



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
DEPARTAMENTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
MODELAGEM EM CIÊNCIAS DA TERRA E DO
AMBIENTE**



**MODELAGEM DO POTENCIAL MINERAL DE BAUXITA A PARTIR
DE DADOS AEROGEOFÍSICOS E DE SENSORIAMENTO REMOTO
NA PORÇÃO SUDESTE DA BAHIA, BRASIL**

RODRIGO SOARES VIEIRA DOS SANTOS

Orientador(a): **Prof. Dr. Washington de Jesus Sant'anna da Franca-Rocha**

Co-Orientador(a): **Dr. Edgar Romeo Herrera de Figueiredo Iza**

Feira de Santana - BA

2023

RODRIGO SOARES VIEIRA DOS SANTOS

**MODELAGEM DO POTENCIAL MINERAL DE BAUXITA A PARTIR
DE DADOS AEROGEOFÍSICOS E DE SENSORIAMENTO REMOTO
NA PORÇÃO SUDESTE DA BAHIA, BRASIL**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem em Ciências da Terra e do Ambiente (PPGM), da Universidade Estadual de Feira de Santana como requisito parcial para obtenção de título de Mestre em Ciências Ambientais.

Orientador(a): Prof. Dr. Washington de Jesus
Sant'anna da Franca-Rocha

Co-Orientador(a): Dr. Edgar Romeo Herrera de
Figueiredo Iza

Feira de Santana – BA

2023

Rodrigo Soares Vieira dos Santos

**"MODELAGEM DO POTENCIAL MINERAL DE BAUXITA A
PARTIR DE DADOS AEROGEOFÍSICOS E DE SENSORIAMENTO
REMOTO NA PORÇÃO SUDESTE DA BAHIA, BRASIL"**

Dissertação apresentada ao Programa de
Pós-Graduação em Modelagem em Ciências da
Terra e do Ambiente da Universidade Estadual
de Feira de Santana.

Orientador: Prof. Dr. Washington de Jesus
Sant'Anna da Franca Rocha.

Linha de Pesquisa: Estudos Ambientais e
Geotecnologias.

Data de aprovação: 17 de março de 2023.

BANCA EXAMINADORA:

Prof. Dr. Washington de Jesus Sant'Anna da Franca Rocha - Orientador
Universidade Estadual de Feira de Santana - UEFS

Dr. Edgar Romeo Herrera de Figueiredo Iza - Coorientador
Companhia de Pesquisa de Recursos Minerais - CPRM

Profª Drª Tati de Almeida
Universidade de Brasília - UNB

Prof. Dr. Rodrigo Nogueira de Vasconcelos
Universidade Estadual de Feira de Santana - UEFS

Ficha Catalográfica - Biblioteca Central Julieta Carteadó - UEFS

Santos, Rodrigo Soares Vieira dos
S238m Modelagem do potencial mineral de bauxita a partir de dados
aerogeofísicos e de sensoriamento remoto na porção sudeste da Bahia, Brasil /
Rodrigo Soares Vieira dos Santos . – 2023.
66 f.: il.

Orientador: Washington de Jesus Sant'anna da Franca-Rocha
Coorientador: Edgar Romeo Herrera de Figueiredo Iza
Dissertação (mestrado) – Universidade Estadual de Feira de Santana,
Programa de Pós-Graduação em Modelagem em Ciências da Terra e do
Ambiente, Feira de Santana, 2023.

1. Recursos minerais. 2. Bauxita. 3. Regolito laterítico. 4. Aerogeofísica.
5. Sensoriamento remoto. I. Franca-Rocha, Washington de Jesus Sant'anna
da, orient. II. Iza, Edgar Romeo Herrera de Figueiredo, coorient.
II. Universidade Estadual de Feira de Santana. IV. Título.

CDU 553.3/.9(814.2)

Daniela Machado Sampaio Costa - Bibliotecária - CRB-5/2077

RESUMO

O presente trabalho visou a construção de um modelo de potencial mineral de bauxita a partir da integração de dados aerogeofísicos e de sensoriamento remoto em uma área que corresponde à porção leste do Cráton São Francisco no sudeste do estado da Bahia, Brasil. Estudos recentes demonstraram um potencial para ocorrências de bauxita e outros minerais supergênicos na região. Esses minerais estão ligados à presença de perfis lateríticos bem desenvolvidos a partir da ação do intemperismo. Para construir o modelo, foram gerados mapas de concentração de K, eTh e eU, que nos permitem identificar a assinatura geoquímica dos materiais associados ao regolito, além do Gradiente Total, que nos fornece informações acerca da susceptibilidade magnética dos materiais. Para estudar o contexto geomorfológico da área, foram gerados um modelo digital de terreno (SRTM) e um mapa de declividade. Imagens *Landsat 8* também foram utilizadas com o intuito de estudar a distribuição espacial dos materiais superficiais a partir da sua composição química, além de um mapa litológico para fornecer informações acerca da rocha-fonte dos materiais residuais. O algoritmo de aprendizado supervisionado *Random Forest* foi aplicado para integrar os mapas de evidências gerados, resultando no modelo de potencial mineral de bauxita para a região. O modelo resultante reduziu consideravelmente a área de pesquisa e obteve medidores de desempenho satisfatórios, atingindo um *F-score* de 0,81 e um coeficiente *kappa* de 0,83. As áreas sugeridas pelo modelo como alvos reforçam o potencial mineral de áreas com ocorrências já conhecidas, além de sugerir novos alvos para prospecção. A metodologia foi implementada na plataforma de processamento digital de imagens na nuvem *Google Earth Engine*, a partir de um ambiente de desenvolvimento integrado (IDE) que utiliza linguagem de programação *JavaScript*. Além de sugerir alvos prospectivos, os resultados ainda indicaram que os mapas de evidência com maior importância na geração do modelo foram os gerados a partir do SRTM, além do urânio e do potássio.

Palavras-chave: Bauxita; Aerogeofísica; Sensoriamento Remoto; Aprendizado Supervisionado; *Random Forest*; *Google Earth Engine*.

ABSTRACT

The present work aimed to build a bauxite mineral potential model by integrating airborne geophysics and remote sensing data in an area that corresponds to the eastern portion of the São Francisco Craton in the southeast of Bahia state, Brazil. Recent studies have demonstrated a potential for bauxite and other supergene minerals in the region, linked to the presence of well-developed lateritic profiles resulting from weathering. To build the model, maps of K, eTh, and eU concentrations were generated to identify the geochemical signature of materials associated with regolith, as well as the Total Gradient, to provide information about the magnetic susceptibility of materials. To study the geomorphological context of the area, a digital elevation model (SRTM) and a slope map were generated. Landsat 8 scenes were also used to study the spatial distribution of surface materials based on their chemical composition, along with a lithological map to provide information about the source rock of the residual materials. The Random Forest supervised learning algorithm was applied to integrate the generated evidence maps, resulting in the bauxite mineral potential model for the region. The resulting model considerably reduced the study area and obtained satisfactory performance measures, achieving an F-score of 0.81 and a kappa coefficient of 0.83. The areas suggested by the model as targets reinforce the mineral potential of areas with known occurrences, as well as suggesting new targets for prospecting. The methodology was implemented on the Google Earth Engine cloud image processing platform using a JavaScript programming language integrated development environment (IDE). In addition to suggesting prospective targets, the results also indicated that the evidence maps with the most importance within the model generation were those from the SRTM, as well as from uranium and potassium.

Palavras-chave: Bauxite; airborne geophysics; Remote Sensing; Supervised Learning; Random Forest; Google Earth Engine.

LISTA DE FIGURAS

Figura 1 - Distribuição geral simplificada dos minerais lateríticos e herdados em um perfil laterítico completo e zoneamento metalogenético esquemático ao longo de um perfil laterítico, de acordo com cada horizonte.....	13
Figura 2 – Representação do espectro eletromagnético mostrando suas regiões em Relação ao comprimento de onda.....	16
Figura 3 – Esquema simplificado do funcionamento do algoritmo Random Forest com n árvores de decisão.....	20
Figura 4 – Um exemplo da geração de amostras bootstrap, que são criadas por substituição, gerando um conjunto de amostras que são sempre distintas entre si.....	21
Figura 5 – A) Localização da área de estudo; B) Mapa geológico simplificado de Souza <i>et al.</i> (2003).....	24
Figura 6 - Dados geoquímicos de Fernandes (1995) (Área do Projeto Terras-Raras), recursos minerais do SGB-CPRM (GeoSGB) e pontos de SEV, realizados por Santos (2010), sobrepostos ao índice laterítico.....	26
Figura 7 – Aspecto geral da área de ocorrência de crosta laterítica aluminosa (bauxita) na região entre os municípios de Jequié e Itagi.....	28
Figura 8 – Modelo conceitual esquematizado com os principais procedimentos realizados até a geração do modelo de potencial mineral.....	30
Figura 9 - Esferas do conhecimento que englobam os fatores críticos para formação e preservação de crostas lateríticas.....	31
Figura 10 - Caracterização do sistema mineral bauxita laterítica. Fatores críticos, processos de formação dos depósitos, critérios mapeáveis e proxies (vetores) em cada esfera de conhecimento.....	32
Figura 11 – Fluxograma padrão de pré-processamento dos dados dos aerolevantamentos geofísicos. Estas etapas são realizadas antes da disponibilização da base de dados para os usuários.....	36

Figura 12 – Integração dos grids de K dos três aerolevantamentos através da extensão <i>grid knitting</i>	37
Figura 13 – Dados gamaespectrométricos (A, B e C) e magnetométricos (D) preparados como mapas de evidência após realização do pré-processamento necessário.....	38
Figura 14 – A) Modelo Digital de Elevação SRTM e B) Mapa de Declividade.....	39
Figura 15 – Fluxograma da metodologia utilizada para construção de um mosaico de imagens <i>Landsat 8</i> sem nuvens para a área de estudo.....	41
Figura 16 – Ilustração da aplicação de um redutor em uma coleção de imagens utilizando a biblioteca <i>ee.Reducer</i> . O valor do pixel no resultado é computado a partir do valor de todas as imagens no ponto.....	42
Figura 17 – Localização dos pontos que formam a base de treinamento e teste do modelo sobreposta ao mosaico <i>Landsat 8</i> , composição 432 (Cor Natural)	43
Figura 18 – Captura de tela do GEE com a lista das bandas contidas na coleção <i>Landsat 8</i> e suas respectivas descrições.....	44
Figura 19 – A) Modelo <i>Random Forest</i> de potencial de bauxita apresentado na forma de classificação binária, onde as áreas em amarelo e preto correspondem às áreas com maior e menor potencial mineral, respectivamente. B) Modelo <i>Random Forest</i> de potencial mineral de bauxita apresentado na forma de probabilidade de classe.....	48
Figura 20 – Gráfico da importância relativa das variáveis utilizadas como mapas de evidência na geração do modelo de potencial mineral de bauxita.....	50
Figura 21 – A) Modelo Random Forest de potencial de bauxita; B) Máscara Litológica; C) Modelo após aplicação da máscara.....	51
Figura 22 – Modelo de potencial mineral de bauxita com a localização das 3 subáreas selecionadas para discussão. subárea 1, em vermelho; subárea 2, em preto e; subárea 3, em amarelo.....	52
Figura 23 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 1. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.....	53
Figura 24 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 2. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.....	54

Figura 25 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 3. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.....55

LISTA DE TABELAS

Tabela 1 - Sumário dos dados disponíveis, especificando sua natureza, fonte, formato e etapa em que serão utilizados.....	34
Tabela 2 – Mapas de evidências produzidos para construção do modelo de potencial mineral de bauxita.....	47
Tabela 3 – Medidores de desempenho do modelo <i>Random Forest</i> utilizando diversos conjuntos de mapas de evidência.....	50

SUMÁRIO

1. INTRODUÇÃO.....	11
2. BASE CONCEITUAL E ÁREA DE ESTUDO.....	12
2.1.Bauxita e Regolito Laterítico.....	12
2.2.Aerogeofísica: Gamaespectrometria e Magnetometria.....	14
2.3.Sensoriamento Remoto.....	16
2.4. <i>Machine Learning</i> na Modelagem de Potencial Mineral.....	18
2.4.1. <i>Mapa de Potencial Mineral</i>	18
2.4.2. <i>Random Forest</i>	20
2.5. <i>Google Earth Engine: Processamento de dados geoespaciais na nuvem</i>	22
2.6.Área de Estudo e Contexto Geológico.....	23
2.6.1. <i>Localização da área de estudo</i>	23
2.6.2. <i>Contexto Geológico</i>	23
2.6.3. <i>Potencial Mineral</i>	25
3. MATERIAIS E MÉTODOS.....	29
3.1.Caracterização do Sistema Mineral.....	29
3.2.Compilação dos Dados.....	33
3.3.Preparação de Mapas de Evidência.....	33
3.3.1. <i>Processamento dos dados aerogeofísicos</i>	33
3.3.2. <i>Modelo digital de elevação</i>	39
3.3.3. <i>Landsat 8</i>	40
3.3.3.1. <i>Geração de Mosaico</i>	40
3.4.Geração de Máscara Litológica.....	45
3.5.Preparação das Amostras de Treinamento.....	45
4. RESULTADOS.....	46
4.1.Modelo de Potencial Mineral de Bauxita.....	46

4.2.Avaliação do desempenho do modelo.....	49
4.3.Importância das variáveis.....	50
4.4.Aplicação de máscara litológica.....	51
5. DISCUSSÕES.....	51
5.1.Modelo de Potencial Mineral.....	51
5.1.1. Subárea 1.....	53
5.1.2. Subárea 2.....	54
5.1.3. Subárea 3.....	55
5.2.Importância Relativa das Variáveis.....	56
6. CONCLUSÕES.....	57
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	58
ANEXO 1 – Links para acesso aos <i>scripts</i> no <i>Google Earth Engine</i>	66

1. INTRODUÇÃO

A porção do Cráton São Francisco que se localiza na região sudeste do estado da Bahia foi alvo de projetos recentes de caracterização da cobertura laterítica e de mapeamento geológico (Projetos Cartografia do Regolito Laterítico e Contendas-Macajuba, SGB/CPRM-Salvador) que indicaram a presença de recursos minerais de origem supergênica (bauxita, óxidos de ferro, níquel laterítico, etc.), levantando a hipótese da área possuir um relevante potencial mineral, especialmente para bauxita. A disponibilização dos dados dos projetos citados, a cobertura de dados aerogeofísicos na região, juntamente com a disponibilização gratuita de imagens de sensoriamento remoto credenciam a área como alvo de pesquisa a partir do emprego de metodologias que visam a integração de dados multifonte, como classificadores supervisionados.

No cenário atual, uma grande parte dos principais depósitos minerais já foram descobertos, diminuindo a taxa de descobrimento de novos depósitos e aumentando o custo por descoberta (SCHODDE, 2014), fazendo com que técnicas que diminuam o custo do projeto e reduzam o tamanho da área de pesquisa se tornem cada vez mais utilizadas. Uma dessas técnicas é a criação de um mapa de potencial mineral. A técnica consiste no delineamento de uma área com alta probabilidade de abrigar depósitos minerais não descobertos com base na integração de diversas camadas de dados que se correlacionam com as ocorrências do mineral alvo (CARRANZA, 2009). Nessa integração será utilizado o algoritmo *Random Forest*, que consiste em um método de modelagem *data-driven* que identifica padrões nos dados a partir de pontos conhecidos de depósitos minerais e mapas de evidência.

Para que essa modelagem seja feita com sucesso, os dados de entrada tem que ser de boa qualidade e escolhidos de acordo com a sua correlação com a mineralização estudada. Para isso, utiliza-se o conceito de Sistemas Minerais (HRONSKY; GROVES, 2008; WYBORN; HEINRICH; JAKES, 1994) que envolve o entendimento dos processos envolvidos na formação e preservação do depósito mineral. No presente estudo, o mineral estudado será a bauxita de origem laterítica, tornando de suma importância o entendimento dos processos de geração de minerais supergênicos como o desenvolvimento de regolitos lateríticos e suas implicações.

Estudos envolvendo a caracterização do regolito para fins de prospecção mineral são realizados em diversas localidades do mundo que possuem as características necessárias para a ocorrência de perfis lateríticos bem desenvolvidos, especialmente na Austrália. No Brasil, o estudo do regolito laterítico tem se concentrado sobretudo na região amazônica, devido à extensa cobertura regolítica e a exploração econômica de depósitos minerais associados (Al, Mn, Fe, etc.). Os principais trabalhos de pesquisa dos regolitos com abordagens de aerogamaespectrometria e dados altimétricos encontram-se também nesta região (CARRINO *et al.*, 2011; HERRERA, 2016; IZA *et al.*, 2018).

