity from 2160 to 2190 nm, giving the crystallinity index from the sample adding the depth of the absorption at 2200 nm ($2160D2190 > 1.005$, [Table 2.4](#)). [Haest et al., \(2012\)](#) use the term “kaolinite crystallinity” referring to the estimation of octahedral vacancy (structural) disorder in kaolinite, which results in less than ~1 are classified as poorly-crystalline and more than ~1 classified as well-crystalline.

The main diagnostic absorption feature of white mica is related to Al-OH and it occurs around 2200 nm. As with other minerals, the position of absorption of white mica will vary based on its composition. According to [Clark et al., \(1990\)](#), features between 2180 and 2190 nm belong to paragonite while features between 2200 and 2208 nm characterize muscovite. Phengite is identified by absorption features between 2216 and 2228 nm. The intermediate wavelength ranges of 2190-2200 nm and 2208-2216 nm determine micas with compositions “tendency” to paragonite and phengite, respectively. The lack of the secondary absorption feature between 2160 and 2185 nm in white micas is used to separate them from the kaolin group and, afterward, the presence of the 2350 nm absorption feature is determined using a four-band ratio ($2160D2190 < 1.005$, $2350D > 0.00035$, [Table 2.4](#)). White mica abundance can be separated from kaolinite abundance using a mask with values above 0.0461, and the composition was calculated using the first derivative of a fitted fourth-order polynomial between 2120 and 2245 nm ([Table 2.4](#)).

Gibbsite (aluminum hydroxide) has a characteristic absorption feature around 2268 nm, as reported by [Clark et al. \(1990\)](#), along with three less prominent features in 1452, 1521, and 1549 nm. This aluminum hydroxide in mixture situations only has strong diagnostic absorption features around 2260 nm. The abundance of gibbsite was calculated from the depth of the continuum removed spectrum at 2260 nm using a radius ranging ± 30 nm (2260D, [Table 4](#)), thus, it was identified using the following condition: $2260D > 0.005$.

Chlorite has a characteristic absorption feature in two diagnostic ranges, based on its composition. Fe-OH bond presence is related to between 2240 and 2265 nm, as well as in 2352 and 2365 nm ([Bishop et al., 2008](#)). For the Mg-OH bond, this mineral group has an absorption feature range at 2320 and 2340 nm. However, there is a transitional range, varying in richer Mg- to Fe-OH between 2340 and 2352 nm, in which the higher the wavelength, the higher is the ferrous chlorite abundance ([Bishop et al., 2008](#)). The abundance was obtained by calculating the depth of the continuum removed spectrum from 2330 to 2360 nm only for samples without any abundance of white mica (2345D, [Table 2.4](#)). The composition of the chlorite was calculated using the depth of the continuum removed spectrum from 2330 to 2360 nm, ranging between 2315 and 2375 nm (2345Wvl, [Table 2.4](#)).

Amphibole minerals occur in some samples and are also characterized by diagnostic features related to Mg-OH at 2320 and 2330 nm, and around 2390 and 2400 nm ([Clark et al., 1990](#)). The parameter from 2375 to 2405, using a range of ± 15 nm and the conditions of $2320D > 0.003$ and $2390D > 0.005$ (2390D, [Table 2.4](#)) defines the abundance of amphibole. To recognize and to differentiate the minerals from the amphibole group can be challenging because of the overlap of their Mg-OH absorptions in these mixed mineral spectra. In the majority of the hornblendes, the main Mg-OH absorption feature is at longer wavelengths than actinolite (around 2384 nm). Wavelength of the continuum removed spectrum from 2375 to 2405 nm was used to define the composition of the amphibole minerals (2390Wvl, [Table 2.4](#)).

Iron oxide identification and abundance in the spectrum was performed using the feature in the visible to near-infrared (VNIR) involving the Fe⁺ in the

slope from 850-1000 nm (Haest et al., 2012). There is a high variance of the data in this region, ranging even outside of the typical Fe+ slop. After defining the depth of the absorption, the index used was a centre in 1007 nm, ranging from 762 to 1252 nm (900D, Table 2.4). In the same range, the position of the absorption can define the composition of the iron oxides, which, according to Townsend (1987), varies from 860 nm for hematite to 920 nm for goethite, so, the wavelength was calculated only in samples with an expressive abundance of iron oxides (900Wvl, Table 2.4).

Results from the reflectance spectroscopy were validated with X-ray diffraction on 12 appropriately chosen representative samples from the Pequizeo area (listed in Table 2.5) where the parental-rock, level, and gold values were considered. Quartz and kaolinite occur in all soil samples as expected. According to XRD, iron oxides (goethite and hematite) occur mainly in the auriferous samples. Smectite was locally present in carbonaceous phyllite and only in the XRD they could be identified and classified as saponite. A sufficient amount of muscovite for its detection was only possible in the quartz-chlorite-carbonate-sericite schist and carbonaceous phyllite, additionally, in one phyllite sample, the chlorite mineral was classified as clinocllore. The gibbsite is not restricted to a specific rock type, and rutile is a trace mineral that was detected only in the XRD.

Table 2.5. Mineralogical interpretation of infrared reflectance spectroscopy (IRS) and X-ray diffraction (XRD) was conducted on selected samples.

Sample	Litho.	Type	Spectral (IRS)	XRD
9B	Carbonaceous phyllite	Mineralized	wm, kln-px	qz, kln, hm, ms, smc (saponite)
11A	Carbonaceous phyllite	Barren	kln-wx, goe, amp, chl	qz, kln, ms, goe, gbs
28C	Carbonaceous phyllite	Barren	wm, goe, (kln-px)	qz, kln, ms
36A	Carbonaceous phyllite	Mineralized	kln-wx, gib, chl	qz, kln, hm, ms, gbs, smc (saponite)
50A	Amphibole schist	Mineralized	kln-wx, goe, amp	qz, kln, goe, gbs
90B	Qz-chl-cb-ser schist	Mineralized	kln-px, hm, chl, gib, amp	qz, kln, hm, gbs
111A	Qz-chl-cb-ser schist	Barren	kln-wx, goe, amp, chl	qz, kln, ms, goe, gbs
125C	Amphibole schist	Mineralized	kln-wx, gib, goe, amp, chl	qz, kln, goe, gbs
132B	Qz-chl-cb-ser schist	Barren	kln-wx, hm, amp, chl	qz, kln, hm, ms, goe, gbs, rt

229B	Metabasalt	Barren	kln-wx, goe, hm, gib	qz, kln, goe, gbs
280C	Metabasalt	Barren	kln-wx, hm, goe, gib, amp, chl	qz, kln, goe, gbs
289C	Carbonaceous phyllite	Mineralized	wm, hm, chl	qz, kln, hm, ms, gbs, chl (clinochlore), rt

Mineralized = Au > 0.2 ppm. Abbreviations (after Warr, 2021): amp = amphibole; cb = carbonate; chl = chlorite; dol = dolomite; gbs = gibbsite; goe = goethite; hm = hematite; ilt = illite; kln-wx = well-ordered kaolinite; kln-px = poorly ordered kaolinite; ms = muscovite; qz = quartz; ser = sericite; smc = smectite; wm = white mica.

Reflectance spectroscopy results of the same samples revealed that kaolinite is predominantly abundant and gibbsite, amphibole, and chlorite occurs in many samples. White mica is more abundant in carbonaceous phyllite, therefore, a histogram of the compositional parameter data was elaborated to visualize the white mica composition variance. The white mica shows a strong frequency in the muscovite zone, with some white mica results tending to phengite composition (ca. 2208 to 2010 nm) (Fig. 2.7a).

The presence of hematite and goethite occurs very often on the samples, but focusing on gold-bearing samples from XRD results only half have iron oxides. The frequency graph from iron oxides showed that those are mostly in a transition zone between hematite and goethite and some in the goethite zone (Fig. 2.7b).

The amphibole from the samples showed a high frequency in the actinolite zone and fewer in the hornblende zone (Fig. 2.7c). To determine the chlorite composition, the IRS results can be separated into four zones according to Mg-OH or Fe-OH in the mineral structure (Fig. 2.7d). Except for muscovite, the other minerals are more abundant in rocks from the metabasic unit.

Comparing the frequency and composition of minerals from IRS results with the occurrence of gold, we found out that even though white mica is mostly classified as muscovite and has high-grade samples, the total gold is higher in samples where the white mica tends to phengite. Nonetheless, in the other minerals histogram the high-grade samples and the amount of gold are in the higher frequency zones: for iron oxides in the hematite-goethite transition zone,

for amphibole in the actinolite zone, and for the chlorites in the Fe-chlorite zone (Fig. 2.7).

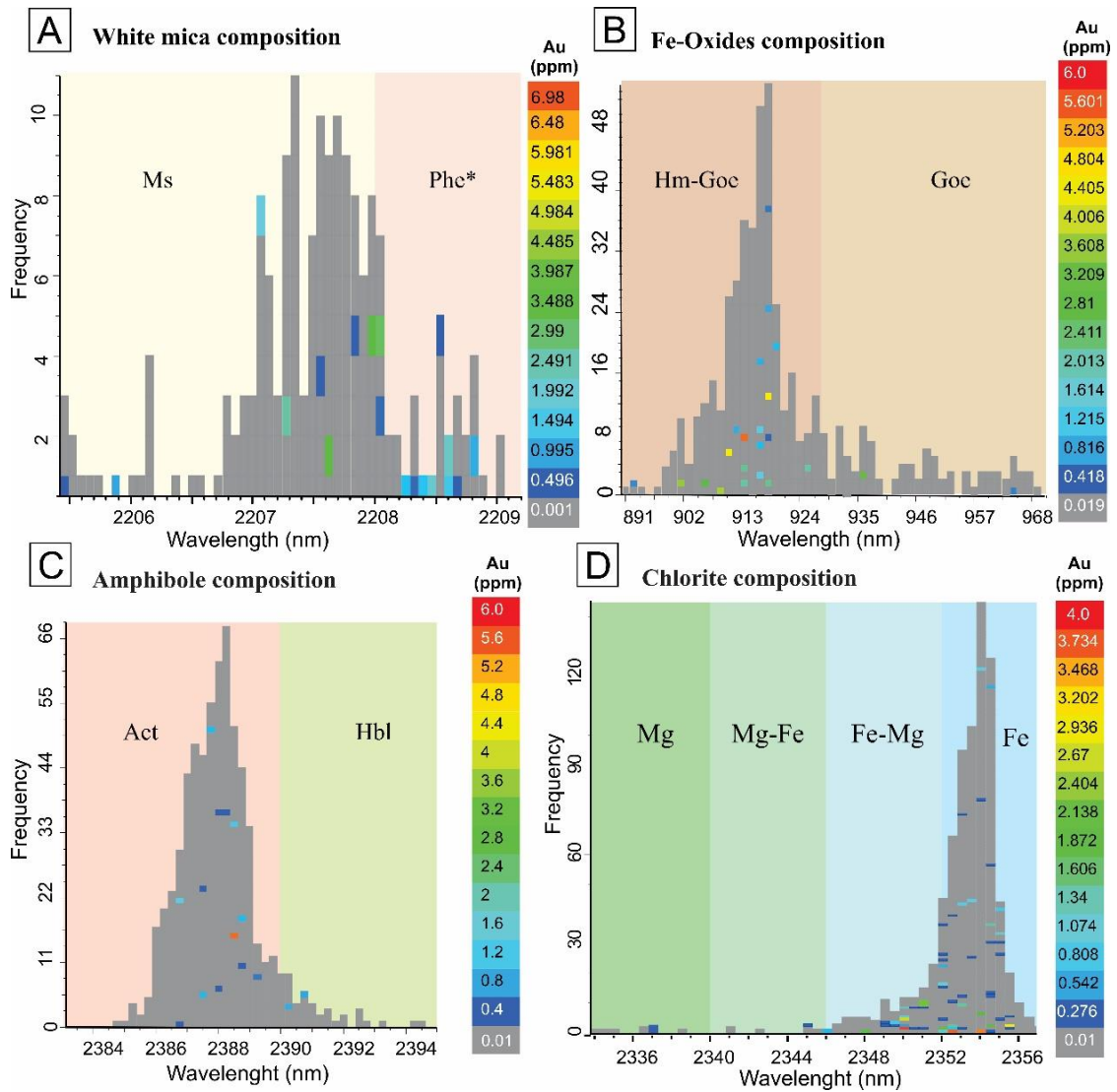


Fig. 2.7. Histograms of the minerals from the reflectance spectral results and showing Au presence; (a) White-mica histogram with 174 samples; (b) Fe-oxides histogram with 487 samples; (c) Amphibole histogram with 640 samples; (d) Chlorite histogram with 779 samples. Abbreviations (after Warr, 2021): act = actinolite; goe = goethite; hbl = hornblende; hm = hematite; ms = muscovite; phe* = tending to phengite.

4.5 FA and PCA results

The factor analysis (FA) and Principal Component Analysis (PCA) were employed for the dataset with 30 variables (centered log-ratios of chemical concentrations).

For the PCA, was performed an eigen-decomposition and, using the scree plot (according to the scree-test by [Cattell, 1966](#)), we were able to reduce the data to the three first components ([Table 2.6](#)) that, together, explain 74.61% of the total data variance. The intensity of correlation between each variable (elements) and a component (PC1, PC2, PC3) is given by its coordinate in relation to the axes (x, y) in a 2D plot from the reduced projection.

Samples were projected on the eigenvectors, this method is used in unrelated factors, being able to distinguish different types of parental rocks in the soil samples (i.e., metabasic unit rocks and metasedimentary unit rocks). The algorithm used explores a combination that maximizes the variance between the loadings of the factors aiding the interpretation, in addition, could effectively differentiate the parental rocks from the soil samples. The factor 1 represents the majority of the data, ca. 61% of the variance and the factor 2 explains 7.6% of the data variance, while the factor 3 explains 5.6%. The factor 2 and 3 together only account for 13.2% of the variance.

Table 2.6. Principal components and the loadings of elements in the first three principal components. Ordered according to the first component. Scores below ± 0.1 were excluded.

	Components		
	1	2	3
Fe	0.978	0.122	
Ba	-0.974		
K	-0.972		
Cr	0.953		
Ni	0.951		
TI	-0.949		
KIn	0.936	-0.145	
Cu	0.914	0.264	
Al	0.908	0.274	
IrOx	0.887	0.171	0.122
Sb	0.882	0.204	0.179
Ti	0.878		
Na	-0.875		-0.165

Mn	0.875	-0.184	
Ga	0.869	0.349	
Amp	0.868		
Co	0.828	-0.257	
Pb	-0.809	0.392	
As	-0.768		0.455
P	0.727	0.148	-0.196
Chl	0.710		0.110
Wm	-0.672	0.188	
Te	-0.656	0.385	0.186
Gbs	0.621	0.194	0.446
Hg	0.595	0.123	
Zn	0.149	0.698	-0.497
Ca	0.410	-0.658	-0.353
Ag	0.196	-0.307	
Mg	-0.156	-0.307	-0.610
Au	-0.344	-0.496	0.571

Abbreviations (after [Warr, 2021](#)): *Amp* = amphibole; *Chl* = chlorite; *Gbs* = gibbsite; *Kln* = kaolinite; *IrOx* = iron oxides (goethite/hematite); *Wm* = white mica.

The results from the geochemical data were analyzed by PCA to enhance the geochemical variance between elements and parental rocks. Using centered log-ratio transformation, the main clusters found on the projection of scores in PC1 vs PC2 biplot ([Fig. 2.8a](#)) were: (a) PC1 is strongly influenced by the high positive values observed for Al, Cr, Cu, Fe, Ni, kaolinite, and iron-oxides, in contrast to the negative loading of Ba, K, Na, Pb, Te, Tl, and white mica; (b) PC2 comprises higher positive loading on Pb, Te, Zn, and white mica, those elements together with muscovite have more influence on the second component, aligned on the negative axis we can find two different clusters of Au-Mg and Ca-Ag in PC2.

We projected PC1 vs PC3 biplot to visualize the high loads of gold on the third component ([Fig. 2.8b](#)), where we can confirm the association of Au, As, and partially Te, with the third component. Minerals such as gibbsite and white mica showed high influence on the PC3, and on the negative side, the presence of Zn, Mg and Ca.

Samples with the content of gold (>0.2 ppm) occurred mostly in the metasedimentary rocks, at the negative axis of PC1 and the positive axis of PC3, demonstrated on the plotted projection of factor-plane 1 x 3 showed on samples in [Fig. 2.8b](#), highlighting the presence of samples from quartz vein clustered

together with mineralized ones. However, samples of the metabasic rocks are strongly represented by the first component where they are often classified as non-mineralized. Based on the PCA and FA, we could extract from the results two geochemistry associations for each factor found in [Table 2.7](#).

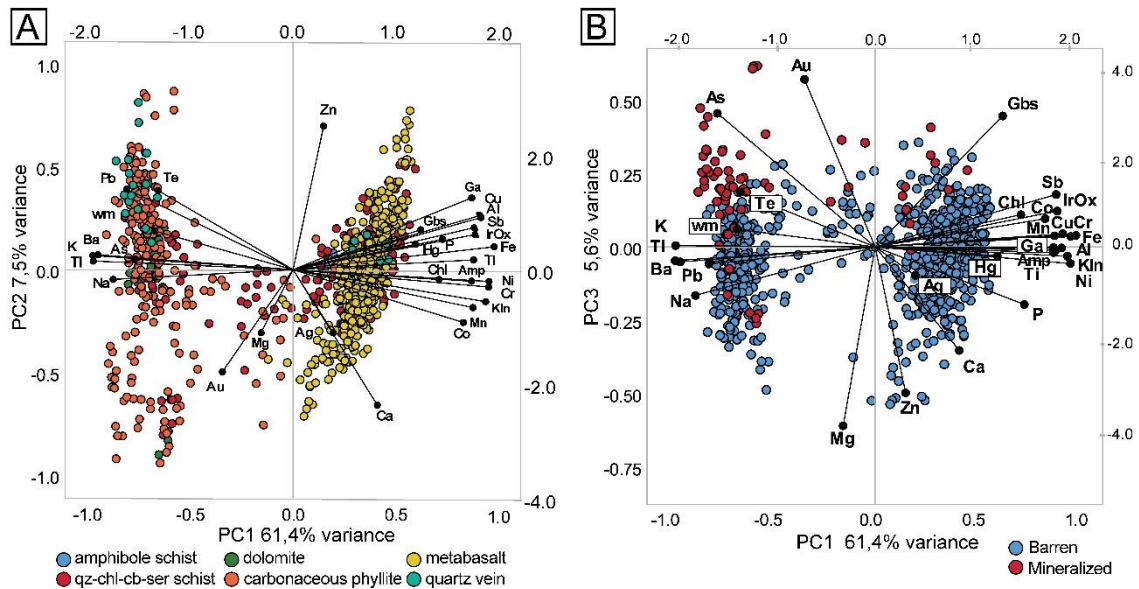


Fig. 2.8. PC and FA loadings of centered log-ratio-transformed $n = 938$ data; (a) The biplot projection of PC1 versus PC2 shows loadings of the elements. Each dot corresponds to an individual sample, in factor-plane 1x2; (b) The biplot projection of PC1 versus PC3 and the factor-plane with mineralized samples ($Au > 0.2$ ppm) in red and barren samples ($Au < 0.2$ ppm) in blue.

The factors scores resulted from the multivariate analysis could be interpolated in three distinct maps where each one is independent and shows different geochemical associations in positive and negative ranges (see [Table 2.7](#)). We use the scores from level B soil samples even though there are certain similarities between levels A and C.

Table 2.7. Factor analysis results with 3 selected factors and their respective geochemical association. Abbreviations (after Warr, 2021): Kln = kaolinite; IrOx = iron oxides; Amp = amphibole; Wm = white mica; Gbs = gibbsite.

Factors	Geochemical association	Mineral mapping	Eigen-values	Variance
Factor 1	Fe-Al-Cr-Ni-Cu (+) K-Na-Ba-Tl (-)	Kln-IrOx-Amp (+) Wm (-)	18.43	61.45
Factor 2	Zn-Pb-Fe-Ga (+) Au-Ag-Ca-Mg (-)		2.27	7.57
Factor 3	Au-As (+) Zn-Ca-Mg (-)	Gbs (+)	1.68	5.59
Total:			22.38	74.61

The first factor showed the highest scores in the positive axis coinciding with one group in the majority mafic samples plotted, and together with the high eigen-values of Fe, Al, Cr, Ni, and Cu, alongside with three minerals, kaolinite, iron-oxides, and amphibole. This discriminated the mafic-like samples, as demonstrated in an interpolated map in Figure 2.9a. In the same factor, in samples located at the negative axis, there is a predominance of white mica and K, Na, Ba, and Tl, related to the metasedimentary rocks (Fig. 2.9a). Based on the sample scores, we delimited four main “pedochemical” domains that can separate, with some precision, the metasedimentary-derived soils at the eastern part of the geological map (pedochemical domain I) and, on the western part with the other three pedochemical domains, the mafic-derived soils representing increasing factor scores (Fig. 2.9b).

The second map, interpolated with the second factor scores, was positively associated with Zn, Pb, Fe, Ga, and negatively with Au, Ag, Ca, and Mg (Fig. 2.9c). Positive regions (>0.03), and their isolines, were plotted on the geological map together with the two NE-SW thrust faults, the western Palmeiras Structure and the eastern Structure IV, so we could estimate its influence on the Factor 2 scores. Note that at the northeast part of the map and nearby those faults there are high scores from this factor, showing an enrichment area with Zn, Pb, Fe, and Ga (Fig. 2.9d).

The last map, of interpolated third factor scores, Au, As, Te, and Sb occur with gibbsite and chlorite in the positive values, in contrast to Zn, Ca, and Mg with negative values (Fig. 2.9e). We plotted the gold-bearing samples on this map to better visualize the influence of the factor with the known mineralized samples. We extracted from this result the positive scores and projected them into the satellite image of the current open pit at Pequizão, showing it is almost inside the pit limits, however, still a trend in the northeast area (emphasizing that the samples were collected before the opening of the pit) (Fig. 2.9f).

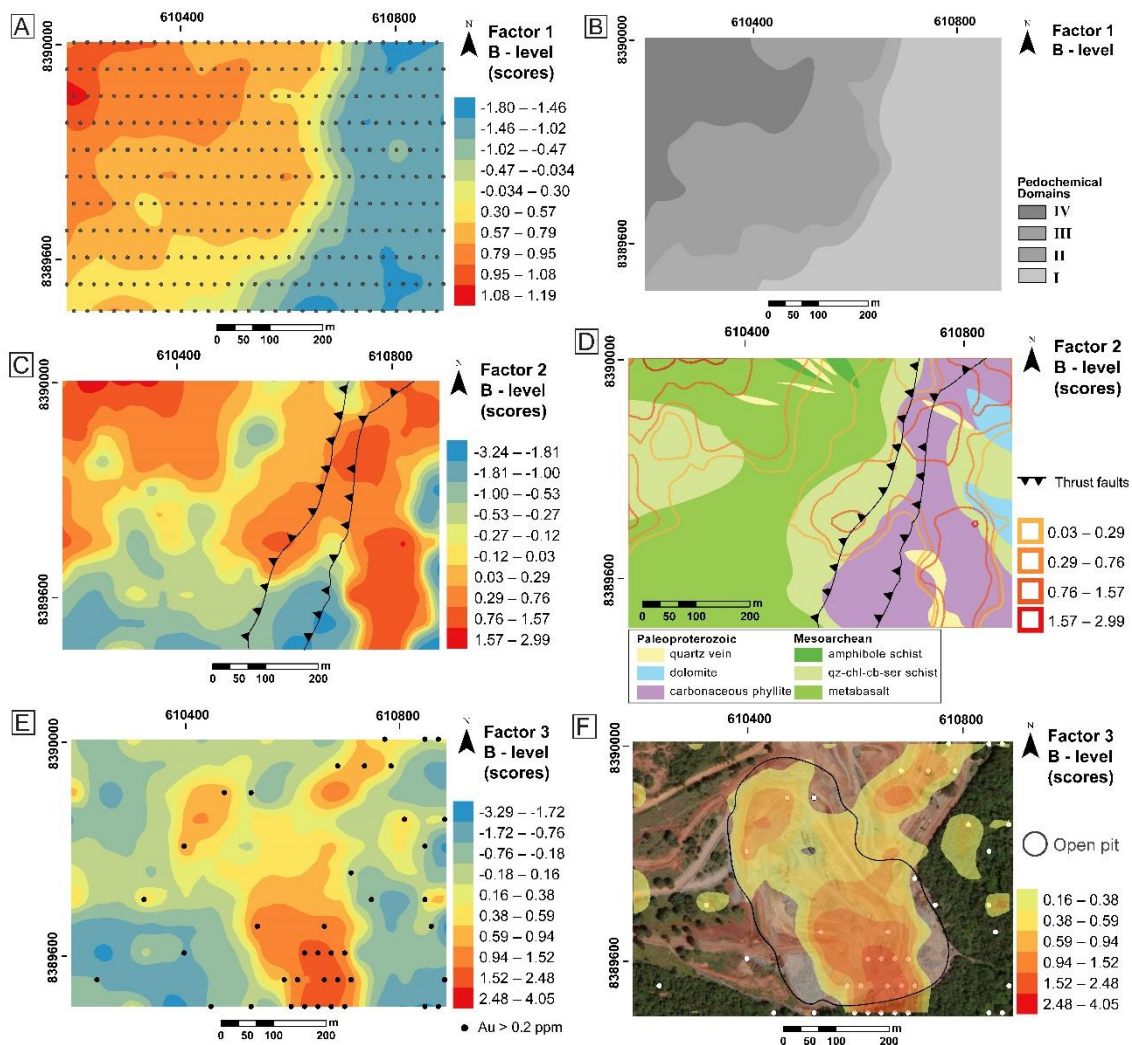


Fig. 2.9. Map of the interpolated scores from factor 1 to 3 in the soil samples from the B level; (a) Factor 1 map, and the black dots are samples location; (b) Map of the four pedochemical domains based on the factor 1 interpolation; (c) Factor 2 map and thrust faults (left: Palmeiras, right: Structure IV); (d) Geological map and isolines from factor 2, highlighting only positive scores; (e) factor 3 map with mineralized soil samples in black dots (Au > 0.2 ppm); (f) Pequizão open pit with the highest factor 3 scores on satellite image also showing high-grade samples in white dots.

5. Discussion

As proposed by [McCuaig and Hronsky \(2014\)](#), the critical elements for ore deposit formation are the fertility of a source region, favorable transient geodynamics and whole-lithosphere architecture, and preservation of the deposit. The origin of orogenic gold deposits is still debated, and it can be formed during the late stages of regional metamorphism from crustal metamorphogenic fluids ([Phillips and Powell, 2010](#)), with fluids sourced from the devolatilization of pyritic sediments; above a subducting oceanic slab, as a result of a change in far field stress due to slab stall ([Groves et al., 2020](#)); or orogenic gold deposits can be even split into distal source-related classes and proximal oxidized intrusion-related ([Witt et al., 2016](#)). First-order architectural controls on the formation of gold deposits are based on their intersection with high-angle accommodation structures, and second-order are dominated by lines of weakness along the belt in which the gold-rich fluid can propagate and, after focusing, promote mineral precipitation ([Goldfarb and Groves, 2015](#); [Wyman et al., 2016](#); [Groves et al., 2018](#)).

The essential for gold exploration is to first understand the gold mineral system, and it includes recognizing alteration patterns, the precipitation mechanisms and the local architecture ([Groves and Phillips, 1987](#)). Fluid-wall rock reaction is one type of gold depositional mechanism forming orogenic Archean gold deposits and this mechanism involves the destabilization of the $\text{Au}(\text{HS})_2^-$ complex. Such destabilization of gold can happen between reactions of fluids rich in sulfur and ferruginous host rocks, or ore fluids reduction and intense CO_2 and Ca metasomatism ([Mikucki, 1998](#); [Fougerouse et al., 2016](#)). The greenstone belts from the Crixás-Goiás Domain have hydrothermal alteration characteristics typical of orogenic deposits with a similar fluid condition, strong structural control, and enrichment of K, Fe, S, and CO_2 ([Goldfarb et al., 2001](#); [Jost et al., 2019](#); [Bogossian et al., 2021](#)).

Once it has formed, many processes could act to exhume, disperse, increase or overprint mineralization on deposits ([Hagemann et al., 2016](#)). Native gold in intensely weathered rocks can occur as millimetric grains and rarely as

gold nuggets (> 4 mm). Studies of this kind were carried out in Australia and those occurrences are usually composed of gold-silver alloys representing an hypogene gold deposit with thermal annealing, including corrosion features and silver depletion due to weathering (Hough et al., 2007). The presence of gold in weathering profile, in this case, is probably hypogenic in origin and concentrated in the soil as the residue of prolonged exposure and weathering of the Crixás orefield, therefore, not by supergene growth.

Previous studies show that gold occurs per differences in chemistry/mineralogy in hydrothermal deposits, and consequently, such alteration patterns provide the potential to find localities of gold mineralization (Neumayr et al., 2008; McCuaig et al., 2010). Direct detection technologies are available at the field scale during an exploration phase, for example, using soil geochemistry grids and detailed geophysical surveys such as Induced Polarization, Electromagnetic survey, ground magnetics, and gravity. However, they are invariably associated with false positive rates (McCuaig et al., 2010). Soil geochemistry grids in soil samples can only indicate the region with enrichment of Au colloidal or retention in a short area, but not the main ore body with an amount of gold that really represents the surface mineralization and can be properly explored (Fig. 2.10).

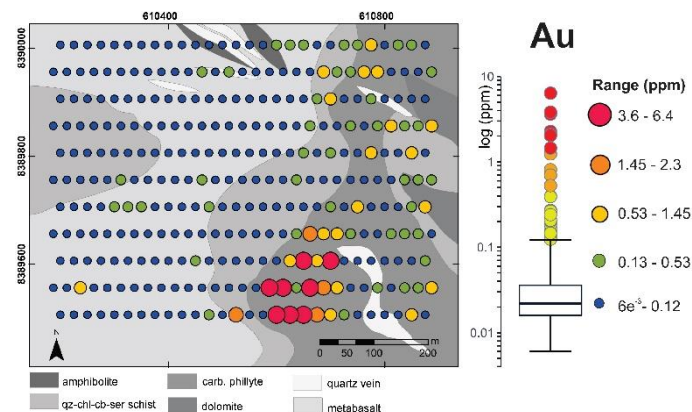


Fig. 2.10. Gold ranges in geochemical analysis from soil samples of level C from Pequizão deposit and the box-plot in log-ratio resulted in three different anomalies. Third-order anomalies in yellow, second-order anomalies in orange and first-order anomalies in red. Green are low priority and blue the background.

Soil samples reflect the rocky substrate with punctual representativeness, and consequently, geochemical soil surveys require a larger sampling density. For this reason, the application of detailed surveys is appropriate to determine and delimit the source of anomalies. Using geochemical data analysis is fundamental as a valuable source of information for mineral exploration. The advances in extraction and processing of data correlated to the anomalies have made significant progress in the last years (e.g., [Shuguang et al., 2015](#); [Levitan et al., 2015](#); [Ilyas et al., 2016](#)). This approach combined with integrated multivariate geochemical analysis of soil data can hold great potential in mineral research. In addition, adopting these techniques enables meaningful results related to mineralization and increases the understanding of many information collected in a survey, as demonstrated by [Wang et al. \(2017\)](#) and [Cevik et al. \(2021\)](#).

Gold occurrences in the Pequizão deposit are usually associated with carbonaceous phyllite as the main ore-rich lithotype, but previous prospective studies in soil samples showed another pattern from the primary alteration in drill cores. At the subsurface, the Pequizão ore body is mostly located in a zone with intense silicification and strongly controlled by the Structure IV, hosted in carbonaceous phyllite. It has disseminated sulfide mineralization that is generally represented by arsenopyrite, pyrrhotite, and subordinate chalcopyrite. Arsenopyrite is an index mineral of Structure IV and has high gold contents, making it one of the main prospective features in drill holes ([Ulrich et al., 2021](#)).

First results from the geochemical soil data clearly point out a hydrothermal alteration in the surface area showing a high concentration of gold-pathfinder elements such as Ag, As, and Sb. Collecting soil samples from three different levels from the soil, we noted a slight change in the concentration of the elements. Therefore, there is always a gradual upward or downward increase of one of the elements between the levels. In general, because samples from the deepest level (level C) were less impacted by weathering and pedogenic processes, they comprise higher concentrations.

Pearson's correlation (r) sustains the geochemical signature of a soil predominantly from basalt and amphibolite, as the associations with Cu, Cr, Fe,

and Ni, and its influences with the clay-like elements. K, P, and Zn, typical from weathering profiles, are more concentrated in the shallow samples at level A. Also, Ga, Hg, Te, and Sb have the same behavior and they are usually related to gold mineralization as hierarchical cluster confirms with the C1 grouping Sb, Te, Tl, Au, Ag, and Hg together. As shown by the dendrogram in [Fig. 2.4](#), Ag and Hg have a higher similarity among themselves rather than with Au which is also part of the group along with Tl and Te. The C3 group is probably clustering elements correlated to mineralized samples derived from metabasalts.

Comparing with a supervised method, as the ensemble learning, to predict the element importance to gold prospecting, we run it on Pequizão samples and put in evidence the presence of the C3 clustering (with Zn, Ga, Pb) and the high importance value as it indicates the metabasic rocks and we can assume that they demonstrate non-mineralized samples. Zinc deficiency is common on leached acid sandy soils and on soils with low Zn availability due to high pH (dolomitic and alkali soils) or high organic matter content (e.g., carbonaceous phyllite) ([Hooda, 2010](#)). On the basis of cluster analysis results, it is also interesting to point out that, although they are part of the same group, Co, Cu, and Ni show a slight degree of dissimilarity with Zn, As, and Pb. The second element with the most importance resulting from ensemble learning was Al, and it is related to the presence of gibbsite in the samples.

Complex mineralogical changes have been mapped, using spectral techniques such as Terra (field spectrometers) and ASTER (airborne spectrometers), that have taken place along weathering profiles above mineral deposits merging visible and infrared imaging spectrometry (e.g., [Ducart et al., 2006](#); [Sícole Seoane et al., 2009](#); [DURING et al., 2016](#); [Carrino et al., 2018](#); [Neal et al., 2018](#); [Laukamp et al., 2021](#)). Although these methods yield good results, they do not provide sufficient detailed mineralogical information of the deposit. Spectral techniques combined with mineralogical, structural, and lithogeochemical analyses can provide the characterization of footprints from gold-related hydrothermal alteration. Kaolinite, white mica, and chlorite display characteristic absorption feature in the SWIR region due to the presence of hydroxyls, and its intensity is relative to the abundances ([Clark et al., 1990](#)). However, studies of distal to proximal hydrothermal alteration vectors in surface

gold deposits in greenstone belts using spectral data are not well-established in the literature.

White micas, clay minerals, and chlorites in sedimentary rocks can be formed during diagenesis, hydrothermal alteration, and metamorphism, as well as being detrital minerals. In contrast with the nature of igneous rocks, sedimentary rocks may develop equivocal mineralogical and chemical compositions that require additional analyses other than whole-rock geochemical composition. Furthermore, weathering influences the physicochemical characteristics of minerals and is important in the formation of new minerals such as clay minerals (Galán, 2006).

The white mica at Pequizão weathering profile is more concentrated in the carbonaceous phyllite and abundant in dolomite (Fig. 2.11a). From an exploration panorama, white mica is relevant to target hydrothermal alteration minerals for spectral analysis because it is almost resistant to chemical weathering. The wavelength position provides information on mineral composition with changes induced by varying cation compositions. The position of the Al-OH in white mica matches with the chemical composition of white mica following Tschermak's exchange $Al_{IV} Al_{VI} \leftrightarrow Si_{IV} (Fe, Mg)_{VI}$, which involves coupled substitution between the tetrahedral and octahedral layers (Duke, 1994; Cloutier et al., 2021). The muscovite composition (high-Al) has a short wavelength and the phengite composition (low-Al) has a long wavelength in this feature. The area with phengite occurrence is relative gold-rich according to the geochemical and spectral analysis and also coincides with the proximity with Structure IV. Chlorite (mainly Fe-chlorite) abundance occurs with intensity in qt-chl-cb-ser schists (Fig. 2.11b), which is the Palmeiras Structure main indicative. Amphibole (mainly actinolite) occurs in wide distribution in the soil, located mainly in the metabasalts (Fig. 2.11c), differently from the distribution of gibbsite, with many abundant samples located in the central area, with a group with high abundance in qt-chl-cb-ser schist (Fig. 2.11d). Kaolinite has a great abundance in mafic samples, but the distribution on the map is random and not meaningful (Fig. 2.11e). The hematite-goethite occurrences are concentrated in the northwest part, distant from the two thrust-faults and the metasedimentary rocks (Fig. 2.11f).

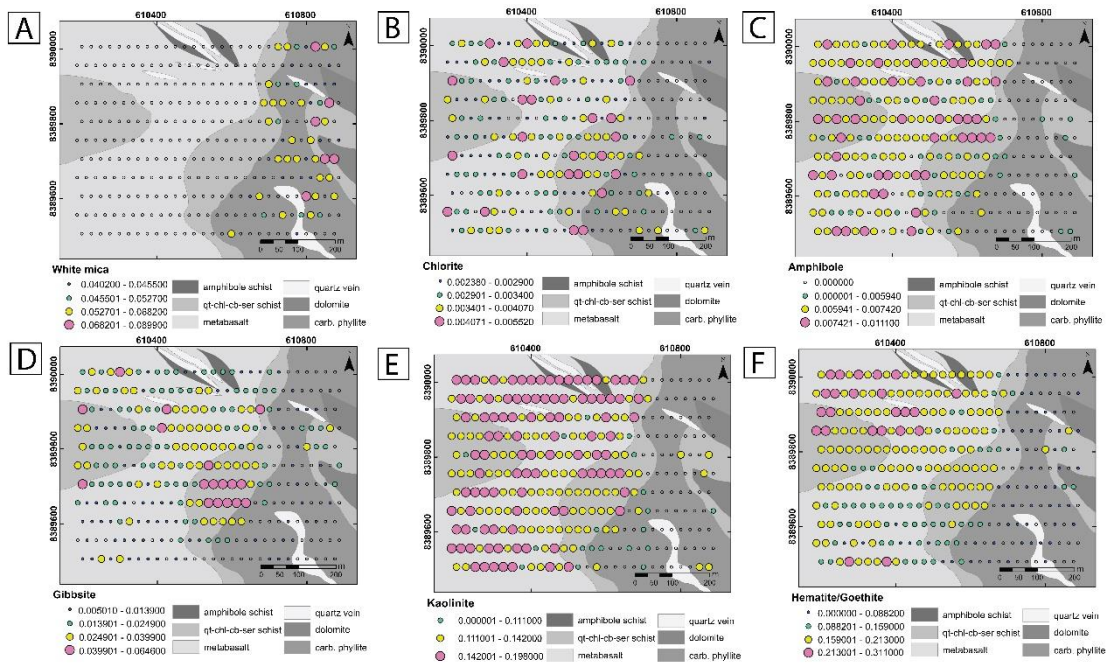
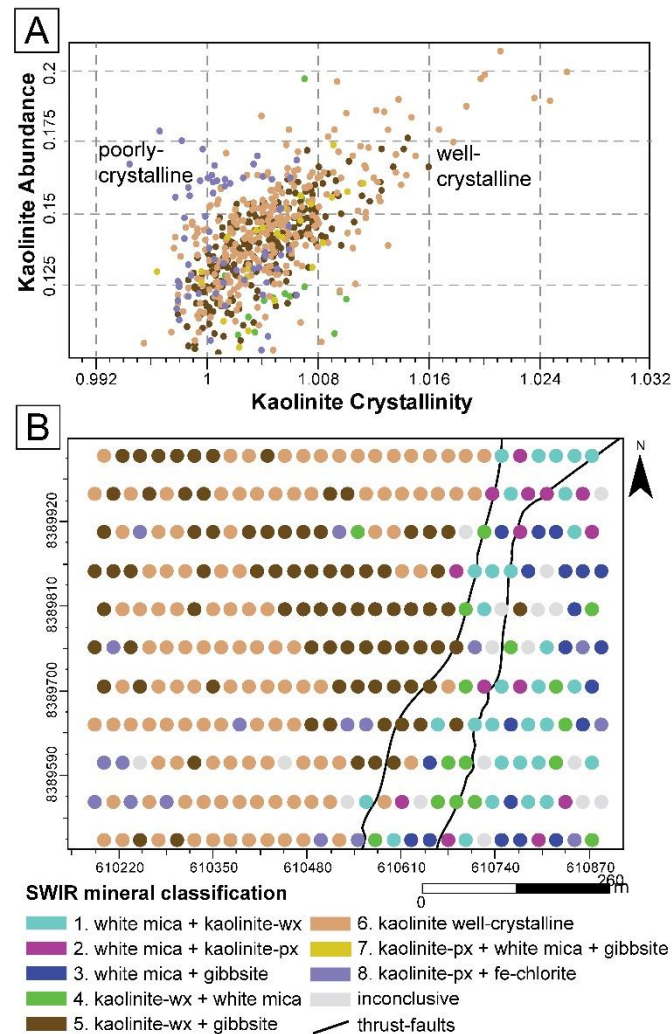


Fig. 2.11. Integrated map with the IRS results and their localization in the geological map, using a scale according to the abundance from: (a) white mica; (b) chlorite; (c) amphibole; (d) gibbsite; (e) kaolinite; (f) hematite/goethite.

Due to the abundance of potential vibrational modes in mineral assemblages, interpretation of spectral results is nontrivial and needs a precise understanding of the potential cause contributing to the reflectance spectra (Laukamp et al., 2021). Binary discrimination diagrams can aid in evaluating mineral composition and also crystallinity. According to Murray (1955), the relative intensity of absorption bands for kaolinite minerals is related to the kaolinite crystallinity. Well-crystallized (wx) kaolinite occurs in the majority of the analyzed samples and it is not restricted to any lithotype, indicating that the soil samples are not transported and have low weathering index. Poorly-crystallized (px) kaolinite is locally distributed in carbonaceous phyllites and rarely in metabasalts (Fig. 2.12a).

We classified the minerals from SWIR results into eight groups according to its abundance in each sample: (1) white mica + kaolinite-wx; (2) white mica + kaolinite-px; (3) white mica + gibbsite; (4) kaolinite-wx + white mica; (5) kaolinite-wx + gibbsite; (6) kaolinite-wx; (7) kaolinite-px + white mica + gibbsite; (8) kaolinite-px + Fe-chlorite. The map with the special analysis of this SWIR minerals classification (Fig. 2.12b) could explain the predominance of kaolinite-wx in the deposit and the association with gibbsite in the central area, near the thrust-zone, which is the same region extracted as the third factor results (Fig. 2.9e), putting in evidence the influence of gibbsite as a target mineral. The abundance of chlorite and amphibole in the soil sample is too low, so they were not grouped here.

Fig. 2.12. Mineral classification at SWIR spectra. (a) kaolin group index and crystallinity plot



for kaolinite crystallinity. (b) Map from the classification of the mineral and highlight the thrust-faults from the deposit.

The pioneer spectral studies in the Crixás greenstone belt ([Campos et al., 2022](#)) have shown contrasting primary alteration patterns at the rocks, represented differently in each regional structure, with epidote, muscovite + paragonite and Fe-chlorite at the Palmeiras Structure, and muscovite, Fe-chlorite and ankerite at the Structure IV. In the Pequizão deposit, we could not precisely define from the soil a mineral zonation that represents the spectral footprint for each structure, but near the thrust-zones, we can see an abundance of muscovite at the Structure IV and gibbsite at the Palmeiras Structure.

Subsurface measurement of samples from fresh rock combined with regolith surface sampling can present great results for gold exploration ([Cudahy, 2016](#); [Wang et al., 2017](#)). The subsurface studies from the Crixás greenstone belt can explain in detail its gold mineralization hydrothermal footprint. [Campos et al. \(2022\)](#) determined the main mineralogical and geochemical characteristics of the hydrothermal alteration haloes associated with both thrust zones that also control the mineralization in the Pequizão deposit: (a) the Palmeiras Structure, is a low-angle thrust fault hosted in the hydrothermalized metabasalt (qz-chl-cb-ser schist), next to the contact with the metasedimentary sequence and exhibits an intense hydrothermal alteration association ([Massucato, 2004](#)). The distal alteration zone from the ore comprises Fe-Mg chlorite, biotite, ankerite, and quartz. The intermediate zone comprises epidote, ankerite, Fe-Mg chlorite and, muscovite, and magnetite, which appear gradually near the proximal halo, and. In the proximal zone, the predominance of Fe-chlorite, muscovite/paragonite, and magnetite occur ([Campos et al., 2022](#)); (b) Structure IV, the main control of the Pequizão deposit, is defined as a high-strain deformation zone hosted in the carbonaceous schist that also displays hydrothermal alteration haloes related. The distal zone is defined by as Fe-Mg chlorite and muscovite enrichment. The intermediate zone has an increase in quartz, ankerite, and muscovite. The proximal zone has a strong presence of quartz, oligoclase, tourmaline, and ankerite, with sulfides as arsenopyrite, pyrrhotite, and some chalcopyrite. In brief, there is silicification, carbonatization, and phyllic alteration occurring ([Campos et al., 2022](#)).

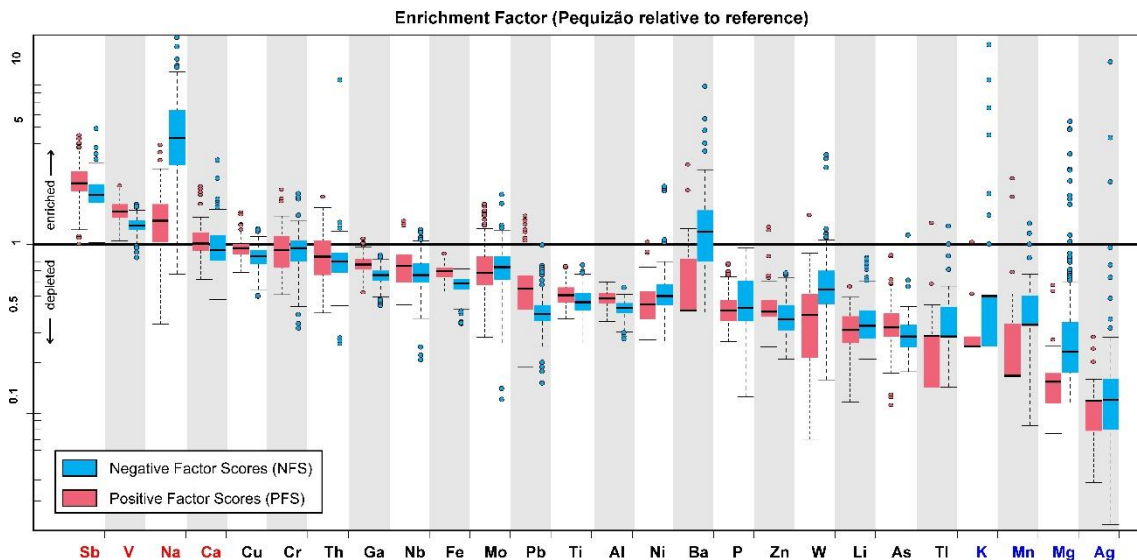
In the lithogeochemical compositions, weathered profiles have the first components from PCA usually capturing relationships of variables that reflect the

lithology underneath. If a component expresses parental rocks, then a map of that component will outline the major lithological variation of the area. When the overburden is thick, linear combinations of variables and the loadings are hard to interpret as they reflect several surficial processes (Grunsky, 2010). The elements from the PC1 results show a clear distinction between mafic samples from the elemental association typical of the clastic silicate components and common immobile elements in hydrothermal settings (e.g., Al, Ti, Zr, etc). PC3 is potentially tracking the enrichment of Au and As and the concurrent depletion of Zn. Lin et al. (2014) confirm that the scores of the third factor can be explained as the surface expression of the concealed mineralization and, consequently, Ag, As, and Zn are the indirect indicators of the mineralization. Therefore, the method can be applied to regional geochemical data for further exploration of minerals of the same kind.

Maps of the scores can help understand geochemical processes. Components that describe rare processes such as mineralization or alteration can likewise be expressed precisely on maps that display the scores as in Fig. 2.9. We put in evidence the hydrothermal alteration in the area in the second factor, where Au, Ag, Ca, and Mg occur as negative scores. Gold occurrences in the Pequizão deposit at the underground mine show that the top of the mineralization occurs in a phyllic alteration zone grading to a potassic alteration zone near the silicification with high-grade ore, and a base of carbonatization zone in the wall-rock of metagraywacke (Sobiesiak, 2011). On the surface, there are quartz veins as a result of the silicification, but the phyllic and carbonatization zone do not clearly appear, and only an enrichment of Zn, Pb, Fe, and Ga is delimited at north in Palmeiras Structure in the area. For this reason, we measured the “enrichment factor” (Fig. 2.13) for these Factor 2 zones to clarify and characterize the differences from the geochemical data between the mineralized zone (near thrust-faults) with the distal zone (northwest side) in metabasalt samples (see Fig. 2.9b). The normalization sample, called here *reference soil sample*, was collected at the top of the drill core sampling from

Campos et al. (2022), where the localization is ca. 2 km from the thrust-zone and the deposit.

Fig. 2.13. Tukey boxplot of enrichment factor in the Pequizão samples from Factor 2 comparing



the Negative Factor Scores (NFS) and the Positive Factor Scores (PSF) relative to the reference sample. The horizontal line at $EF = 1$ represents equivalence between the Pequizão and reference ratios. Elements with different X_{Peq}/X_{Ref} ratios between Pequizão and reference soil samples are labeled in bold, with significantly enriched elements in red, and significantly depleted elements in blue.

The enrichment factor was calculated for Pequizão using the Factor 2 from FA (Fig. 2.9c) to select the samples of negative factor scores (NFS), at the southwest area, and the positive factor scores (PFS) from metabasalt samples. This method was employed to better understand the Factor 2 results and to compare the variation between the elements. In this case, mobile elements should display stair-stepping patterns (different from 1), and immobile elements should have identical portions near 1. Elements with EFs greater than 1 from NFS samples were Na, Sb, V, Ba, and from PFS were Sb, V, Na, Ba, in decreasing order (Fig. 2.13). The most depleted elements in the set, EFs less than 0.5, were Pb, Ti, Al, Ni, Ba, P, Zn, W, Li, As, Tl, K, Mn, Mg, and Ag, with Ba from NFS samples having EF larger than 1. The NFS samples are gold-related in Factor 2, in contrast to the PFS samples that are distant to thrust-faults and with no mineralized samples. Comparing the NFS with PFS samples, the most significant differences were the increase of Na, Ba, W, Tl, and Mg at NFS samples.

Analyzing the enrichment result in Pequizão, we can clearly point out the variance of Na, V, Tl, Sb, and Ba is similar to those found in the Palmeiras Structure footprint (Campos et al., 2022). Although several trace elements are enriched in the Pequizão soil samples relative to the surrounding soil sample, other trace elements, including those described as pathfinders in gold deposits, as As, Ag, and W, are not significantly different and depleted between the Pequizão soil and reference soil sample.

The element mobility is controlled by the mineralogical changes that take place during alteration and the composition, temperature and relative volume of the migrating fluid. According to Pearce (1983), usually, the low field strength group (LILE) are incompatible elements, therefore, fluid-mobile (e.g., K, Ba, and Pb), and the high field strength elements (HFSE, e.g., REE, Th, Zr, Ti, Nb, and P) are immobile, but its mobility is dependent on conditions of the fluid, such as temperature, pH and oxygen fugacity, and its fluid to rock ratio (Perrouty et al., 2019). There is no simple correlation with the degree of mobility of HFSE and rock type, highlighting the importance of local control of fluids and minerals. For example, REE are stable in the presence of an aqueous fluid, but can be more easily released from a glassy basalt during alteration than from a crystalline rock and that REE can be mobilized by mineralizing fluids carbonate- or halogen-rich in a rock (Humphries, 1984; Rollinson and Pease, 2021).

The Pequizão orebody, in the subsurface, is mostly characterized by underground exploration using drill cores sampling. A cross-section simplifying the geological architecture in the study area could show how the massive mineralization is parallel to the lithologies and basically controlled by the thrust-faults (Fig. 2.14a,b). That known mineralization at the subsurface has a different signature than the mineralization at the surface. The soil formation above these ore fields can concentrate a great amount of ore and adsolves important elements that facilitate prospecting targets near mines. According to Bogossian et al. (2021), in Crixás the hydrothermal fluid was channeled in thrust faults, developing Au-bearing quartz ± carbonate veins, and only localized flow was originated by later shear-zones. The resulting alteration typically overprints metamorphic assemblages, but distal alteration can be subtle and hardly distinguished from the metamorphic host rocks. The orebodies are typically parallel to foliation and

adjacent to metabasalts and carbonaceous phyllites contact (Fig. 2.14a,b). The fluids that occurred in Crixás were highly saline in the H₂O-CO₂-NaCl-KCl-CH₄-N₂ system and with temperatures ranging from 350° to 450° C with pressures of 2 to 3 kbars (Fortes, 1996; Fortes et al., 2000). Elements that can best define the Pequizão footprint, besides Au, there are Ag, As, Ba, Na, Sb, Te, Tl, and W (not in order of importance). The spectral analysis pointed out the high content of muscovite that tends to phengite in gold-rich samples from the Structure IV zone. In the same region, there is a kaolinite-wx + gibbsite occurrence that marks the proximal area in the Palmeiras Structure and it is the main mineral footprint of ore occurrence. Due to the atmospheric deposition of S (mainly SO₂), there is soil acidification and this increases the Al solubility in soils forming soluble Al(OH)SO₄. Sulfides oxidation present in the soil, like pyrite/arsenopyrite, produces H₂SO₄, interacting with other minerals to yield Al ions, which then fills exchange sites (Kabata-Pendias, 2011). Of the elements related to gold, As is an example of effective footprint and, together with gibbsite, can point out great targets. The As is mobilized, under reducing conditions, by the reductive dissolution of iron or manganese oxides and it can be readsorbed due to the sorptive properties of the soil. The gibbsite is especially efficient in adsorbing or readsorbing As, as is the remaining unreduced fraction of the iron oxides (Mello et al., 2006).

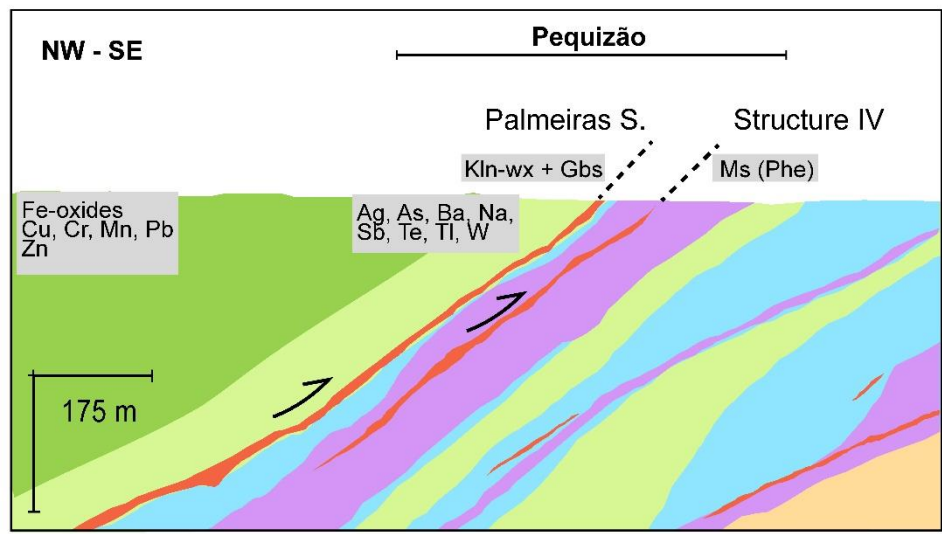
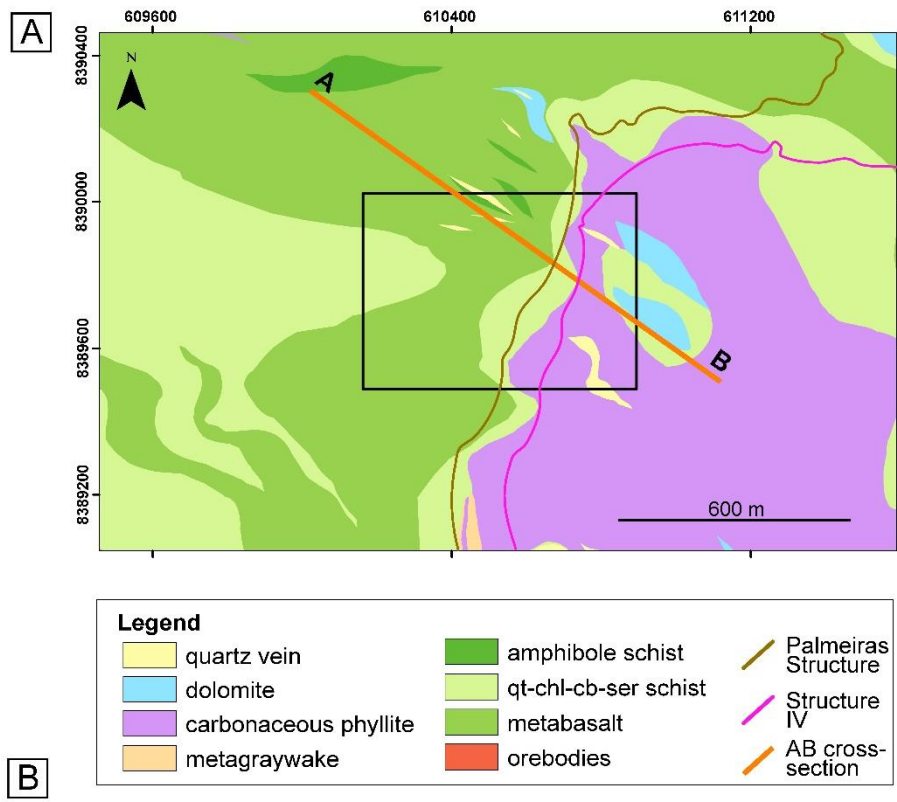


Fig. 2.14. (a) Geological map from Crixás greenstone belt highlighting the Pequizão deposit, scale 1:10.000; (b) Cross-section (A-B) from the deposit, according to previous studies in the area using drill cores, modified after [Borges et al., 2021b](#). Abbreviations (after [Warr, 2021](#)): Gbs = gibbsite; Kln-wx = Well-crystalline kaolinite; Ms (Phe) = muscovite tending to phengite.

6. Conclusion

The Pequizão deposit in the Crixás greenstone belt is inserted in a typical orogenic gold terrain. This is a pioneering exploratory studying of reflectance spectroscopy, machine learning, and multivariate soil analysis on soil samples enhancing the knowledge discovery from mineral maps and geochemical data providing the delineation of new exploration vectors. The integration of clustering, ensemble learning, reflectance spectroscopy, and factor analysis of the weathering profile can classify them as a potential gold source in the Crixás greenstone belt.

On the whole, the spectral data reveal that the mineralization at Pequizão is associated with gibbsite, white mica tending to phengite, and well-crystallized kaolinite as pathfinder minerals in the brownfield explorations.

Hierarchical clustering of geochemical data was able to differentiate elements associated with metabasic and metasedimentary rocks and highlight gold-bearing elements. Combined with ensemble learning, Zn, Ga, and Pb stand out as elements of greatest importance in predicting Au content, with Factor Analysis demonstrating that these elements are related to gold absence. However, the importance of the Al is related to the gibbsite.

The Principal Component Analysis allowed the identification of carbonaceous phyllite and dolomite as the main mineralized parental rock. PC1 captured the characteristics of the parental rocks units and successfully separated them. PC3 clarified the mineralization signatures and representative elements of the hydrothermal alteration in the surface as As, Ag, Te, Ca, and Mg. Comparing Pequizão samples with distal and non-mineralized soil, we could infer Sb, V, Na, Ba, and W enrichment.

It is concluded that the characterization of surface mineralized zones using those integrated methods can also delineate supposedly gold-rich areas on the basis of this footprint from the Pequizão deposit. Besides, to elevate the prospecting accuracy, it would be recommended to use these methods on Pequizão drill holes and directly compare with our soil-based results. Otherwise,

a spectral analysis in greenfield areas allied with the brownfield techniques could greatly improve district-wise exploration.

Acknowledgments

This project is part of the Agreement for Cooperation and Technical Scientific Support between UnB and the Anglogold Ashanti - Serra Grande Unit - SGU (Number: 23106.006453/2016-32 - Geosciences Institute / UnB - Code 9961). We are grateful to the SGU for providing the soil samples and geochemical data and supporting the project. A.M. Silva and C.L.B. Toledo acknowledge the CNPq for their scientific grants. This work also was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (Grant Number: 88882.347173/2019-01).

Supplementary data

Supplementary data to this article can be found in the dissertation appendices.

7. References

- Abdi, H., and Williams, L.J., 2010, Principal component analysis: Wiley Interdisciplinary Reviews: Computational Statistics, v. 2, p. 433–459, doi:10.1002/wics.101.
- Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, London.
- Almeida, F.F.M., Hasui, Y., de Brito Neves, B.B., and Fuck, R.A., 1981, Brazilian structural provinces: An introduction: Earth Science Reviews, v. 17, p. 1–29, doi:10.1016/0012-8252(81)90003-9.
- Almeida, F. R., 2021. Caracterização Da Assinatura Da Mineralização Polimetálica Do Greenstone Belt De Faina (GO). Dissertação de Mestrado. Instituto de Geociências, Universidade de Brasília, 104 p.
- Anand, R.R., Cornelius, M., and Phang, C., 2007, Use of vegetation and soil in mineral exploration in areas of transported overburden, Yilgarn Craton, Western Australia: A contribution towards understanding metal transportation processes: Geochemistry: Exploration, Environment, Analysis, v. 7, p. 267–288, doi:10.1144/1467-7873/07-142.

- Arndt, N.T., Teixeira, N.A., and White, W.M., 1989, Bizarre geochemistry of komatiites from the Crixás greenstone belt, Brazil: *Contributions to Mineralogy and Petrology*, v. 101, p. 187–197, doi:10.1007/BF00375305.
- Bacha, R.R.B., Toledo, C.L.B., Silva, A.M., and Mungall, J.E., 2021, Alkali element mobility as a vector for gold mineralization: An example from Central Brazil: *Geochemistry*, v. 81, p. 125808, doi:10.1016/j.chemer.2021.125808.
- Beghelli Junior, L.P., 2012, Charnockitos e ortognaisses da porção Centro-Oeste do Bloco Arqueano de Goiás: Dados geoquímicos e isotópicos: Universidade de Brasília.
- Bishop, J.L., Lane, M.D., Dyar, M.D., and Brown, A.J., 2008, Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas: *Clay Minerals*, v. 43, p. 35–54, doi:10.1180/claymin.2008.043.1.03.
- Bogossian, J., Kemp, A.I.S., and Hagemann, S.G., 2021, Linking gold systems to the crust-mantle evolution of archaic crust in central Brazil: *Minerals*, v. 11, doi:10.3390/min11090944.
- Bohachevsky, I.O., Johnson, M.E., Stein, M.L., 1986. Generalized Simulated Annealing for Function Optimization. *Technometrics*, v. 28, n. 3, p. 209–217. <https://doi.org/10.2307/1269076>
- Borges, C.C.A., Toledo, C.L.B., A.M., Silva, Kirk, J., Ruiz, J., Chemale Jr., F., Souza, R. G., Santos, B.A., Campos, M.P., Campos, L.M., Santos, A.M., 2021b, Archaean to Paleoproterozoic evolution of the Crixás greenstone belt, Central Brazil: Insights from two contrasting assemblages of metaigneous rocks: *Lithos*, v. 404–405, p. 106493, doi:10.1016/j.lithos.2021.106493.
- Borges, C.C.A., Toledo, C.L.B., Silva, A.M., Chemale, F., dos Santos, B.A., Figueiredo, F.L., and Zacchi, É.N.P., 2021a, Unraveling a hidden Rhyacian magmatic arc through provenance of metasedimentary rocks of the Crixás greenstone belt, Central Brazil: *Precambrian Research*, v. 353, doi:10.1016/j.precamres.2020.106022.
- Breiman, L. 1996. Bagging predictors. *Machine Learning*, v. 24, p. 123–140. <https://doi.org/10.1007/BF00058655>
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., and Reinsch, T.G., 2006, Global soil characterization with VNIR diffuse reflectance spectroscopy: *Geoderma*, v. 132, p. 273–290, doi:10.1016/j.geoderma.2005.04.025.
- Campos, L.M., Toledo, C.L.B., Silva, A.M., Ducart, D.F., Santos, B. A., Campos, M.P., Borges, C.C.A., 2022, The Hydrothermal Footprint of Crixás Orogenic Gold Deposit, Central Brazil. *Ore Geology Reviews*, (accepted and in minor reviews).
- Carranza, E.J.M., 2010, Mapping of anomalies in continuous and discrete fields of stream sediment geochemical landscapes: *Geochemistry: Exploration, Environment, Analysis*, v. 10, p. 171–187, doi:10.1144/1467-7873/09-223.
- Carrino, T.A., Crósta, A.P., Toledo, C.L.B., and Silva, A.M., 2018, Hyperspectral remote sensing applied to mineral exploration in southern Peru: A multiple data integration approach in the Chapi Chiara gold prospect: *International Journal of Applied Earth Observation and Geoinformation*, v. 64, p. 287–300, doi:10.1016/j.jag.2017.05.004.