A partir disso temos como principais objetivos do trabalho:

Objetivo geral:

- Desenvolvimento de uma metodologia de modelagem de potencial mineral de bauxita que seja confiável e replicável em outras áreas de interesse que possuam características similares às da área de estudo;

Objetivos específicos:

- Elaboração de um modelo conceitual baseado nos preceitos de caracterização de sistemas minerais para orientar a construção do modelo;

- Construção de uma base de dados de qualidade e que represente com fidelidade os diversos processos envolvidos na geração de depósitos lateríticos;

- Buscar os dados (mapas de evidência) que melhor se correlacionam com as mineralizações da área, assim produzindo um modelo mais confiável e um melhor entendimento do tipo de mineralização para servir como referência e fornecer *insights* para trabalhos futuros;

- Implementar os procedimentos na plataforma *Google Earth Engine*, uma plataforma gratuita e acessível, construindo um algoritmo que permita uma replicação mais rápida da metodologia.

2. BASE CONCEITUAL E ÁREA DE ESTUDO

Os conceitos aplicados para o desenvolvimento do trabalho envolvem o entendimento dos processos de geração e preservação dos depósitos minerais de bauxita, como formação do regolito, intemperismo e seus materiais associados; geologia da área de estudo; base teórica e aplicações da gamaespectrometria e do sensoriamento remoto, além do conhecimento acerca dos métodos de integração de dados de diversas fontes utilizando algoritmos de *machine learning*.

2.1. Bauxita e Regolito Laterítico

Os depósitos de bauxita podem ser classificados em lateríticos, sedimentares e cársticos, de acordo com suas estruturas, composições e condições de formação (BÁRDOSSY; ALEVA, 1990; BOGATYREV; ZHUKOV; TSEKHOVSKY, 2009). Devido às características da área de estudo, o presente trabalho aborda a categoria dos depósitos lateríticos de bauxita, que tem origem supergênica e estão geralmente relacionados com crostas lateríticas formadas em regiões quentes e úmidas, que foram intensamente

intemperizadas (SUN *et al.*, 2020). Estes depósitos representam a maior parte das reservas de bauxita no mundo (BLATT, 1992; BOGATYREV; ZHUKOV; TSEKHOVSKY, 2009).

A formação de depósitos de origem supergênica está intimamente ligada aos conceitos relacionados ao regolito laterítico e desenvolvimento do perfil laterítico (Figura 1). O regolito é representado por todo material inconsolidado ou secundariamente recimentado sobreposto ao embasamento inalterado, que tenha sido formado pelo intemperismo, erosão, transporte e deposição a partir de materiais pretéritos (EGGLETON, 2001). A lateritização é um dos importantes processos envolvidos na formação do regolito (solos, crostas lateríticas, aluvião, etc.), no enriquecimento supergênico e no modelamento da paisagem (FREYSSINET *et al.*, 2005).

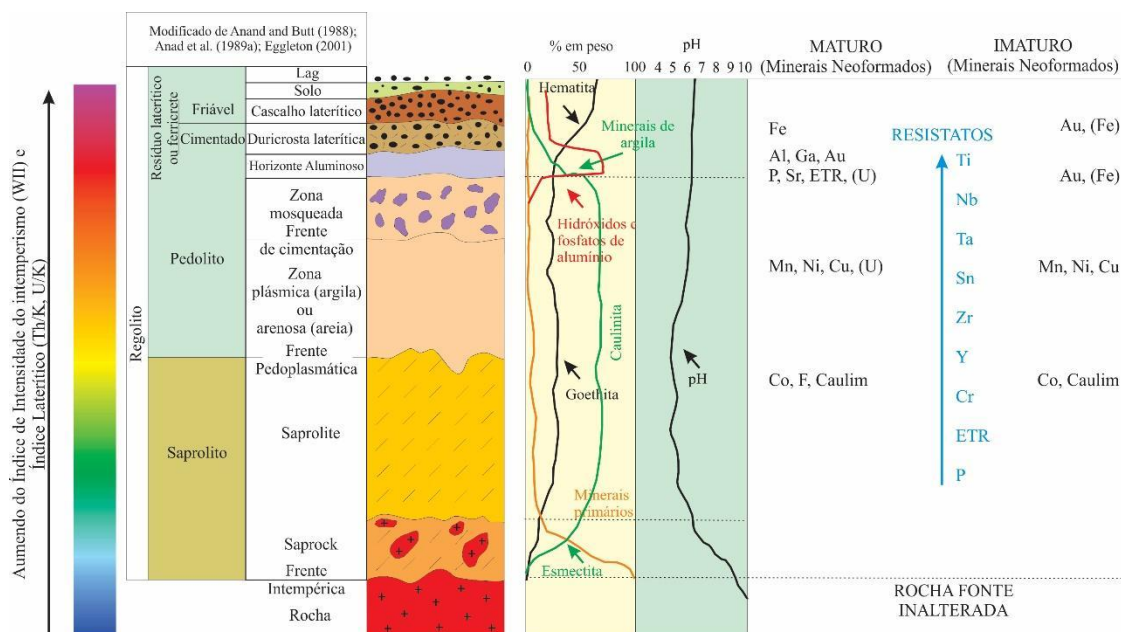


Figura 1 - Distribuição geral simplificada dos minerais lateríticos e herdados em um perfil laterítico completo e zoneamento metalogenético esquemático ao longo de um perfil laterítico, de acordo com cada horizonte. Modificado de Anand (1988), Anand *et al* (1989), Eggleton (2001) e Costa (2007).

Regolitos altamente intemperizados, que apresentam horizontes ferruginosos e/ou bauxíticos, ocorrem principalmente em regiões associadas ao cinturão intertropical (latitudes entre 30° N e 30° S), onde os processos de lateritização são mais efetivos devido às altas temperaturas e pluviosidade (FREYSSINET *et al.*, 2005; HORBE; ANAND, 2011). Em diversas regiões do mundo, como Austrália, América do Sul e Sul da Ásia, foram formados materiais associados ao regolito sobrepostos às rochas do embasamento, representando o intemperismo acumulado por um longo período (ANAND; PAINE, 2002; BOGATYREV; ZHUKOV; TSEKHOVSKY, 2009). O estudo do regolito tem permitido a compreensão dos processos supergênicos, pedológicos, geomorfológicos e contribuído fortemente para prospecção mineral, por ser fonte de bauxita e de outros minerais metálicos (Fe, Mn, Ni, Cu, Pb, Au, ETR, etc.), além de agregados para construção civil (TAYLOR; EGGLETON, 2001).

Diversas variáveis estão envolvidas na formação e composição do regolito e de seus materiais associados, sendo o clima, o relevo, a rocha-fonte, o tempo, a fauna e a flora as principais delas, que em conjunto, são responsáveis pelos diversos produtos que formam o regolito (LICHT, 2007; THOMAS, 1974). A partir disso podemos partir do princípio que os principais fatores críticos para formação de perfis lateríticos são as características do relevo associado à formação e preservação dos perfis; os aspectos climáticos que estão ligados com a intensidade e a velocidade de alteração e a composição da rocha-fonte que desempenha papel importante na evolução geoquímica do perfil e na resposta da rocha ao processo intempérico (IZA; SANTOS; FILHO, 2020). É importante salientar aqui, que uma rocha com teores mais altos de minerais metálicos na sua composição, tem mais potencial para geração de mineralizações supergênicas com teores econômicos.

2.2. Aerogeofísica: Gamaespectrometria e Magnetometria

Um levantamento gamaespectrométrico mede a radiação natural emitida pelos materiais superficiais (até 30 cm de profundidade) da Terra (DICKSON; SCOTT, 1997). A emissão de radiação gama ocorre quando o núcleo atômico instável de um isótopo se desintegra para uma forma mais estável, emitindo partículas de energia chamadas de radiação nuclear. Esse processo é conhecido como decaimento nuclear ou desintegração (IAEA, 2003).

Embora muitos elementos que ocorrem na natureza possuem isótopos radioativos na sua composição, apenas as séries de decaimento radioativo do potássio, tório e urânio emitem radiação com intensidade suficiente para ser captada por um espectrômetro de raios gama. Isso ocorre porque estes elementos são relativamente abundantes na crosta terrestre (2-2,5% de potássio, 2-3 ppm de urânio e 8-12 ppm de tório, em média) (IAEA, 2003). Devido a essa abundância na crosta, esses elementos são encontrados na composição da maioria dos materiais superficiais e podem ser utilizados para nos permitir estudar a distribuição desses materiais na superfície terrestre através de mapas gerados com os dados de um levantamento aéreo. Os mapas de concentração de K (%), eTh (ppm) e eU (ppm) indicam a variação lateral da concentração desses radioelementos na parte mais superficial da crosta, visto que os aerolevantamentos gamaespectrométricos têm a capacidade de investigar apenas alguns centímetros de profundidade (DICKSON; SCOTT, 1997).

A aerogamaespectrometria é utilizada largamente na pesquisa mineral, mapeamento geológico, mapeamento do regolito e das crostas lateríticas devido a sua capacidade de diferenciar assinaturas radiométricas de diferentes litologias (ARHIN *et al.*, 2015; CARRINO *et al.*, 2009, 2011; EBELE; OFOEGBU; NUR, 2021; HERRERA, 2016; IZA *et al.*, 2018; MCQUEEN; CRAIG, 1995; MINTY, 2011; WILFORD, 2012; WILFORD; BIERWIRTH; CRAIG, 1997). Por ser um método com baixa penetração no solo, a aerogamaespectrometria se torna um método muito eficaz no mapeamento de assinaturas de coberturas superficiais,

além de se integrar muito bem com outros métodos superficiais como as imagens de sensoriamento remoto e também com os mapas geológicos.

No presente estudo se faz importante ter o conhecimento do comportamento dos elementos químicos nos processos supergênicos relacionados à formação de depósitos. Esse comportamento varia de acordo com diversos fatores como tempo, pH, oxirredução, clima, entre outros (COSTA, 2007; WEDEPOHL; CORRENS, 1969). Quando consideramos os elementos presentes na composição das rochas e nos diversos horizontes do perfil laterítico, observam-se tendências gerais.

Em ambiente supergênico, de forma geral, os radioelementos que compõem a rocha são liberados ou redistribuídos, podendo ser incorporados a argilas, óxidos de ferro, matéria orgânica e água subterrânea (WEDEPOHL; CORRENS, 1969; WILFORD; BIERWIRTH; CRAIG, 1997). Devido à sua alta solubilidade, a concentração de K geralmente diminui com o aumento do intemperismo, tendendo a ser lixiviado de minerais primários (micas, K-feldspatos, etc.) ou ser absorvido em argilominerais. Já o U e o Th são considerados menos móveis que o K. Quando liberado de minerais pelo intemperismo, o U é mais móvel em condições oxidantes, mas precipita em condições redutoras, voltando a ser móvel quando o pH é menor do que 4. Segundo Dickson e Scott (1997), a mobilidade do U pode ser modificada pela adsorção em argilominerais, colóides, óxidos e hidróxidos de ferro, tendo forte relação com a formação de crostas lateríticas ferruginosas. Embora possa se tornar móvel em combinação com complexos orgânicos em solos e água subterrânea, o Th tende a se concentrar no regolito, por meio de minerais resistentes e óxidos de ferro (WILFORD; BIERWIRTH; CRAIG, 1997).

Para serem utilizados de forma ainda mais efetiva, os dados gamaespectrométricos muitas vezes são apresentados na forma de razões gamaespectrométricas que podem ser calculadas de acordo com a característica a ser realçada (CARRINO *et al.*, 2009; IZA; SANTOS; FILHO, 2020; WILFORD; BIERWIRTH; CRAIG, 1997). Dauth (1997) afirma que a razão eTh/K é útil na identificação de materiais intempéricos. Minty (2011) considera que a utilização de razões diminui o efeito de artefatos ambientais como vegetação e umidade do solo. Iza (IZA, 2017) propõe o uso do índice laterítico (razão $eTh * eU/K^2$) para mapear áreas com maior grau de intemperismo visto que a razão realça áreas ricas em Th e U e pobres em K, assinatura geralmente esperada em áreas que sofreram maior grau de laterização.

Dados magnetométricos, por sua vez, são utilizados para investigar a geologia através de anomalias no Campo Magnético Terrestre causadas pela variação da composição das rochas em subsuperfície (KEAREY; BROOKS; HILL, 2002). Devido a essa capacidade de diferenciar materiais com diferentes teores de minerais magnéticos, encontram-se na literatura diversas referências que utilizam dados magnetométricos para estudos geológicos, tanto em metodologias de geração de mapas litológicos preditivos (COSTA; TAVARES; OLIVEIRA, 2019), como no estudo do potencial de ocorrência de minerais supergênicos (IZA; SANTOS; FILHO, 2020).

2.3. Sensoriamento Remoto

O Sensoriamento Remoto é a técnica que utiliza sensores para captação e registro à distância, sem contato direto, da energia refletida ou absorvida pela superfície terrestre (FITZ, 2008). Estes registros podem ser registrados de formas diversas de acordo com o contexto de utilização, como imagens, gráficos, tabelas, entre outros, e, por sua vez, são passíveis de serem armazenados, manipulados e analisados por meio de *softwares* e plataformas específicas.

Os materiais presentes na superfície da Terra interagem com a radiação solar, exibindo comportamentos que variam em função dos comprimentos de onda ao longo do espectro eletromagnético (Figura 2). Essa variação do comportamento espectral dos materiais está relacionada com a sua composição química e estrutura molecular (CHAVES; ROCHA, 2006).

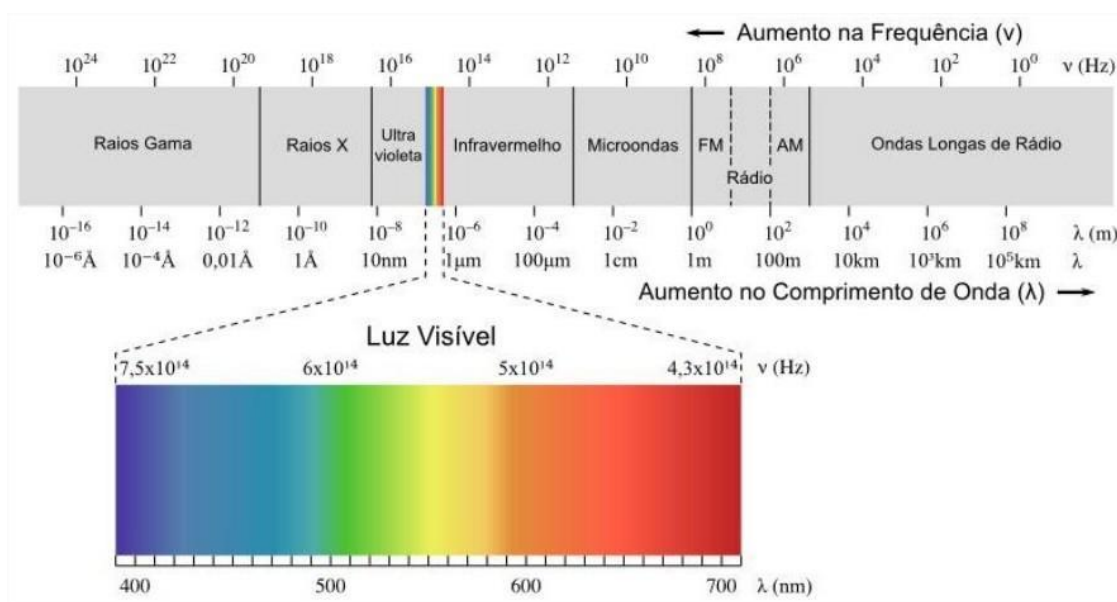


Figura 2 – Representação do espectro eletromagnético mostrando suas regiões em relação ao comprimento de onda. Destaque para a região do espectro visível. (DAMASCENO; BASSINI, 2021).

Os materiais absorvem ou refletem quantidades variáveis de energia ao longo do espectro, exibindo diferentes comportamentos espectrais. Quando esses comportamentos são medidos por sensores remotos, como sensores a bordo de satélites, podemos utilizar essas informações para estudar a composição e a distribuição desses materiais na superfície terrestre. As regiões do espectro eletromagnético mais utilizadas são o visível (VIS), o infravermelho próximo (NIR), o infravermelho de ondas curtas (SWIR) e o infravermelho termal (TIR) (CHAVES; ROCHA, 2006).

As imagens provenientes dos diferentes tipos de sensores nos fornece informações muito úteis para caracterização dos materiais superficiais e seus ambientes de ocorrência. Imagens de sensores ópticos, por exemplo, nos fornecem a reflectância espectral de uma camada superficial, nos permitindo inferir a composição mineral dos materiais, enquanto imagens provenientes de radares de abertura sintética (SAR) nos fornecem informações acerca da estrutura e morfologia do terreno.

As imagens de sensores imageadores ópticos são de grande utilidade e larga aplicação em trabalhos de mapeamento de materiais associados ao intemperismo, sendo utilizadas de forma rotineira no estudo do regolito em diversas localidades do mundo como Austrália (ANAND; PAINE, 2002; CRAIG *et al.*, 1993; GEOLOGICAL SURVEY OF WESTERN AUSTRALIA, 2013; WILFORD, 2012, 2014), Brasil (IZA; SANTOS; FILHO, 2020), África (ARHIN; NUDE, 2009; METELKA *et al.*, 2018), Índia (GUHA *et al.*, 2013; TRIPATHI; GOVIL, 2020), entre outros.