- Cattell, R. B. 1966. The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. doi:10.1207/s15327906mbr0102_10
- Cevik, I.S., Olivo, G.R., and Ortiz, J.M., 2021, A combined multivariate approach analyzing geochemical data for knowledge discovery: The Vazante – Paracatu Zinc District, Minas Gerais, Brazil: *Journal of Geochemical Exploration*, v. 221, p. 1–18, doi:10.1016/j.gexplo.2020.106696.
- Chayes, F., 1960, On Correlation between Variables of Constant Sum, *Journal of Geophysical Research*, v.65, n12, p. 4185-4193, doi:10.1029/JZ065i012p04185
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., 2021. xgboost: Extreme Gradient Boosting. R package version 1.4.1.1. <https://CRAN.R-project.org/package=xgboost>
- Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7:e623. <https://doi.org/10.7717/peerj-cs.623>
- Christy, C.D., 2008, Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy: *Computers and Electronics in Agriculture*, v. 61, p. 10–19, doi:10.1016/j.compag.2007.02.010.
- Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., and Vergo, N., 1990, High spectral resolution reflectance spectroscopy of minerals: *Journal of Geophysical Research*, v. 95, doi:10.1029/jb095ib08p12653.
- Clark, R.N., Swayze, G.A., Wise, R.A., Livo, K.E., Hoefen, T.M., Kokaly, R.F. and Sutley, S.J., 2007. USGS digital spectral library splib06a (No. 231). US Geological Survey. doi: 10.3133/ds231
- Cloutier, J., Piercey, S.J., and Huntington, J., 2021, Mineralogy, mineral chemistry and swir spectral reflectance of chlorite and white mica: *Minerals*, v. 11, doi:10.3390/min11050471.
- Comas-Cufi, M., Thió-Henestrosa, S., 2011, CoDaPack 2.0: a stand-alone, multiplatform compositional software. In: Egozcue, J.J., Tolosana-Delgado, R., Ortego, M.I. (Eds), *Proceedings of the CoDaWork'11: 4th International Workshop on Compositional Data Analysis*, Sant Feliu de Guíxols.
- Cordeiro, P.F. de O., and Oliveira, C.G. de, 2017, The Goiás Massif: Implications for a pre-Columbia 2.2–2.0 Ga continent-wide amalgamation cycle in central Brazil: *Precambrian Research*, v. 298, p. 403–420, doi:10.1016/j.precamres.2017.06.021.
- Cudahy, T.J., and Ramanaidou, E.R., 1997, Measurement of the hematite: Goethite ratio using field visible and near-infrared reflectance spectrometry in channel iron deposits, Western Australia: *Australian Journal of Earth Sciences*, v. 44, p. 411–420, doi:10.1080/08120099708728322.

- Cudahy, T., 2016, Mineral mapping for exploration: An Australian journey of evolving spectral sensing technologies and industry collaboration: *Geosciences (Switzerland)*, v. 6, doi:10.3390/geosciences6040052.
- Ducart, D. F., Crósta, A. P., Filho, C. R. S., & Coniglio, J. 2006. Alteration mineralogy at the Cerro La Mina epithermal prospect, Patagonia, Argentina: Field mapping, short-wave infrared spectroscopy, and ASTER images. *Economic Geology*. 101(5), 981-996
- Duke, E.F., 1994, reaction progress: Implications for remote sensing:, doi:10.1130/0091-7613(1994)022<0621.
- Duuring, P., Hassan, L., Zelic, M., and Gessner, K., 2016, Geochemical and spectral footprint of metamorphosed and deformed VMS-style mineralization in the Quinns district, Yilgarn craton, Western Australia: *Economic Geology*, v. 111, p. 1411–1438, doi:10.2113/econgeo.111.6.1411.
- Ferreira, B.F., Marques, J.C., Frantz, J.C., de Souza, R.G., Campos, M.P., Figueiredo, F.L.A. de R., and Padilha, N.L., 2021, Stratigraphy and U–Pb detrital zircon provenance, Crixás Greenstone Belt, Goiás-Brazil: Gold-bearing vs barren siliciclastic rocks: *Journal of South American Earth Sciences*, v. 105, doi:10.1016/j.jsames.2020.102994.
- Filzmoser, P., Hron, K., and Templ, M., 2018, Applied compositional data analysis: with worked examples in R: v. 89, 3165–3165 p., doi:10.1080/00949655.2019.1628880.
- Fortes, P. de T.F. de O., 1996, Metalogenia dos Depósitos Auríferos Mina III, Mina Nova e Mina Inglesa, Greenstone Belt de Crixás-GO: Universidade de Brasília, 242 p.
- Fortes, P. de T.F. de O., Coelho, R.F., Giuliani, G., 2000, Au/Ag Ratio Variations At Mina III, Mina Nova and Mina Inglesa Gold Deposits, Crixás Greenstone Belt, Brazil I: *Revista Brasileira de Geociências*, v. 30, p. 246–250.
- Fortes, P.D.T.F.D.O., Coelho, R.F., And Giuliani, G., 2000, Au/Ag Ratio Variations At Mina III, Mina Nova and Mina Inglesa Gold Deposits, Crixás Greenstone Belt, Brazil: *Revista Brasileira de Geociências*, v. 30, p. 246–250, doi:10.25249/0375-7536.2000302246250.
- Fougerouse, D. et al., 2016, Gold remobilisation and formation of high grade ore shoots driven by dissolution-reprecipitation replacement and Ni substitution into auriferous arsenopyrite: *Geochimica et Cosmochimica Acta*, v. 178, p. 143–159, doi:10.1016/j.gca.2016.01.040.
- Freund, Y., Schapire, R.E., 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, v. 55, n. 1, p. 119–139. <https://doi.org/10.1006/jcss.1997.1504>
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, v. 29, n. 5, p. 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Fuck, R.A., Dantas, E.L., Pimentel, M.M., Botelho, N.F., Armstrong, R., Laux, J.H., Junges, S.L., Soares, J.E., and Praxedes, I.F., 2014, Paleoproterozoic crust-formation and reworking events in the Tocantins Province, central Brazil: A contribution for Atlantica supercontinent reconstruction: *Precambrian Research*, v. 244, p. 53–74, doi:10.1016/j.precamres.2013.12.003.
- Galán, E., 2006, Chapter 14 Genesis of Clay Minerals: *Developments in Clay Science*, v. 1, p. 1129–1162, doi:10.1016/S1572-4352(05)01042-1.

- Gielar, A., Helios-Rybicka, E., Möller, S., and Einax, J.W., 2012, Multivariate analysis of sediment data from the upper and middle Odra River (Poland): *Applied Geochemistry*, v. 27, p. 1540–1545, doi:10.1016/j.apgeochem.2012.04.004.
- Gladwell, D.R., Lett, R.E., and Lawrence, P., 1983, Application of reflectance spectrometry to mineral exploration using portable radiometers (Abitibi, Ontario): *Economic Geology*, v. 78, p. 699–710, doi:10.2113/gsecongeo.78.4.699.
- Goldfarb, R.J., and Groves, D.I., 2015, Orogenic gold: Common or evolving fluid and metal sources through time: *Lithos*, v. 233, p. 2–26, doi:10.1016/j.lithos.2015.07.011.
- Goldfarb, R.J., Groves, D.I., and Gardoll, S., 2001, Orogenic gold and geologic time: A global synthesis: *Ore Geology Reviews*, v. 18, p. 1–75, doi:10.1016/S0169-1368(01)00016-6.
- Greenwell, B.M, Boehmke, B.C., 2020. Variable Importance Plots—An Introduction to the vip Package. *The R Journal*, v. 12, n. 1, 343–366. <https://doi.org/10.32614/RJ-2020-013>.
- Groves, D.I., and Phillips, G.N., 1987, The genesis and tectonic control on archaean gold deposits of the Western Australian Shield - A metamorphic replacement model: *Ore Geology Reviews*, v. 2, p. 287–322, doi:10.1016/0169-1368(87)90009-6.
- Groves, D.I., Santosh, M., Deng, J., Wang, Q., Yang, L., and Zhang, L., 2020, A holistic model for the origin of orogenic gold deposits and its implications for exploration: *Mineralium Deposita*, v. 55, p. 275–292, doi:10.1007/s00126-019-00877-5.
- Groves, D.I., Santosh, M., Goldfarb, R.J., and Zhang, L., 2018, Structural geometry of orogenic gold deposits: Implications for exploration of world-class and giant deposits: *Geoscience Frontiers*, v. 9, p. 1163–1177, doi:10.1016/j.gsf.2018.01.006.
- Grunsky, E.C., 2010, The interpretation of geochemical survey data: *Geochemistry: Exploration, Environment, Analysis*, v. 10, p. 27–74, doi:10.1144/1467-7873/09-210.
- Haest, M., Cudahy, T., Laukamp, C., and Gregory, S., 2012, Quantitative mineralogy from infrared spectroscopic data. I. Validation of mineral abundance and composition scripts at the rocklea channel iron deposit in Western Australia: *Economic Geology*, v. 107, p. 209–228, doi:10.2113/econgeo.107.2.209.
- Hagemann, S.G., Lisitsin, V.A., and Huston, D.L., 2016, Mineral system analysis: Quo vadis: *Ore Geology Reviews*, v. 76, p. 504–522, doi:10.1016/j.oregeorev.2015.12.012.
- Harraden, C.L., McNulty, B.A., Gregory, M.J., and Lang, J.R., 2013, Shortwave infrared spectral analysis of hydrothermal alteration associated with the Pebble porphyry copper-gold-molybdenum deposit, Iliamna, Alaska: *Economic Geology*, v. 108, p. 483–494, doi:10.2113/econgeo.108.3.483.
- Herrmann, W., Blake, M., Doyle, M., Huston, D., Kamprad, J., Merry, N., and Pontual, S., 2001, Short Wavelength Infrared (SWIR) spectral analysis of hydrothermal alteration zones associated with base metal sulfide deposits at Rosebery and Western Tharsis, Tasmania, and Highway-Reward, Queensland: *Economic Geology*, v. 96, p. 939–955, doi:10.2113/gsecongeo.96.5.939.
- Hooda, P.S., 2010. Trace elements in soils, first ed. John Wiley & Sons Ltd., Chichester, United Kingdom. doi:10.1002/9781444319477

- Hough, R.M., Butt, C.R.M., Reddy, S.M., and Verrall, M., 2007, Gold nuggets: Supergene or hypogene? *Australian Journal of Earth Sciences*, v. 54, p. 959–964, doi:10.1080/08120090701488289.
- Humphris, S.E., 1984, The mobility of the rare earth elements in the crust.: *Rare Earth Element Geochemistry*, p. 317–342, doi:10.1016/b978-0-444-42148-7.50014-9.
- Ilyas, A., Kashiwaya, K., and Koike, K., 2016, Ni grade distribution in laterite characterized from geostatistics, topography and the paleo-groundwater system in Sorowako, Indonesia: *Journal of Geochemical Exploration*, v. 165, p. 174–188, doi:10.1016/j.gexplo.2016.03.002.
- Jolliffe, I. T., 2002, *Principal Component Analysis*, 2 ed., Springer Series in Statistics. 488 p. doi:10.1007/b98835
- Jolliffe, I.T., and Cadima, J., 2016, Principal component analysis: A review and recent developments: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, doi:10.1098/rsta.2015.0202.
- Jost, H., and De Oliveira, A.M., 1991, Stratigraphy of the greenstone belts, Crixás region, Goiás, central Brazil: *Journal of South American Earth Sciences*, v. 4, p. 201–214, doi:10.1016/0895-9811(91)90031-F.
- Jost, H., Carmelo, A.C., Meneses, P., 1994a, Litologias, relações de contato e estrutura do Bloco Caiamar, Crixás, Goiás, in SBG, *Simpósio de Geologia do Centro-Oeste IV*, p. 18–20.
- Jost, H., Vargas, M.C., Fuck, R.A., Kuyumjian, R.M., Pires, A.C.B., 1994b, Relações de contato, litologias, geologia estrutural e geofísica do Bloco Arqueano do Moquém, Crixás, Goiás, in SBG, *Simpósio de Geologia do Centro-Oeste IV*, p. 21–23.
- Jost, H., Theodoro, S.M.D.C.H., Figueiredo, A.M.G., And Boaventura, G.R., 1996, Propriedades Geoquímicas E Proveniência De Rochas Metassedimentares Detríticas Arqueanas Dos Greenstone Belts De Crixás E Guarinos, Goiás: *Revista Brasileira de Geociências*, v. 26, p. 151–166, doi:10.25249/0375-7536.1996151166.
- Jost, H. et al., 2001, Geologia De Terrenos Arqueanos E Proterozóicos Da Região De Crixás-Cedrolina, Goiás: *Revista Brasileira de Geociências*, v. 31, p. 315–328, doi:10.25249/0375-7536.2001313315328.
- Jost, H., and Fortes, P.T.F.O., 2001, Gold deposits and occurrences of the Crixás Goldfield, Central Brazil: *Mineralium Deposita*, v. 36, p. 358–376, doi:10.1007/s001260100171.
- Jost, H., Bennel, M.R., Massucatto, A.J. 2005. Sistemas de alteração hidrotermal contrastantes no distrito aurífero de Crixás, Goiás. In: *I Simpósio Brasileiro de Metalogenia*. Gramado, RS. Resumos expandidos do I Simpósio Brasileiro de Metalogenia, p. 1-3.
- Jost, H., Chemale, F., Dussin, I.A., Tassinari, C.C.G., and Martins, R., 2010, A U-Pb zircon Paleoproterozoic age for the metasedimentary host rocks and gold mineralization of the Crixás greenstone belt, Goiás, Central Brazil: *Ore Geology Reviews*, v. 37, p. 127–139, doi:10.1016/j.oregeorev.2010.01.003.
- Jost, H., Junior, F.C., Fuck, R.A., and Dussin, I.A.Ô., 2013, Uvá Complex, The Oldest Orthogneisses Of The Archean-Paleoproterozoic Terrane Of Central Brazil: *Journal of South American Earth Sciences*, v. 47, p. 201–212, doi:10.1016/j.jsames.2013.07.002.

- Jost, H., Carvalho, M.J. de, Rodrigues, V.G., and Martins, R., 2014, Metalogenia dos Greenstone Belts De Goiás, in Silva, M. da G. da, Rocha Neto, M.B., Jost, H., and Kuyumjian, R.M. eds., *Metalogênese das províncias tectônicas brasileiras*, Belo Horizonte, CPRM, p. 141–168.
- Jost, H., Hofmann Apollo, J.F., Weber, W., Salles, R. dos R., Marques, J.C., Massucatto, A.J., Costa, D.A., and Araújo dos Santos, B., 2019, Stratigraphic update, paleotectonic, paleogeographic, and depositional environments of the Crixás Greenstone Belt, Central Brazil: *Journal of South American Earth Sciences*, v. 96, p. 102329, doi:10.1016/j.jsames.2019.102329.
- Kabata-Pendias, A., 2011, *Trace elements in soils and plants* (4th ed.). CRC press, v. 4, 548 p., doi:10.1201/b10158
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, p. 3146–3154. <https://dl.acm.org/doi/10.5555/3294996.3295074>
- Kirkpatrick, S., Gelatt, D., Vecchi, M., 1983. Optimization by Simulated Annealing. *Science*, v. 220, n. 4598, p. 671–680., <https://doi.org/10.1126/science.220.4598.671>
- Kokaly, R.F., Clark, R.N., Swayze, G.A., Livo, K.E., Hoefen, T.M., Pearson, N.C., Wise, R.A., Benzel, W.M., Lowers, H.A., Driscoll, R.L., and Klein, A.J., 2017, USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035, 61 p., <https://doi.org/10.3133/ds1035>.
- Kuhn, M., 2021. finetune: Additional Functions for Model Tuning. R package version 0.1.0. <https://CRAN.R-project.org/package=finetune>
- Kuhn, M., Wickham, H., 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. <https://www.tidymodels.org>
- Kuyumjian, R.M., and Jost, H., 2006, Low- and high-alumina komatiites of Goiás, Central Brazil: *Journal of South American Earth Sciences*, v. 20, p. 315–326, doi:10.1016/j.jsames.2005.11.003.
- Lacerda Filho, J.V. de, 1999, Programa Levantamentos Geológicos Básicos do Brasil. *Geologia e Recursos Minerais do Estado de Goiás e Distrito Federal: CPRM*, v. Projeto de, p. 200.
- Lampinen, H.M., Laukamp, C., Occhipinti, S.A., and Hardy, L., 2019, Mineral footprints of the Paleoproterozoic sediment-hosted Abra Pb-Zn-Cu-Au deposit Capricorn Orogen, Western Australia: *Ore Geology Reviews*, v. 104, p. 436–461, doi:10.1016/j.oregeorev.2018.11.004.
- Laukamp, C., Rodger, A., Legras, M., Lampinen, H., Lau, I.C., Pejčić, B., Stromberg, J., Francis, N., and Ramanaidou, E., 2021, Mineral physicochemistry underlying feature-based extraction of mineral abundance and composition from shortwave, mid and thermal infrared reflectance spectra: *Minerals*, v. 11, doi:10.3390/min11040347.
- Levitan, D.M., Zipper, C.E., Donovan, P., Schreiber, M.E., Seal, R.R., Engle, M.A., Chermak, J.A., Bodnar, R.J., Johnson, D.K., and Aylor, J.G., 2015, Statistical analysis of soil geochemical data to identify pathfinders associated with mineral deposits: An example from the Coles Hill uranium deposit, Virginia, USA: *Journal of Geochemical Exploration*, v. 154, p. 238–251, doi:10.1016/j.gexplo.2014.12.012.

- Lin, X., Wang, X., Zhang, B., and Yao, W., 2014, Multivariate analysis of regolith sediment geochemical data from the Jinwozi gold field, north-western China: *Journal of Geochemical Exploration*, v. 137, p. 48–54, doi:10.1016/j.gexplo.2013.11.006.
- Loh, W.-L., 2014. Fifty Years of Classification and Regression Trees. *International Statistical Review*, v. 82, n. 3, p. 329–348. <https://doi.org/10.1111/insr.12016>
- Marques, J.C., Jost, H., Creaser, R.A., Frantz, J.C., and Osorio, R.G., 2013, Age of Arsenopyrite Gold-Bearing Massive Lenses of the Mina III and its Implication on Exploration, Crixás Greenstone Belt, Goiás, Brazil, in III Simpósio Brasileiro de Metalogenia, Gramado (RS).
- Martins-Ferreira, M.A.C., Dias, A.N.C., Chemale, F., Campos, J.E.G., Seraine, M., and Novais-Rodrigues, E., 2020, Multi-stage crustal accretion by magmatic flare-up and quiescence intervals in the western margin of the São Francisco Craton: U-Pb-Hf and geochemical constraints from the Almas Terrane: *Gondwana Research*, v. 85, p. 32–54, doi:10.1016/j.gr.2020.04.005.
- Massucato, A.J., 2004, Relatório de Geologia Estrutural—Aspectos Estruturais do Greenstone Belt de Crixás—GO—Anglogold Ashanti, Crixás-GO; AngloGold Ashanti: Johannesburg, South Africa, p. 30, Unpublished Internal Report.
- McCuaig, T.C., and Hronsky, J., 2017, The mineral systems concept: the key to exploration targeting: *Applied Earth Science*, v. 126, p. 77–78, doi:10.1080/03717453.2017.1306274.
- McCuaig, T.C., Beresford, S., and Hronsky, J., 2010, Translating the mineral systems approach into an effective exploration targeting system: *Ore Geology Reviews*, v. 38, p. 128–138, doi:10.1016/j.oregeorev.2010.05.008.
- Medina, C.M., Ducart, D.F., Passos, J.S., and de Oliveira, L.R., 2021, Exploration vectoring from the white mica spectral footprint in the atypical auriferous Lavra Velha deposit, San Francisco Craton, Brazil: *Ore Geology Reviews*, v. 139, p. 104438, doi:10.1016/j.oregeorev.2021.104438.
- Mello, J.W.V., Roy, W.R., Talbott, J.L., and Stucki, J.W., 2006, Mineralogy and arsenic mobility in arsenic-rich Brazilian soils and sediments: *Journal of Soils and Sediments*, v. 6, p. 9–19, doi:10.1065/jss2005.09.144.
- Metsalu, T., and Vilo, J., 2015, ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap: *Nucleic Acids Research*, v. 43, p. W566–W570, doi:10.1093/nar/gkv468.
- Mikucki, E.J., 1998, Hydrothermal transport and depositional processes in Archean lode-gold systems: A review: *Ore Geology Reviews*, v. 13, p. 307–321, doi:10.1016/S0169-1368(97)00025-5.
- Murray, H.H., 1955, Correlation of Paper-Coating Quality with Degree of Crystal Perfection of Kaolinite: *Clays and Clay Minerals*, v. 4, p. 31–40, doi:10.1346/ccmn.1955.0040105.
- Neal, L.C., Wilkinson, J.J., Mason, P.J., and Chang, Z., 2018, Spectral characteristics of propylitic alteration minerals as a vectoring tool for porphyry copper deposits: *Journal of Geochemical Exploration*, v. 184, p. 179–198, doi:10.1016/j.gexplo.2017.10.019.
- Neumayr, P., Walshe, J., Hagemann, S., Petersen, K., Roache, A., Frikken, P., Horn, L., and Halley, S., 2008, Oxidized and reduced mineral assemblages in greenstone belt rocks of

- the St. Ives gold camp, Western Australia: Vectors to high-grade ore bodies in Archaean gold deposits? *Mineralium Deposita*, v. 43, p. 363–371, doi:10.1007/s00126-007-0170-2.
- Opitz, D., Maclin, R., 1999. Popular Ensemble Methods: An Empirical Study. *Journal of Artificial Intelligence Research*, v. 11, p. 169–198. <https://doi.org/10.1613/jair.614>
- Pearce, J. A., 1983, Role of the sub-continental lithosphere in magma genesis at active continental margins. In: Hawkesworth, C.J. and Norry, M.J. eds. *Continental basalts and mantle xenoliths*, Nantwich, Cheshire: Shiva Publications, pp. 230-249.
- Pearson, K., 1897, Mathematical contributions to the theory of evolution - On a form of spurious correlation which may arise when indices are used in the measurement of organs: *Proceedings of the Royal Society of London*, v. 60, p. 489–498, doi:10.1098/rspl.1896.0076.
- Pereira, H. S., 2017. Integração De Dados Geológicos, Geoquímicos, Espectrorradiométricos E De Propriedades Físicas De Rocha: O Estudo De Caso Do Corpo N5S, Província Mineral De Carajás. Dissertação de Mestrado. Instituto de Geociências, Universidade de Brasília, 150 p.
- Perrouty, S., Linnen, R.L., Leshner, C.M., Olivo, G.R., Piercey, S.J., Gaillard, N., Clark, J.R., and Enkin, R.J., 2019, Expanding the size of multi-parameter metasomatic footprints in gold exploration: utilization of mafic dykes in the Canadian Malartic district, Québec, Canada: *Mineralium Deposita*, v. 54, p. 761–786, doi:10.1007/s00126-018-0829-x.
- Petersen Junior, K.J., 2003, Estudo das mineralizações auríferas do corpo IV e V da estrutura IV do Greenstone Belt de Crixás (GO): Universidade de São Paulo, 195 p.
- Phillips, G.N., and Powell, R., 2010, Formation of gold deposits: A metamorphic devolatilization model: *Journal of Metamorphic Geology*, v. 28, p. 689–718, doi:10.1111/j.1525-1314.2010.00887.x.
- Pimentel, M., Fuck, R., Jost, H., Filho, C., and Araújo, S., 2000, The Basement of the Brasília Fold Belt and the Goiás Magmatic Arc: Tectonic evolution of South America, p. 195–229.
- Pimentel, M.M., Rodrigues, J.B., DellaGiustina, M.E.S., Junges, S., Matteini, M., and Armstrong, R., 2011, The tectonic evolution of the Neoproterozoic Brasília Belt, central Brazil, based on SHRIMP and LA-ICPMS U-Pb sedimentary provenance data: A review: *Journal of South American Earth Sciences*, v. 31, p. 345–357, doi:10.1016/j.jsames.2011.02.011.
- Pontual, S., Merry, N., and Gamson, P., 2008a, Epithermal alteration systems. GMEX. Spectral analysis guides for mineral exploration: Victoria, AusSpec International Pty. Ltd., 61 p.
- Pontual, S., Merry, N., Gamson, P., 2008b. Spectral Analysis guides for mineral exploration. Spectral interpretation. Field manual, third ed. AusSpec International Ltd., Victoria.
- Prado, E.M.G., Silva, A.M., Ducart, D.F., Toledo, C.L.B., and de Assis, L.M., 2016, Reflectance spectroradiometry applied to a semi-quantitative analysis of the mineralogy of the N4ws deposit, Carajás Mineral Province, Pará, Brazil: *Ore Geology Reviews*, v. 78, p. 101–119, doi:10.1016/j.oregeorev.2016.03.007.
- Queiroz, C.L., Jost, H., da Silva, L.C., and McNaughton, N.J., 2008, U-Pb SHRIMP and Sm-Nd geochronology of granite-gneiss complexes and implications for the evolution of the Central

- Brazil Archean Terrain: *Journal of South American Earth Sciences*, v. 26, p. 100–124, doi:10.1016/j.jsames.2007.07.007.
- R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reimann, C., Filzmoser, P., and Garrett, R.G., 2002, Factor analysis applied to regional geochemical data: Problems and possibilities: *Applied Geochemistry*, v. 17, p. 185–206, doi:10.1016/S0883-2927(01)00066-X.
- Reimann, C., Filzmoser, P., and Garrett, R.G., 2005, Background and threshold: Critical comparison of methods of determination: *Science of the Total Environment*, v. 346, p. 1–16, doi:10.1016/j.scitotenv.2004.11.023. Rokach, L., 2010. Ensemble-based classifiers. *Artificial Intelligence Review*, v. 33, p. 1–39. <https://doi.org/10.1007/s10462-009-9124-7>
- Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. *Statistical Data Analysis Explained. Applied Environmental Statistics with R*. Wiley, Chichester, UK, 343 pp.
- Rollinson, H., & Pease, V., 2021, *Using Geochemical Data: To Understand Geological Processes (2nd ed.)*. Cambridge: Cambridge University Press. doi:10.1017/9781108777834
- Rose, A.W., Hawkes, H.E. & Webb, J.S., 1979, *Geochemistry in Mineral Exploration*. 2nd edn. Academic Press.
- Saboia, L.A., and Teixeira, N.A., 1980, Lavas Ultrabásicas da Unidade Basal do Greenstone Belt de Crixás (GO): Uma Nova Classe de Rochas Ultrabásicas no Estado de Goiás: *Revista Brasileiro de Geociências*, v. 10, p. 28–42.
- Saboia, L.A.; Teixeira, N. A.; Castro, J. H. G.; Teixeira, A.S., 1981, Geologia do greenstone belt de Crixás (GO) e suas implicações geotectônicas, in SBG, *Simpósio Sobre o Cráton do São Francisco e Suas Faixas Marginais*, p. 41–49.
- Sadeghi, M., Billay, A., and Carranza, E.J.M., 2015, Analysis and mapping of soil geochemical anomalies: Implications for bedrock mapping and gold exploration in Giyani area, South Africa: *Journal of Geochemical Exploration*, v. 154, p. 180–193, doi:10.1016/j.gexplo.2014.11.018.
- Schapire, R.E., 1990. The Strength of Weak Learnability. *Machine Learning*, v. 5, p. 197–227. <https://doi.org/10.1007/BF00116037>
- Schapire, R.E., 2003. The Boosting Approach to Machine Learning: An Overview. In: Denison, D.D., Hansen, M.H., Holmes, C.C., Mallick, B., Yu., B. (eds). *Nonlinear Estimation and Classification. Lecture Notes in Statistics*, vol 171. Springer, New York. https://doi.org/10.1007/978-0-387-21579-2_9
- Shuguang, Z., Kefa, Z., Yao, C., Jinlin, W., and Jianli, D., 2015, Exploratory data analysis and singularity mapping in geochemical anomaly identification in Karamay, Xinjiang, China: *Journal of Geochemical Exploration*, v. 154, p. 171–179, doi:10.1016/j.gexplo.2014.12.007.
- Sícoli Seoane, J.C., Castro, N.A., Osako, L.S., Baars, F.J., 2009. Multispectral Imagery Applied to Nickel Laterite Exploration: The Conceição do Araguaia Discovery. *Reviews in Economic Geology*, 16, pp.109-122.

- Simpson, M.P., and Christie, A.B., 2019, Hydrothermal alteration mineralogical footprints for New Zealand epithermal Au-Ag deposits: *New Zealand Journal of Geology and Geophysics*, v. 62, p. 483–512, doi:10.1080/00288306.2019.1577278.
- Sobiesiak, M.S., 2011, Caracterização de Depósito Aurífero no Corpo Pequizão, Crixás - GO; Trabalho de Conclusão de Curso; Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.
- Sonntag, I., Laukamp, C., and Hagemann, S.G., 2012, Low potassium hydrothermal alteration in low sulfidation epithermal systems as detected by IRS and XRD: An example from the Co-O mine, Eastern Mindanao, Philippines: *Ore Geology Reviews*, v. 45, p. 47–60, doi:10.1016/j.oregeorev.2011.08.001.
- Tarvainen, T., Sapon, S., and Jarva, J., 2019, Applying heatmaps in interpretation of geochemical baseline data on urban soils in Finland: *Journal of Geochemical Exploration*, v. 205, p. 106345, doi:10.1016/j.gexplo.2019.106345.
- Tassinari, C.C.G., Jost, H., Santos, J.C., Nutman, A.P., Bennell, M.R., 2006. Pb and Nd isotope signatures and SHRIMP U-Pb geochronological evidence of Paleoproterozoic age for Mina III gold mineralization. In: *South American Symposium on Isotope Geology*, pp. 615–617.
- Templ, M., Filzmoser, P., and Reimann, C., 2008, Cluster analysis applied to regional geochemical data: Problems and possibilities: *Applied Geochemistry*, v. 23, p. 2198–2213, doi:10.1016/j.apgeochem.2008.03.004.
- Theodoro, S. de C.H., and Jost, H., 1996, The Discovery of Primary Sedimentary Structures in Archean Marbles of Crixás, Central Brazil, in *Symposium Archean Terranes of the South American Platform*, Brasília, p. 41–42.
- Townsend, T. E. (1987), Discrimination of iron alteration minerals in visible and near-infrared reflectance data, *J. Geophys. Res.*, 92(B2), 1441–1454, doi:10.1029/JB092iB02p01441.
- Ulrich, S., Hageman, S., Marques, J.C., Figueiredo, F.L.A.R., Ramires, J.E.F., Frantz, J.C., and Petersen, K., 2021, The Orogenic Crixás Gold Deposit, Goiás, Brazil: A Review and New Constraints on the Structural Control of Ore Bodies: *Minerals*, v. 11, p. 1050, doi:10.3390/min11101050.
- Van der Meer, F., 2004, Analysis of spectral absorption features in hyperspectral imagery: *International Journal of Applied Earth Observation and Geoinformation*, v. 5, p. 55–68, doi:10.1016/j.jag.2003.09.001.
- Varmuza, K., and Filzmoser, P., 2009, *Introduction to Multivariate Statistical Analysis in Chemometrics (1st ed.)*. CRC Press. doi:10.1201/9781420059496
- Wang, R. et al., 2017, White mica as a hyperspectral tool in exploration for the sunrise dam and Kanowna belle gold deposits, Western Australia: *Economic Geology*, v. 112, p. 1153–1176, doi:10.5382/econgeo.2017.4505.
- Warr, L.N., 2021, IMA–CNMNC approved mineral symbols: *Mineralogical Magazine*, v. 85, p. 291–320, doi:10.1180/mgm.2021.43.

- Witt, W.K., Hagemann, S.G., Villanes, C., Vennemann, T., Zwingmann, H., Laukamp, C., and Spangenberg, J.E., 2016, Multiple gold mineralizing styles in the Northern Pataz District, Peru: *Economic Geology*, v. 111, p. 355–394, doi:10.2113/econgeo.111.2.355.
- Wyman, D.A., Cassidy, K.F., and Hollings, P., 2016, Orogenic gold and the mineral systems approach: Resolving fact, fiction and fantasy: *Ore Geology Reviews*, v. 78, p. 322–335, doi:10.1016/j.oregeorev.2016.04.006.
- Xue, J., Lee, C., Wakeham, S.G., and Armstrong, R.A., 2011, Using principal components analysis (PCA) with cluster analysis to study the organic geochemistry of sinking particles in the ocean: *Organic Geochemistry*, v. 42, p. 356–367, doi:10.1016/j.orggeochem.2011.01.01
- Zuo, R., 2011, Identifying geochemical anomalies associated with Cu and Pb-Zn skarn mineralization using principal component analysis and spectrum-area fractal modeling in the Gangdese Belt, Tibet (China): *Journal of Geochemical Exploration*, v. 111, p. 13–22, doi:10.1016/j.gexplo.2011.06.012.

3. Discussões e conclusões

Esta dissertação teve como objetivo caracterizar o *footprint* superficial da zona de alteração hidrotermal do depósito Pequizão que é controlado estruturalmente pela Estrutura IV e que corresponde a uma das principais estruturas que controla a mineralização aurífera na região do *greenstone belt* Crixás. A metodologia usada neste trabalho correspondeu a integração técnicas robustas para o processamento, e análise de dados de geoquímica de solo e espectrorradiometria.

A origem dos depósitos de ouro orogênico ainda é debatida, podendo ser formada durante os estágios finais do metamorfismo regional a partir de fluidos crustais (Phillips e Powell, 2010), oriundos da devolatização de sedimentos piritosos, acima de uma subducção de crosta oceânica como resultado de uma mudança no estresse compressional (Groves et al., 2020), ou podem ser divididos em classes relacionadas a uma fonte distal e a uma intrusão proximal oxidada (Witt et al., 2016).

Os controles estruturais de primeira ordem sobre a formação desses depósitos são baseados em sua intersecção com estruturas de alto ângulo, e a segunda ordem é dominada por linhas de fraqueza ao longo da faixa na qual o fluido mineralizante pode se propagar e, após focalizar, promover a precipitação mineral (Goldfarb e Groves, 2015; Wyman et al., 2016; Groves et al., 2018).

É essencial para a exploração tentar compreender previamente o sistema mineralizante em que está inserido o depósito, e isso inclui o reconhecimento dos padrões de alteração, dos mecanismos de precipitação e da arquitetura local (Groves e Phillips, 1987). Os *greenstone belts* do Domínio Crixás-Goiás têm características de alteração hidrotermais típicas de depósitos orogênicos, um forte controle estrutural e enriquecimento de K, Fe, S e CO₂ (Goldfarb et al., 2001; Jost et al., 2019).

O ouro ocorre por diferenças química/mineralógica em um depósito hidrotermal e, consequentemente, tais padrões de alteração fornecem um potencial para encontrar alvos mineralizados (Neumayr et al., 2008; McCuaig et al., 2010). Após sua formação, muitos processos podem obliterar, exumar, dispersar, concentrar ou sobrepor a mineralização dos depósitos (Hagemann et al., 2016). A presença de ouro no perfil de solo, neste caso, é provavelmente de origem hipogênica e concentrada no solo como resíduo da exposição prolongada e intemperismo, portanto, não sendo de crescimento supergênico. O impacto

que o intemperismo pode ter no *footprint* mineral nunca foi devidamente abordado, com a consequência de que este processo pode gerar localmente anomalias geoquímicas que podem ser confundidas como alvos exploratórios, levando a um número crescente de falsos alvos positivos.

As ocorrências de ouro no depósito do Pequizão são geralmente associadas ao filito carbonoso como o principal litotipo mineralizado. Em subsuperfície, o corpo de minério do Pequizão está localizado em uma zona de silicificação intensa e fortemente controlado pela Estrutura IV. Ocorre a presença de sulfetos disseminados que são geralmente representados por arsenopirita, pirrotita e calcopirita subordinada. O arsenopirita é um mineral índice da estrutura e tem alto teor, tornando-o uma das principais características prospectivas em furos de sondagem (Ulrich et al., 2021).

Os primeiros resultados dos dados geoquímicos do solo apontam para uma assinatura da alteração hidrotermal em superfície, mostrando uma concentração de elementos associados ao Au, tais como Ag, As e Sb. Entre os três níveis diferentes do solo, notamos uma ligeira alteração na concentração dos elementos. Com isso, há sempre um aumento gradual ascendente ou descendente dos elementos entre os níveis, como por exemplo o aumento de Ag com a profundidade. Em geral, como as amostras do nível mais profundo (nível C) são mais preservadas elas possuem concentrações mais elevadas.

A correlação de Pearson (r) indica uma assinatura geoquímica de um solo predominantemente de rochas máficas, como as associações com Cu, Cr, Fe, e Ni, e a influência de elementos argilosos (i.e., Al, Ba e Mn). K, P e Zn, típicos de regolitos e são mais concentrados nas amostras rasas no nível A. Além disso, Ga, Hg, Te e Sb têm o mesmo comportamento e normalmente estão relacionados com a mineralização do ouro, como o agrupamento de Sb, Te, Tl, Au, Ag e Hg pelo cluster hierárquico. Comparando com o *ensemble learning*, é interessante notar a presença do agrupamento Zn, Ga, e Pb, e o alto valor de importância para a predição do conteúdo de Au. No entanto, estes elementos estão fortemente associados com a ausência da mineralização. A deficiência de Zn é comum em solos arenosos lixiviados e com baixa disponibilidade de Zn devido ao alto pH (solos dolomíticos e alcalinos) ou alto teor de matéria orgânica (por exemplo, filito carbonosos) (Hooda, 2010).

A utilização de espectrômetros de campo em amostras de solo pode ser extremamente rápida e bem-sucedida para a exploração da superfície (Velasco et al., 2005). Técnicas

espectrais combinadas com análises mineralógicas, estruturais e litogeoquímicas podem fornecer a caracterização de *footprints* associados com o ouro. Os resultados obtidos dentro desta pesquisa mostram que a mica branca no depósito Pequizão é mais concentrada no filito carbonoso e na dolomita. A muscovita (alto Al) tem um comprimento de onda curto e a fengita (baixo Al) tem um comprimento de onda longo. A área com ocorrência de fengita é relativamente rica em ouro, de acordo com a análise geoquímica e espectral, e também coincide com a proximidade com a Estrutura IV. A abundância de clorita (principalmente Fe-clorita) ocorre com intensidade em qt-clr-cb-sericita xistos, que é o principal indicativo da Estrutura Palmeiras. O anfibólio (principalmente actinolita) ocorre em ampla distribuição no solo, localizado nos metabasaltos, diferentemente da distribuição do gibsita, com muitas amostras abundantes localizadas na área central.

Classificamos os minerais dos resultados SWIR em oito grupos de acordo com a sua abundância em cada amostra: (1) mica branca + caulinita bem cristalizada; (2) mica branca + caulinita má cristalizada; (3) mica branca + gibsita; (4) caulinita bem cristalizada + mica branca; (5) caulinita bem cristalizada + gibsita; (6) caulinita bem cristalizada; (7) caulinita com baixa cristalinidade + mica branca + gibsita; (8) caulinita com baixa cristalinidade + Fe-clorita. O mapa mineral explica a predominância da caulinita bem cristalizada no depósito e a associação com o gibsita na área central, que é a mesma região extraída dos resultados do fator 3.

A análise estatística multivariada dos dados geoquímica de composição (adequadamente transformados por *centered log-ratio* – clr) obtidos do regolito de superfície no depósito do Pequizão mostrou alguns resultados promissores. Os elementos dos resultados do PC1 mostram uma clara distinção entre rochas máficas da associação elementar típica dos componentes do silicato clássico e elementos imóveis comuns em ambientes hidrotermais (por exemplo, Al, Ti, Zr, etc.). PC3 está potencialmente rastreando o enriquecimento de Au e As e o empobrecimento concomitante de Zn. Lin et al. (2014) confirmam que os escores do terceiro fator podem ser explicados como a expressão superficial da mineralização cega e, conseqüentemente, Ag, As e Zn são os indicadores indiretos da mineralização.

Medimos o "fator de enriquecimento" para zonas do fator 2 para esclarecer e caracterizar as diferenças dos dados geoquímicos entre a zona mineralizada, próxima à estrutura (escores de fatores negativos, ou NFS), com a zona mais distal a noroeste

(escores de fatores positivos, ou PFS) em amostras de metabasalto. A amostra de normalização, aqui chamada amostra de solo de referência, foi coletada no topo da sondagem de Campos et al. (2022), onde a localização está a cerca de 2 km a leste do depósito. Comparando as amostras, as diferenças mais significativas foram o aumento de Na, Ba, W, Tl e Mg nas amostras NFS, relacionadas a mineralização. Analisando o resultado do enriquecimento em Pequizão, podemos apontar claramente que a variância de Na, V, Tl, Sb e Ba é semelhante às encontradas no footprint da Estrutura Palmeiras (Campos et al., 2022).

Por fim, os elementos que melhor definem a assinatura do depósito Pequizão, além de Au, correspondem aos elementos Ag, As, Ba, Na, Sb, Te, Tl e W (não em ordem de importância). A análise espectral apontou o alto conteúdo de muscovita que tende a fengita em amostras ricas em ouro da zona da Estrutura IV. Na mesma região, há uma ocorrência de caulinita bem cristalizada + gibsita que marca a área proximal na Estrutura Palmeiras e é a principal vetor associado com a mineralização.

Os objetivos propostos para realização do trabalho foram alcançados e a utilização da metodologia aplicada mostrou eficácia na caracterização da alteração hidrotermal no solo, ou seja, do *footprint* superficial. Com isso, as principais considerações obtidas durante o desenvolvimento deste projeto e algumas sugestões para trabalhos futuros estão listadas abaixo:

(1). A aplicação da análise univariada como primeira abordagem permitiu observar o comportamento dos elementos e suas correlações em cada nível amostrado, evidenciando um aumento da concentração de elementos farejadores (*pathfinders*) com a profundidade e uma maior retenção de ouro no nível B. Adicionalmente, mapeia a característica geoquímica de cada unidade geológica no regolito;

(2). A análise de agrupamento (*clustering*) mapeia os elementos associados ao ouro (i.e., As, Sb, Te, Tl) e destaca os conjuntos relacionados às rochas metabásicas e metassedimentares. Contudo, para melhor entender a importância de cada elemento na predição de ouro, a técnica de *machine learning* (*ensemble learning*) utilizada conseguiu evidenciar o Zn, Ga e Pb, como um grupo importante para a prospecção mineral de ouro neste contexto. A interpretação desse resultado, integrado com a análise de fatores e componentes principais conseguiu responder que alta concentração desses elementos corresponde a ausência de mineralização;

(3). A espectroscopia de reflectância é uma ferramenta extremamente eficaz na realização de mapeamento mineral e, dessa forma, a zonas mineralizadas em superfície no *greenstone belt* Crixás. Foi possível determinar, como a assinatura mineralógica, a gibbsita e a mica branca como minerais indicativos da alteração hidrotermal para as estruturas Palmeiras e Estrutura IV, respectivamente;

(4). A integração dos resultados de espectroscopia de reflectância com a análise geoquímica multivariada possibilitou a interpretação de mapas mineralógicos e de fatores, definindo assim alvos favoráveis para se mapear a mineralização de ouro;

(5). Deve-se ampliar o investimento e colocar como parte da rotina a caracterização geoquímica e mineralógica do solo (e sua evolução) por empresas, além da academia e o governo, para auxiliar na exploração mineral brasileira. A geoquímica do regolito traz um conjunto de ferramentas complementares de valor a ser considerado, em conjunto com dados geológicos, geofísicos e geocronológicos. O presente estudo demonstra um fluxo de trabalho viável para uma análise prospectiva desde estudos locais até regionais.

4. Referências Bibliográficas

- Abdi, H., and Williams, L.J., 2010, Principal component analysis: Wiley Interdisciplinary Reviews: Computational Statistics, v. 2, p. 433–459, doi:10.1002/wics.101.
- Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, London.
- Almeida, F. R., 2021. Caracterização Da Assinatura Da Mineralização Polimetálica Do Greenstone Belt De Faina (GO). Dissertação de Mestrado. Instituto de Geociências, Universidade de Brasília, 104 p.
- Ames, D.E., Galley, A.G., Kjarsgaard, I.M., Tardif, N., and Taylor, B.T., 2016, Hanging-wall vectoring for buried volcanogenic massive sulfide deposits, paleoproterozoic flin flon mining camp, Manitoba, Canada: Economic Geology, v. 111, p. 963–1000, doi:10.2113/econgeo.111.4.963.
- Anhaeusser, C. R. 2014. Archaean Greenstone Belts and Associated Granitic Rocks—A Review. Journal of African Earth Sciences, V. 100, P. 684-732
- Arndt N.T., Teixeira N.A., White, W.M. 1989. Bizarre Geochemistry Of Komatiites From The Crixás Greenstone Belt. Contr. Mineral. Petrol. 101:187-197
- Bacha, R.R.B., Toledo, C.L.B., Silva, A.M., and Mungall, J.E., 2021, Alkali element mobility as a vector for gold mineralization: An example from Central Brazil: Geochemistry, p. 125808, doi:10.1016/j.chemer.2021.125808.
- Beghelli Junior, L.P., 2012, Charnockitos e ortogneisses da porção Centro-Oeste do Bloco Arqueano de Goiás: Dados geoquímicos e isotópicos: Universidade de Brasília.
- Bogossian, J., Kemp, A.I.S., and Hagemann, S.G., 2021, Linking gold systems to the crust-mantle evolution of archaean crust in central Brazil: Minerals, v. 11, doi:10.3390/min11090944.
- Bohachevsky, I.O, Johnson, M.E., Stein, M.L., 1986. Generalized Simulated Annealing for Function Optimization. Technometrics, v. 28, n. 3, p. 209–217. <https://doi.org/10.2307/1269076>
- Borges, C. C. A., Toledo, C. L. B., Silva, A. M., Junior, F. C., Jost, H., & De Carvalho Lana, C. 2017. Geochemistry And Isotopic Signatures Of Metavolcanic And Metaplutonic Rocks Of The Faina And Serra De Santa Rita Greenstone Belts, Central Brazil: Evidences For A Mesoarchean Intraoceanic Arc. Precambrian Research, 292, 350-377.
- Borges, C.C.A., Toledo, C.L.B., A.M., Silva, Kirk, J., Ruiz, J., Chemale Jr., F., Souza, R. G., Santos, B.A., Campos, M.P., Campos, L.M., Santos, A.M., 2021b, Archean to Paleoproterozoic evolution of the Crixás greenstone belt, Central Brazil: Insights from two contrasting assemblies of metaigneous rocks: Lithos, v. 404–405, p. 106493, doi:10.1016/j.lithos.2021.106493.
- Borges, C.C.A., Toledo, C.L.B., Silva, A.M., Chemale, F., dos Santos, B.A., Figueiredo, F.L., and Zacchi, É.N.P., 2021a, Unraveling a hidden Rhyacian magmatic arc through provenance of metasedimentary rocks of the Crixás greenstone belt, Central Brazil: Precambrian Research, v. 353, doi:10.1016/j.precamres.2020.106022.
- Breiman, L., 1996. Bagging predictors. Machine Learning, v. 24, p. 123–140. doi:10.1007/BF00058655
- Breiman, L., 2001. Random Forests. Machine Learning, v. 45, p. 5–32. doi:10.1023/A:1010933404324
- Campos, D.S. de, Silva, A.M., Toledo, C.L.B., Carvalho, M.J. de, Rodrigues, V.G., and Araujo, K., 2017, Prospectivity analysis of gold and iron oxide copper-gold-(silver) mineralizations from the Faina Greenstone Belt, Brazil, using multiple data sets: Brazilian Journal of Geology, v. 47, p. 561–590, doi:10.1590/2317-4889201720170012.
- Campos, L.M., 2019, Caracterização Das Zonas De Alteração Hidrotermal Associadas À Mineralização De Ouro No Greenstone Belt Crixás – GO, Dissertação de Mestrado. Instituto de Geociências, Universidade de Brasília, 165 p.

- Campos, L.M., Toledo, C.L.B., Silva, A.M., Ducart, D.F., Santos, B. A., Campos, M.P., Borges, C.C.A., 2022. The Hydrothermal Footprint of Crixás Orogenic Gold Deposit, Central Brazil. *Ore Geology Reviews*, (submetido e em menores revisões).
- Carranza, E.J.M., 2011, Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values: *Journal of Geochemical Exploration*, v. 110, p. 167–185. doi:10.1016/j.gexplo.2011.05.007.
- Carrino, T.A., Crósta, A.P., Toledo, C.L.B., and Silva, A.M., 2018, Hyperspectral remote sensing applied to mineral exploration in southern Peru: A multiple data integration approach in the Chapi Chiara gold prospect: *International Journal of Applied Earth Observation and Geoinformation*, v. 64, p. 287–300, doi:10.1016/j.jag.2017.05.004.
- Cattell, R. B. 1966. The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. doi:10.1207/s15327906mbr0102_10
- Cevik, I.S., Olivo, G.R., and Ortiz, J.M., 2021, A combined multivariate approach analyzing geochemical data for knowledge discovery: The Vazante – Paracatu Zinc District, Minas Gerais, Brazil: *Journal of Geochemical Exploration*, v. 221, p. 1–18, doi:10.1016/j.gexplo.2020.106696.
- Chayes, F., 1960, On Correlation between Variables of Constant Sum, *Journal of Geophysical Research*, v.65, n12, p. 4185-4193, doi:10.1029/JZ065i012p04185
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., 2021. xgboost: Extreme Gradient Boosting. R package version 1.4.1.1. <https://CRAN.R-project.org/package=xgboost>
- Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7:e623. <https://doi.org/10.7717/peerj-cs.623>
- Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., and Vergo, N., 1990, High spectral resolution reflectance spectroscopy of minerals: *Journal of Geophysical Research*, v. 95, doi:10.1029/jb095ib08p12653.
- Clark, R.N., Swayze, G.A., Wise, R.A., Livo, K.E., Hoefen, T.M., Kokaly, R.F. and Sutley, S.J., 2007. USGS digital spectral library splib06a (No. 231). US Geological Survey. doi:10.3133/ds231
- Cloutier, J., Piercey, S.J., and Huntington, J., 2021, Mineralogy, mineral chemistry and swir spectral reflectance of chlorite and white mica: *Minerals*, v. 11, doi:10.3390/min11050471.
- Comas-Cufi, M., Thió-Henestrosa, S., 2011, CoDaPack 2.0: a stand-alone, multiplatform compositional software. In: Egozcue, J.J., Tolosana-Delgado, R., Ortego, M.I. (Eds), *Proceedings of the CoDaWork'11: 4th International Workshop on Compositional Data Analysis*, Sant Feliu de Guíxols.
- Cooke, D.R. et al., 2018, Porphyry Indicator Minerals (PIMS) and Porphyry Vectoring and Fertility Tools (PVFTS)-Indicators of Mineralization Styles and Recorders of Hypogene Geochemical Dispersion Halos: *Journal of Geochemical Exploration*, v. 190, p. 109–129, <https://doi.org/10.1016/j.gexplo.2018.02.015>.
- Cordeiro, P.F. de O., and Oliveira, C.G. de, 2017, The Goiás Massif: Implications for a pre-Columbia 2.2–2.0 Ga continent-wide amalgamation cycle in central Brazil: *Precambrian Research*, v. 298, p. 403–420, doi:10.1016/j.precamres.2017.06.021.
- Cudahy, T.J., and Ramanaidou, E.R., 1997, Measurement of the hematite: Goethite ratio using field visible and near-infrared reflectance spectrometry in channel iron deposits, Western Australia: *Australian Journal of Earth Sciences*, v. 44, p. 411–420, doi:10.1080/08120099708728322.

- Cudahy, T., 2016, Mineral mapping for exploration: An Australian journey of evolving spectral sensing technologies and industry collaboration: *Geosciences* (Switzerland), v. 6, doi:10.3390/geosciences6040052.
- Ducart, D.F., Crósta, A. P., Filho, C. R. S., & Coniglio, J. 2006. Alteration mineralogy at the Cerro La Mina epithermal prospect, Patagonia, Argentina: Field mapping, short-wave infrared spectroscopy, and ASTER images. *Economic Geology*. 101(5), 981-996
- Duke, E.F., 1994, reaction progress: Implications for remote sensing:, doi:10.1130/0091-7613(1994)022<0621.
- Duuring, P., Hassan, L., Zelic, M., and Gessner, K., 2016, Geochemical and spectral footprint of metamorphosed and deformed VMS-style mineralization in the Quinns district, Yilgarn craton, Western Australia: *Economic Geology*, v. 111, p. 1411–1438, doi:10.2113/econgeo.111.6.1411.
- Ferreira, B.F., Marques, J.C., Frantz, J.C., de Souza, R.G., Campos, M.P., Figueiredo, F.L.A. de R., and Padilha, N.L., 2021, Stratigraphy and U–Pb detrital zircon provenance, Crixás Greenstone Belt, Goiás-Brazil: Gold-bearing vs barren siliciclastic rocks: *Journal of South American Earth Sciences*, v. 105, doi:10.1016/j.jsames.2020.102994.
- Filzmoser, P., Hron, K., and Templ, M., 2018, Applied compositional data analysis: with worked examples in R: v. 89, 3165–3165 p., doi:10.1080/00949655.2019.1628880.
- Fortes, P. de T.F. de O., 1996, Metalogenia dos Depósitos Auríferos Mina III, Mina Nova e Mina Inglesa, Greenstone Belt de Crixás-GO: Universidade de Brasília, 242 p.
- Fortes, P. de T.F. de O., Coelho, R.F., Giuliani, G., 2000, Au/Ag Ratio Variations At Mina III, Mina Nova and Mina Inglesa Gold Deposits, Crixás Greenstone Belt, Brazil I: *Revista Brasileira de Geociências*, v. 30, p. 246–250.
- Fortes, P. de T.F. de O., Pimentel, M.M., Santos, R.V., Junges, S. 2003. Sm-Nd Study Of The Crixás Greenstone Belt, Brazil: Implications For The Age Of Deposition Of The Upper Sedimentary Rocks And Associated Au Mineralization. *Journal Of South American Earth Sciences* 16, 503-512.
- Freund, Y., Schapire, R.E., 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, v. 55, n. 1, p. 119–139. <https://doi.org/10.1006/jcss.1997.1504>
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, v. 29, n. 5, p. 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Fuck, R.A., Dantas, E.L., Pimentel, M.M., Botelho, N.F., Armstrong, R., Laux, J.H., Junges, S.L., Soares, J.E., and Praxedes, I.F., 2014, Paleoproterozoic crust-formation and reworking events in the Tocantins Province, central Brazil: A contribution for Atlantica supercontinent reconstruction: *Precambrian Research*, v. 244, p. 53–74, doi:10.1016/j.precamres.2013.12.003.
- Galán, E., 2006, Chapter 14 Genesis of Clay Minerals: *Developments in Clay Science*, v. 1, p. 1129–1162, doi:10.1016/S1572-4352(05)01042-1.
- Gielar, A., Helios-Rybicka, E., Möller, S., and Einax, J.W., 2012, Multivariate analysis of sediment data from the upper and middle Odra River (Poland): *Applied Geochemistry*, v. 27, p. 1540–1545, doi:10.1016/j.apgeochem.2012.04.004.
- Gisbert, G., Tornos, F., Losantos, E., Pons, J.M., and Videira, J.C., 2021, Vectors to ore in replacive volcanogenic massive sulfide (VMS) deposits of the northern Iberian Pyrite Belt: Mineral zoning, whole rock geochemistry, and application of portable X-ray fluorescence: *Solid Earth*, v. 12, p. 1931–1966, doi:10.5194/se-12-1931-2021.
- Gladwell, D.R., Lett, R.E., and Lawrence, P., 1983, Application of reflectance spectrometry to mineral exploration using portable radiometers (Abitibi, Ontario): *Economic Geology*, v. 78, p. 699–710, doi:10.2113/gsecongeo.78.4.699.

- Goldfarb, R.J., Groves, D.I., and Gardoll, S., 2001, Orogenic gold and geologic time: A global synthesis: *Ore Geology Reviews*, v. 18, p. 1–75, doi:10.1016/S0169-1368(01)00016-6.
- Goldfarb, R.J., and Groves, D.I., 2015, Orogenic gold: Common or evolving fluid and metal sources through time: *Lithos*, v. 233, p. 2–26, doi:10.1016/j.lithos.2015.07.011.
- Goldfarb, R.J., 2021, *Lode Gold Deposits in Time and Space*: Elsevier Inc., 663–679 p., doi:10.1016/b978-0-08-102908-4.00006-0.
- Greenwell, B.M., Boehmke, B.C., 2020. Variable Importance Plots—An Introduction to the vip Package. *The R Journal*, v. 12, n. 1, 343–366. <https://doi.org/10.32614/RJ-2020-013>.
- Groves, D.I., and Phillips, G.N., 1987, The genesis and tectonic control on archaean gold deposits of the Western Australian Shield - A metamorphic replacement model: *Ore Geology Reviews*, v. 2, p. 287–322, doi:10.1016/0169-1368(87)90009-6.
- Groves, D.I., Santosh, M., Deng, J., Wang, Q., Yang, L., and Zhang, L., 2020, A holistic model for the origin of orogenic gold deposits and its implications for exploration: *Mineralium Deposita*, v. 55, p. 275–292, doi:10.1007/s00126-019-00877-5.
- Groves, D.I., Santosh, M., Goldfarb, R.J., and Zhang, L., 2018, Structural geometry of orogenic gold deposits: Implications for exploration of world-class and giant deposits: *Geoscience Frontiers*, v. 9, p. 1163–1177, doi:10.1016/j.gsf.2018.01.006.
- Groves, D.I., Santosh, M., Deng, J., Wang, Q., Yang, L., and Zhang, L., 2020, A holistic model for the origin of orogenic gold deposits and its implications for exploration: *Mineralium Deposita*, v. 55, p. 275–292, doi:10.1007/s00126-019-00877-5.
- Grunsky, E.C., 2010, The interpretation of geochemical survey data: *Geochemistry: Exploration, Environment, Analysis*, v. 10, p. 27–74, doi:10.1144/1467-7873/09-210.
- Haest, M., Cudahy, T., Laukamp, C., and Gregory, S., 2012, Quantitative mineralogy from infrared spectroscopic data. I. Validation of mineral abundance and composition scripts at the rocklea channel iron deposit in Western Australia: *Economic Geology*, v. 107, p. 209–228, doi:10.2113/econgeo.107.2.209.
- Hagemann, S.G., Lisitsin, V.A., and Huston, D.L., 2016, Mineral system analysis: Quo vadis: *Ore Geology Reviews*, v. 76, p. 504–522, doi:10.1016/j.oregeorev.2015.12.012.
- Harraden, C.L., McNulty, B.A., Gregory, M.J., and Lang, J.R., 2013, Shortwave infrared spectral analysis of hydrothermal alteration associated with the Pebble porphyry copper-gold-molybdenum deposit, Iliamna, Alaska: *Economic Geology*, v. 108, p. 483–494, doi:10.2113/econgeo.108.3.483.
- Hasui, Y., 2012, Sistema Orogênico Tocantins. In: Hasui, Y.; Carneiro, C.D.R.; Almeida, F.F.M.; Bartorelli, A. *Geologia Do Brasil*. São Paulo: Beca, 2012. Cap Iv-14a, P. 289-325.
- Herrmann, W., Blake, M., Doyle, M., Huston, D., Kamprad, J., Merry, N., and Pontual, S., 2001, Short Wavelength Infrared (SWIR) spectral analysis of hydrothermal alteration zones associated with base metal sulfide deposits at Rosebery and Western Tharsis, Tasmania, and Highway-Reward, Queensland: *Economic Geology*, v. 96, p. 939–955, doi:10.2113/gsecongeo.96.5.939.
- Hooda, P.S., 2010. *Trace elements in soils*, first ed. John Wiley & Sons Ltd., Chichester, United Kingdom. doi:10.1002/9781444319477
- Hofer, G., Wagreich, M., and Neuhuber, S., 2013, Geochemistry of fine-grained sediments of the upper Cretaceous to Paleogene Gosau Group (Austria, Slovakia): Implications for paleoenvironmental and provenance studies: *Geoscience Frontiers*, v. 4, p. 449–468, doi:10.1016/j.gsf.2012.11.009.
- Hollis, S.P., Foury, S., Caruso, S., Johnson, S., Barrote, V., and Pumphrey, A., 2021, Lithochemical and hyperspectral halos to ag-zn-au mineralization at nimbus in the eastern goldfields superterrane, Western Australia: *Minerals*, v. 11, p. 1–35, doi:10.3390/min11030254.

- Hough, R.M., Butt, C.R.M., Reddy, S.M., and Verrall, M., 2007, Gold nuggets: Supergene or hypogene? *Australian Journal of Earth Sciences*, v. 54, p. 959–964, doi:10.1080/08120090701488289.
- Humphris, S.E., 1984, The mobility of the rare earth elements in the crust.: *Rare Earth Element Geochemistry*, p. 317–342, doi:10.1016/b978-0-444-42148-7.50014-9.
- IBGE, 2015, Manual Técnico de Pedologia, Folha SD-22 Goiás, Coordenação de Recursos Naturais e Estudos Ambientais – Instituto Brasileiro de Geografia e Estatística, 3. ed., Rio de Janeiro, 430 p.
- Ilyas, A., Kashiwaya, K., and Koike, K., 2016, Ni grade distribution in laterite characterized from geostatistics, topography and the paleo-groundwater system in Sorowako, Indonesia: *Journal of Geochemical Exploration*, v. 165, p. 174–188, doi:10.1016/j.gexplo.2016.03.002.
- Jolliffe, I. T., 2002, *Principal Component Analysis*, 2 ed., Springer Series in Statistics. 488 p. doi:10.1007/b98835
- Jolliffe, I.T., and Cadima, J., 2016, Principal component analysis: A review and recent developments: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, doi:10.1098/rsta.2015.0202.
- Jost, H., and De Oliveira, A.M., 1991, Stratigraphy of the greenstone belts, Crixás region, Goiás, central Brazil: *Journal of South American Earth Sciences*, v. 4, p. 201–214, doi:10.1016/0895-9811(91)90031-F.
- Jost, H., Carmelo, A.C., Meneses, P., 1994a, Litologias, relações de contato e estrutura do Bloco Caiamar, Crixás, Goiás, in SBG, Simpósio de Geologia do Centro-Oeste IV, p. 18–20.
- Jost, H., Vargas, M.C., Fuck, R.A., Kuyumjian, R.M., Pires, A.C.B., 1994b, Relações de contato, litologias, geologia estrutural e geofísica do Bloco Arqueano do Moqué, Crixás, Goiás, in SBG, Simpósio de Geologia do Centro-Oeste IV, p. 21–23.
- Jost, H., & Fortes, P.D.T.O., 2001, Gold Deposits And Occurrences Of The Crixás Goldfield, Central Brazil. *Mineralium Deposita*, 36(3-4), 358-376.
- Jost, H. et al., 2001, Geologia De Terrenos Arqueanos E Proterozóicos Da Região De Crixás-Cedrolina, Goiás: *Revista Brasileira de Geociências*, v. 31, p. 315–328, doi:10.25249/0375-7536.2001313315328.
- Jost, H., Dussin, I. A., Chemale Jr., F., Tassinari, C. C. G., & Junges, S. 2008. U-Pb And Sm-Nd Constraints For The Paleoproterozoic Age Of The Metasedimentary Sequences Of The Goiás Archean Greenstone Belts. In *Book Of Abstracts*. [Buenos Aires]: Ingeis - Instituto De Geocronología Y Geología Isotópica.
- Jost, H., Chemale Jr., F., Dussin, I.A., Tassinari, C.C.G., Martins, R. 2010. A U–Pb Zircon Paleoproterozoic Age For The Metasedimentary Host Rocks And Gold Mineralization Of The Crixás Greenstone Belt, Goiás, Central Brazil. *Ore Geol. Reviews* 37, 127–139
- Jost, H., Carvalho, M.J., Rodrigues, V.G., Martins, R., 2014, Metalogênese Dos Greenstone Belts De Goiás, In: Silva, M.G., Neto, M.B.R., Jost, H., Kuyumjian, R.M. (Orgs.), *Metalogênese Das Províncias Tectônicas Brasileiras*, Belo Horizonte, Cprm, Pp. 141-168.
- Jost, H., Apollo, J. F. H., Weber, W., Dos Reis Salles, R., Marques, J. C., Massucatto, A. J., & Dos Santos, B. A. 2019. Stratigraphic Update, Paleotectonic, Paleogeographic, And Depositional Environments Of The Crixás Greenstone Belt, Central Brazil. *Journal Of South American Earth Sciences*, 96, 102329.
- Kabata-Pendias, A., 2011, Trace elements in soils and plants (4th ed.). CRC press, v. 4, 548 p., doi:10.1201/b10158
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Proceedings of the 31st International Conference on*

- Neural Information Processing Systems (NIPS'17), p. 3146–3154. <https://dl.acm.org/doi/10.5555/3294996.3295074>
- Kirkpatrick, S., Gelatt, D., Vecchi, M., 1983. Optimization by Simulated Annealing. *Science*, v. 220, n. 4598, p. 671–680. <https://doi.org/10.1126/science.220.4598.671>
- Kuhn, M., 2021. finetune: Additional Functions for Model Tuning. R package version 0.1.0. <https://CRAN.R-project.org/package=finetune>
- Kuhn, M., Wickham, H., 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. <https://www.tidymodels.org>
- Kusky, T., Windley, B.F., Polat, A., Wang, L., Ning, W., and Zhong, Y., 2021, Archean dome-and-basin style structures form during growth and death of intraoceanic and continental margin arcs in accretionary orogens: *Earth-Science Reviews*, v. 220, p. 103725, doi:10.1016/j.earscirev.2021.103725.
- Kuyumjian, R.M., and Jost, H., 2006, Low- and high-alumina komatiites of Goiás, Central Brazil: *Journal of South American Earth Sciences*, v. 20, p. 315–326, doi:10.1016/j.jsames.2005.11.003.
- Lacerda Filho, J.V. de, 1999, Programa Levantamentos Geológicos Básicos do Brasil. *Geologia e Recursos Minerais do Estado de Goiás e Distrito Federal: CPRM*, v. Projeto de, p. 200.
- Lampinen, H.M., Laukamp, C., Occhipinti, S.A., and Hardy, L., 2019, Mineral footprints of the Paleoproterozoic sediment-hosted Abra Pb-Zn-Cu-Au deposit Capricorn Orogen, Western Australia: *Ore Geology Reviews*, v. 104, p. 436–461, doi:10.1016/j.oregeorev.2018.11.004.
- Laukamp, C., Rodger, A., Legras, M., Lampinen, H., Lau, I.C., Pejčic, B., Stromberg, J., Francis, N., and Ramanaidou, E., 2021, Mineral physicochemistry underlying feature-based extraction of mineral abundance and composition from shortwave, mid and thermal infrared reflectance spectra: *Minerals*, v. 11, doi:10.3390/min11040347.
- Leshner, M. et al., 2017, Integrated Multi-Parameter Exploration Footprints of the Canadian Malartic Disseminated Au, McArthur River-Millennium Unconformity U, and Highland Valley Porphyry Cu Deposits: Preliminary Results from the NSERC-CMIC Mineral Exploration Footprints Research: *Proceedings of Exploration*, v. 17, p. 8–12, www.dmec.ca.
- Levitani, D.M., Zipper, C.E., Donovan, P., Schreiber, M.E., Seal, R.R., Engle, M.A., Chermak, J.A., Bodnar, R.J., Johnson, D.K., and Aylor, J.G., 2015, Statistical analysis of soil geochemical data to identify pathfinders associated with mineral deposits: An example from the Coles Hill uranium deposit, Virginia, USA: *Journal of Geochemical Exploration*, v. 154, p. 238–251, doi:10.1016/j.gexplo.2014.12.012.
- Lin, X., Wang, X., Zhang, B., and Yao, W., 2014, Multivariate analysis of regolith sediment geochemical data from the Jinwozi gold field, north-western China: *Journal of Geochemical Exploration*, v. 137, p. 48–54, doi:10.1016/j.gexplo.2013.11.006.
- Loh, W.-L., 2014. Fifty Years of Classification and Regression Trees. *International Statistical Review*, v. 82, n. 3, p. 329–348. <https://doi.org/10.1111/insr.12016>
- Marques, J.C., Jost, H., Creaser, R.A., Frantz, J.C., and Osorio, R.G., 2013, Age of Arsenopyrite Gold-Bearing Massive Lenses of the Mina III and its Implication on Exploration, Crixás Greenstone Belt, Goiás, Brazil, in III Simpósio Brasileiro de Metalogenia, Gramado (RS).
- Martins-Ferreira, M.A.C., Dias, A.N.C., Chemale, F., Campos, J.E.G., Seraine, M., and Novais-Rodrigues, E., 2020, Multi-stage crustal accretion by magmatic flare-up and quiescence intervals in the western margin of the São Francisco Craton: U-Pb-Hf and geochemical constraints from the Almas Terrane: *Gondwana Research*, v. 85, p. 32–54, doi:10.1016/j.gr.2020.04.005.