O sucesso no uso de imagens de sensores ópticos para caracterização e mapeamento do regolito laterítico está ligado à possibilidade de identificação da assinatura de alguns materiais dentro das bandas de imagens que representam regiões específicas do espectro eletromagnético. Na literatura encontram-se diversas técnicas para realçar essas assinaturas, envolvendo muitas vezes o cálculo de índices espectrais.

Um exemplo da aplicação de imagens provenientes de sensores ópticos é a razão *Gozzard* (GOZZARD; TAPLEY, 1994). Essa técnica tem o intuito de realçar as assinaturas de materiais que permitem caracterizar o regolito e identificar diferentes materiais associados (GEOLOGICAL SURVEY OF WESTERN AUSTRALIA, 2013). Utiliza-se uma imagem multispectral proveniente do satélite *Landsat 8* e consiste na geração de uma composição RGB de três razões de bandas: 6/7 (argilas e minerais com hidroxilas); 5/7 (óxido de ferro e argilas) e 5/3 (óxido de ferro e goethita), respectivamente.

Além de imagens provenientes de sensores ópticos, imagens provenientes de radares de abertura sintética também são largamente utilizadas em aplicações geológicas (BAGHDADI *et al.*, 2005). Em geral, utiliza-se a faixa de frequência das microondas, visto que essa porção do espectro sofre menos interferência da atmosfera terrestre. A nebulosidade, por exemplo, não interfere na qualidade da imagem gerada, o que a torna uma informação de boa qualidade mesmo em regiões de clima tropical úmido (FITZ, 2008).

As imagens de radar podem ser obtidas em diversas resoluções e nos fornece informações acerca das estruturas geológicas e de características morfológicas da área de estudo, que são especialmente importantes quando se trata de minerais supergênicos. Modelos digitais de terreno e seus derivados são utilizados com frequência em trabalhos de mapeamento de potencial mineral em conjunto com dados geofísicos e/ou de sensores ópticos (ABDELKAREEM *et al.*, 2018; IZA; SANTOS; FILHO, 2020; METELKA *et al.*, 2018; PORWAL *et al.*, 2015).

2.4. Machine Learning na Modelagem de Potencial Mineral

A exploração mineral é geralmente composta de quatro grandes etapas: (1) seleção de áreas, (2) geração de alvos, (3) avaliação dos recursos, e (4) definição de reserva (MCKAY; HARRIS, 2016). A etapa de seleção de áreas é descrita como a delimitação das unidades geológicas favoráveis à formação do depósito, enquanto que a etapa de geração de alvos consiste na utilização de diversas informações como dados geoquímicos, geofísicos, entre outros dados de diferentes fontes que se correlacionem com o tipo de depósito. Se forem delimitados alvos com características interessantes economicamente, o projeto segue para as fases de avaliação de recursos (furos de sondagem, geofísica terrestre, etc.) e definição de reserva (definição e classificação precisa do depósito).

Como citado no primeiro capítulo, nas últimas décadas observou-se uma diminuição da taxa de descobrimento de novos depósitos e um aumento do custo por descoberta (SCHODDE, 2014), o que acarretou na popularização da utilização de novas metodologias que diminuem de forma significativa o custo e o tempo dos projetos de prospecção mineral. O grande volume de dados de diversas fontes levantados pela indústria e pelos governos (serviços geológicos e agências), além da disponibilidade de dados de sensoriamento remoto, permitiu a criação de mapas de prospectividade mineral, que possuem a capacidade de integrar diversos bancos de dados geocientíficos em ambiente GIS, se tornando uma metodologia largamente utilizada para as duas primeiras etapas da exploração mineral.

2.4.1 Mapa de Potencial Mineral

O mapa de potencial mineral consiste na análise de diversas camadas de dados espaciais (mapas de evidências) de diversas naturezas (geofísica, geoquímica, sensoriamento remoto, litologia, geologia estrutural, geomorfologia, entre outras) com o objetivo de delimitar os alvos com maior potencial para descoberta de novos depósitos do tipo estudado dentro da área (BONHAM-CARTER; BONHAM-CARTER, 1994; CARRANZA, 2009), assim diminuindo a área de pesquisa. As diferentes metodologias envolvidas na geração de mapas de prospectividade podem ser divididas em dois grandes grupos: modelos *knowledge-driven* e modelos *data-driven*. Nos modelos *knowledge-driven*, o grau de importância (peso) atribuído a cada camada de informação é definido pelo *expert*, a partir do julgamento do grau de associação de cada variável com o tipo de depósito estudado. Os modelos *data-driven*, por sua vez, utilizam a correlação espacial entre os dados de depósitos já descobertos na área e os mapas de evidências para definir os critérios utilizados para classificar e delimitar os alvos prospectivos. Carranza e Laborte (2015a) destacam que os modelos *knowledge-driven* normalmente são indicados para a modelagem em áreas com pouco ou nenhum conhecimento de depósitos minerais descobertos (*greenfield*), enquanto que

os modelos *data-driven* são mais efetivos em áreas com bom ou moderado conhecimento prévio de depósitos e ocorrências (*brownfield*). Como a área de pesquisa do presente trabalho já possui ocorrências conhecidas de bauxita, as técnicas que serão empregadas na modelagem de potencial mineral serão algoritmos de *machine learning*, que são métodos de modelagem *data-driven*.

Aprendizado de máquina, ou *machine learning*, é a ciência da programação de computadores para “aprender” com os dados (SAMUEL, 1959). Os algoritmos são capazes de se adaptar e aprender sem seguir instruções explícitas, sendo uma ferramenta de grande valor para o reconhecimento automático de padrões em dados de diversas fontes, acelerando a interpretação e diminuindo os erros. Os métodos se mostram capazes de auxiliar na prospecção mineral, principalmente em áreas que possuam ocorrências minerais conhecidas e onde há disponibilidade de imagens orbitais e aerogeofísicas, já que funcionam de forma bastante efetiva em bancos de dados estruturados. A partir desses dados, os modelos são treinados para definir as características de áreas com depósitos e áreas sem depósitos. Para integrar os pontos de ocorrências de mineralizações e os vários mapas de evidências, diversas técnicas de *machine learning* são comumente utilizadas, como *Random Forest* (CARRANZA; LABORTE, 2015b, 2015a; FORD, 2020; HARIHARAN *et al.*, 2017; MCKAY; HARRIS, 2016; RODRIGUEZ-GALIANO; CHICA-OLMO; CHICA-RIVAS, 2014; ZHANG; ZUO; XIONG, 2016), *Logistic Regression* (CARRANZA, 2001; HARRIS; PAN, 1999; REDDY; BONHAM-CARTER, 1991); *Support Vector Machines* (LIN *et al.*, 2021; PRADO *et al.*, 2020; RODRIGUEZ-GALIANO *et al.*, 2015), *Artificial Neural Networks* (HARRIS; PAN, 1999; LIN *et al.*, 2021; PORWAL; CARRANZA; HALE, 2003; RODRIGUEZ-GALIANO *et al.*, 2015), entre outras.

Um exemplo da utilização de métodos de *machine learning* no estudo de materiais associados ao regolito é o trabalho de Metelka *et al.* (2018). Eles utilizaram redes neurais artificiais com o objetivo de mapear unidades regolíticas com crostas lateríticas ferruginosas em Burkina Faso, no oeste da África. O melhor modelo obteve coeficiente *kappa* de 0.92 utilizando uma base de dados composta de dados gamaespectrométricos, SRTM e imagens *Landsat*. Este trabalho é pioneiro na demonstração do potencial de métodos de *machine learning* no mapeamento de coberturas lateríticas.

Carranza & Laborte (2015b) utilizaram o algoritmo *Random Forest* para gerar um modelo de prospectividade para ouro no distrito de Baguio, Filipinas. Além do modelo de prospectividade, o trabalho tinha como objetivo examinar a sensibilidade do algoritmo à utilização de diferentes conjuntos de dados de treinamento, onde o modelo *Random Forest* se mostrou estável e reproduzível. O modelo conceitual para ouro epitermal foi baseado na abordagem de sistemas minerais e utilizou uma base de dados composta de mapa litológico, mapa de falhas e fraturas e localização dos depósitos descobertos.

Como exemplo de um trabalho comparativo de precisão de diferentes métodos, Rodriguez-Galiano *et al.* (2015) compararam a precisão de diversos algoritmos de *machine learning* na modelagem de potencial mineral de ouro epitermal no distrito mineral de

Rodalquilar, sudeste da Espanha. Foram utilizados como critérios na comparação a precisão no delineamento de áreas prospectivas, a sensibilidade aos hiperparâmetros, a sensibilidade ao tamanho do conjunto de dados de treinamento e a interpretabilidade dos parâmetros do modelo. O modelo *Random Forest* foi o que teve melhor desempenho, demonstrando estabilidade e robustez mesmo com variações dos parâmetros de treinamento, além de demonstrar maior precisão.

2.4.2 *Random Forest*

O modelo de potencial mineral de bauxita deste trabalho foi gerado a partir do algoritmo *Random Forest* (BREIMAN, 2001), que é um método *ensemble* de classificação supervisionada baseado na construção de diversas árvores de decisão (Figura 3). Os métodos *ensemble* consistem na combinação de vários modelos para realizar a classificação, geralmente possuindo performance melhor do que classificadores individuais (GÉRON, 2019). No caso do *Random Forest*, cada árvore de decisão fornece um resultado para a classificação, que será entendido como um voto. O algoritmo considera os votos realizados pelas árvores de decisão e seleciona a classe mais votada.

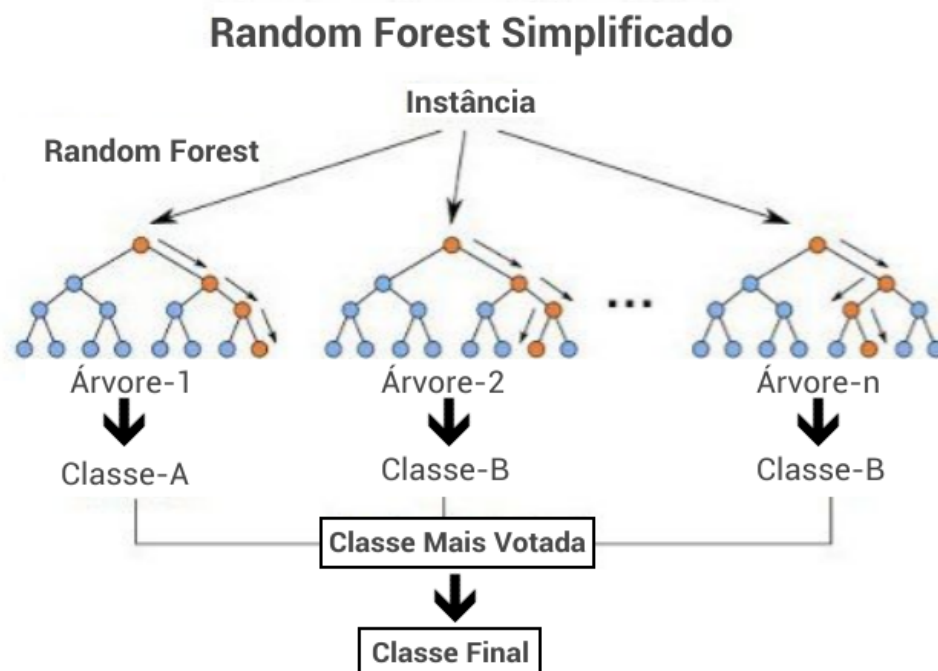


Figura 3 – Esquema simplificado do funcionamento do algoritmo *Random Forest* com n árvores de decisão. Cada árvore fornece um voto e a classe final do algoritmo é representada pela maioria dos votos. Modificado de Furquim (2018).

Para realizar a classificação, o algoritmo cria árvores de decisão não relacionadas. Para isso, utiliza o princípio do *Bagging*, também conhecido como *Bootstrapping Aggregation*. O *Bagging* consiste na geração de amostras aleatórias dos dados de treinamento a partir da substituição (Figura 4). Cada amostra gerada (*bootstrap*) contém parte dos dados de treinamento e é utilizada para treinar uma árvore de forma independente, com o resultado final consistindo no voto da maioria das árvores em um processo chamado de *aggregation*.



Figura 4 – Um exemplo da geração de amostras *bootstrap*, que são criadas por substituição, gerando um conjunto de amostras que são sempre distintas entre si. Fonte: Researchgate, 2020.

Além de ser um classificador robusto, o *Random Forest* nos permite calcular a importância das variáveis, que pode fornecer informações de grande valor quando o estudo visa, além da geração do modelo, um melhor entendimento da relação entre as variáveis utilizadas e o problema estudado. O cálculo da importância das variáveis nos indica quais variáveis tem o maior poder de previsão no modelo e quais as variáveis com o menor impacto, nos permitindo inclusive optar por omitir do modelo algumas variáveis, tornando-o mais simples e rápido. Na prática essa informação é uma medida do quanto remover uma variável diminui na precisão do modelo, ou o quanto adicionar uma variável aumenta na precisão do modelo.

O *random Forest* é um modelo relativamente rápido, simples e versátil e suas especificidades implicam em diversas vantagens na sua aplicação (SRUTHI, 2021):

- Demonstra boa performance tanto utilizando variáveis contínuas quanto utilizando variáveis categóricas;
- Demonstra boa performance mesmo em bancos de dados que contém valores nulos;
- Tende a não sofrer com *overfitting*, visto que a classificação final é obtida através do voto da maioria das árvores, onde cada uma delas é construída de forma independente (nenhuma árvore tem acesso a todo o banco de dados);
- Pelo mesmo motivo anterior é pouco sensível à ruídos na base de dados de treinamento;
- É relativamente rápido no processamento de grandes bases de dados;

- Ao contrário de outros algoritmos, poucos hiperparâmetros precisam ser ajustados, o que simplifica o entendimento e a reprodução da metodologia;

- Além de ser utilizado como classificador, nos permite calcular a importância das variáveis.

Com relação às desvantagens, pode-se citar que a classificação pode ser lenta em problemas que necessitem de um número muito alto de árvores para atingir uma boa precisão, sendo necessário buscar alguma solução computacional para possibilitar o trabalho, como serviços de processamento em nuvem ou utilização de um hardware robusto. Também é importante ressaltar que, por ser uma ferramenta de modelagem preditiva, os resultados não contêm respostas descritivas acerca do relacionamento dos dados de entrada.

2.5. *Google Earth Engine*: Processamento de dados geoespaciais na nuvem

O presente estudo visa a geração de um modelo de potencial de bauxita integrando dados geoespaciais multifonte de uma área com dimensão considerável utilizando um algoritmo de classificação supervisionada. Nesse contexto, diversos desafios operacionais se apresentaram, como a necessidade de acesso à diversos *softwares* específicos para processar os diferentes tipos de dados, ter acesso a um *hardware* de capacidade suficiente para lidar com um volume de informação considerável, além do tempo necessário para desenvolver todos os passos para geração do modelo, como o pré-processamento dos mapas de evidência e integração pelo algoritmo. Para lidar com esses desafios e possibilitar o desenvolvimento do trabalho, o *Google Earth Engine* foi escolhido como plataforma principal para realizar o processamento dos dados geoespaciais utilizados, além da geração do modelo propriamente dito.

O *Google Earth Engine* (GEE) é uma plataforma baseada em nuvem de análise e processamento de dados geoespaciais em escala global que utiliza o gigante potencial computacional do *Google* para permitir a realização de trabalhos em diversas áreas do conhecimento, como estudos do desflorestamento, uso e cobertura do solo, monitoramento climático, proteção ambiental, entre outros. Além de ser uma plataforma de processamento, o GEE também possui um catálogo extenso de dados de sensoriamento remoto em diversos níveis de processamento e prontos para uso em análises mais elaboradas. O acesso à plataforma é realizado no próprio navegador, fornecendo ao usuário uma interface de programação (API) e um ambiente de desenvolvimento integrado (IDE) que permitem a geração e visualização de resultados de forma rápida e eficiente sem a necessidade de instalação de *software* (GORELICK *et al.*, 2017).

Diversos trabalhos espalhados pelo mundo já foram realizados através da plataforma em aplicações como distribuição espacial de espécies (DONG *et al.*, 2016), uso e cobertura do solo (MEDEIRO, 2018; OLIVEIRA, 2016; SOUZA *et al.*, 2020; WANG *et al.*, 2020),

crescimento urbano (PATEL *et al.*, 2015), geomorfologia (BOOTHROYD *et al.*, 2021), entre outros. No Brasil, destaque para o Projeto MAPBIOMAS (Projeto de Mapeamento Anual da Cobertura e Uso do Solo do Brasil) que utiliza o potencial do GEE para mapear mudanças de uso e cobertura da terra no Brasil através de seus diferentes biomas (SOUZA *et al.*, 2020). O projeto conta com a participação de diversas instituições, entre elas a Universidade Estadual de Feira de Santana (UEFS).

Embora possua muitas vantagens na sua utilização, o GEE ainda não é utilizado em larga escala para trabalhos envolvendo a geração de modelos de potencial mineral. No contexto do presente trabalho, a plataforma foi imprescindível para a execução dos objetivos propostos, sendo utilizada em diversas etapas como o acesso aos dados de modelo digital de terreno, acesso aos dados e geração de mosaico de imagens do satélite *Landsat 8*, pré-processamento dos mapas de evidência e dados de treinamento até a implementação do algoritmo *Random Forest* e verificação da precisão do modelo. Os links para acesso aos scripts gerados neste trabalho estão presentes como Anexos.

2.6 Área de Estudo e Contexto Geológico

2.6.1 Localização da área de estudo

A área de pesquisa (Figura 5A), contendo aproximadamente 106.200 km², se localiza na região sudeste estado da Bahia e tem como principais cidades (> de 50.000 habitantes): Ipirá, Feira de Santana, Itaberaba, Santo Antônio de Jesus, Valença, Jequié, Vitória da Conquista, Itabuna, Ilhéus e Itapetinga. As altitudes variam desde 0 m no litoral, a leste, até 1.501 m na região da Chapada Diamantina.

Nesta região ocorrem dois importantes biomas (Mata Atlântica e a Caatinga) que caracterizam diferentes aspectos físicos (relevo, clima, vegetação), econômicos, sociais e culturais. O primeiro predomina na porção leste da área e apresenta clima quente e úmido, temperaturas entre 24 e 35°C e índice pluviométrico entre 1.000 e 1.500 mm anuais. O segundo bioma predomina na porção oeste da área e tem clima semiárido caracterizado por temperaturas médias similares ao bioma anterior mas com precipitação média anual entre 300 e 800 mm.

2.6.2 Contexto Geológico

O contexto geológico da área de estudo (Figura 5B) abrange, principalmente, a porção sudeste do Cráton do São Francisco (CSF), estabilizado no final da orogenia riaciana (~2,0

Ga), e parte das Zonas Externa e Interna da Faixa de Dobramento Araçuaí implantada, a sul do CSF, durante o ciclo Brasileiro do Neoproterozoico (~600 Ma) (ALMEIDA, 1977).

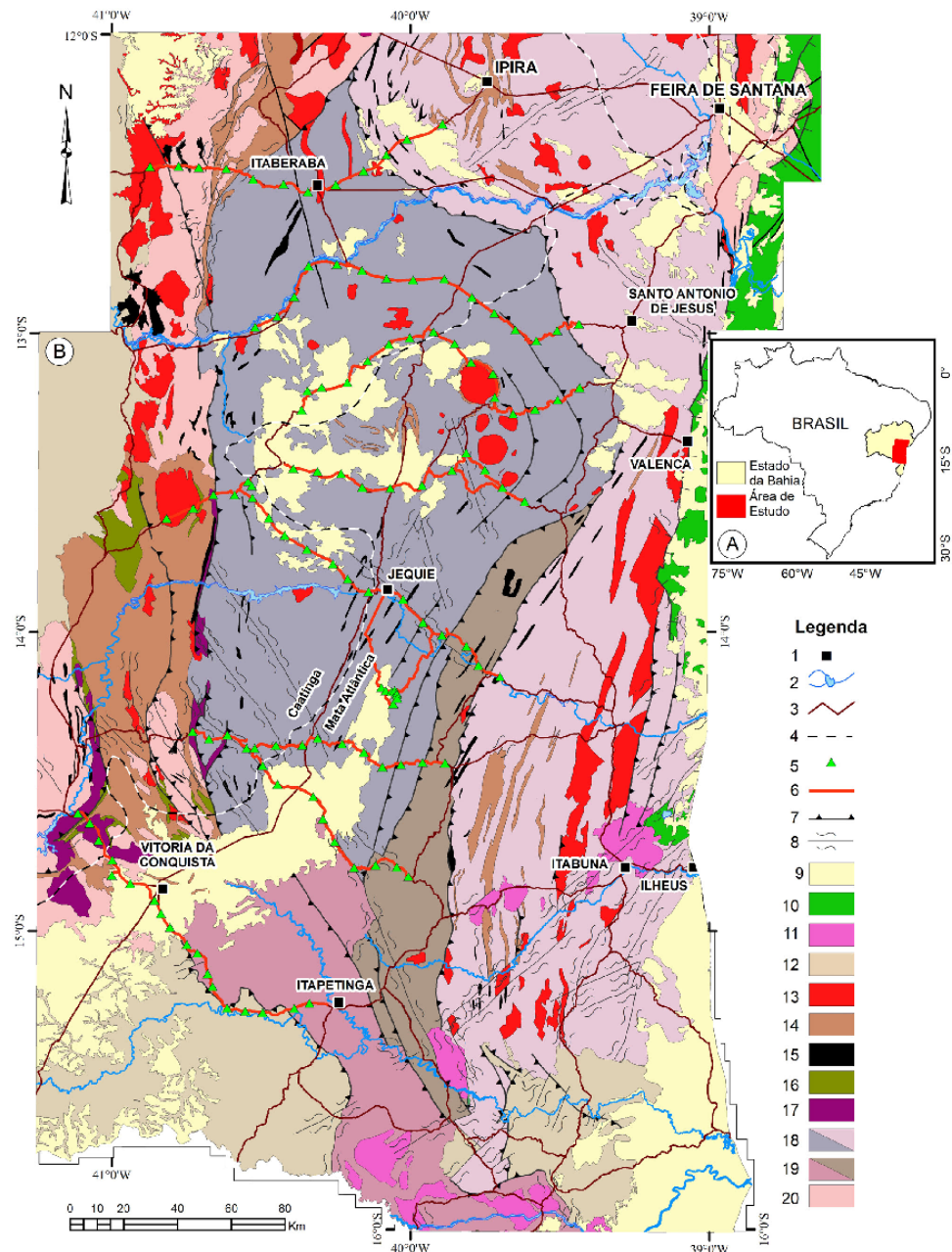


Figura 5– A) Localização da área de estudo; B) Mapa geológico simplificado de Souza *et al.* (2003).
 Legenda: 1. Cidades; 2. Rios; 3. Estradas; 4. Limite entre biomas; 5. Pontos visitados; 6. Caminhamentos; 7. Falhas de empurrão; 8. Zonas de cisalhamento. 9. Coberturas cenozoicas; 10. Coberturas mesozoicas; 11. Granitoides neoarqueanos; 12. Coberturas meso-neoproterozoicas; 13. Granitoides paleoproterozoicos; 14. Sequências metassedimentares neoarqueanas a paleoproterozoicas; 15. Corpos máficos-ultramáficos arqueanos e paleoproterozoicos; 16. Sequências metavolcanossedimentares neoarqueanas; 17. Granitoides neoarqueanos; 18. Granulitos dos blocos Itabuna-Salvador-Curaçá e Jequiê; 19. Gnaisses da Banda Ipiaú e do Complexo Itapetinga; 20. Gnaisses-migmatíticos dos blocos Gavião (oeste) e Serrinha (leste).

O embasamento do CSF corresponde a unidade mais expressiva da área. Ele se encontra compartimentado em segmentos crustais gnáissico-migmatíticos (3,2-2,6 Ga; blocos Gavião, a oeste, e Serrinha, a nordeste), granulíticos (2,8-2,3 Ga; blocos Jequié, na porção central, e Itabuna-Salvador-Curaçá), e gnáissico (~2,7 Ga; Banda Ipiaú). Associados a esses terrenos, ocorrem sequências metavulcanossedimentares arqueanas e metassedimentares paleoproterozoicas. As rochas encontram-se intrudidas por granitoides (granitos, granodioritos e sienitos), corpos máficos e máfico-ultramáficos (vulcânicas máficas, gabro, anortosito, dunito, peridotito) com dimensões e idades distintas (neoarqueanas a paleoproterozoicas).

Na extremidade noroeste da área, localizam-se as coberturas plataformais do CSF, representadas por rochas metassedimentares clásticas meso-neoproterozoicas (1,75-1,0 Ga; Supergrupo Espinhaço) e clasto-carbonáticas neoproterozoicas (1.00-700 Ma; Supergrupo São Francisco).

A Faixa Araçuaí (FA) é formada pelo Complexo Itapetinga (ortognaisses da fácies anfibolito, com restos de supracrustais), que corresponde à parte do embasamento do CSF que foi intensamente deformado durante a Orogênese Brasileira, além de granitoides anorogênicos tonianos nele intrudidos (Suíte Salto da Divisa). Na FA, próximo a margem sudeste da CSF, ocorrem rochas metassedimentares neoproterozoicas dos grupos Rio Pardo (sedimentos clásticos) e Macaúbas (sedimentos clasto-carbonáticos) e da Formação Caboclo (sedimentos clásticos), além de intrusivas graníticas e alcalinas (Suíte Intrusiva Itabuna), de idade neoproterozoica.

As coberturas fanerozoicas, representadas por sedimentos mesozoicos das bacias do Recôncavo-Tucano e Camamú-Almada, foram implantadas durante a abertura do Oceano Atlântico, assim como os sedimentos cenozoicos dos planaltos de Vitória da Conquista e Maracás, da costa atlântica e dos aluviões.

2.6.3 *Potencial Mineral*

Além de possuir um contexto climático e geomorfológico favorável para o desenvolvimento de materiais oriundos do intemperismo, trabalhos anteriores (Figura 6) reforçam o potencial mineral da região para minerais supergênicos (ALVES, 2018; FERNANDES, 1996; IZA; SANTOS; FILHO, 2020; MARTINS; SANTOS, 1997; SANTOS, 2010). Hayston (2012) apresentou estudo que discute o potencial da região como uma província mineral de bauxita.

O Projeto Cartografia do Regolito Laterítico (IZA; SANTOS; FILHO, 2020) foi desenvolvido na área para caracterizar e estudar o regolito e seus materiais correlatos. O trabalho demonstra que os principais recursos minerais da área (Figura 6) estão na sua parte central (subárea 1) e estão relacionados com minerais supergênicos (Al, Fe, Ni e Mn).

Nota-se a sobreposição dos recursos minerais com os altos valores de índice laterítico, principalmente

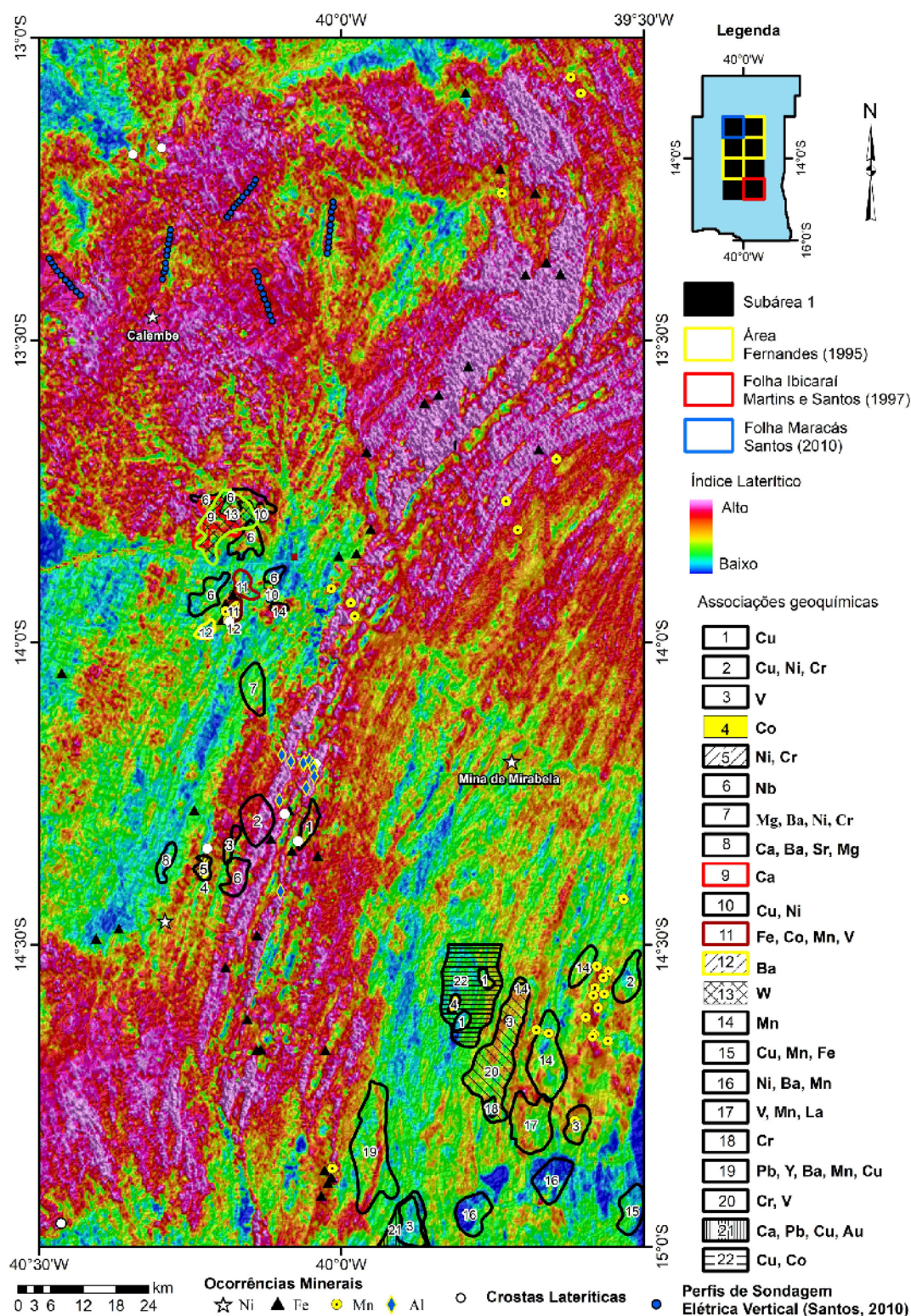


Figura 6 - Dados geoquímicos de Fernandes (1995) (Área do Projeto Terras-Raras), recursos minerais do SGB-CPRM (GeoSGB) e pontos de SEV, realizados por Santos (2010), sobrepostos ao índice laterítico. Iza *et al.* (IZA; SANTOS; FILHO, 2020).

as ocorrências de Al e Fe. A Figura 7 ilustra o aspecto geral de uma ocorrência de bauxita localizada na região.

Santos (2010) realizou a avaliação do aquífero do Alto da Bacia do Rio Jiquiriçá através de métodos geoeletricos com o intuito de avaliar o potencial hídrico da região. Dentre os principais resultados, destaca-se a identificação de perfis lateríticos com espessuras entre 9 e 50 m, que foram validados através de dados de poço. Esses perfis estão localizados em regiões de alto índice laterítico (Figura 6, em azul), reforçando o potencial da gamaespectrometria para identificar a assinatura de áreas fortemente intemperizadas.

Alves (2018) realizou a caracterização do protominério de níquel laterítico de Calembe (Figura 6, representado por uma estrela branca), localizado no município de Lajedo do Tabocal, Bahia. A partir da descrição de testemunhos de furos de sondagem, o autor identificou a espessura do regolito *in situ* de 20 a 40 m. Esse resultado corrobora os resultados de Santos (2010) e também coincidem com regiões que apresentam altos valores de índice laterítico.