- Massucato, A.J., 2004, Relatório de Geologia Estrutural—Aspectos Estruturais do Greenstone Belt de Crixás—GO—Anglogold Ashanti, Crixás-GO; AngloGold Ashanti: Johannesburg, South Africa, p. 30, Unpublished Internal Report.
- McCuaig, T.C., and Hronsky, J., 2017, The mineral systems concept: the key to exploration targeting: *Applied Earth Science*, v. 126, p. 77–78, doi:10.1080/03717453.2017.1306274.
- McCuaig, T.C., Beresford, S., and Hronsky, J., 2010, Translating the mineral systems approach into an effective exploration targeting system: *Ore Geology Reviews*, v. 38, p. 128–138, doi:10.1016/j.oregeorev.2010.05.008.
- Medina, C.M., Ducart, D.F., Passos, J.S., and de Oliveira, L.R., 2021, Exploration vectoring from the white mica spectral footprint in the atypical auriferous Lavra Velha deposit, San Francisco Craton, Brazil: *Ore Geology Reviews*, v. 139, p. 104438, doi:10.1016/j.oregeorev.2021.104438.
- Mello, J.W.V., Roy, W.R., Talbott, J.L., and Stucki, J.W., 2006, Mineralogy and arsenic mobility in arsenic-rich Brazilian soils and sediments: *Journal of Soils and Sediments*, v. 6, p. 9–19, doi:10.1065/jss2005.09.144.
- Metsalu, T., and Vilo, J., 2015, ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap: *Nucleic Acids Research*, v. 43, p. W566–W570, doi:10.1093/nar/gkv468.
- Mikucki, E.J., 1998, Hydrothermal transport and depositional processes in Archean lode-gold systems: A review: *Ore Geology Reviews*, v. 13, p. 307–321, doi:10.1016/S0169-1368(97)00025-5.
- Murray, H.H., 1955, Correlation of Paper-Coating Quality with Degree of Crystal Perfection of Kaolinite: *Clays and Clay Minerals*, v. 4, p. 31–40, doi:10.1346/ccmn.1955.0040105.
- Neal, L.C., Wilkinson, J.J., Mason, P.J., and Chang, Z., 2018, Spectral characteristics of propylitic alteration minerals as a vectoring tool for porphyry copper deposits: *Journal of Geochemical Exploration*, v. 184, p. 179–198, doi:10.1016/j.gexplo.2017.10.019.
- Neumayr, P., Walshe, J., Hagemann, S., Petersen, K., Roache, A., Frikken, P., Horn, L., and Halley, S., 2008, Oxidized and reduced mineral assemblages in greenstone belt rocks of the St. Ives gold camp, Western Australia: Vectors to high-grade ore bodies in Archaean gold deposits? *Mineralium Deposita*, v. 43, p. 363–371, doi:10.1007/s00126-007-0170-2.
- Okada, K., 2021, A Historical Overview of the Past Three Decades of Mineral Exploration Technology: *Natural Resources Research*, v. 30, p. 2839–2860, doi:10.1007/s11053-020-09721-4.
- Opitz, D., Maclin, R., 1999. Popular Ensemble Methods: An Empirical Study. *Journal of Artificial Intelligence Research*, v. 11, p. 169–198. <https://doi.org/10.1613/jair.614>
- Pearce, J. A., 1983, Role of the sub-continental lithosphere in magma genesis at active continental margins. In: Hawkesworth, C.J. and Norry, M.J. eds. *Continental basalts and mantle xenoliths*, Nantwich, Cheshire: Shiva Publications, pp. 230-249.
- Pearson, K., 1897, Mathematical contributions to the theory of evolution - On a form of spurious correlation which may arise when indices are used in the measurement of organs: *Proceedings of the Royal Society of London*, v. 60, p. 489–498, doi:10.1098/rspl.1896.0076.
- Pereira, H. S., 2017. Integração De Dados Geológicos, Geoquímicos, Espectrorradiométricos E De Propriedades Físicas De Rocha: O Estudo De Caso Do Corpo N5S, Província Mineral De Carajás. Dissertação de Mestrado. Instituto de Geociências, Universidade de Brasília, 150 p.
- Perrouy, S., Linnen, R.L., Leshner, C.M., Olivo, G.R., Piercey, S.J., Gaillard, N., Clark, J.R., and Enkin, R.J., 2019, Expanding the size of multi-parameter metasomatic footprints in gold exploration: utilization of mafic dykes in the Canadian Malartic district, Québec, Canada: *Mineralium Deposita*, v. 54, p. 761–786, doi:10.1007/s00126-018-0829-x.

- Petersen Junior, K.J., 2003, Estudo das mineralizações auríferas do corpo IV e V da estrutura IV do Greenstone Belt de Crixás (GO): Universidade de São Paulo, 195 p.
- Phillips, G.N., and Powell, R., 2010, Formation of gold deposits: A metamorphic devolatilization model: *Journal of Metamorphic Geology*, v. 28, p. 689–718, doi:10.1111/j.1525-1314.2010.00887.x.
- Pimentel, M., Fuck, R., Jost, H., Filho, C., and Araújo, S., 2000, The Basement of the Brasília Fold Belt and the Goiás Magmatic Arc: Tectonic evolution of South America, p. 195–229.
- Pimentel, M.M., Jost, H. and Fuck, R.A., 2004. O embasamento da Faixa Brasília eo arco magmático de Goiás. *Geologia do continente Sul-americano: evolução da obra de Fernando Flávio Marques de Almeida*, pp.355-358.
- Pontual, S., Merry, N., and Gamson, P., 1997, GMEX - Apectral analysis guides for mineral exploration—spectral interpretation field manual: Kew, Victoria 3101, Australia, AusSpec International Pty. Ltd., v. 1, 169 p.
- Pontual, S., Merry, N., and Gamson, P., 2008a, Epithermal alteration systems. GMEX. Spectral analysis guides for mineral exploration: Victoria, AusSpec International Pty. Ltd., 61 p.
- Pontual, S., Merry, N., Gamson, P., 2008b. Spectral Analysis guides for mineral exploration. Spectral interpretation. Field manual, third ed. AusSpec International Ltd., Victoria.
- Prado, E.M.G., Silva, A.M., Ducart, D.F., Toledo, C.L.B., and de Assis, L.M., 2016, Reflectance spectroradiometry applied to a semi-quantitative analysis of the mineralogy of the N4ws deposit, Carajás Mineral Province, Pará, Brazil: *Ore Geology Reviews*, v. 78, p. 101–119, doi:10.1016/j.oregeorev.2016.03.007.
- Queiroz, C.L., Jost, H., da Silva, L.C., and McNaughton, N.J., 2008, U-Pb SHRIMP and Sm-Nd geochronology of granite-gneiss complexes and implications for the evolution of the Central Brazil Archean Terrain: *Journal of South American Earth Sciences*, v. 26, p. 100–124, doi:10.1016/j.jsames.2007.07.007.
- R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reimann, C., Filzmoser, P., and Garrett, R.G., 2002, Factor analysis applied to regional geochemical data: Problems and possibilities: *Applied Geochemistry*, v. 17, p. 185–206, doi:10.1016/S0883-2927(01)00066-X.
- Rezaei, M., Nikbakht, M., and Shakeri, A., 2017, Geochemistry and sources of fluoride and nitrate contamination of groundwater in Lar area, south Iran: *Environmental Science and Pollution Research*, v. 24, p. 15471–15487, doi:10.1007/s11356-017-9108-0.
- Rokach, L., 2010. Ensemble-based classifiers. *Artificial Intelligence Review*, v. 33, p. 1–39. <https://doi.org/10.1007/s10462-009-9124-7>
- Rollinson, H., & Pease, V., 2021, *Using Geochemical Data: To Understand Geological Processes* (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/9781108777834
- Rose, A.W., Hawkes, H.E. & Webb, J.S., 1979, *Geochemistry in Mineral Exploration*. 2nd edn. Academic Press.
- Saboia, L.A., and Teixeira, N.A., 1980, Lavas Ultrabásicas da Unidade Basal do Greenstone Belt de Crixás (GO): Uma Nova Classe de Rochas Ultrabásicas no Estado de Goiás: *Revista Brasileiro de Geociências*, v. 10, p. 28–42.
- Saboia, L.A.; Teixeira, N. A.; Castro, J. H. G.; Teixeira, A.S., 1981, Geologia do greenstone belt de Crixás (GO) e suas implicações geotectônicas, in SBG, Simpósio Sobre o Cráton do São Francisco e Suas Faixas Marginais, p. 41–49.

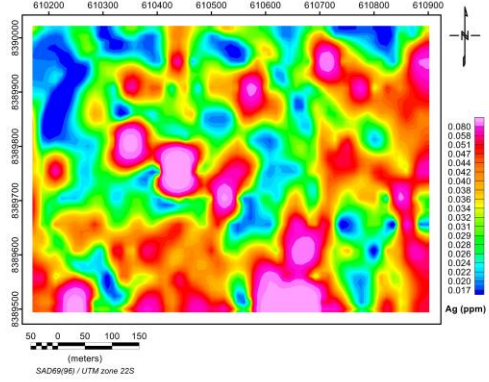
- Saboia, L.A.; Teixeira, N. 1983. Ultramafic Flows of The Crixas Greenstone Belt, Goiás — Brazil. *Precambrian Research*, V. 22, N. 1-2, p. 23-40.
- Schapire, R.E., 1990. The Strength of Weak Learnability. *Machine Learning*, v. 5, p. 197–227. <https://doi.org/10.1007/BF00116037>
- Schapire, R.E., 2003. The Boosting Approach to Machine Learning: An Overview. In: Denison, D.D., Hansen, M.H., Holmes, C.C., Mallick, B., Yu., B. (eds). *Nonlinear Estimation and Classification. Lecture Notes in Statistics*, vol 171. Springer, New York. https://doi.org/10.1007/978-0-387-21579-2_9
- Sícoli Seoane, J.C., Castro, N.A., Osako, L.S., Baars, F.J., 2009. Multispectral Imagery Applied to Nickel Laterite Exploration: The Conceição do Araguaia Discovery. *Reviews in Economic Geology*, 16, pp.109-122.
- Shuguang, Z., Kefa, Z., Yao, C., Jinlin, W., and Jianli, D., 2015, Exploratory data analysis and singularity mapping in geochemical anomaly identification in Karamay, Xinjiang, China: *Journal of Geochemical Exploration*, v. 154, p. 171–179, doi:10.1016/j.gexplo.2014.12.007.
- Simpson, M.P., and Christie, A.B., 2019, Hydrothermal alteration mineralogical footprints for New Zealand epithermal Au-Ag deposits: *New Zealand Journal of Geology and Geophysics*, v. 62, p. 483–512, doi:10.1080/00288306.2019.1577278.
- Sobiesiak, M.S., 2011, Caracterização de Depósito Aurífero no Corpo Pequizado, Crixás - GO; Trabalho de Conclusão de Curso; Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.
- Sonntag, I., Laukamp, C., and Hagemann, S.G., 2012, Low potassium hydrothermal alteration in low sulfidation epithermal systems as detected by IRS and XRD: An example from the Co-O mine, Eastern Mindanao, Philippines: *Ore Geology Reviews*, v. 45, p. 47–60, doi:10.1016/j.oregeorev.2011.08.001.
- Tappert, M.C., Rivard, B., Giles, D., Tappert, R., and Mauger, A., 2013, The mineral chemistry, near-infrared, and mid-infrared reflectance spectroscopy of phengite from the Olympic Dam IOCG deposit, South Australia: *Ore Geology Reviews*, v. 53, p. 26–38, doi:10.1016/j.oregeorev.2012.12.006.
- Tarvainen, T., Sapon, S., and Jarva, J., 2019, Applying heatmaps in interpretation of geochemical baseline data on urban soils in Finland: *Journal of Geochemical Exploration*, v. 205, p. 106345, doi:10.1016/j.gexplo.2019.106345.
- Tassinari, C. C. G., Jost, H., Santos, J. C., Nutman, A. P., & Bennell, M. R. 2006. Pb And Nd Isotope Signatures And Shrimp U–Pb Geochronological Evidence Of Paleoproterozoic Age For Mina Iii Gold Mineralization, Crixás District, Central Brazil. In 5th South American Symposium On Isotope Geology, Punta Del Este, Uruguay, Short Papers Volume (Pp. 527-529).
- Templ, M., Filzmoser, P., and Reimann, C., 2008, Cluster analysis applied to regional geochemical data: Problems and possibilities: *Applied Geochemistry*, v. 23, p. 2198–2213, doi:10.1016/j.apgeochem.2008.03.004.
- Theodoro, S. de C.H., and Jost, H., 1996, The Discovery of Primary Sedimentary Structures in Archean Marbles of Crixás, Central Brazil, in *Symposium Archaean Terranes of the South American Platform*, Brasília, p. 41–42.
- Townsend, T. E. (1987), Discrimination of iron alteration minerals in visible and near-infrared reflectance data, *J. Geophys. Res.*, 92(B2), 1441– 1454, doi:10.1029/JB092iB02p01441.
- Ulrich, S., Hageman, S., Marques, J.C., Figueiredo, F.L.A.R., Ramires, J.E.F., Frantz, J.C., and Petersen, K., 2021, The Orogenic Crixás Gold Deposit, Goiás, Brazil: A Review and New Constraints on the Structural Control of Ore Bodies: *Minerals*, v. 11, p. 1050, doi:10.3390/min11101050.

- UNCOVER Project, 2012, Searching the deep Earth: A vision for exploration geoscience in Australia, Australian Academy of Science, p. 47. <https://www.science.org.au/supporting-science/science-policy-and-sector-analysis/reports-and-publications/searching-deep-earth>, acesso em 25 de novembro de 2021.
- Van der Meer, F., 2004, Analysis of spectral absorption features in hyperspectral imagery: *International Journal of Applied Earth Observation and Geoinformation*, v. 5, p. 55–68, doi:10.1016/j.jag.2003.09.001.
- Van der Meer, F., 2004, Analysis of spectral absorption features in hyperspectral imagery: *International Journal of Applied Earth Observation and Geoinformation*, v. 5, p. 55–68, doi:10.1016/j.jag.2003.09.001.
- Varmuza, K., & Filzmoser, P., 2009, *Introduction to Multivariate Statistical Analysis in Chemometrics* (1st ed.). CRC Press. <https://doi.org/10.1201/9781420059496>
- Velasco, F., Alvaro, A., Suarez, S., Herrero, J.M., and Yusta, I., 2005, Mapping Fe-bearing hydrated sulphate minerals with short wave infrared (SWIR) spectral analysis at San Miguel mine environment, Iberian Pyrite Belt (SW Spain): *Journal of Geochemical Exploration*, v. 87, p. 45–72, doi:10.1016/j.gexplo.2005.07.002.
- Wang, R. et al., 2017, White mica as a hyperspectral tool in exploration for the sunrise dam and Kanowna belle gold deposits, Western Australia: *Economic Geology*, v. 112, p. 1153–1176, doi:10.5382/econgeo.2017.4505.
- Warr, L.N., 2021, IMA–CNMNC approved mineral symbols: *Mineralogical Magazine*, v. 85, p. 291–320, doi:10.1180/mgm.2021.43.
- Witt, W.K., Hagemann, S.G., Villanes, C., Vennemann, T., Zwingmann, H., Laukamp, C., and Spangenberg, J.E., 2016, Multiple gold mineralizing styles in the Northern Pataz District, Peru: *Economic Geology*, v. 111, p. 355–394, doi:10.2113/econgeo.111.2.355.
- Wyman, D.A., Cassidy, K.F., and Hollings, P., 2016, Orogenic gold and the mineral systems approach: Resolving fact, fiction and fantasy: *Ore Geology Reviews*, v. 78, p. 322–335, doi:10.1016/j.oregeorev.2016.04.006.
- Xue, J., Lee, C., Wakeham, S.G., and Armstrong, R.A., 2011, Using principal components analysis (PCA) with cluster analysis to study the organic geochemistry of sinking particles in the ocean: *Organic Geochemistry*, v. 42, p. 356–367, doi:10.1016/j.orggeochem.2011.01.012.
- Zhao, J., Chen, S., and Zuo, R., 2017, Identification and mapping of lithogeochemical signatures using staged factor analysis and fractal/multifractal models: *Geochemistry: Exploration, Environment, Analysis*, v. 17, p. 239–251, doi:10.1144/geochem2016-013.

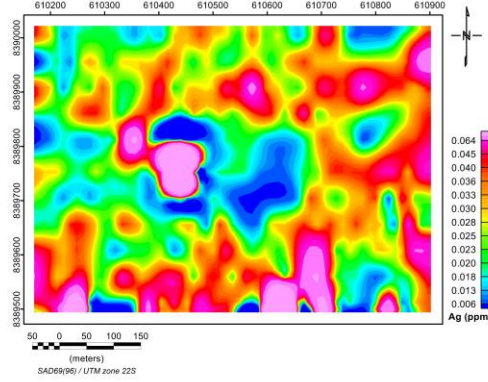
Anexo 1 – Mapas Unielementares Interpolados

Ag

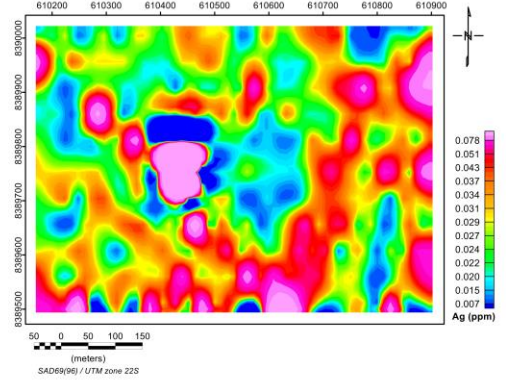
A



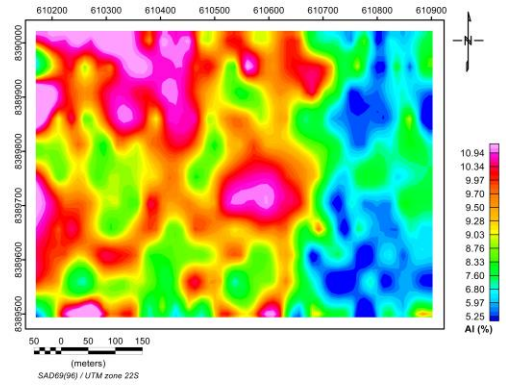
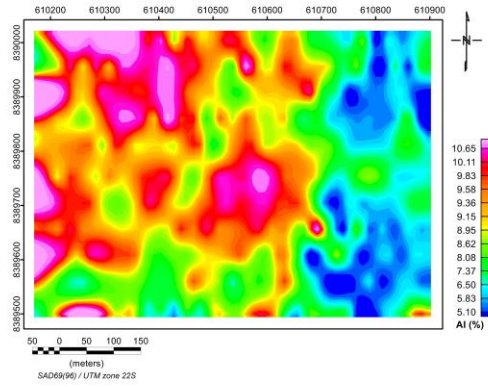
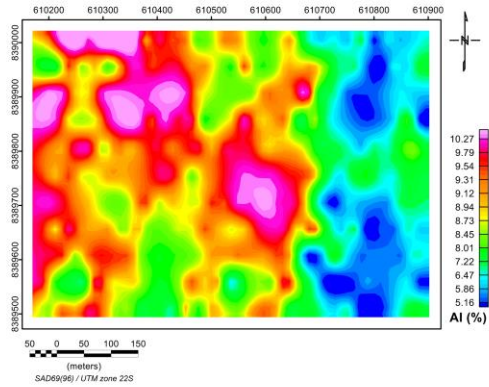
B



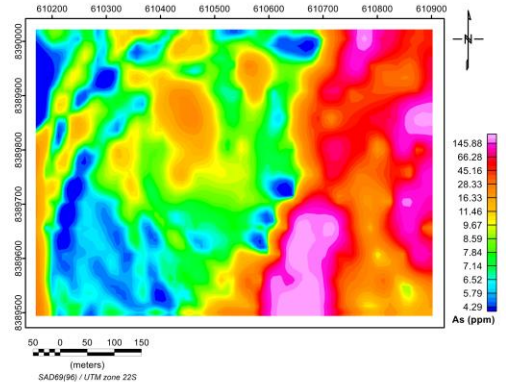
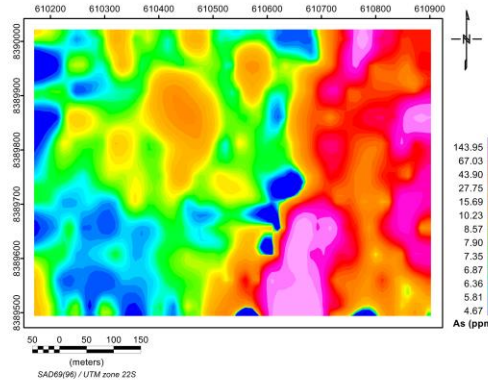
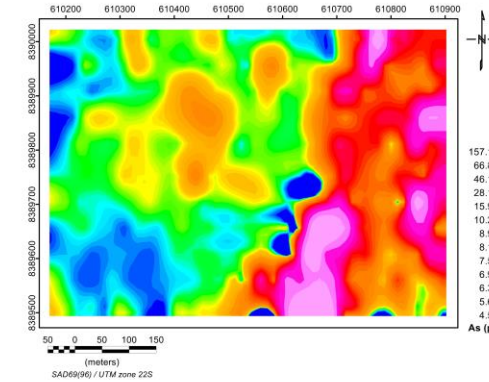
C



Al



As

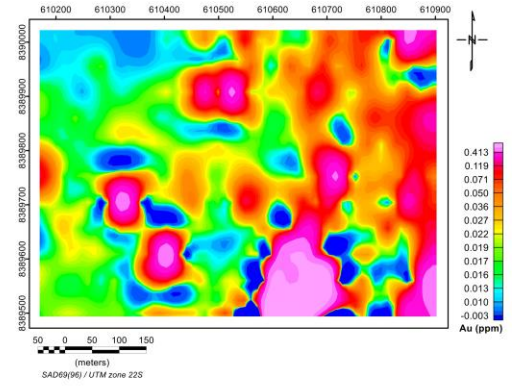
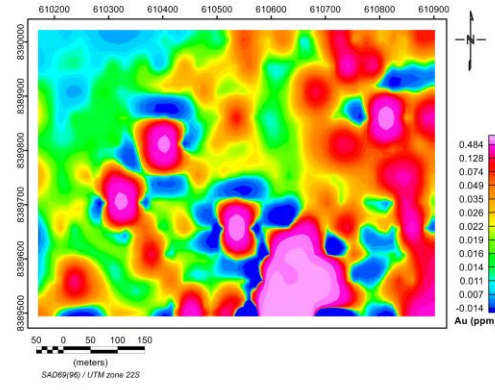
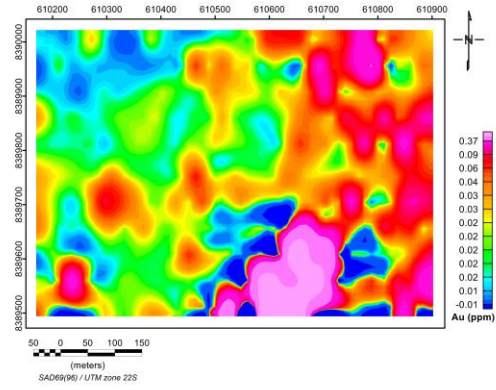


A

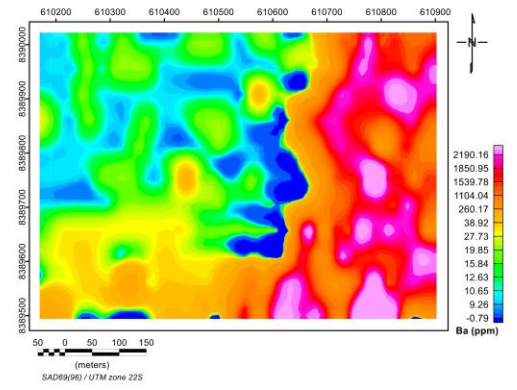
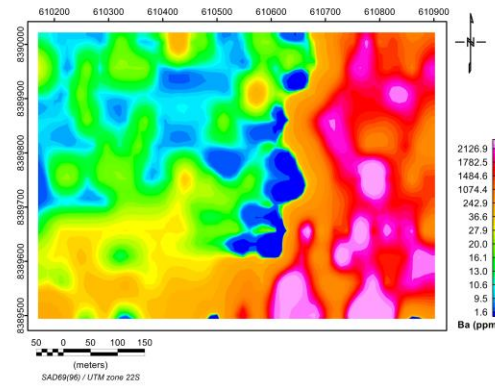
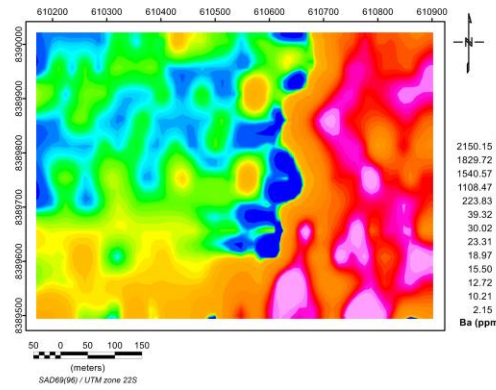
B

C

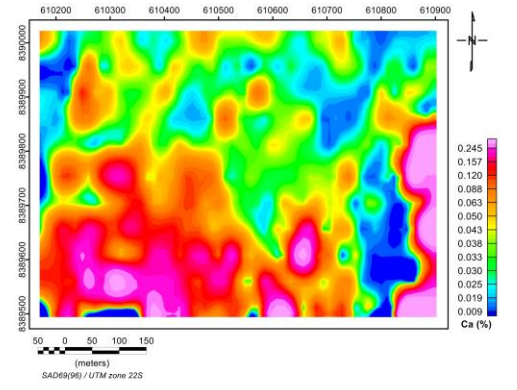
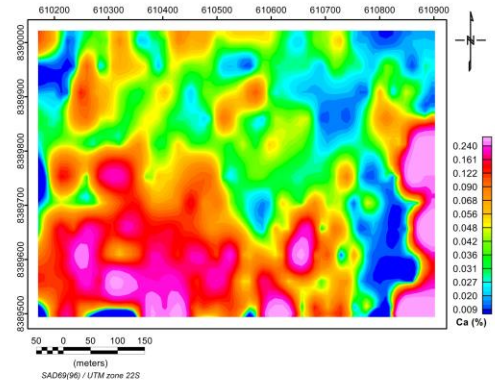
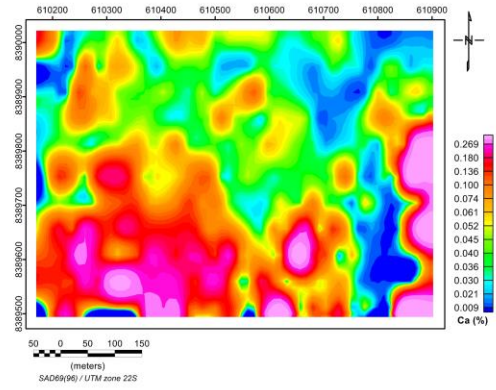
Au



Ba



Ca

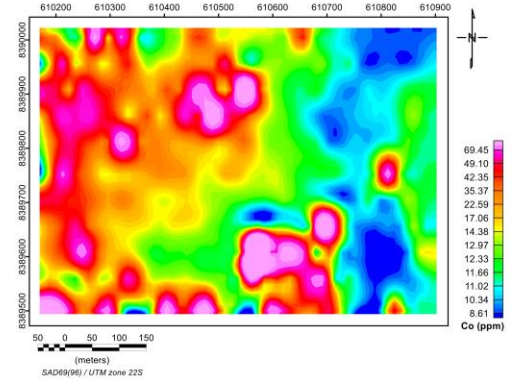
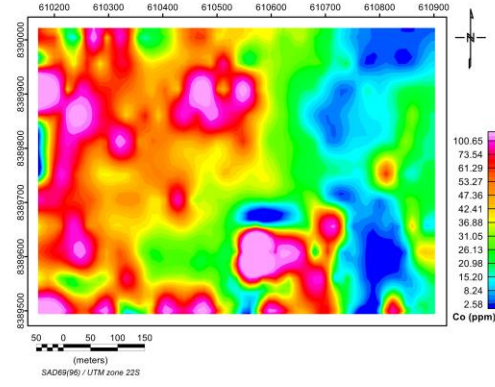
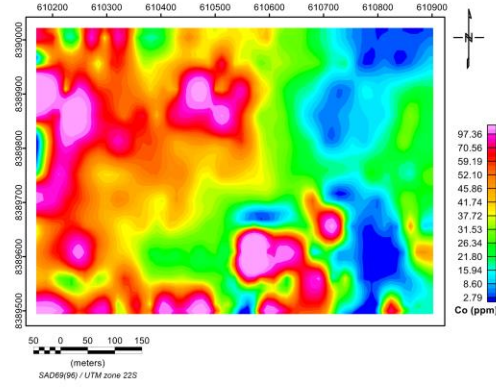


A

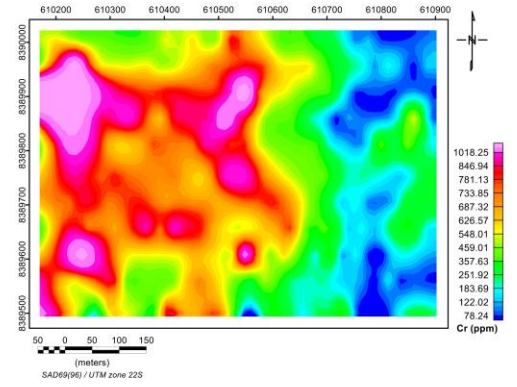
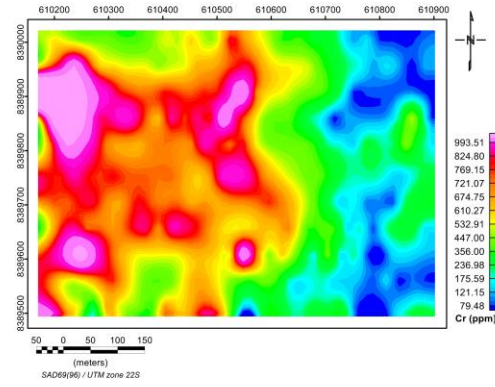
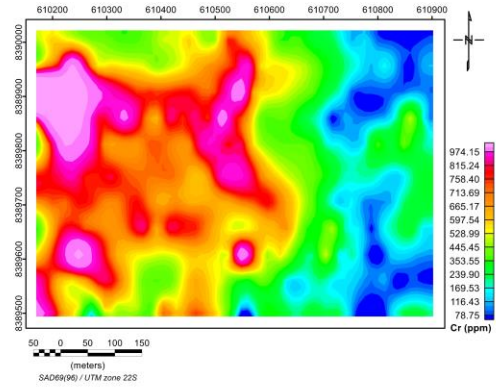
B

C

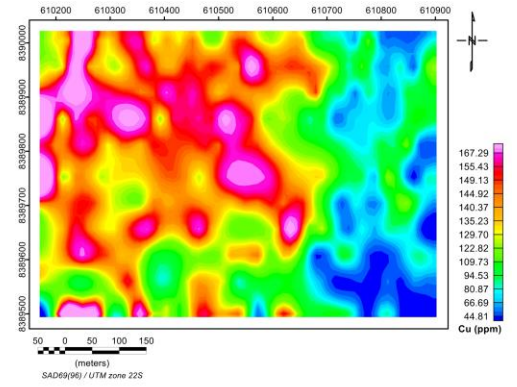
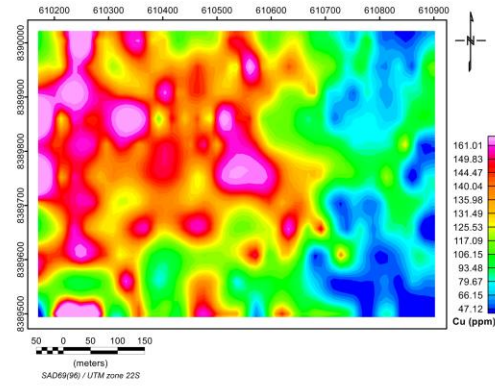
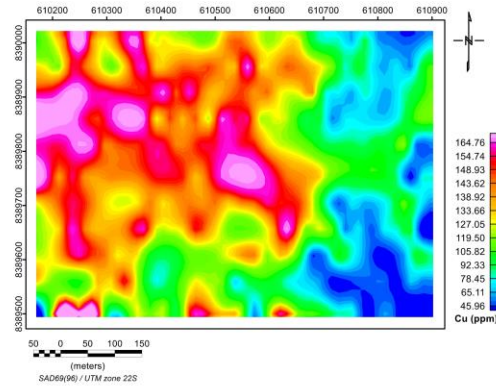
Co