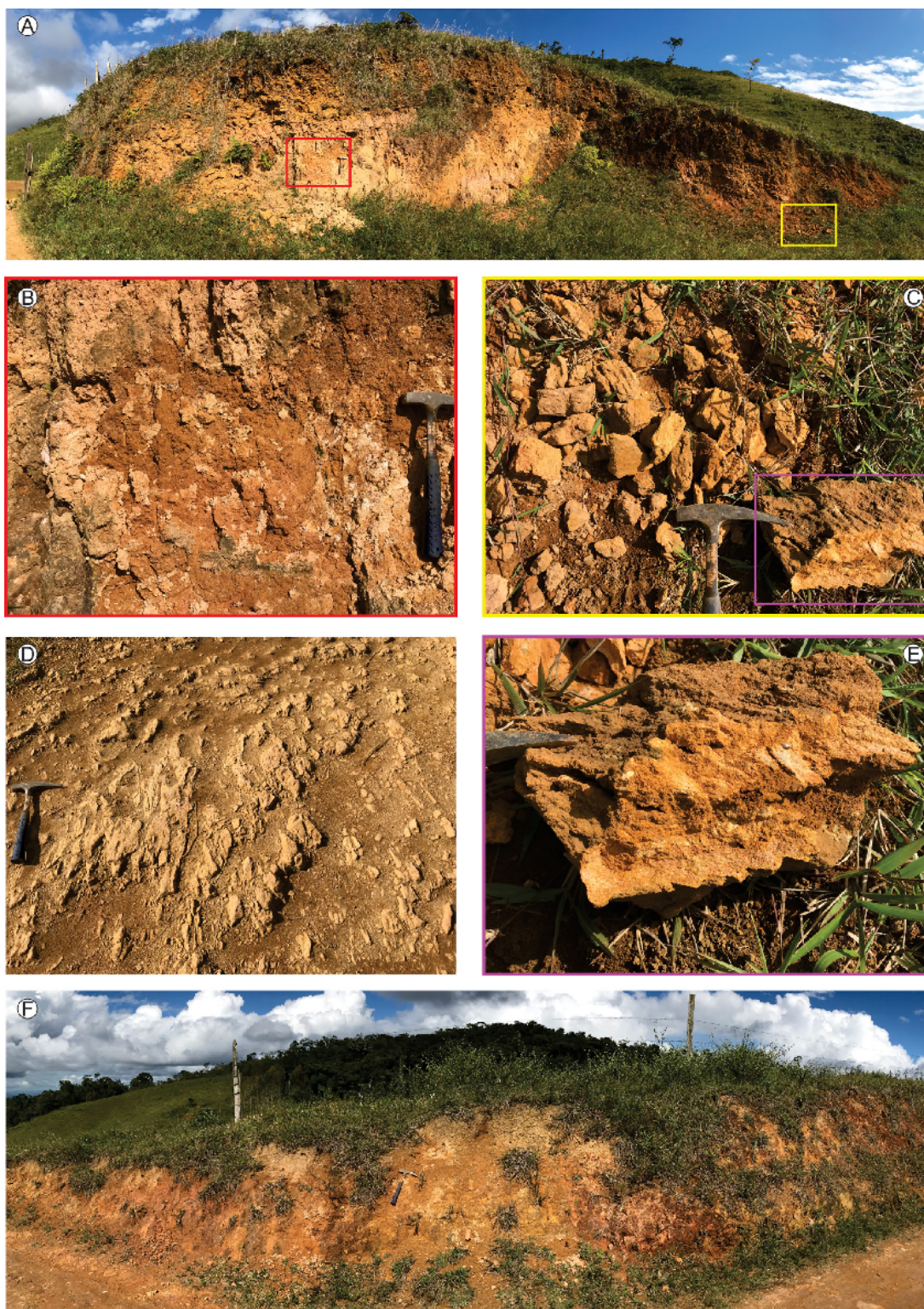


Figura 7 – Aspecto geral da área de ocorrência de crosta laterítica aluminosa (bauxita) na região entre os municípios de Jequié e Itagi. A) Perfil laterítico com pelo menos 2 m aflorante. B) Solo vermelho em tons rosados. C) Fragmentos centimétricos de crosta aluminosa. D) Crosta aluminosa (substrato da estrada). E) Detalhe da crosta aluminosa maciça.

3. MATERIAIS E MÉTODOS

Como citado anteriormente, o mapa de prospectividade mineral consiste no delineamento de áreas com grande probabilidade de conter depósitos não descobertos a partir de diversas camadas de informações e conhecimento prévio de ocorrências já conhecidas do tipo de mineral procurado (BONHAM-CARTER; BONHAM-CARTER, 1994; CARRANZA, 2009). O objetivo principal é a tradução de um conhecimento teórico e empírico do tipo de mineralização, em um mapa de potencial mineral, que pode ser utilizado como base para um projeto de exploração mais eficiente e menos custoso.

De maneira geral, a construção de um modelo de potencial mineral de sucesso envolve uma série de passos que incorporam o bom entendimento do sistema mineral estudado, compilação de uma base de dados confiáveis, que representem os fatores críticos de geração do depósito, pré-processamento dos dados para permitir a utilização dos mesmos de forma correta no treinamento e na geração do modelo final. Embora, neste trabalho, os dados de localização de depósitos conhecidos sejam utilizados como base para treinamento e geração do modelo (*data-driven*), é de suma importância a compreensão e conhecimento acerca dos processos envolvidos na geração e preservação dos depósitos, que nos permite uma escolha bem fundamentada dos dados a serem utilizados.

A metodologia utilizada no presente trabalho pode ser vista de forma esquematizada no modelo conceitual ilustrado na Figura 8.

3.1. Caracterização do Sistema Mineral

O primeiro passo é a definição e caracterização do sistema mineral proposto, no caso deste trabalho, bauxita de origem supergênica. Um sistema mineral é o conjunto de processos envolvidos na formação e preservação de depósitos minerais. O conceito é utilizado com o objetivo de identificar similaridades entre os depósitos já conhecidos associados ao sistema e criar um modelo que englobe os fatores críticos dentro das esferas de conhecimento que envolvem os processos de geração e preservação dos depósitos. Além de ser o passo inicial para construção do mapa de potencial mineral, essa etapa é de suma importância no avanço do entendimento da geração da mineralização, que baseia a escolha dos dados utilizados como mapas de evidência, além de servir como referência para trabalhos futuros. O conceito de sistema mineral é largamente utilizado como base para produção de mapas de prospectividade (CAMPOS *et al.*, 2017; CARRANZA; LABORTE, 2015b, 2015a; DAVIES *et al.*, 2020; LAWLEY *et al.*, 2021; PARSA; MAGHSOUDI, 2021).

Para compreender de forma satisfatória o sistema mineral estudado e possibilitar um mapeamento preditivo eficiente dos depósitos minerais, foi utilizada a abordagem de quatro passos proposta por McCuaig *et al.* (2010):

1. Identificação dos fatores críticos do sistema mineral;
2. Identificação dos processos associados com a formação dos depósitos;
3. Definição de evidências mapeáveis resultantes desses processos de formação;
4. Conversão das evidências em *proxies* (vetores) que podem ser integrados em ambiente GIS.

Depósitos minerais são resultado de uma conjunção de processos globais que ocorrem na Terra e uma forma de identificar os fatores críticos para a formação desses depósitos é compreender as esferas de conhecimento (Figura 9) que englobam os conceitos que baseiam esses fatores. Para o depósito de bauxita ser formado, é necessário um conjunto de condições que envolvem a composição da rocha-fonte, as condições geomorfológicas para que esse material seja exposto e posteriormente preservado, além das condições climáticas favoráveis. Partindo desses princípios conceituais e utilizando a abordagem dos quatro passos, foi gerado o modelo conceitual (Figura 10) para o sistema mineral estudado que será utilizado como referência na formação do banco de dados utilizado para geração do modelo.



Figura 9 - Esferas do conhecimento que englobam os fatores críticos para formação e preservação de crostas lateríticas.

FATORES CRÍTICOS	PROCESSOS	CRITÉRIOS MAPEÁVEIS	PROXIES
<p>EXPOSIÇÃO E PRESERVAÇÃO</p> <p>Geomorfologia que favoreça a exposição ao intemperismo físico e químico.</p>	<p>A formação de depósitos de bauxita laterítica possui uma forte componente geomorfológica, com as principais ocorrências correlacionadas à cotas e morfologia específicas</p>	<p>Topografia; características morfológicas do terreno.</p>	<ul style="list-style-type: none"> • Modelo digital de terreno (SRTM); • Mapa de declividade.
<p>ROCHA-FONTE</p> <p>Rocha-Fonte que já possua um teor de minerais metálicos considerável e seja susceptível a sofrer alteração.</p>	<p>Forte intemperismo químico e físico promove a formação de perfis lateríticos que tem composição intimamente ligada à composição da rocha-fonte.</p>	<p>Unidades geológicas com composição inicial favorável a geração de crostas lateríticas com teor econômico.</p>	<ul style="list-style-type: none"> • Mapa Litológico; • Gamaespectrometria (K, eTh e eU); • Magnetometria (Gradiente Total).
<p>INTEMPERISMO</p> <p>Condições de clima e tempo para que a rocha-fonte seja intensamente laterizada, permitindo o desenvolvimento do depósito.</p>	<p>Os elementos químicos se comportam de forma característica em ambiente supergênico, nos permitindo inferir o grau de intemperismo a partir da composição da camada aflorante.</p>	<p>Teores de elementos químicos e minerais associados ao intemperismo e à mineralização.</p>	<ul style="list-style-type: none"> • Gamaespectrometria (K, eTh e eU); • Magnetometria (Gradiente Total); • Imagens de Sensoriamento remoto.

Figura 10 - Caracterização do sistema mineral bauxita laterítica. Fatores críticos, processos de formação dos depósitos, critérios mapeáveis e proxies (vetores) em cada esfera de conhecimento.

3.2. Compilação dos Dados

Após a caracterização do sistema mineral, deve-se conduzir uma compilação dos dados para que eles posteriormente possam ser trabalhados em ambiente GIS, no caso deste trabalho, a plataforma *Google Earth Engine*. Essa é uma etapa que exige bastante atenção, visto que a qualidade e confiabilidade dos dados refletem na qualidade do modelo final. Erros no banco de dados podem ser propagados levando a um modelo pouco confiável. O banco de dados (Tabela 1) foi formado por bases que permitiram a geração dos *proxies* definidos na etapa de criação do modelo conceitual (Figura 10). Para isso, foram compiladas bases de dados provenientes de diversas fontes públicas e privadas. O banco de dados que será utilizado como material nesta dissertação de mestrado compreende dados aerogamaespectrométricos, modelo digital de elevação, mapa litológico, coleção de imagens do satélite *Landsat 8*, além de pontos de campo descritos que foram coletados na área em trabalhos prévios.

De acordo com sua natureza, parte dos dados são utilizados como dados de treinamento e avaliação do modelo (pontos de campo), enquanto os dados aerogeofísicos e de sensoriamento remoto são base para a construção dos mapas de evidência. Por fim, o mapa litológico é utilizado para geração de uma máscara litológica. Os procedimentos para aquisição e pré-processamento dos dados serão detalhados nos próximos itens.

3.3. Preparação de Mapas de Evidência

A fase de preparação de mapas evidência consiste na geração de mapas integráveis em ambiente GIS a partir dos dados brutos compilados. Nesta fase, diversos processamentos são realizados e ferramentas aplicadas para corrigir erros e problemas permitindo que o mapa de evidência seja da melhor qualidade possível.

Neste item, serão discutidas as etapas de pré-processamento dos dados de diferentes naturezas. Como cada base de dados tem suas próprias características e formatos, diferentes procedimentos foram necessários para cada tipo de informação. Um resumo dos principais procedimentos realizados para preparação dos mapas de evidências a partir do banco de dados compilado pode ser visto na Figura 8. Antes de serem incorporados ao modelo, todos os mapas de evidência passaram por um processo de diminuição da resolução para 125 m (para que todos os mapas possuam a mesma escala do dado de menor resolução, os dados gamaespectrométricos e magnetométricos, no caso) e uma normalização para valores entre 0 e 1 (para evitar problemas no modelo relativos a uma maior influência de variáveis que contém valores numéricos de maior magnitude).

INFORMAÇÃO	NATUREZA	FONTE	FORMATO	ETAPA DE UTILIZAÇÃO
Gamaespectrometria (K, eTh e eU) e Magnetometria (Gradiente Total)	Geofísica	CBPM e SGB/CPRM	Tabela	Geração de mapas de evidências
Pontos de Campo e Base RecMin (SGB/CPRM)	Recursos Minerais	SGB/CPRM	Tabela	Treinamento e avaliação do modelo
SRTM – Modelo Digital de Elevação	Sensoriamento Remoto	GEE	Raster	Geração de Mapas de Evidências
Mapa Litológico	Geologia	SGB/CPRM	Shape	Geração de máscara litológica.
Coleção <i>Landsat 8</i>	Sensoriamento Remoto	GEE	Raster	Geração de Mapas de Evidências

Tabela 1 - Sumário dos dados disponíveis, especificando sua natureza, fonte, formato e etapa em que serão utilizados.

3.3.1. *Processamento dos dados aerogeofísicos*

A área de estudo é recoberta por três projetos aerogeofísicos, sendo dois deles realizados pela Companhia Baiana de Pesquisa Mineral – CBPM e um por convênio entre a CBPM e o Serviço Geológico do Brasil – CPRM. Os levantamentos são constituídos por perfis aeromagnéticos e aerogamaespectrométricos de alta resolução e foram integrados em um único banco de dados, são eles: o Projeto Aerogeofísico Ruy Barbosa – Vitória da Conquista; Projeto Aerogeofísico Cândido Sales – Mascote e o Projeto Aerogeofísico Ipirá – Ilhéus, que juntos abrangem 100% da área.

O Projeto Aerogeofísico Ruy Barbosa – Vitória da Conquista (CBPM, 2007) abrange um bloco de aproximadamente 41.911 km². O levantamento alcançou cerca de 92.103 km de perfis aeromagnéticos e aerogamaespectrométricos de alta resolução, com linhas de voo com espaçamento de 500 m.

O Projeto Aerogeofísico Cândido Sales – Mascote (CBPM, 2009) abrange uma área de aproximadamente 24.430 km², onde foram feitos cerca de 53.831 km de perfis com linhas de voo com espaçamento de 500 m.

O Projeto Aerogeofísico Ipirá – Ilhéus (CBPM; CPRM, 2011) foi realizado por meio de convênio entre a CBPM e o SGB-CPRM, entre os anos de 2010 e 2011, compreendendo um bloco de aproximadamente 40.077 km². Cerca de 86.629 km de perfis foram sobrevoados, também, com espaçamento entre linhas de 500 m.

Os dados levantados recebem diversas correções no intuito de gerar os produtos utilizados no presente estudo com a menor quantidade de erros possíveis. As correções foram necessárias para compensar erros causados por diversos fatores, tanto naturais quanto causados pelos próprios equipamentos de medição e pelas variações de voo da aeronave. Os tratamentos foram aplicados pelas empresas contratadas para realizar os levantamentos utilizando diversos parâmetros obtidos durante a fase de levantamento. A Figura 11 destaca os principais procedimentos realizados nos dados aerogeofísicos antes de serem disponibilizados para uso. Maiores detalhes podem ser encontrados nos relatórios citados dos aerolevantamentos.

Os bancos de dados gamaespectrométricos e magnetométricos foram tratados por meio do *software Oasis Montaj 9.2 (Bentley)*. Os dados gamaespectrométricos obtidos no aerolevantamento são provenientes de uma relação direta de intensidade de radiação captada pelo equipamento. Quando ocorre a conversão dos valores de cps (contagem por segundo) para a concentração equivalente do radioelemento, algumas inconsistências podem surgir no banco de dados, como a presença de valores negativos de concentração. Desse modo, antes da geração dos *grids*, o banco de dados de cada projeto aerogeofísico foi corrigido por meio da exclusão dos valores negativos ou pela substituição dos valores negativos por uma constante igual a 5% da média dos valores do canal a ser corrigido (IAEA, 2003).

Após a correção, os canais K, eTh e eU dos projetos foram gridados utilizando o método de interpolação *Minimum Curvature* (BRIGGS, 1974) com tamanho de célula de 125 m (1/4 da distância entre as linhas de voo). Como os dados foram adquiridos em levantamentos realizados em condições distintas, os *grids* inicialmente foram nivelados utilizando a relação linear entre os dados das áreas de sobreposição entre os diferentes aerolevantamentos. Posteriormente, os *grids* nivelados foram integrados (Figura 12) utilizando a extensão *Grid Knitting* do *Oasis Montaj*, para gerar os mosaicos de cada radioelemento que englobam toda a área de estudo.

Os dados magnetométricos dos projetos, por sua vez, foram gridados através do método de interpolação *bi-directional grid* com os mesmos 125 m de tamanho de célula, posteriormente sendo integrados de forma similar aos dados gamaespectrométricos. Posteriormente, foi aplicada o filtro *Analytic Signal*, disponível no menu *Grid and Image* do *Oasis Montaj*, para geração do mapa do Gradiente Total (NABIGHIAN, 1972), que consiste em um método que centraliza a anomalia acima da fonte magnética e retira a natureza dipolar do *grid* original.

Para permitir a leitura dos dados na plataforma *Google Earth Engine*, os grids dos canais individuais de K, eTh, eU (Figura 13A, B e C) e do gradiente total (Figura 13D) foram posteriormente transformados para o formato *.tiff* por meio da extensão do *geosoft* em ambiente *ArcGIS* (ESRI).

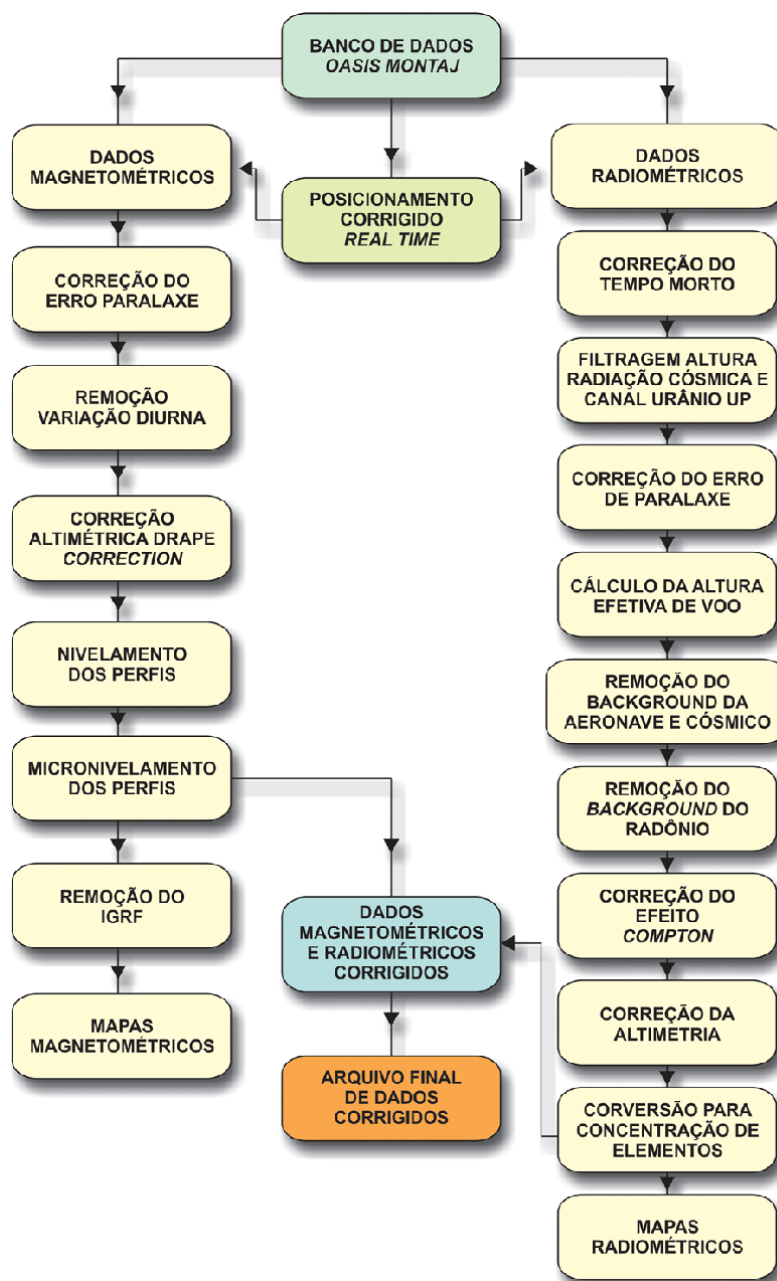


Figura 11 – Fluxograma padrão de pré-processamento dos dados dos aerolevantamentos geofísicos. Estas etapas são realizadas antes da disponibilização da base de dados para os usuários. Fonte: CBPM (2007).

Grid Knitting

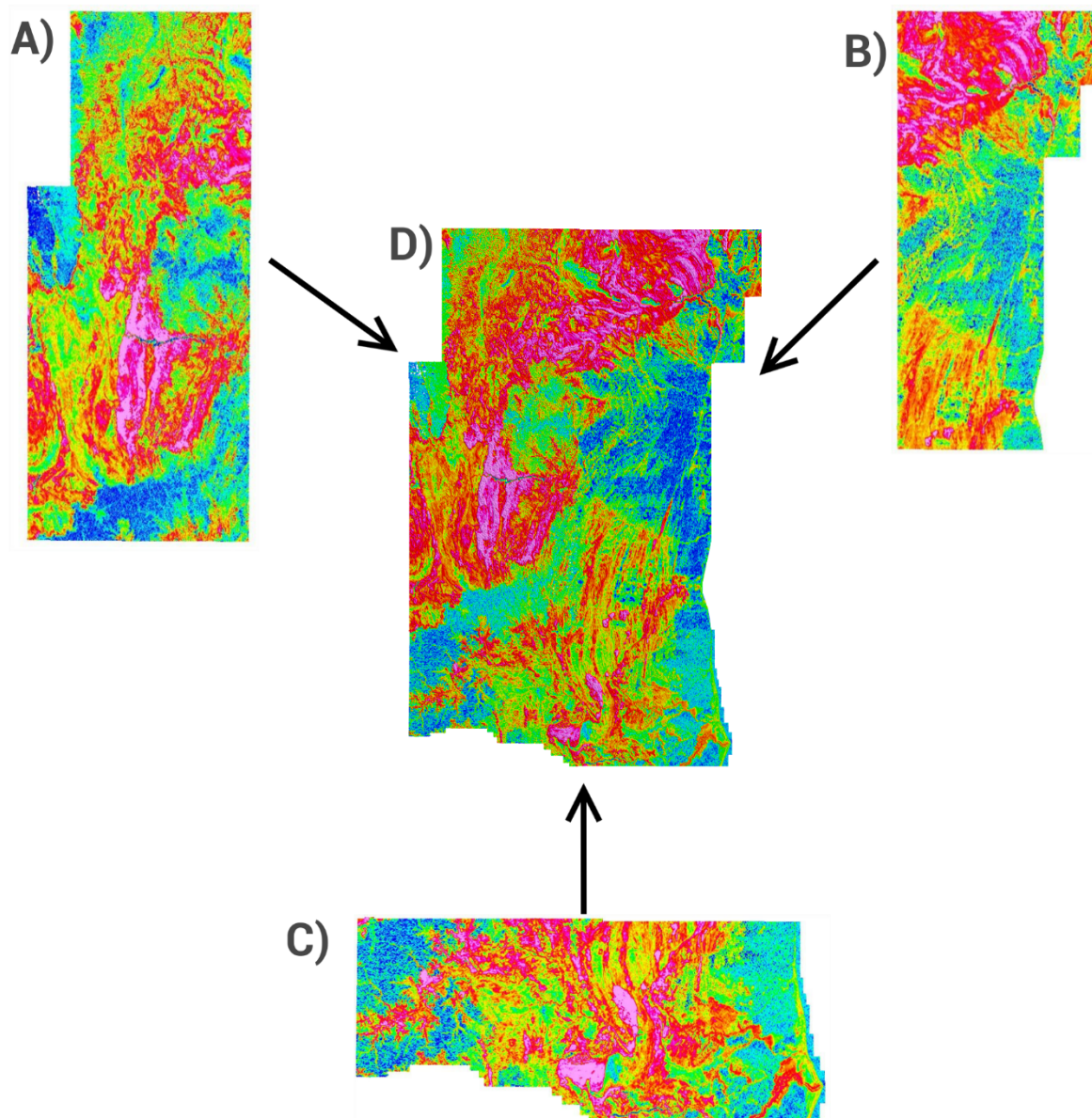


Figura 12 – Integração dos *grids* de K dos três aerolevantamentos através da extensão *grid knitting*. A) Projeto Ruy Barbosa-Vitória da Conquista. B) Projeto Ipirá-Ilhéus. C) Projeto Cândido Sales – Mascote. D) Mapa do K da área de estudo após a integração dos três projetos aerogeofísicos.

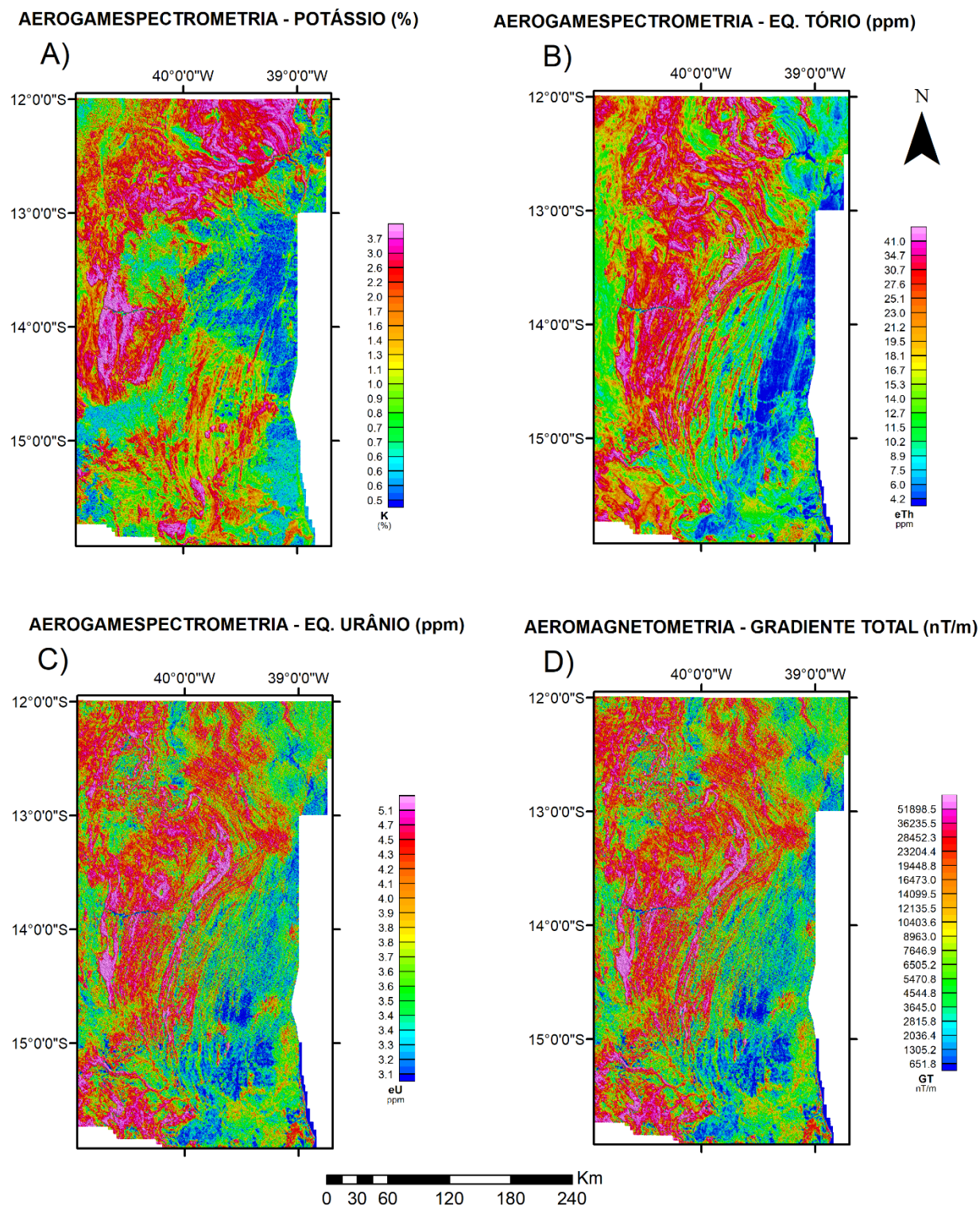


Figura 13 – Dados gamaespectrométricos (A, B e C) e magnetométricos (D) preparados como mapas de evidência após realização do pré-processamento necessário.

3.3.2. Modelo digital de elevação

A formação dos depósitos de origem supergênica está intimamente relacionada com as características geomorfológicas do terreno, sugerindo que um modelo digital de elevação (MDE) e/ou seus derivados fossem utilizados como mapas de evidência neste trabalho. O MDE utilizado foi o proveniente da missão SRTM (*Shuttle Radar Topography Mission*). O projeto foi um esforço internacional de cooperação entre a NASA e as agências espaciais da Itália e Alemanha. O sistema de radar voou a bordo do ônibus espacial *Endeavour* em fevereiro do ano 2000.

A imagem SRTM processada (FARR *et al.*, 2007) foi adquirida através da plataforma *Google Earth Engine*, e fornece informação acerca da elevação do terreno à uma resolução espacial de 30 m. Para utilização na área de estudo, a imagem foi cortada através da função *.clip* (mais detalhes acerca do algoritmo podem ser visto no Anexo I) utilizando um arquivo vetorial que representa o contorno da área.

Além da imagem SRTM, também foi gerado o mapa de declividade (Figura 14), ou *slope*, a partir do próprio modelo de elevação. O mapa de declividade é um mapa topográfico que mostra a variação da elevação em uma área e é calculado a partir de uma operação de vizinhança entre *pixels* utilizando a biblioteca *Terrain* do *Google Earth Engine*.

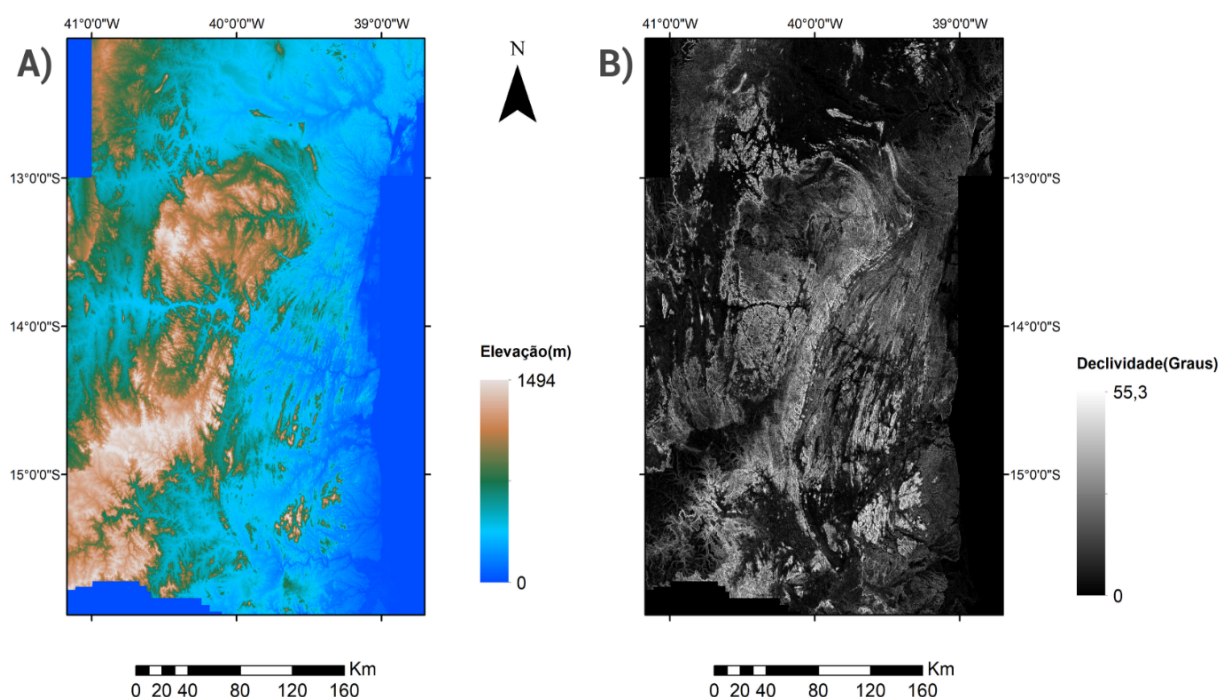


Figura 14 – A) Modelo Digital de Elevação SRTM e B) Mapa de Declividade.

3.3.3. Landsat 8

No estudo de mineralizações associadas à formações superficiais, as imagens de sensoriamento remoto fornecem informações de grande valor, visto que nos permitem estudar a composição dos materiais na superfície da Terra através das suas interações com a radiação eletromagnética.

Embora exista um grande volume de imagens de sensores remotos disponíveis de forma gratuita, pesquisadores usuários desses dados muitas vezes encontram dificuldades de encontrar metodologias para gerar imagens em áreas extensas, heterogêneas e muito fragmentadas. (VASCONCELOS; NOVO, 2004). Para superar essas questões e produzir um dado de qualidade para ser utilizado no modelo, foi escolhido como plataforma para o processamento digital de imagens (PDI) o *Google Earth Engine*, devido ao alto poder computacional que a plataforma fornece, além do extenso banco de dados de imagens de sensoriamento remoto disponíveis para uso. As imagens escolhidas para uso no trabalho foram as da coleção de imagens *Landsat 8 Surface Reflectance, Collection 2*, disponibilizadas na plataforma *Google Earth Engine*. Este banco de dados contém cenas produzidas pelos sensores *Landsat 8 OLI/TIRS* obtidas entre 2013 e 2022, corrigidas para superfície de reflectância e com resolução espacial de 30 m. Cada cena contém 3 bandas do visível, uma banda do infravermelho próximo (VNIR), duas bandas do infravermelho curto (SWIR) e uma banda do infravermelho termal (TIR), além de bandas específicas para controle de qualidade.

3.3.3.1. Geração de Mosaico

Para gerar um mapa de evidência preparado para ser utilizado no modelo de potencial mineral, foram realizados alguns passos (Figura 15) com o intuito de obter um mosaico de imagens que cubra toda a área de estudo. O *script* escrito no GEE para geração do mosaico pode ser acessado a partir do link no Anexo I. A coleção utilizada contém todas as cenas *Landsat 8* adquiridas em todo planeta desde o ano de 2013 (lançamento do satélite) até o momento de realização do trabalho, um volume de dados que além de conter muitas imagens que não seriam utilizadas no trabalho (cenas de outras localidades), pode interferir na eficiência do processo. Para reduzir esse volume de dados, foram aplicados filtros na coleção com o intuito de restringir o número de cenas de acordo com os parâmetros de preferência, aumentando a velocidade do processo e mantendo apenas as imagens mais adequadas para o nosso estudo.

Primeiramente, foi aplicado um filtro de região com o intuito de manter na coleção apenas as cenas que possuem *pixels* dentro da área de estudo. Para isso foi utilizada a função *filterBounds* a partir de um arquivo vetorial que contém os limites da área de estudo.