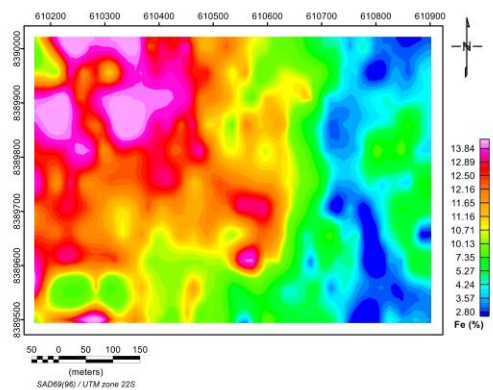
Cr



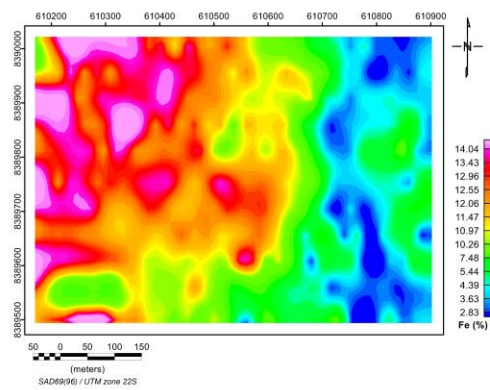
Cu



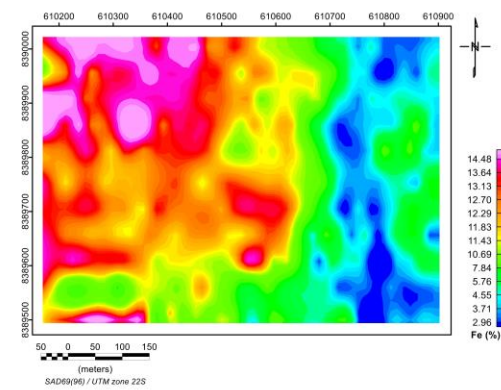
A



B

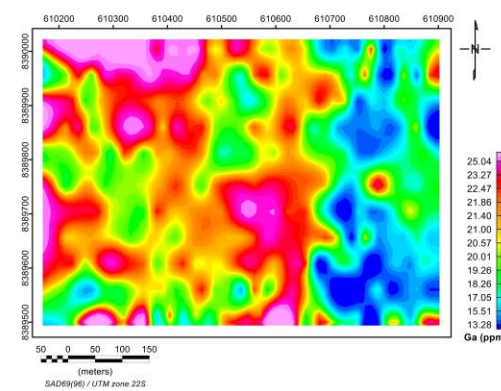
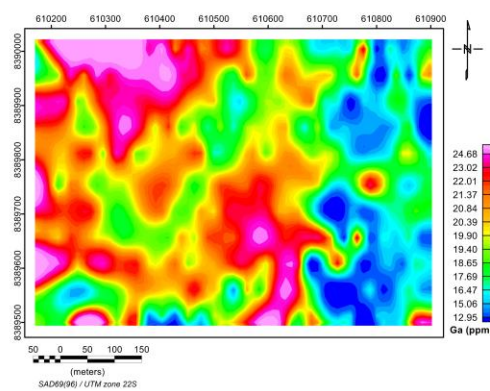
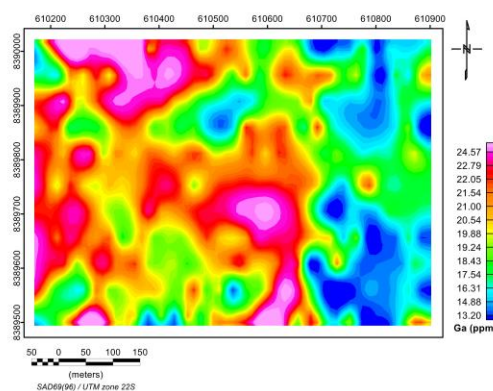


C

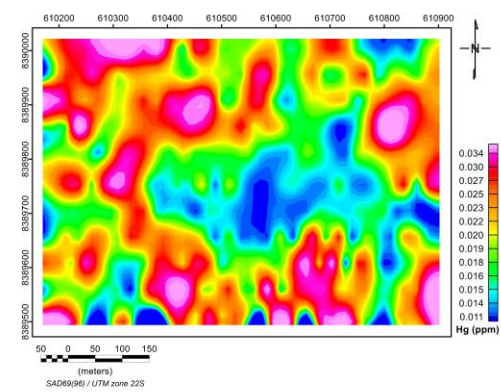
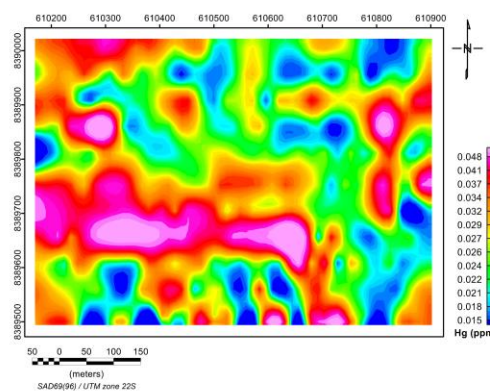
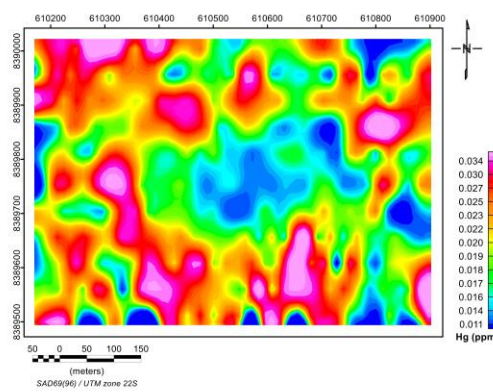


Fe

Ga



Hg

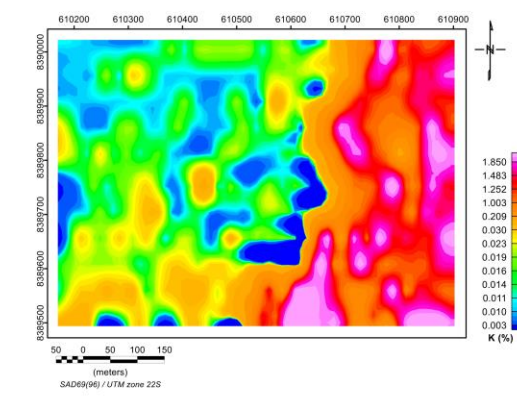
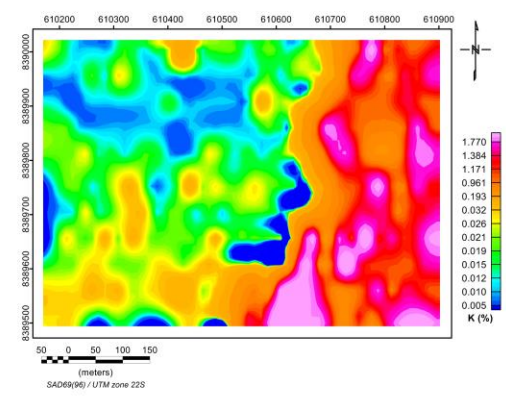
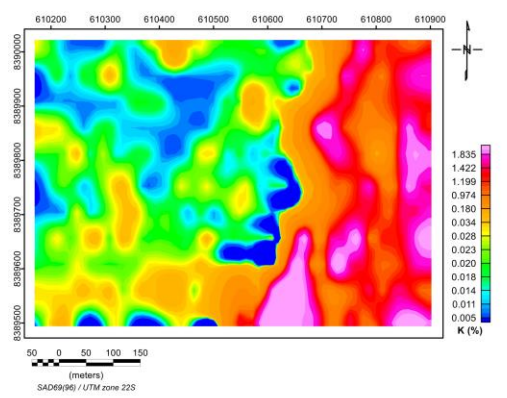


A

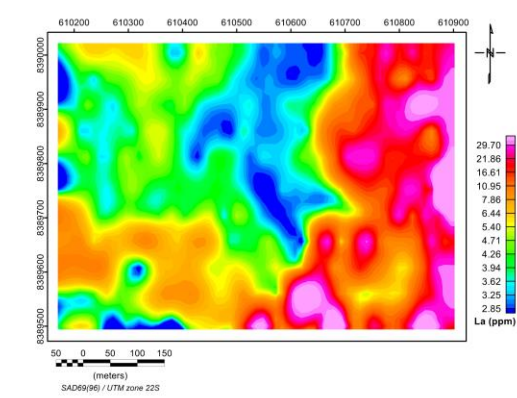
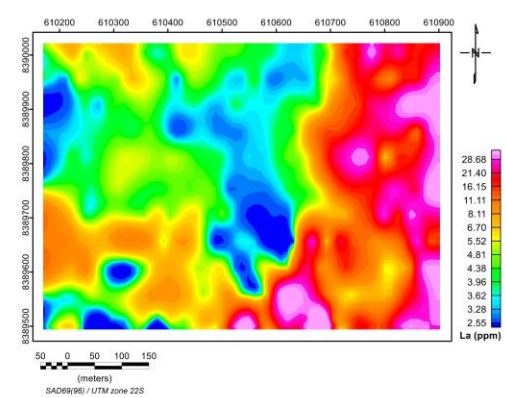
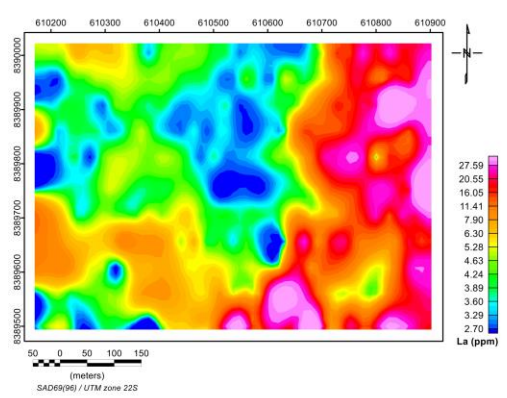
B

C

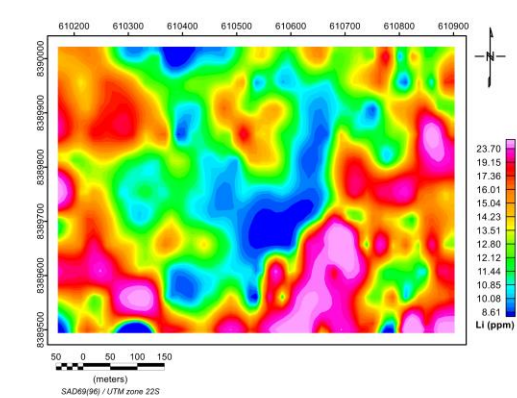
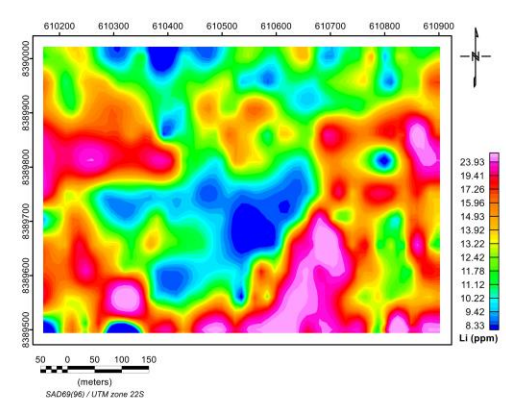
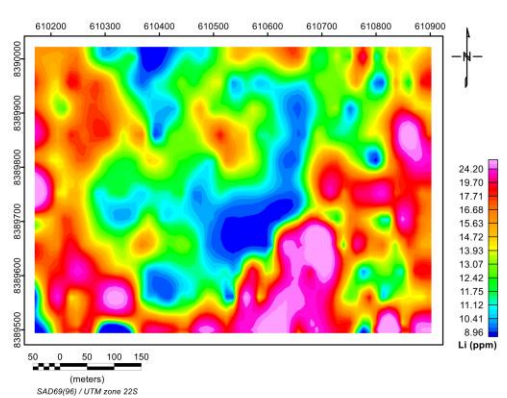
K



La



Li

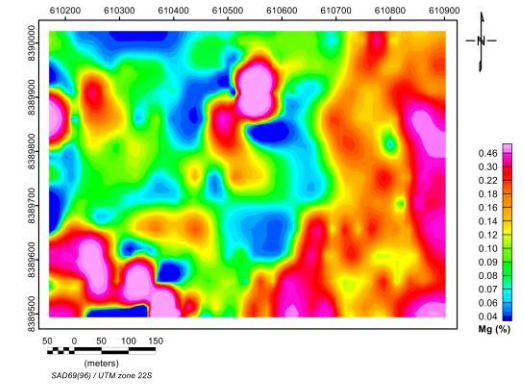
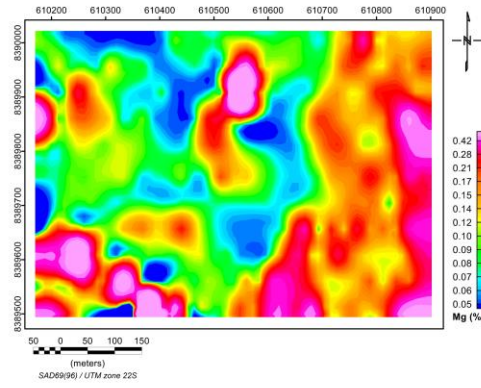
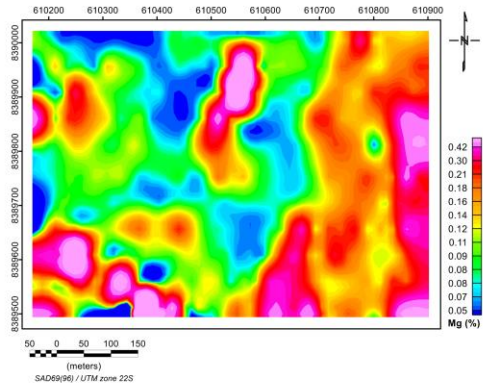


A

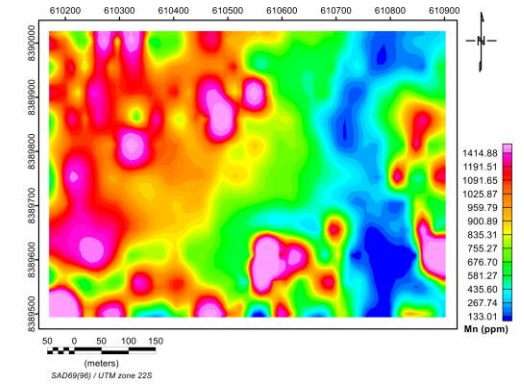
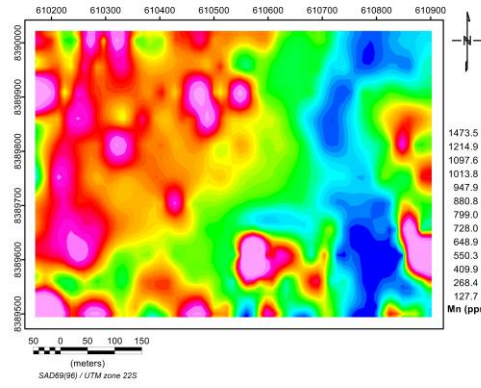
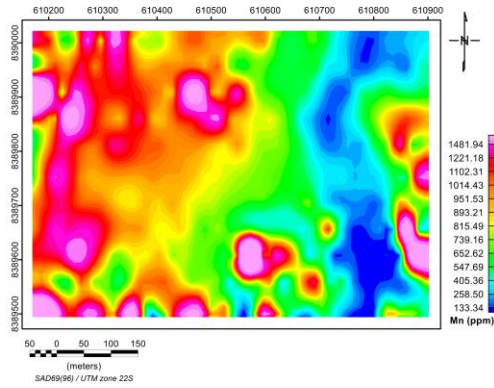
B

C

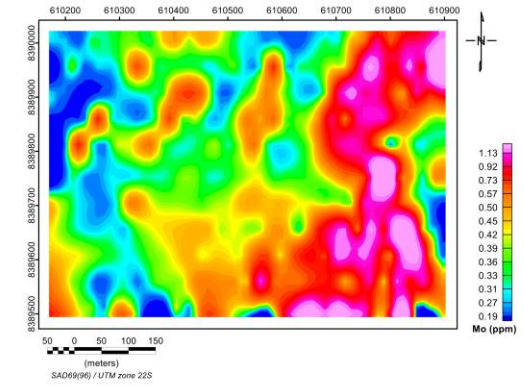
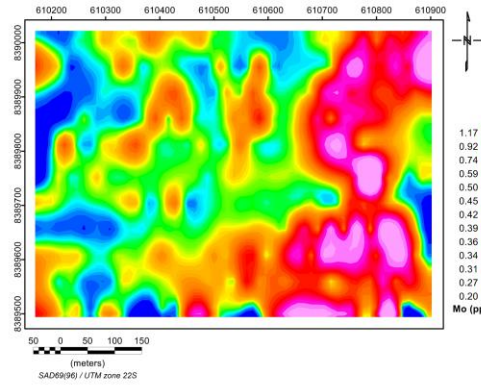
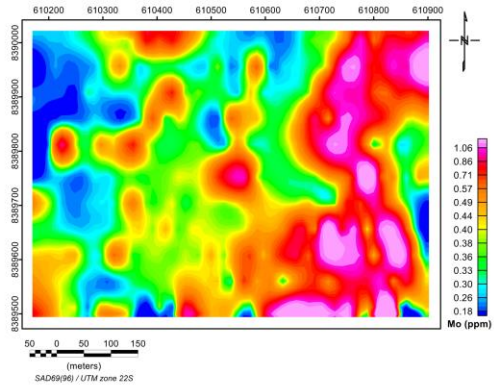
Mg



Mn

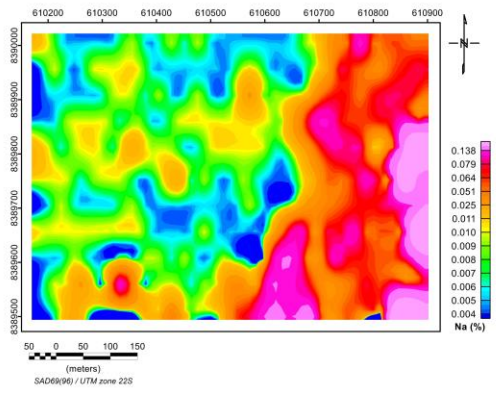


Mo

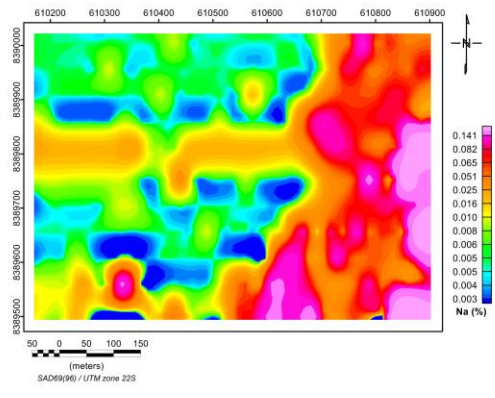


Na

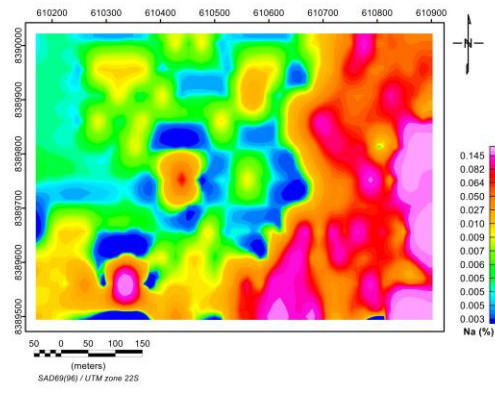
A



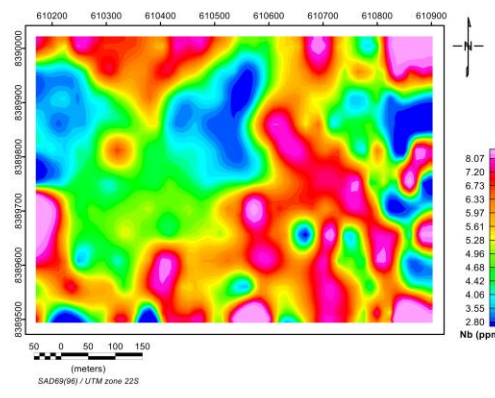
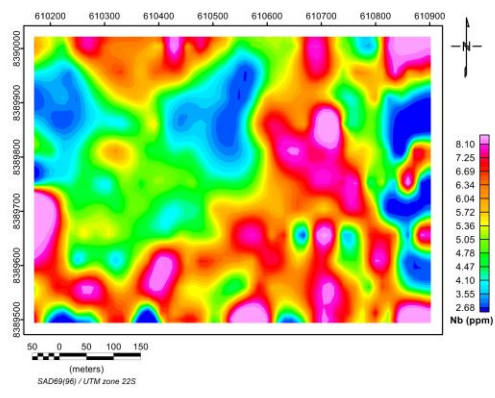
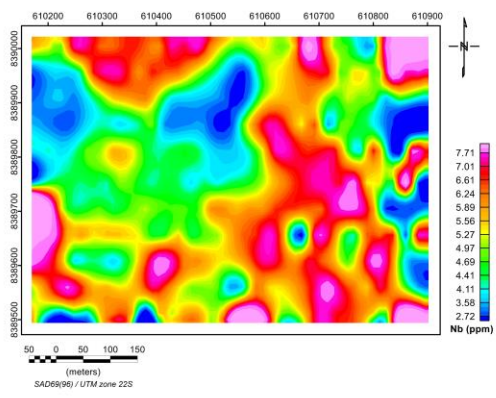
B



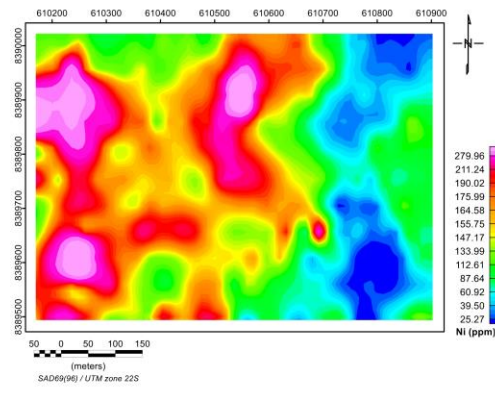
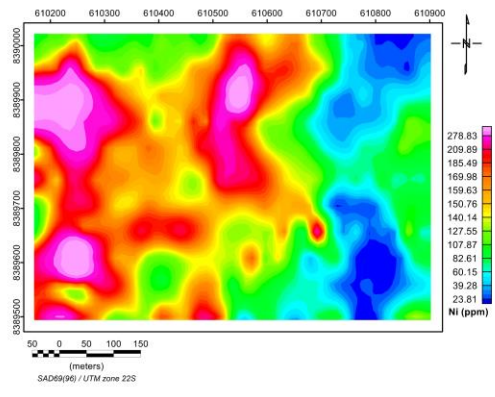
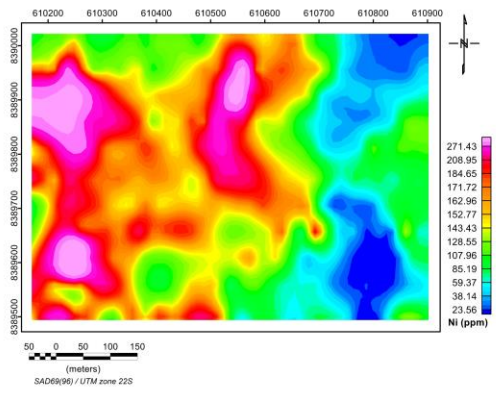
C



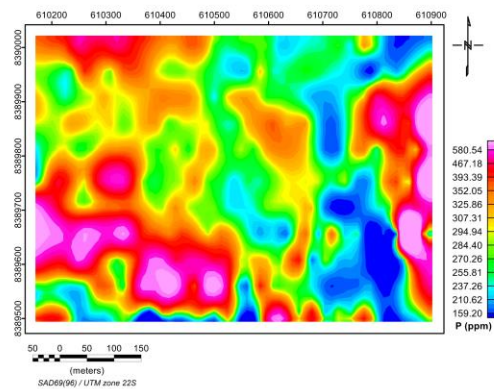
Nb



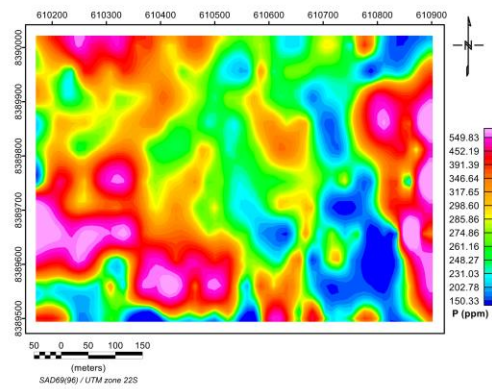
Ni



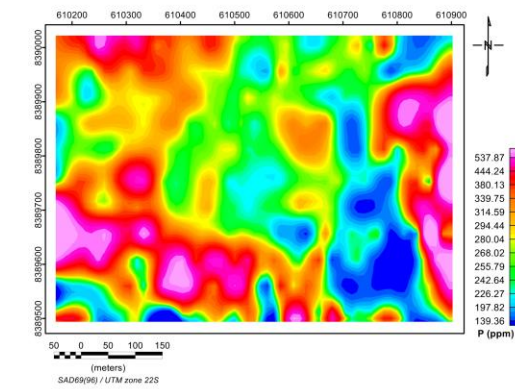
A



B

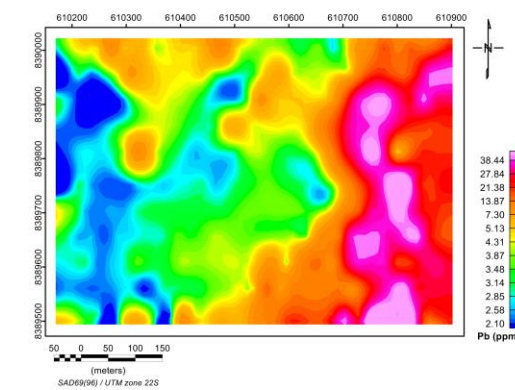
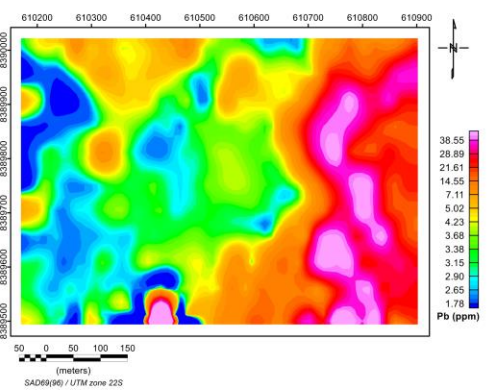
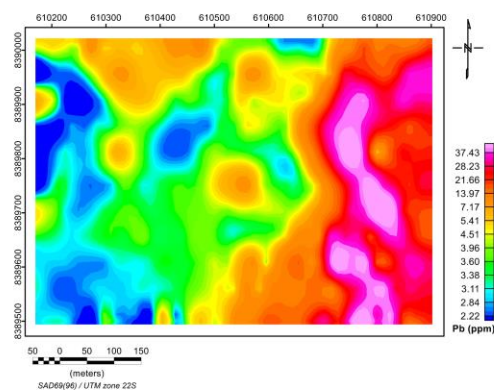


C

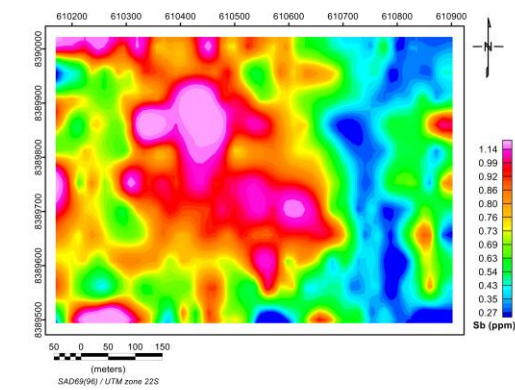
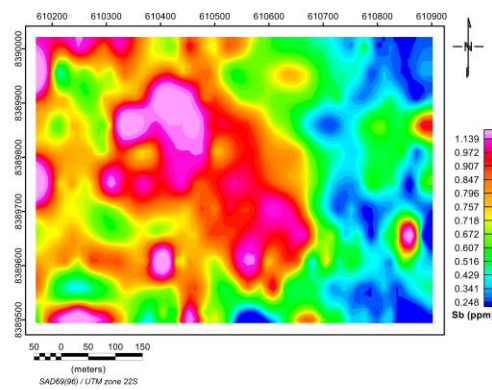
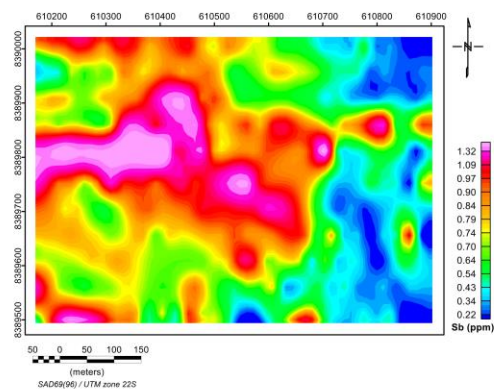


P

Pb



Sb

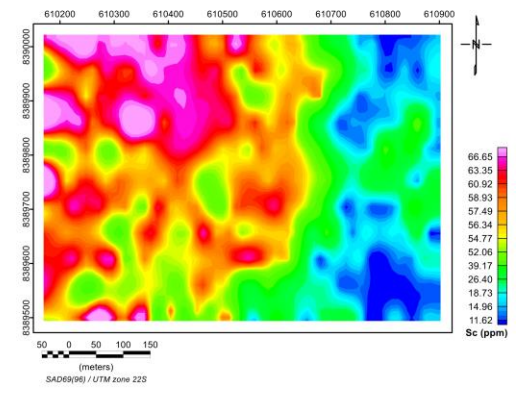
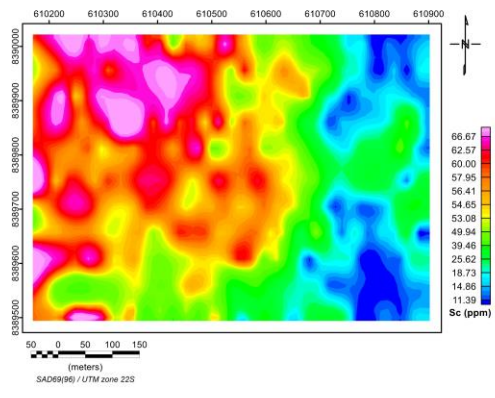
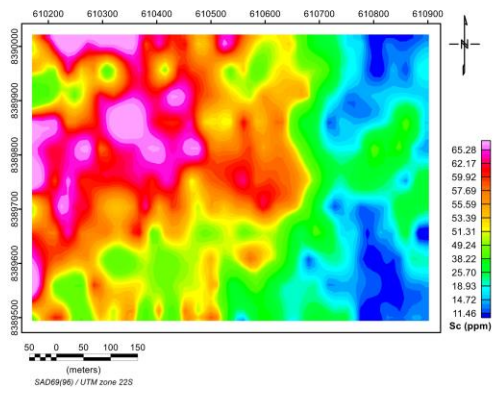


A

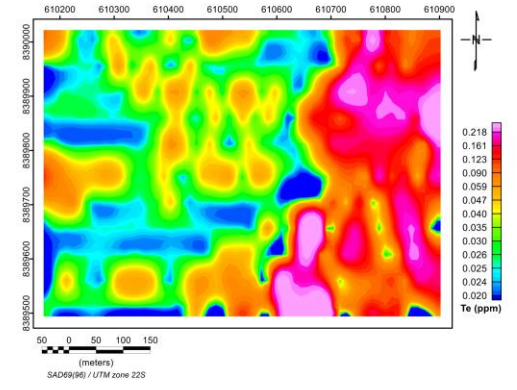
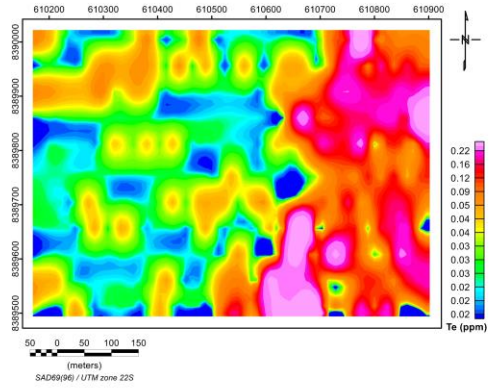
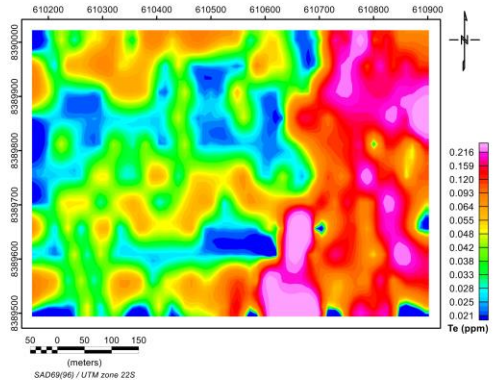
B