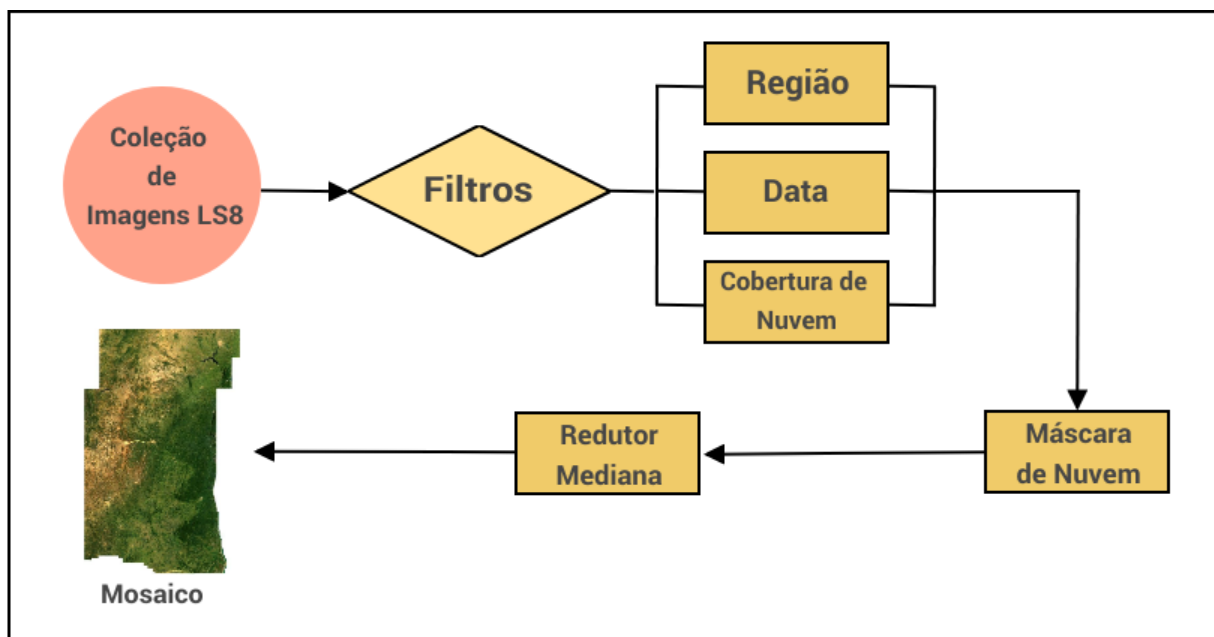


Figura 15 – Fluxograma da metodologia utilizada para construção de um mosaico de imagens Landsat 8 sem nuvens para a área de estudo.

Na sequência foi aplicado um filtro temporal para reduzir o número de cenas a partir da escolha do período de aquisição das imagens utilizando a função *filterDate*. Devido às dimensões e características climáticas heterogêneas da área de estudo, foi preciso utilizar uma amplitude temporal relativamente larga para gerar um mosaico de boa qualidade e com baixa presença de nuvens. Selecionamos as imagens adquiridas entre 1 de janeiro de 2017 e 31 de dezembro de 2021, que forneceram dados suficientes para a geração de um mosaico de qualidade.

Além dos filtros de região e temporal, foi aplicado um filtro de cobertura de nuvens. As cenas da coleção contém nos seus metadados a informação do percentual de *pixels* cobertos por nuvens. A partir dessa informação, foi aplicado um filtro através da função *filterMetadata* para manter na coleção apenas imagens com uma cobertura de nuvens abaixo de um valor de corte. Após alguns testes, foi verificado que o valor de corte ótimo era de 50%. A aplicação dos filtros resultou em um banco de dados que contém 633 cenas que foram utilizadas na geração do mosaico.

Para evitar ruídos na implementação do modelo, antes da geração do mosaico, foi aplicada uma máscara de nuvem e sombra de nuvem a partir da banda de controle de qualidade “QA_PIXEL”, disponível em cada imagem. Essa banda é gerada por um algoritmo que classifica cada *pixel* como um atributo, onde *pixels* com os valores 3 e 5 representam as classes nuvem e sombra de nuvem, respectivamente.

Um dos métodos disponíveis no *Google Earth Engine* para geração de um mosaico a partir de uma coleção de imagens é o uso de um redutor estatístico. Para evitar valores muito

altos (nuvens) ou muito baixos (sombras), a geração do mosaico foi realizada a partir da aplicação do redutor estatístico mediana utilizando a biblioteca *ee.Reducer*. A metodologia consiste na redução de uma coleção de imagens em uma imagem individual (Figura 16), computada *pixel a pixel*, onde o valor de cada *pixel* corresponde ao valor da mediana de todas as imagens naquele ponto. O resultado pode ser visto na figura 17. O mosaico resultante possui as diversas bandas (Figura 18) contidas em cada imagem da coleção *Landsat 8* utilizada. Para uso como mapas de evidência, foram mantidas apenas as bandas 2 (Azul), 3 (Verde), 4 (Vermelho), 5 (Infravermelho próximo – NIR), 6 (Infravermelho Curto 1 – SWIR 1) e 7 (Infravermelho Curto 2 – SWIR 2).

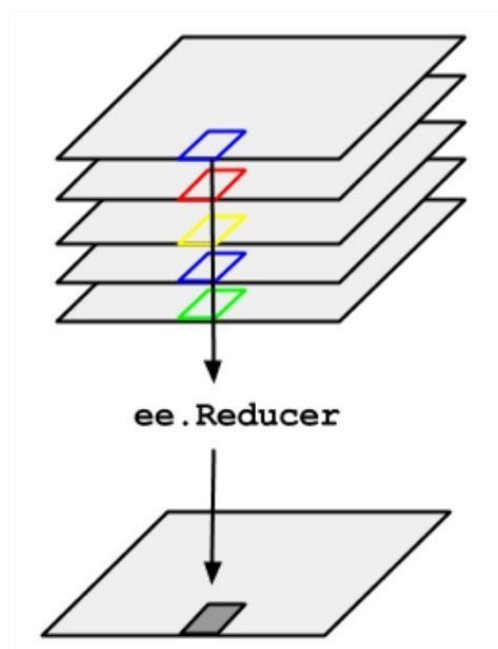


Figura 16 – Ilustração da aplicação de um redutor em uma coleção de imagens utilizando a biblioteca *ee.Reducer*. O valor do pixel no resultado é computado a partir do valor de todas as imagens no ponto.
Fonte: <https://developers.google.com/earth-engine/> .

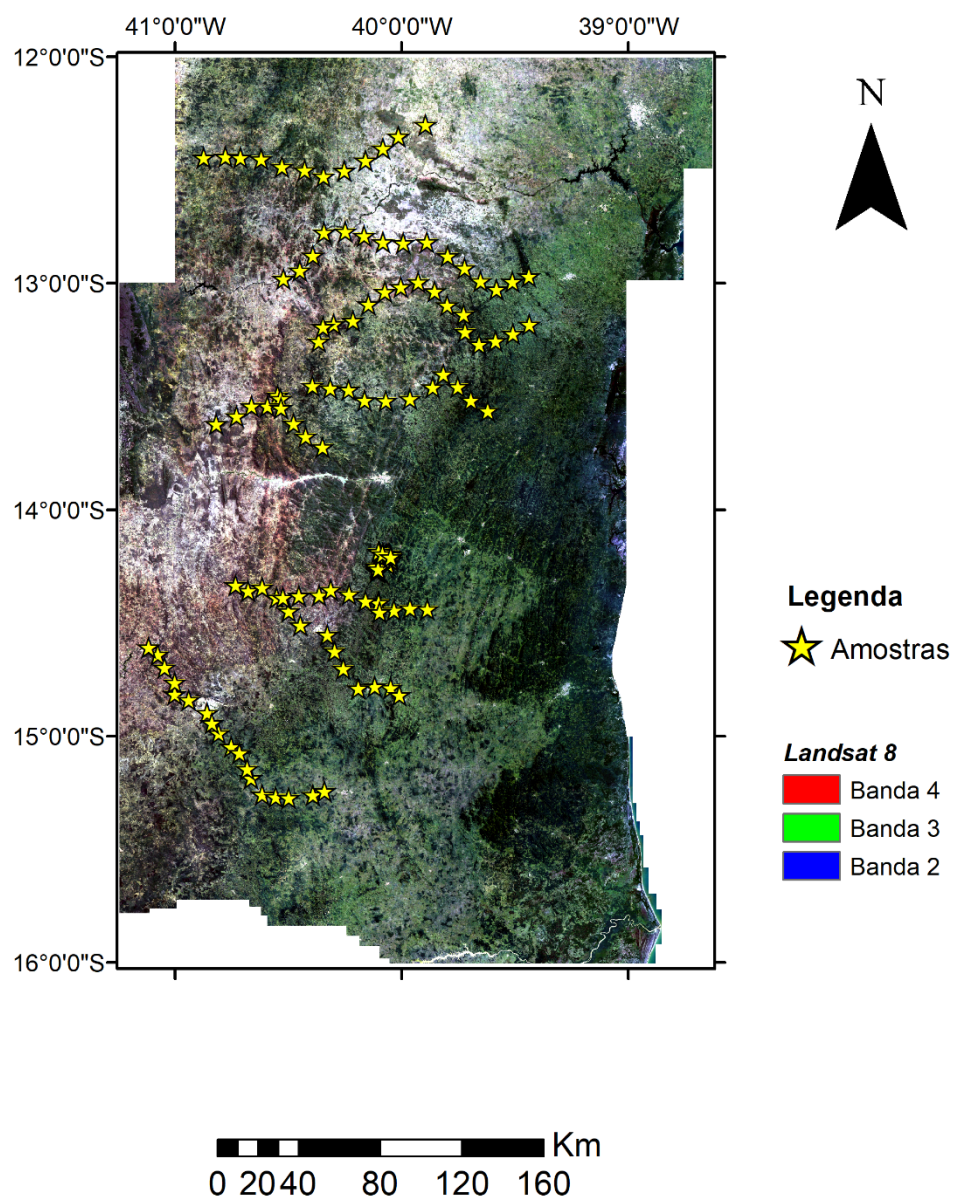


Figura 17 – Localização dos pontos que formam a base de treinamento e teste do modelo sobreposta ao mosaico *Landsat 8*, composição 432 (Cor Natural).

<div>Resolution</div> 30 meters							
Bands							
Name	Units	Min	Max	Scale	Offset	Wavelength	Description
SR_B1		1	65455	2.75e-05	-0.2	0.435-0.451 μm	Band 1 (ultra blue, coastal aerosol) surface reflectance
SR_B2		1	65455	2.75e-05	-0.2	0.452-0.512 μm	Band 2 (blue) surface reflectance
SR_B3		1	65455	2.75e-05	-0.2	0.533-0.590 μm	Band 3 (green) surface reflectance
SR_B4		1	65455	2.75e-05	-0.2	0.636-0.673 μm	Band 4 (red) surface reflectance
SR_B5		1	65455	2.75e-05	-0.2	0.851-0.879 μm	Band 5 (near infrared) surface reflectance
SR_B6		1	65455	2.75e-05	-0.2	1.566-1.651 μm	Band 6 (shortwave infrared 1) surface reflectance
SR_B7		1	65455	2.75e-05	-0.2	2.107-2.294 μm	Band 7 (shortwave infrared 2) surface reflectance
SR_QA_AEROSOL							
Aerosol attributes							
Bitmask for SR_QA_AEROSOL							
ST_B10	Kelvin	0	65535	0.00341802	149	10.60-11.19 μm	Band 10 surface temperature. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_ATRAN		0	10000	0.0001			Atmospheric Transmittance. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_CDIST	km	0	24000	0.01			Pixel distance to cloud. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_DRAD	$\text{W}/(\text{m}^2\text{s}\text{sr}\mu\text{m})/\text{DN}$	0	28000	0.001			Downwelled Radiance. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_EMIS		0	10000	0.0001			Emissivity of Band 10 estimated from ASTER GED. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_ENSD		0	10000	0.0001			Emissivity standard deviation. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.
ST_QA	K	0	32767	0.01			Uncertainty of the Surface Temperature band. If 'PROCESSING_LEVEL' is set to 'L2SR', this band is fully masked out.

Figura 18 – Captura de tela do GEE com a lista das bandas contidas na coleção *Landsat 8* e suas respectivas descrições.

3.4. Geração de Máscara Litológica

Embora as características geomorfológicas da região e a exposição a intensa laterização sejam fatores críticos para a formação de um perfil laterítico bem desenvolvido, a composição da rocha-fonte é decisiva para que os materiais gerados pelo intemperismo possuam teor econômico de bauxita, sendo importante o uso de uma base litológica para indicar os locais que possuem maior potencial para desenvolvimento de minerais supergênicos com teor econômico.

A base litológica utilizada foi extraída do GEOSGB®, a base de dados pública do Serviço Geológico do Brasil (<https://geosgb.cprm.gov.br/>). Ela consiste em um arquivo vetorial no formato *shapefile* com escala de 1:250.000 contendo 107 litologias distintas.

Devido a grande extensão da área, a base litológica é formada por diversos projetos de mapeamento geológico realizados em diferentes épocas, integrados para abranger a totalidade da área. Como o dado não possui uma resolução que entrega o mesmo nível de detalhe dos mapas de evidências, optou-se por usar a base litológica como uma máscara ao final da modelagem, com o intuito de preservar apenas áreas potenciais que coincidem de forma regional com as litologias mais propícias para desenvolvimento da mineralização. Os litotipos foram selecionados a partir do conhecimento adquirido nas campanhas de campo (pontos de campo) e em trabalhos anteriores na área, citados no item 2.6.3, correspondendo basicamente às rochas que estão no contexto do Bloco Jequié, além das regiões mapeadas como coberturas detrito-lateríticas ou indiferenciadas.

Para geração da máscara, a base litológica foi submetida a um processo de rasterização utilizando a função *Polyline to Raster* da *toolbox Conversion Tools* disponível no *software ArcMap* 10.3.1. O mapa resultante possui um valor numérico para cada *pixel*, variando de 1 a 107 de acordo com a litologia representada. Posteriormente, o *raster* foi reclassificado utilizando a ferramenta *reclassify* em um *raster* binário, onde os valores correspondentes aos litotipos mais propícios foram transformados em 1 e os demais litotipos foram transformados em 0. O mapa foi posteriormente exportado no formato *tiff* para permitir a utilização no GEE.

3.5. Preparação das Amostras de Treinamento

Para realização de uma classificação supervisionada, é necessária a utilização de amostras para treinamento e teste do modelo (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Dados provenientes de descrições de campo são as melhores opções para compor a base de dados de treinamento, visto que são observações diretas acerca do material aflorante. A base de dados utilizada para compor as amostras de treinamento foi composta por pontos de visita de afloramento adquiridos nos últimos 3 anos em dois projetos do Serviço Geológico do

Brasil: O Projeto Cartografia do Regolito Laterítico (IZA; SANTOS; FILHO, 2020) e o Projeto Contendas-Macajuba (Em Andamento). A base nos fornece um total de 142 pontos (Figura 17), sendo 25 de presença de mineralização (caso positivo) e 117 pontos de ausência de mineralização (caso negativo).

A preparação da base de dados para utilização no treinamento e teste do modelo consiste na realização de alguns passos para que a tabela de atributos contenha todas as informações necessárias. O primeiro passo foi adicionar uma coluna que contenha a informação acerca da classe da amostra. Para amostras de caso positivo, foi atribuído o valor “1”, enquanto que nas amostras de caso negativo foi atribuído o valor “0”. O segundo passo foi a extração dos valores dos mapas de evidência nos locais dos pontos, para possibilitar que o modelo seja treinado e identifique os padrões das variáveis dentro de cada classe. O último passo foi a separação da base de dados em dados de treinamento (60%) e dados de teste (40%).

Após a aplicação dos procedimentos descritos neste capítulo, o material está pronto para ser utilizado na geração do modelo de potencial mineral de bauxita.

4. RESULTADOS

Este capítulo apresenta os resultados gerados após a execução dos procedimentos descritos no capítulo 3 e ilustrados de forma esquematizada na figura 8.

4.1. Modelo de Potencial Mineral de Bauxita

Para geração do modelo de potencial mineral de bauxita, foram gerados 12 mapas de evidências (Tabela 2) baseados no modelo conceitual proposto no capítulo 3 (Figura 10), além de uma base de amostras de treinamento e teste.

O conjunto de amostras de treinamento foi utilizado para treinar o modelo *Random Forest* utilizando o classificador *smileRandomForest* no ambiente de desenvolvimento do *Google Earth Engine* (para mais detalhes acerca do código ver Anexo I). Foi ajustado um modelo com 100 árvores de decisão, onde o resultado final da classificação corresponde ao resultado da maioria das árvores, como visto no capítulo 2. O número de árvores foi definido após uma série de testes que verificou que o modelo não melhora sua performance quando aumentamos o número de árvores para valores ainda maiores, apenas diminuindo a velocidade de obtenção dos resultados.

MAPA EVIDÊNCIA	DE	TEMA	ESFERA DE CONHECIMENTO
Potássio (K%)		Gamaespectrometria	Intemperismo e Rocha-fonte
Equivalente de Tório		Gamaespectrometria	Intemperismo e Rocha-fonte
Equivalente de Urânio		Gamaespectrometria	Intemperismo e Rocha-fonte
Gradiente Total		Magnetometria	Intemperismo e Rocha-fonte
SRTM		Relevo	Exposição e Preservação
Mapa de Declividade		Relevo	Exposição e Preservação
Landsat 8 – Bandas 2 a 7		Sensoriamento Remoto	Intemperismo

Tabela 2 – Mapas de evidências produzidos para construção do modelo de potencial mineral de bauxita.