C

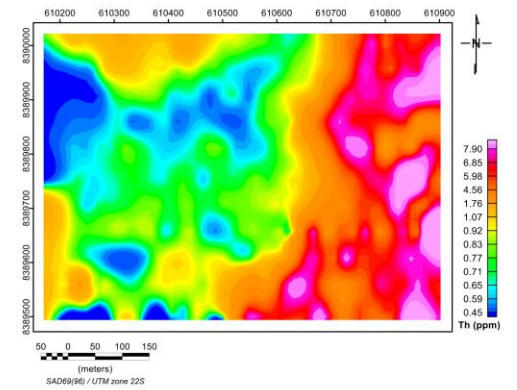
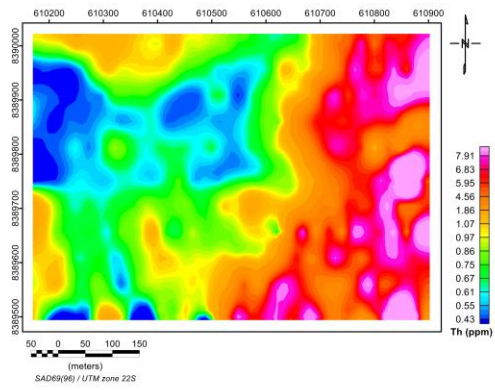
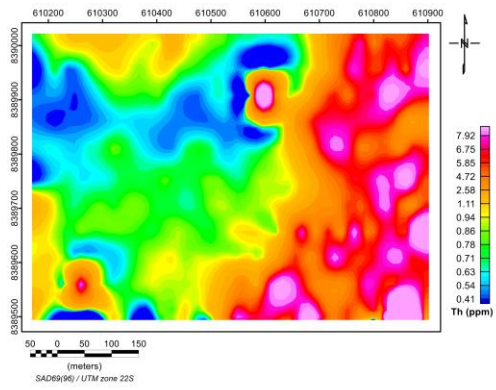
Sc



Te



Th

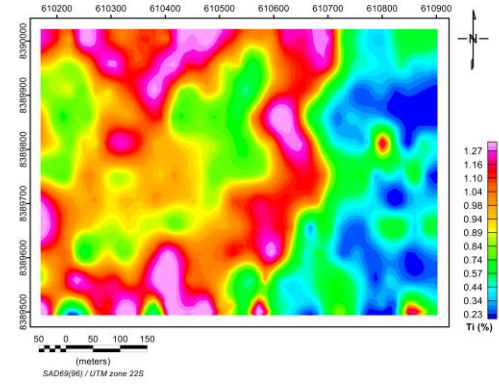
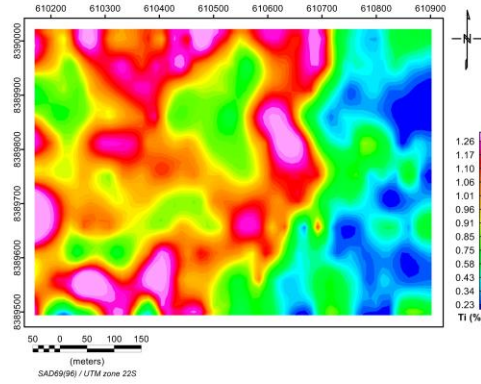
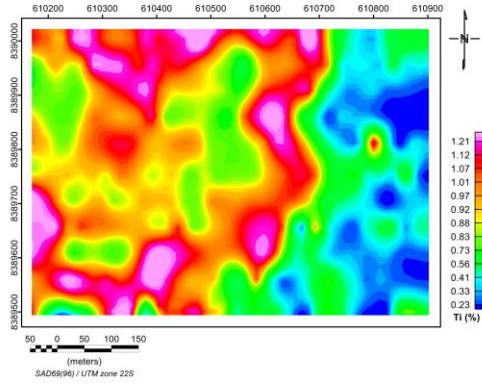


A

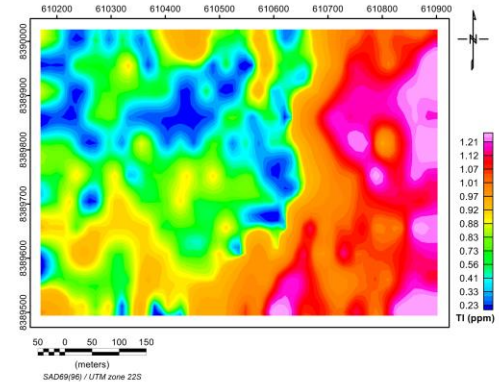
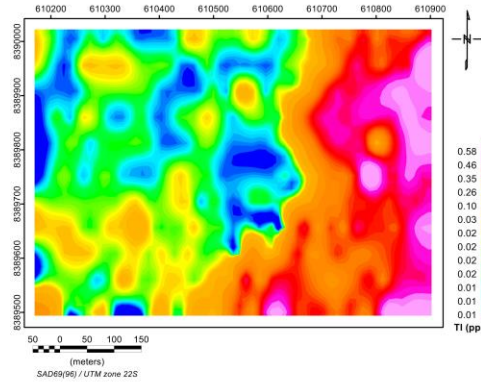
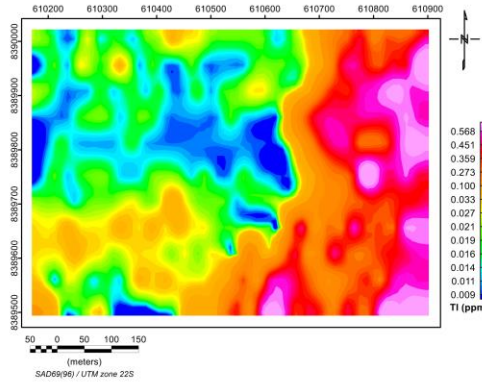
B

C

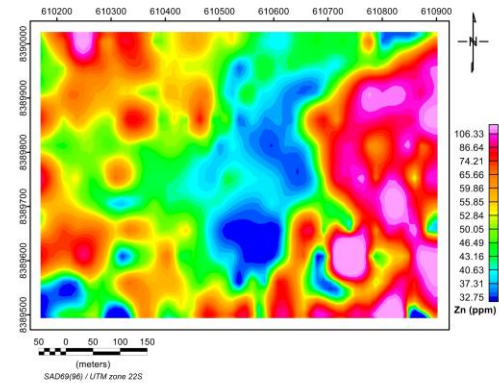
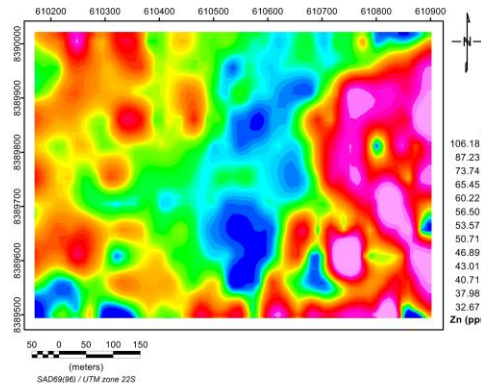
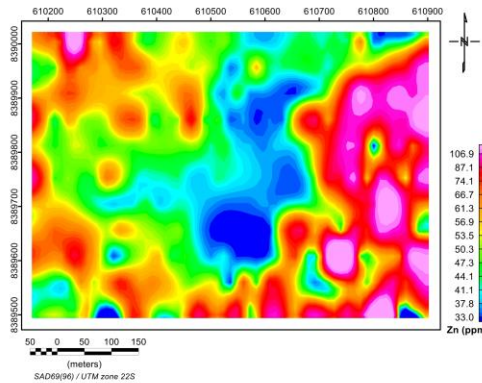
Ti



Ti

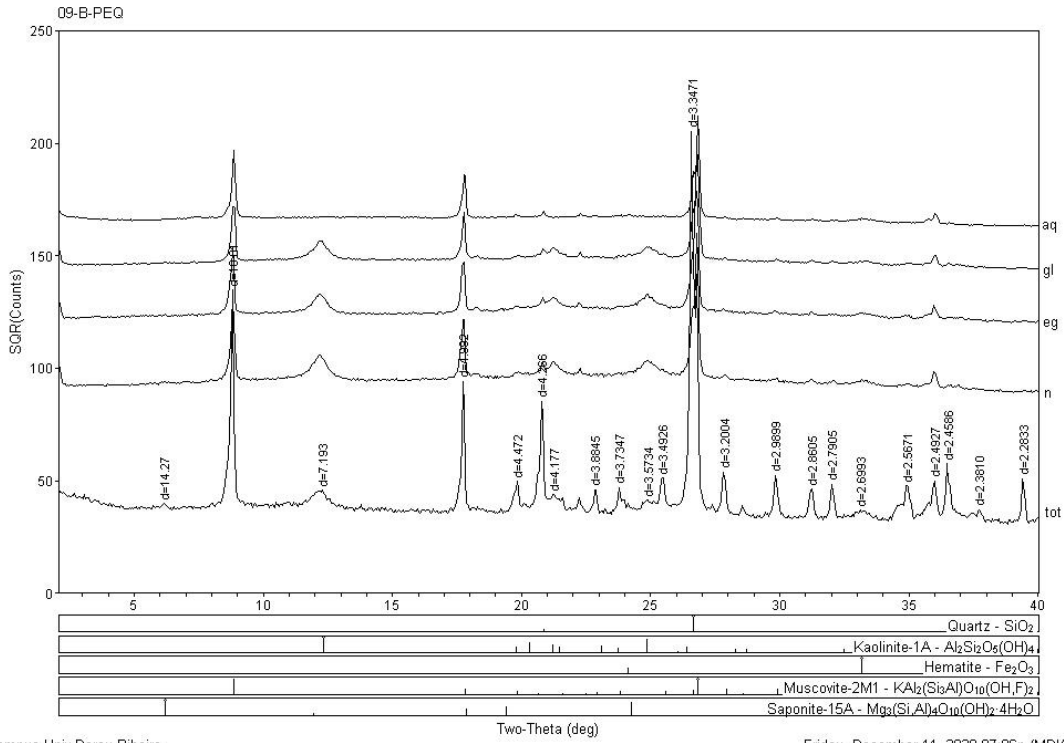


Zn



Anexo 2 – Difractometria de Raio-X em amostras de solo

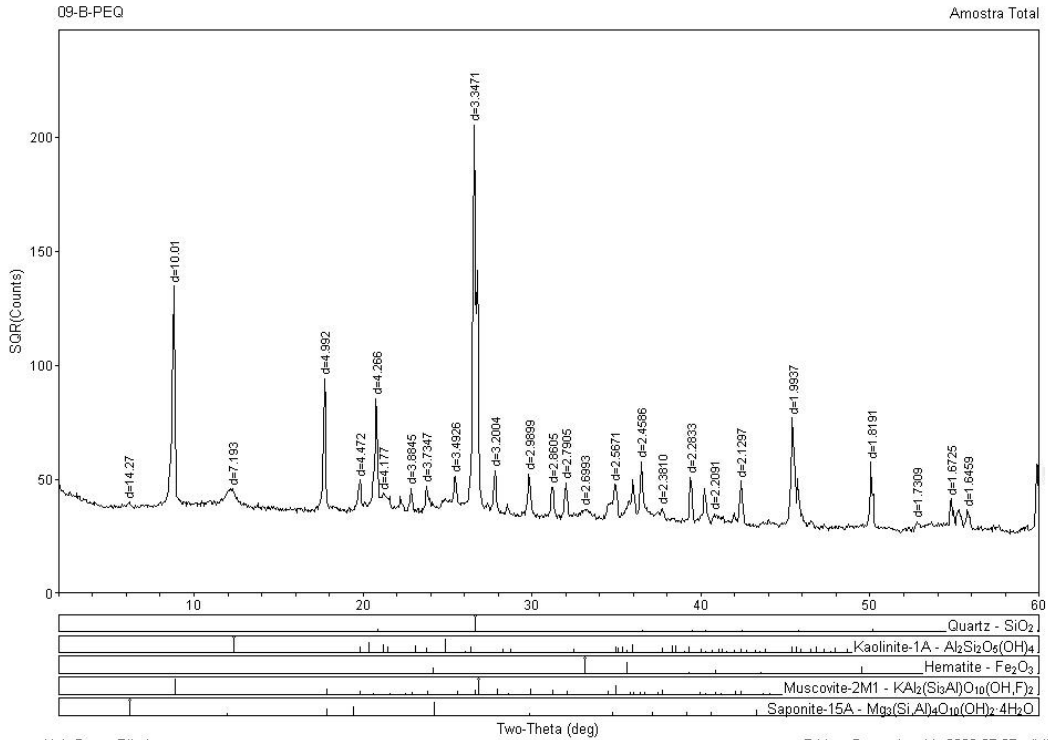
Amostra PEQS009B - Argilas



Campus Univ Darcy Ribeiro

Friday, December 11, 2020 07:06p (MDI/JADE9)

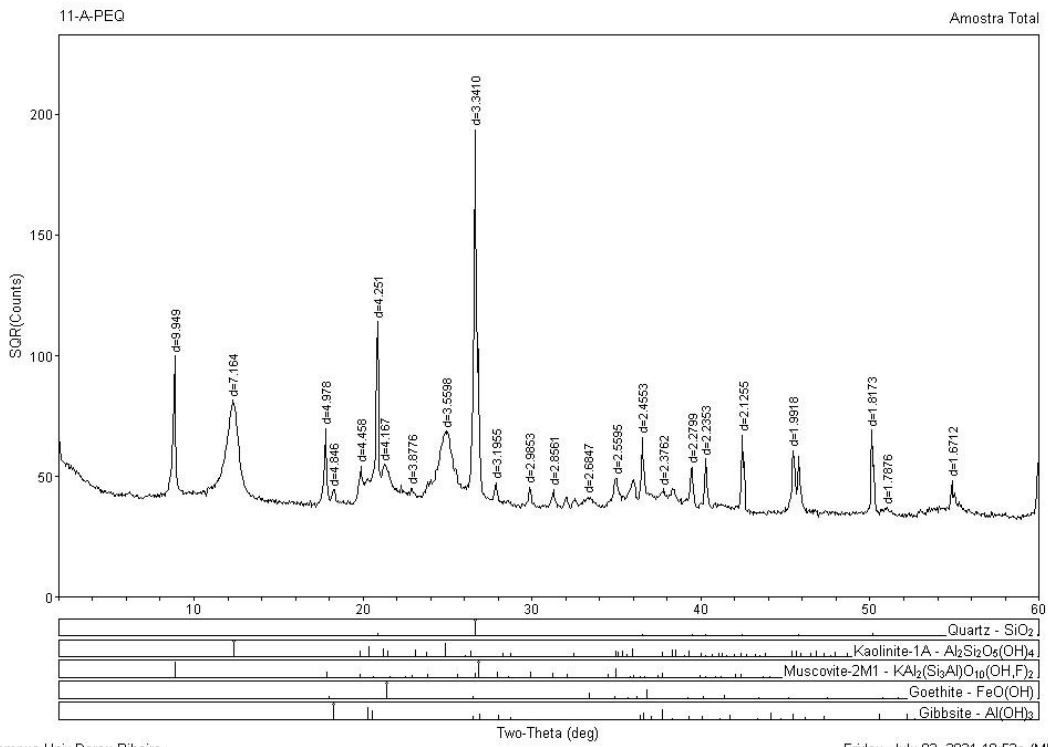
Amostra PEQS009B - Total



Campus Univ Darcy Ribeiro

Friday, December 11, 2020 07:07p (MDI/JADE9)

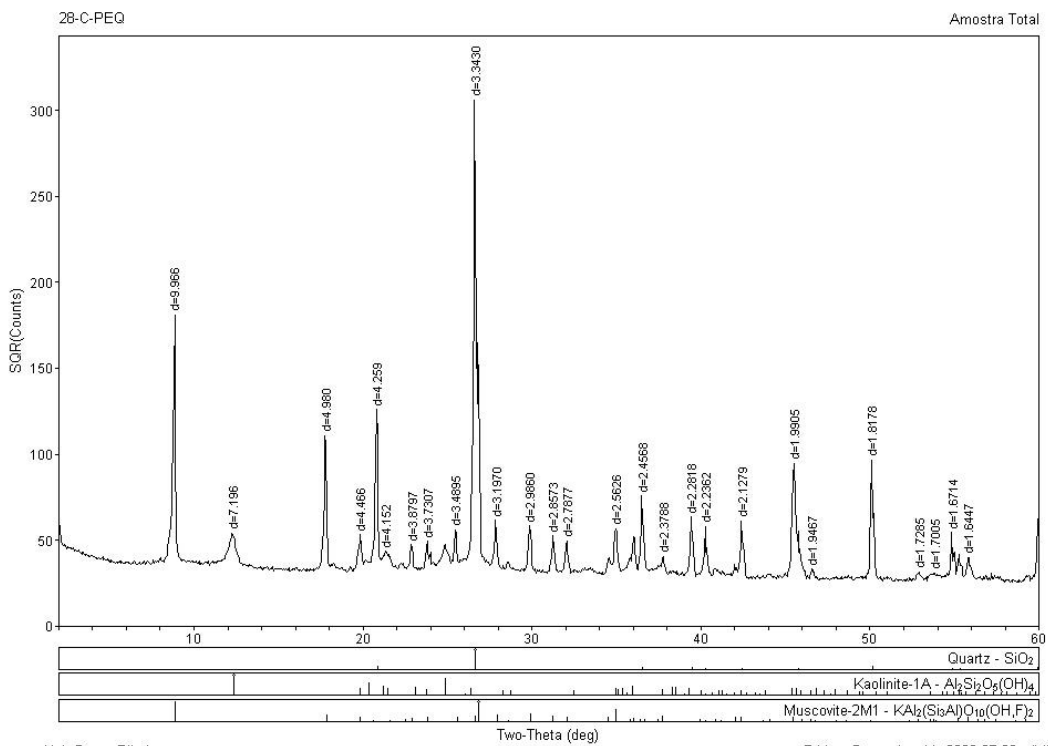
Amostra PEQSO011A - Total



Campus Univ Darcy Ribeiro

Friday, July 02, 2021 10:53a (MDI/JADE9)

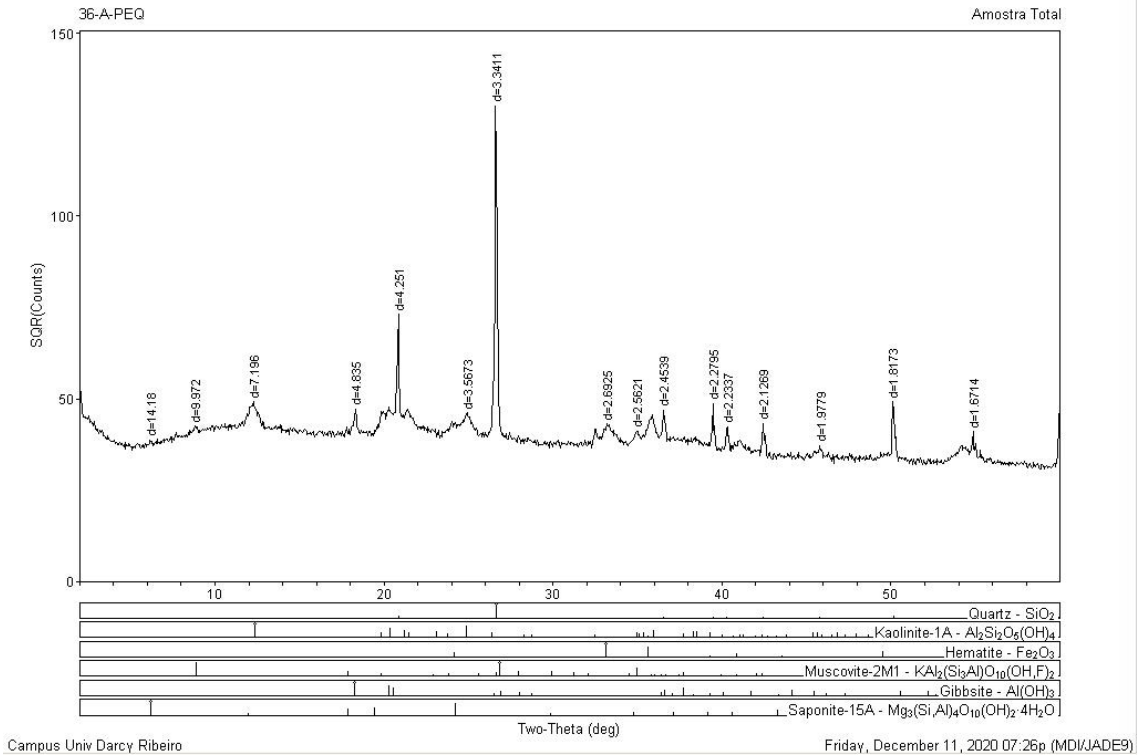
Amostra PEQSO028C - Total



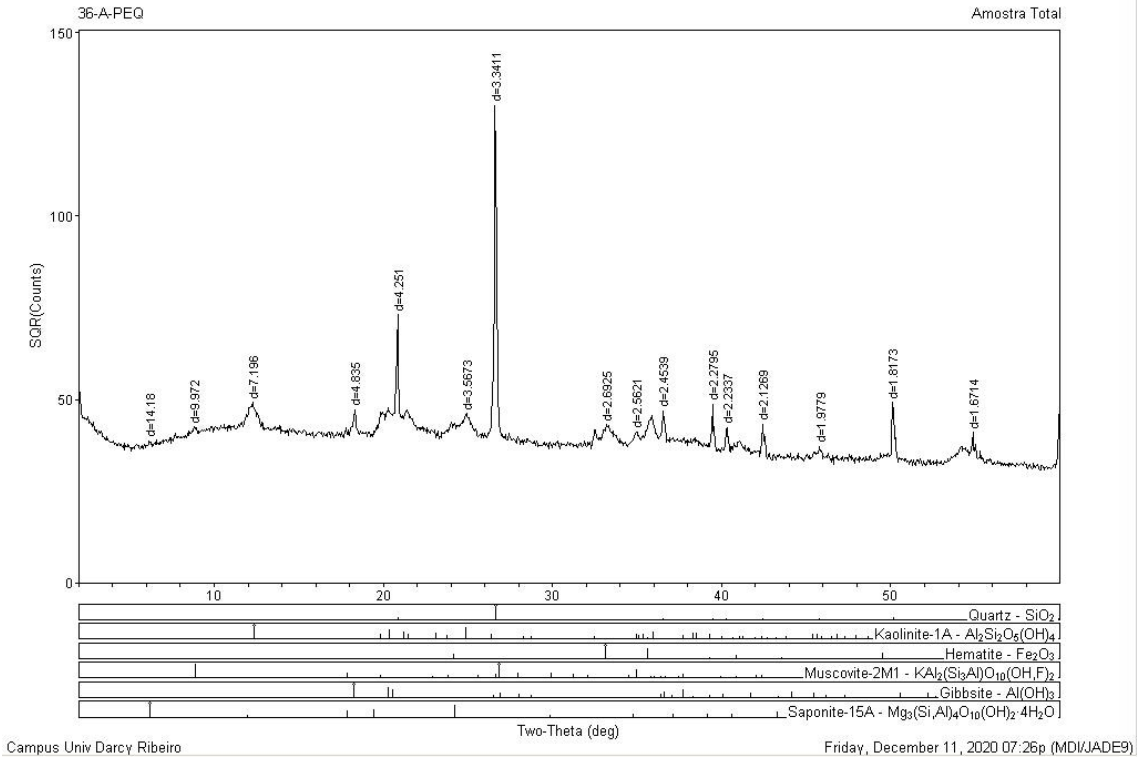
Campus Univ Darcy Ribeiro

Friday, December 11, 2020 07:22p (MDI/JADE9)

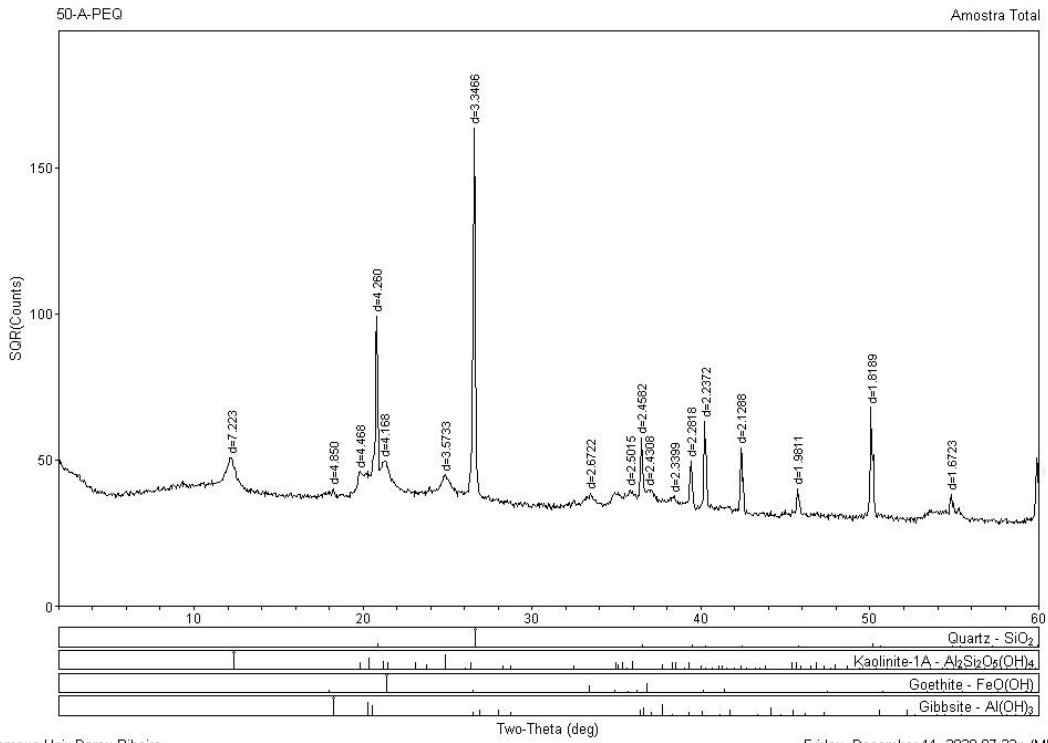
Amostra PEQS0036A - Argila



Amostra PEQS0036A - Total



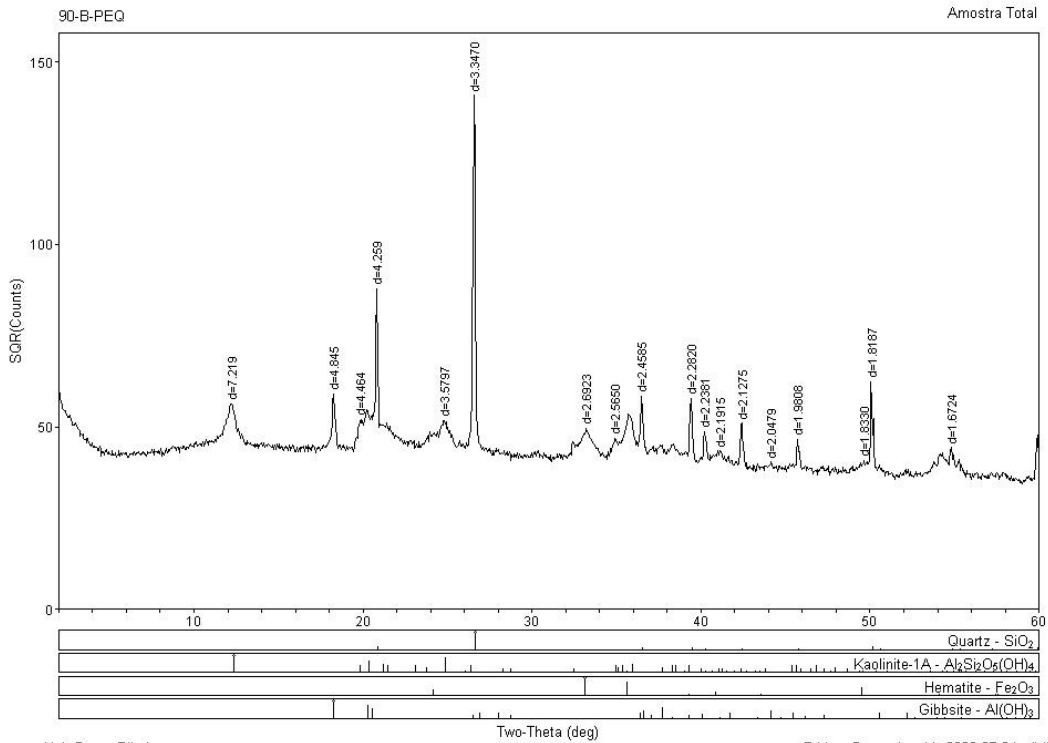
Amostra PEQSO050A - Total



Campus Univ Darcy Ribeiro

Friday, December 11, 2020 07:23p (MDI/JADE9)

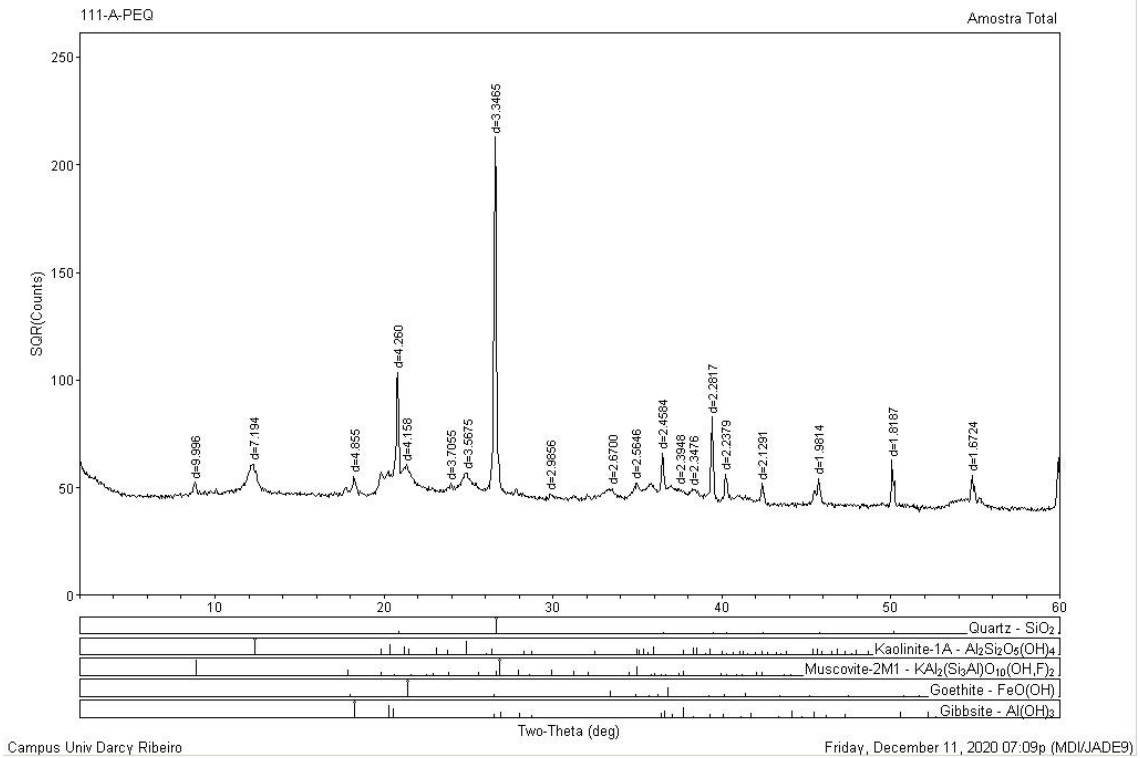
Amostra PEQSO090B - Total



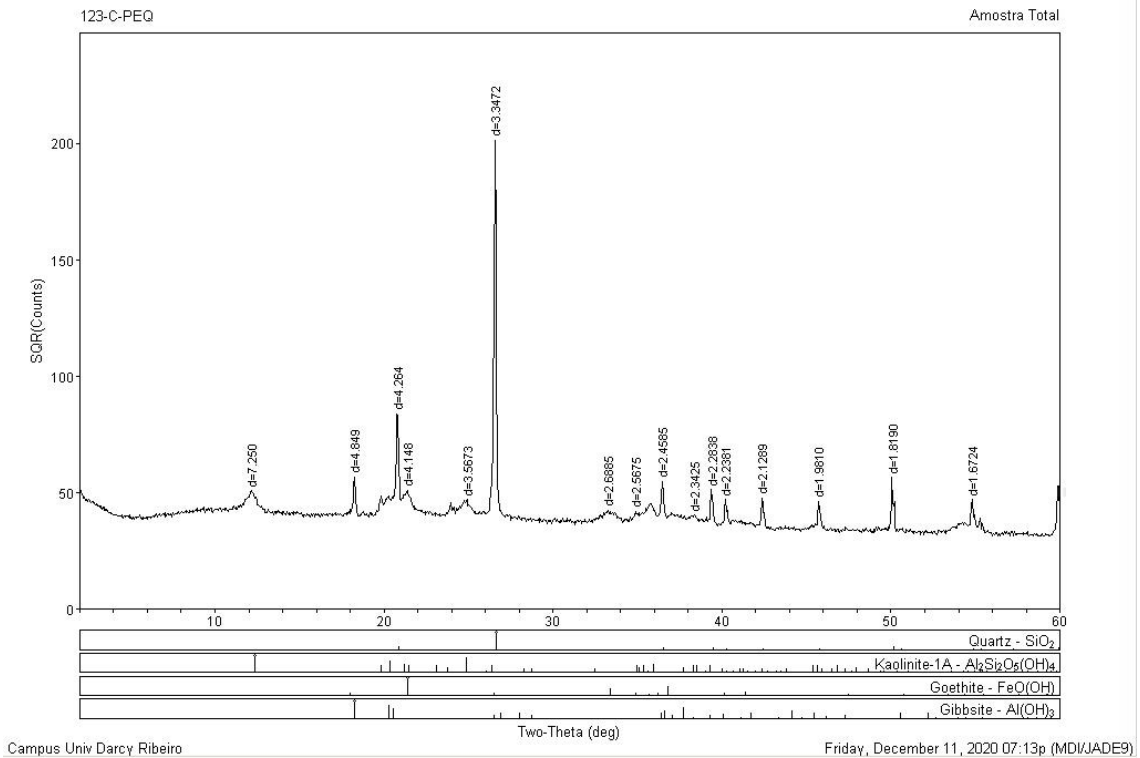
Campus Univ Darcy Ribeiro

Friday, December 11, 2020 07:24p (MDI/JADE9)

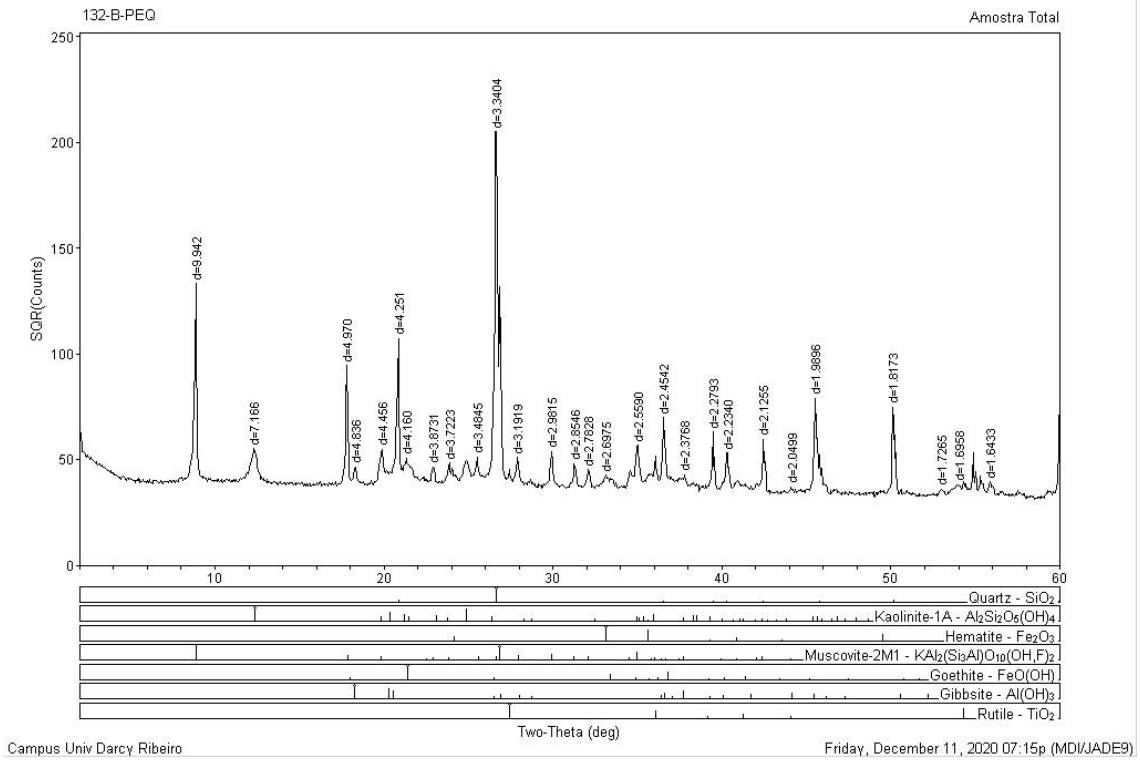
Amostra PEQSO111A - Total



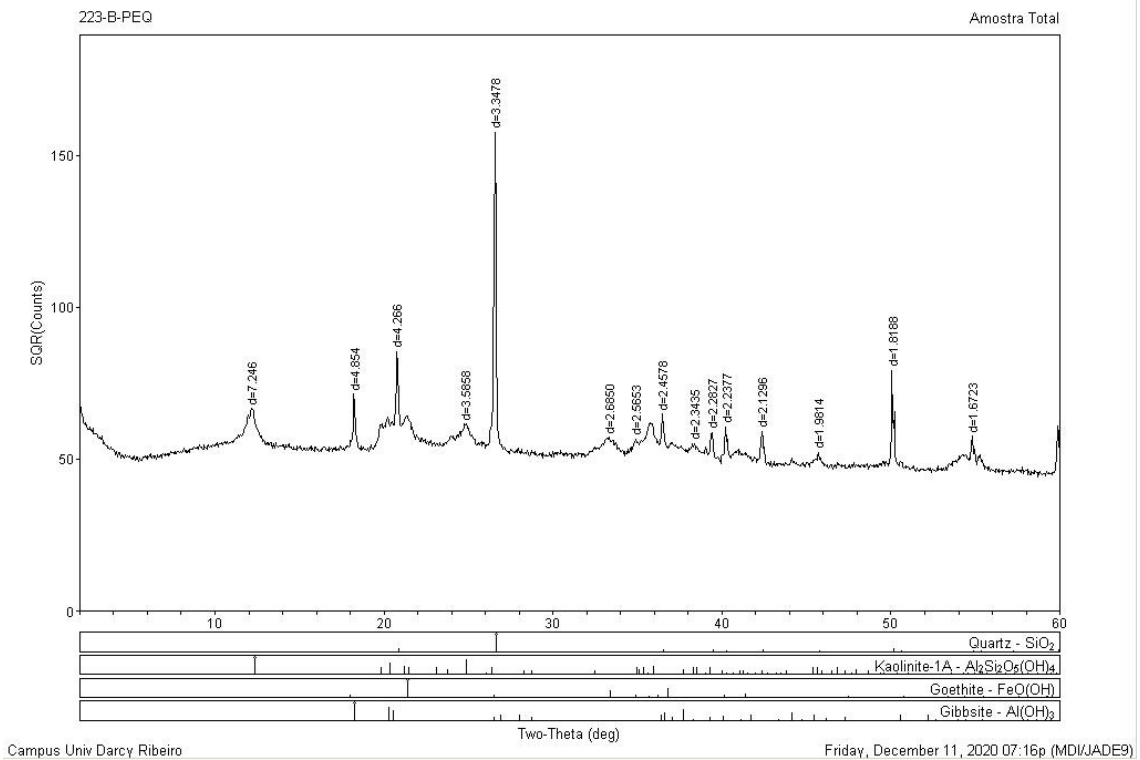
Amostra PEQSO125C - Total



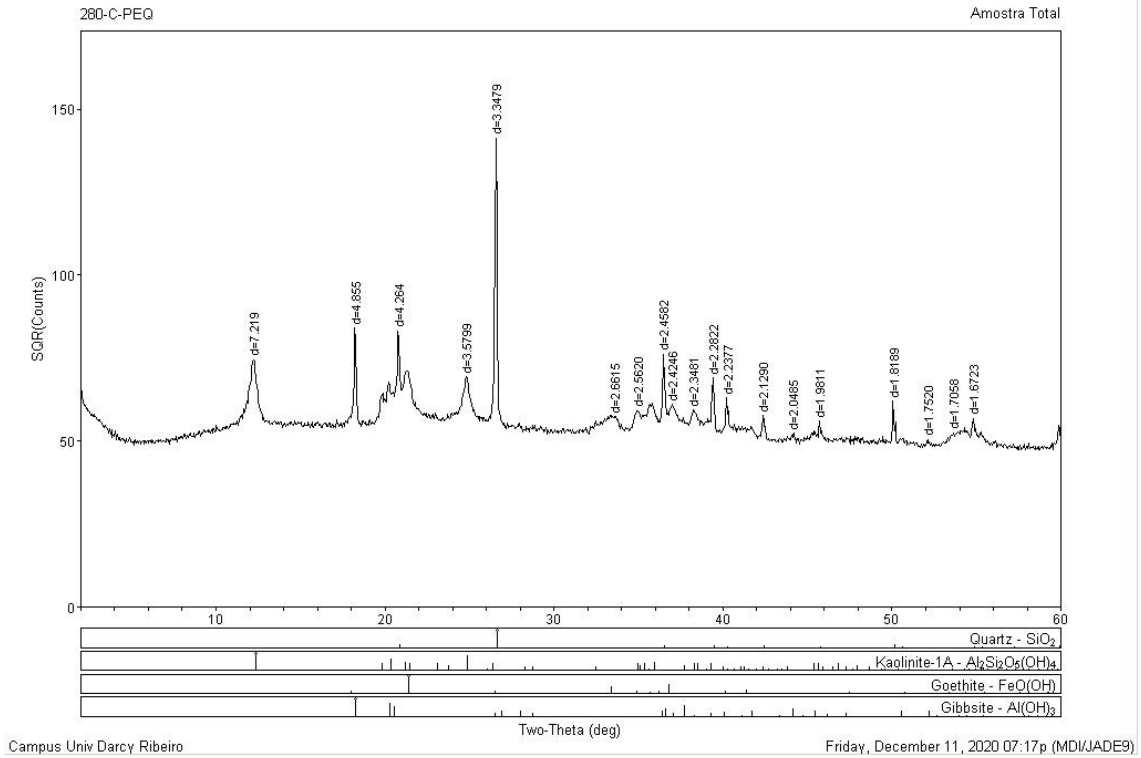
Amostra PEQSO132B - Total



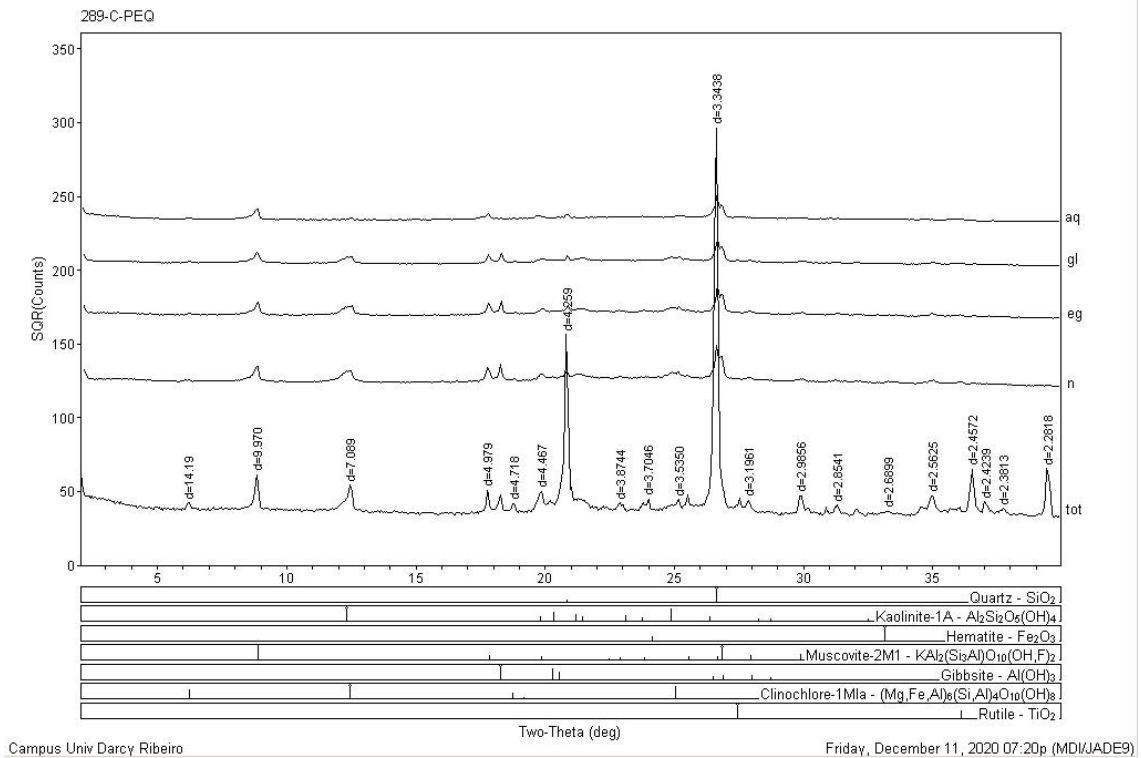
Amostra PEQSO229B - Total



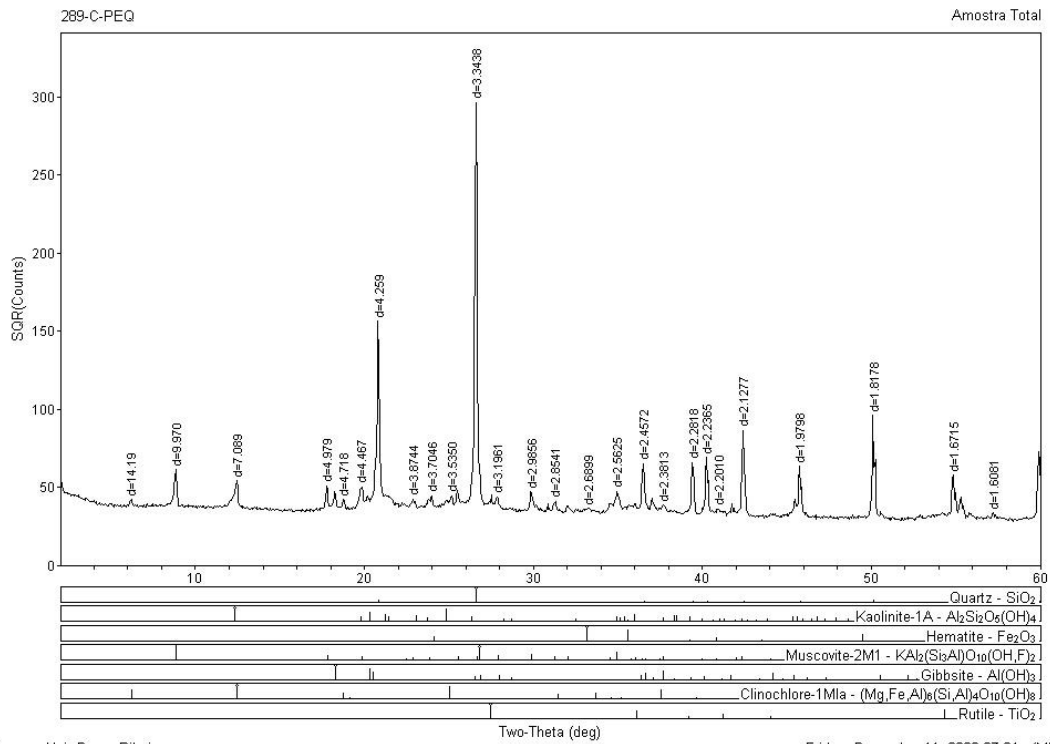
Amostra PEQSO280C - Total



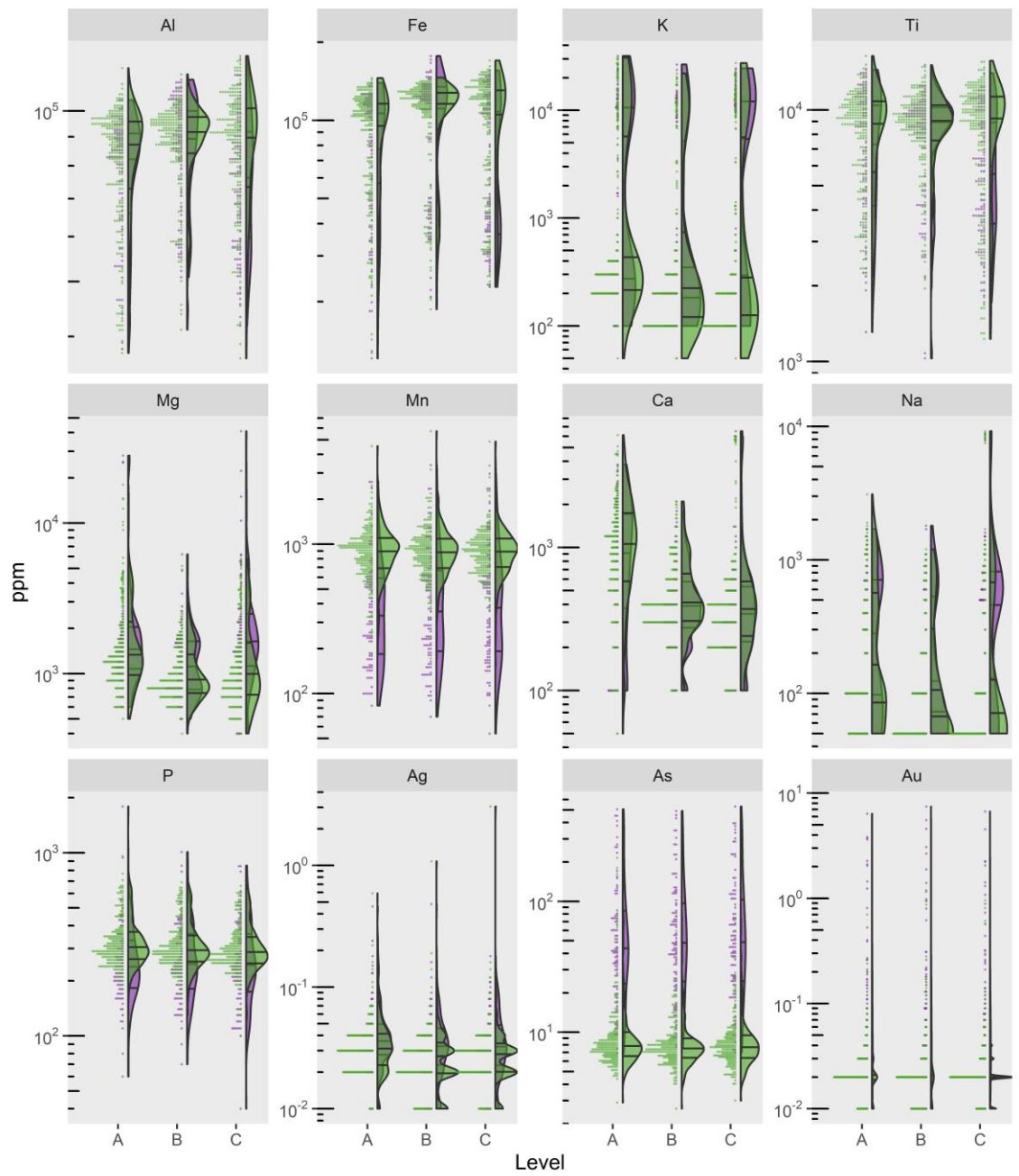
Amostra PEQSO289C - Argilas

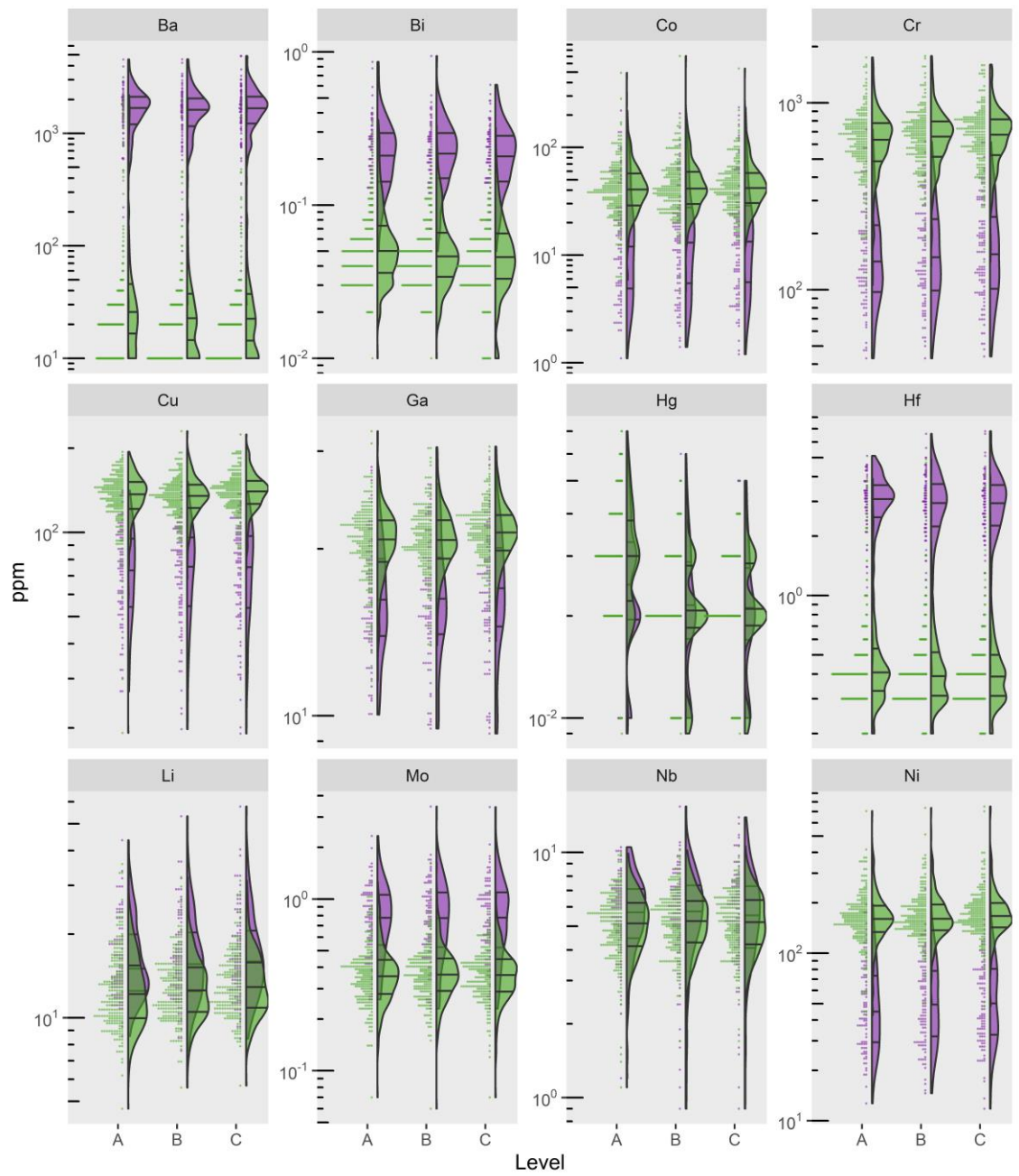


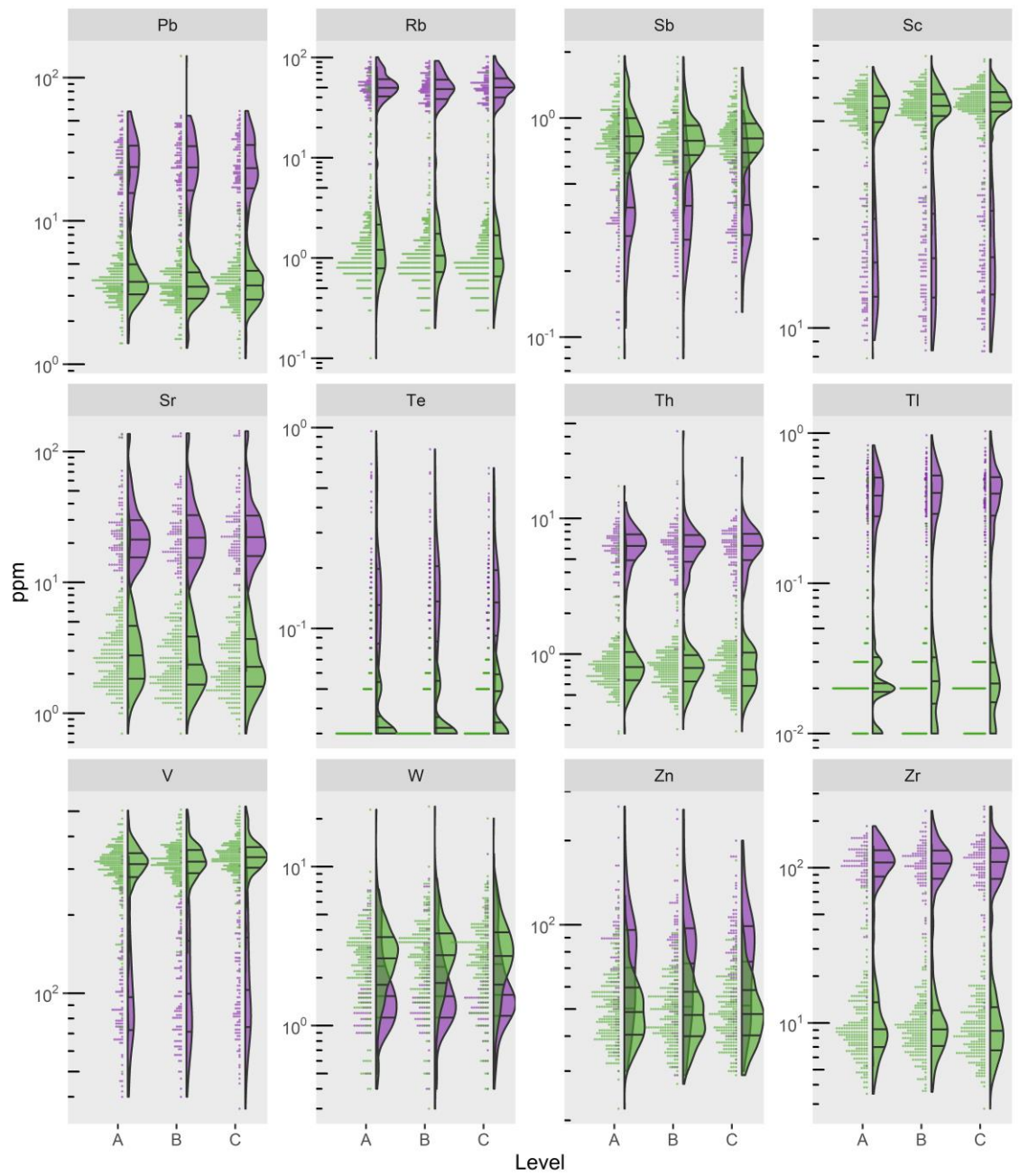
Amostra PEQSO289C - Total



Supplementary Data 1 - Violin-plots for A, B and C level

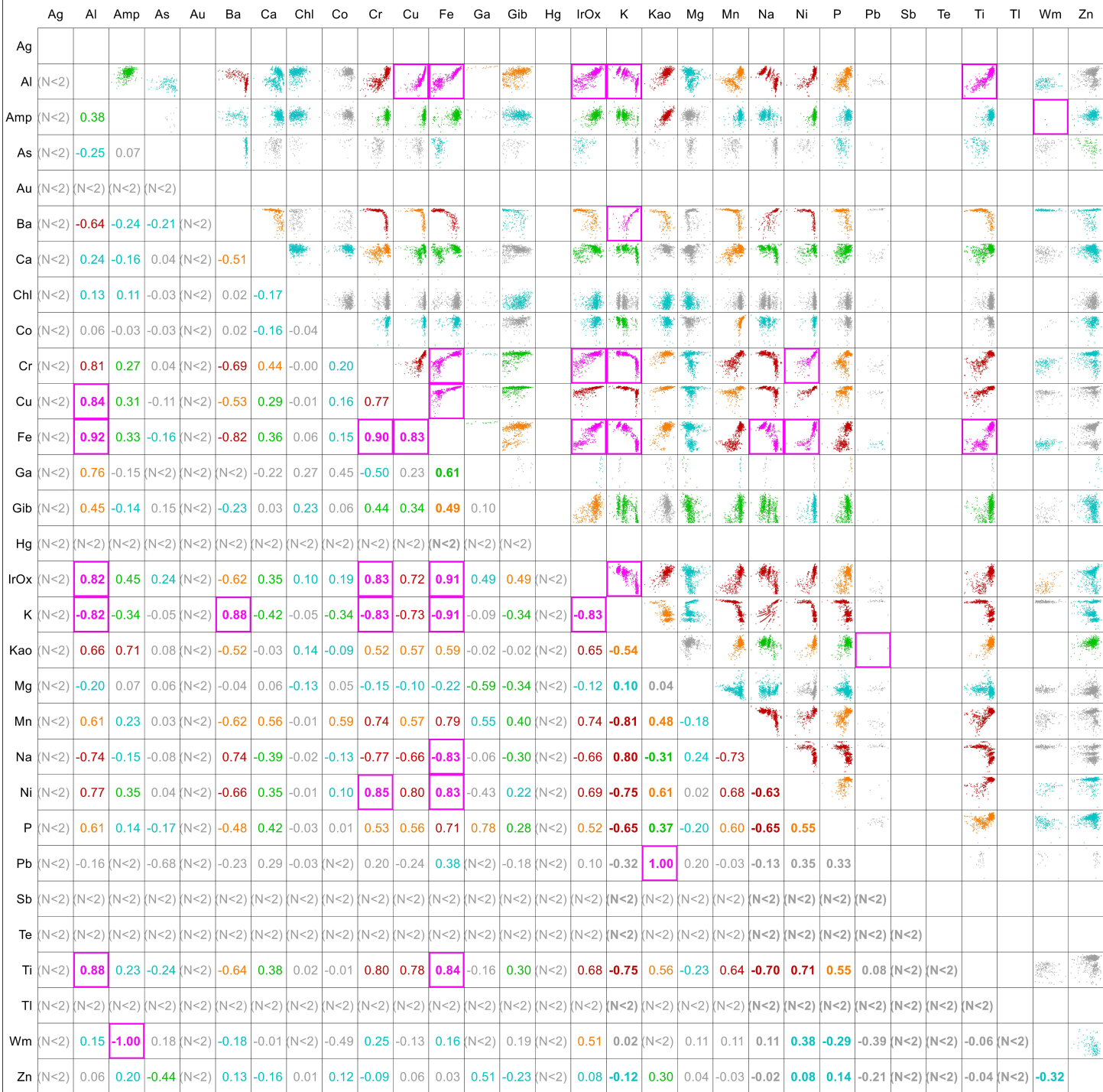






Level

Supplementary Data 2 - Correlation matrix for A, B and C level



Colour Coding (significance 0.95)

- very strong
- strong
- moderate
- weak
- very weak
- null

Sample Size 938

Appendix 1. Summary of geochemical data.

Element (ppm)	Variation coeff. (%)	Mean	Minimum	1st Quartile	Median	3rd Quartile	2nd Order Threshold (Q3+1,5[Q3-Q1])	1st Order Threshold (Q3+3[Q3-Q1])	Maximum
Ti	39.87	8274	1030	5520	8985	10700	18470	26240	16400
Al	21.17	84707	36600	73900	89000	97200	132150	167100	125000
Fe	41.91	94960	12100	51900	110250	125500	235900	346300	177000
Mg	145	1739	400	800	1200	1700	3050	4400	40800
Ca	121	823	50	300	500	1000	2050	3100	6500
Na	248	440	50	50	100	500	1175	1850	9200
K	150	4798	50	200	300	10100	24950	39800	31900
Ag	278	0.04	0.01	0.02	0.03	0.04	0.07	0.10	3.05
As	201	33.60	2.600	7.0	8.900	26.40	55.50	84.60	531
Au	438	0.132	0.005	0.016	0.021	0.039	0.074	0.1080	7.470
Ba	152	590	10	10	30	1260	3135	5010	4880
Be	75.83	0.6290	0.210	0.30	0.38	0.93	1.875	2.820	2.980
Bi	98.08	0.1126	0.005	0.04	0.06	0.15	0.3150	0.48	0.94
Co	106	43.16	1.100	20.40	36.05	53.60	103	153	706
Cr	61.03	523	43	219	542	735	1509	2283	1780
Cu	31.76	118	19.10	94.10	127	145	220	296	230
Ga	19.18	19.77	9.290	17.55	20.20	22.20	29.17	36.15	32.60
Hf	105	1.287	0.200	0.300	0.500	2.30	5.3	8.300	6.80
Hg	42.34	0.024	0.003	0.017	0.022	0.029	0.047	0.065	0.074
Li	37.59	14.63	4.70	11	13.35	16.70	25.25	33.80	57.70
Mn	63.43	819	54	528	789	1020	1758	2496	5710
Mo	65.95	0.53	0.06	0.31	0.41	0.62	1.085	1.550	3.460
Nb	32.83	5.507	0.90	4.30	5.30	6.50	9.8	13.10	15.40
Ni	59.77	140	11.80	79.70	144	178	325	473	751
P	42.66	313	40	240	280	360	540	720	1790
Pb	117	11.24	1.100	3.200	4.400	17.10	37.95	58.80	142
Rb	140	18.30	0.100	0.900	1.700	40.40	99.65	159.00	104
Sb	42.06	0.7219	0.080	0.500	0.740	0.890	1.475	2.06	1.920
Sr	165	12.03	0.70	1.90	3.70	16.70	38.90	61.10	143
Te	124	0.08	0.03	0.03	0.05	0.09	0.1875	0.285	0.960
Th	122	2.78	0.26	0.70	0.98	4.98	11.40	17.82	43.90
Tl	134	0.15	0.01	0.02	0.03	0.30	0.720	1.140	1.030
V	43.30	260	36	154	299	340	619	898	522
W	71.82	2.747	0.30	1.50	2.40	3.40	6.250	9.100	23.90
Zn	46.58	59.62	22	42	52	68	107	146	265
Zr	119	42.90	2.800	7.900	11.90	84.70	200	315	248