Após treinar o modelo, podemos implementar o algoritmo para toda a área de interesse através da função *classify*, utilizando como dados de entrada o conjunto de mapas de evidências gerados. O modelo treinado identifica os padrões das classes (depósito e não-depósito) e classifica *pixel a pixel* a área de estudo, atribuindo o valor “1” para cada *pixel* classificado como depósito e valor “0” para cada *pixel* classificado como não-depósito.

O modelo resultante pode ser visualizado por dois tipos de abordagem. A primeira abordagem (Figura 19A) é na forma binária, onde só existem dois valores possíveis (0 ou 1). Na segunda abordagem (Figura 19B), o resultado é ilustrado na forma de probabilidade de classe (depósito, no caso), possibilitando diferenciar de forma mais precisa as áreas de potencial de bauxita a partir dos valores contínuos de probabilidade.

4.2. Avaliação do desempenho do modelo

Existem diversos modelos de avaliação da performance de um classificador. O método mais simples de ser compreendido é o simples cálculo da acurácia, que consiste na razão entre o total de acertos e o total de pontos. Infelizmente esta medida de desempenho não é confiável para problemas de classificação, especialmente para problemas de classificação com dados de treinamento assimétricos (uma classe com significativamente mais amostras do que outra) (GÉRON, 2019), visto que o desempenho do classificador na classe maior terá um peso muito grande na medida.

Modelo de Potencial Mineral de Bauxita

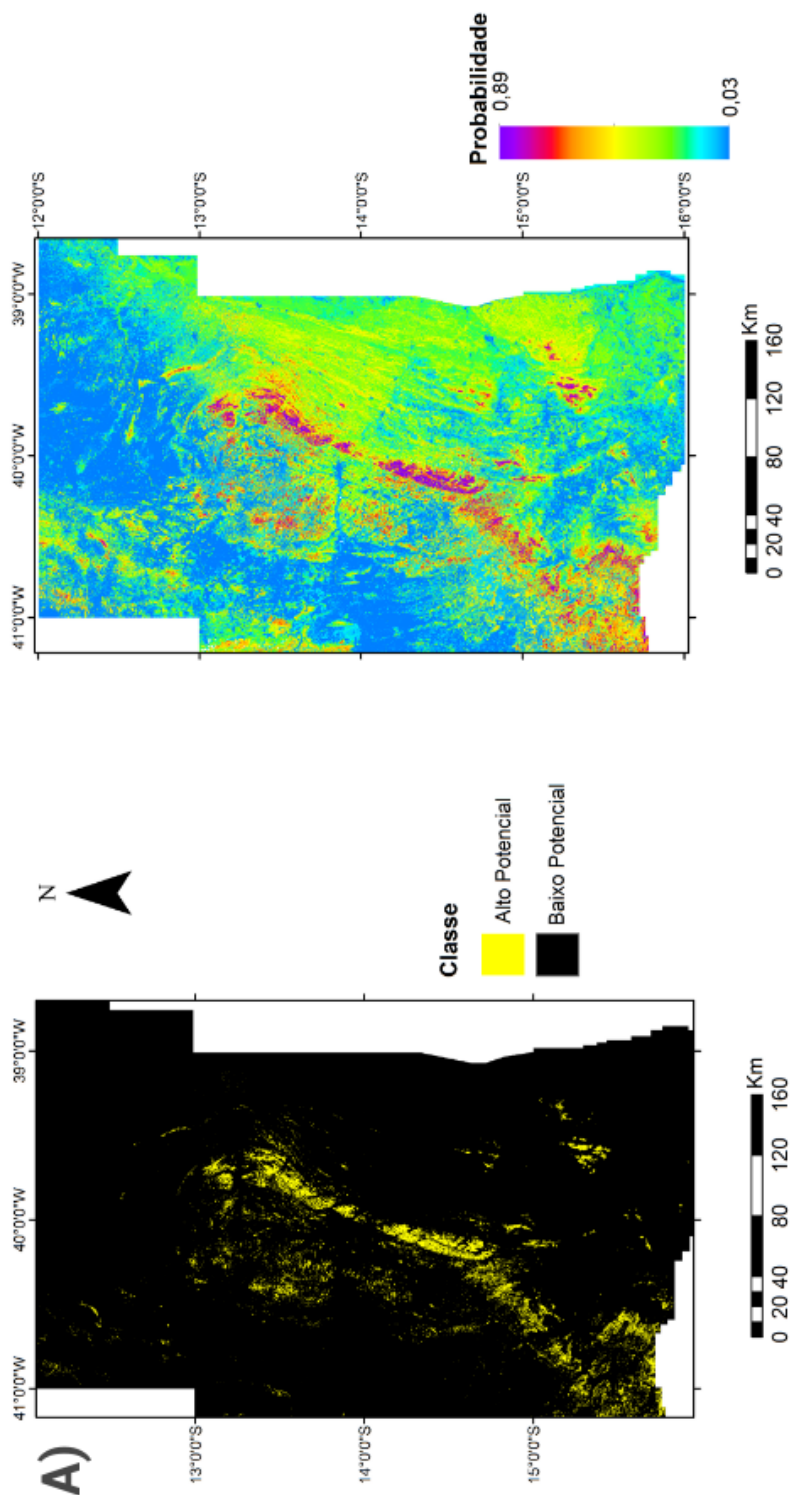


Figura 19 – A) Modelo Random Forest de potencial de bauxita apresentado na forma de classificação binária, onde as áreas em amarelo e preto correspondem às áreas com maior e menor potencial mineral, respectivamente. B) Modelo Random Forest de potencial mineral de bauxita apresentado na forma de probabilidade de classe. O potencial mineral varia de forma crescente do azul (áreas com menor potencial) até o roxo (áreas de maior potencial).

Uma outra medida de desempenho bastante utilizada é o coeficiente de concordância *kappa* (COHEN, 1960). Essa medida é calculada a partir da matriz de confusão (de ordem 2X2 no caso binário), onde as linhas representam as classes reais e as colunas representam as classes preditas. Assim, temos a quantidade de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. O coeficiente *kappa* pode ser calculado a partir da equação:

$$k = \frac{n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i+} x_{+i}}{n^2 - \sum_{i=1}^c x_{i+} x_{+i}}$$

onde x_{ii} é o valor na linha i e coluna i , x_{i+} é a soma da linha i ; x_{+i} é a soma da coluna i da matriz de confusão, n é o número total de amostras e c é o número de classes (COHEN, 1960). Embora seja uma medida de desempenho mais robusta que a acurácia, o *kappa* também pode não ser tão confiável em problemas com dados assimétricos.

Para obter uma medida de performance do modelo que seja confiável mesmo em bases de dados assimétricas, geralmente utiliza-se as medidas *precision*, *recall* e *F1 score* (GÉRON, 2019). *Precision* ou precisão corresponde à razão entre o número de acertos de uma classe e todas as previsões da classe. *Recall* corresponde à razão entre o número de acertos de uma classe e o número real de pontos daquela classe. O *F1 score* ou *F-score* é calculado a partir da combinação entre a precisão e o *recall*. Se as duas medidas forem altas, o *F-score* também será alto. Se as duas medidas forem baixas, o *F-score* também será baixo. Finalmente se uma medida for alta e a outra baixa, o *F-score* será baixo, ou seja, o *F-score* é um bom indicador para verificar se um classificador realmente é confiável ou se ele apenas demonstra bons números de acurácia e *kappa* porque classifica bem apenas a classe com maior número de amostras. É calculado pela razão entre a multiplicação e a soma da precisão e do *recall*, multiplicado por 2.

Nesse trabalho utilizamos o coeficiente *kappa* e o *F-score* como medidores de desempenho no modelo. Foram realizados testes gerando modelos *Random Forest* com cada conjunto de variáveis (mapas de evidências) relativos a cada tema (gamaespectrometria, magnetometria, terreno e *Landsat 8*) e, posteriormente, testes adicionando os temas com o intuito de verificar se o desempenho do modelo realmente melhorava com a adição de mais mapas de evidências (Tabela 3).

CONJUNTOS DE VARIÁVEIS	<i>KAPPA</i>	<i>F-SCORE</i>
Gamaespectrometria	0,4	0,47
Magnetometria	0,33	0,48
Landsat 8	0,57	0,63
Terreno (SRTM + Declividade)	0,52	0,64
Gamaespectrometria + Magnetometria + Landsat 8 + Relevo	0,81	0,83

Tabela 3 – Medidores de desempenho do modelo *Random Forest* utilizando diversos conjuntos de mapas de evidência.

4.3. Importância das variáveis

Uma das grandes vantagens da utilização de um modelo *Random Forest* é a possibilidade de rapidamente calcular a importância das variáveis ou importância da característica. De forma geral é calculada analisando o quanto a impureza reduz quando uma variável é utilizada na árvore de decisão. No *Google Earth Engine* podemos rapidamente ter acesso a esses valores e os resultados da importância relativa de cada variável na geração do nosso modelo pode ser vista no gráfico da Figura 20. Quanto maior a importância da variável, maior foi o seu peso na melhora de desempenho do modelo.



Figura 20 – Gráfico da importância relativa das variáveis utilizadas como mapas de evidência na geração do modelo de potencial mineral de bauxita.

4.4. Aplicação de máscara litológica

Uma outra alternativa para visualização do modelo é a aplicação de uma máscara litológica com o intuito de restringir ainda mais a área de pesquisa. Após a aplicação, são mantidas no modelo (Figura 21) apenas as áreas classificadas como de alto potencial para a mineralização que estão localizadas em áreas correspondentes aos litotipos mais propícios.

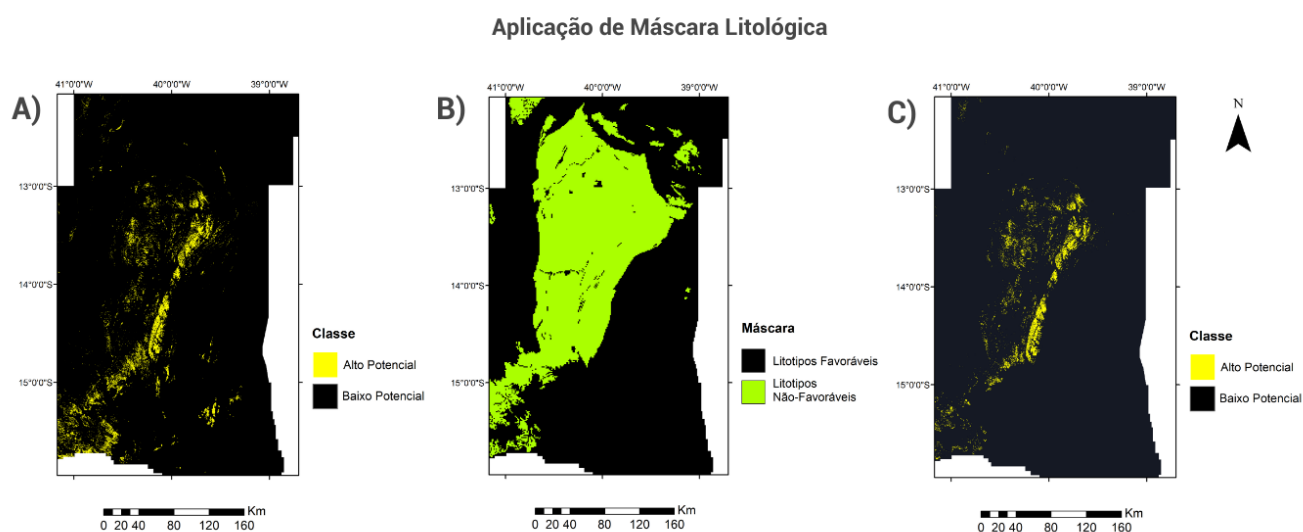


Figura 21 – A) Modelo *Random Forest* de potencial de bauxita; B) Máscara Litológica; C) Modelo após aplicação da máscara.

5. DISCUSSÕES

5.1. Modelo de Potencial Mineral

O modelo de potencial mineral de bauxita na porção leste do Cráton São Francisco nos trás uma indicação das áreas com maior probabilidade da descoberta de novos depósitos. Um dos principais objetivos da geração de uma mapa de prospectividade é a redução da área de estudo, objetivo que foi alcançado visto que o modelo considera como região de grande potencial para ocorrência de bauxita aproximadamente 3,2% da área total (2% da área total, se considerar o modelo após a aplicação da máscara litológica). Essa restrição da área de pesquisa pode reduzir significativamente o custo e tempo de uma empreitada na região.

Para realizar a análise do resultado, as áreas de maior potencial mineral foram divididas em subáreas 1, 2 e 3 (Figura 22), por estarem em contextos específicos e graus de potencialidade distintos.

Mapa de Potencial Mineral de Bauxita

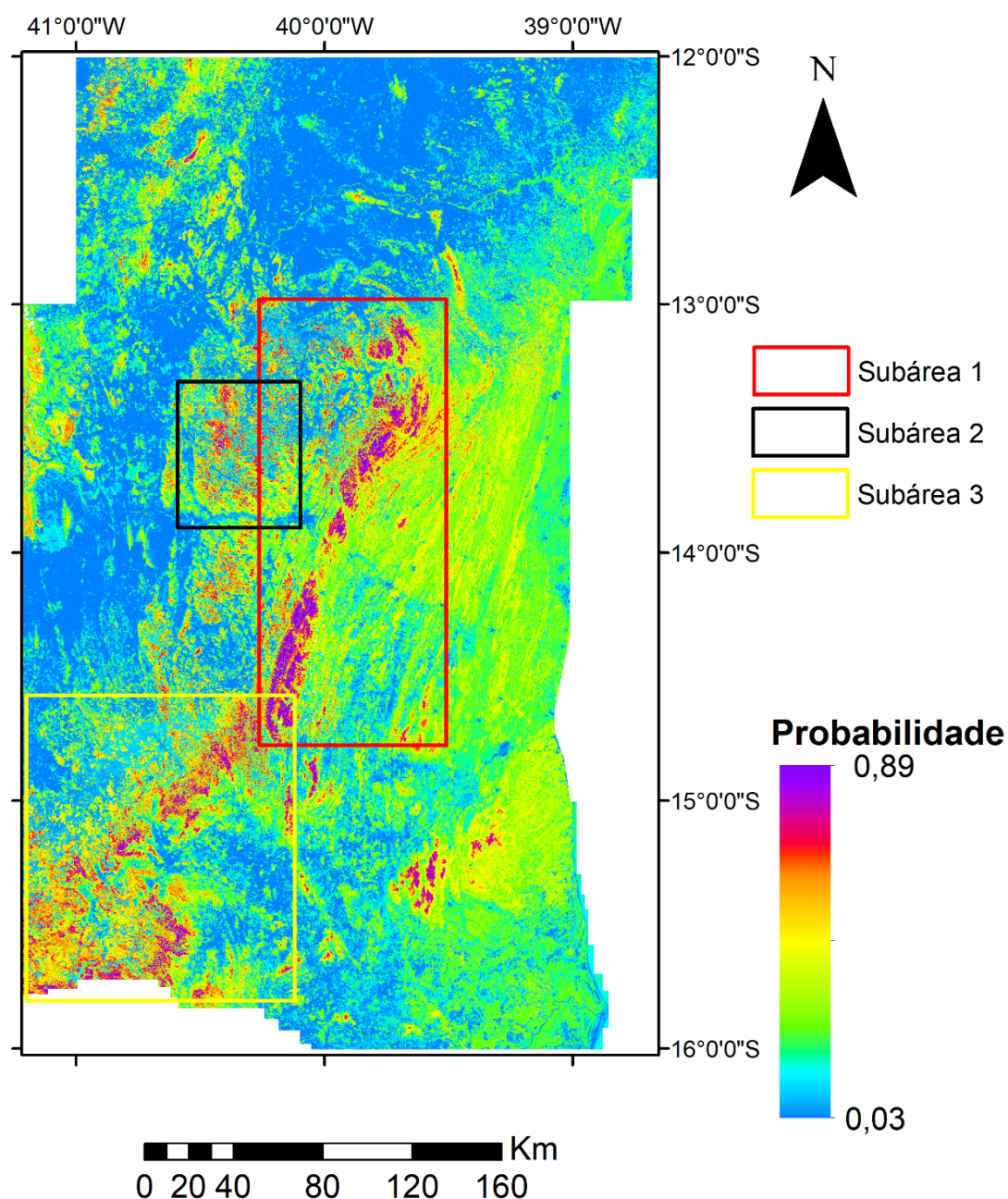


Figura 22 – Modelo de potencial mineral de bauxita com a localização das 3 subáreas selecionadas para discussão. Subárea 1, em vermelho; Subárea 2, em preto e; Subárea 3, em amarelo.

5.1.1. Subárea 1

A subárea 1 representa a região com maior potencial para a ocorrência de bauxita. É a região que contém os alvos com os mais altos valores de probabilidade de classe (até 89%), em especial ao longo do *trend* SW-NE que corta a zona central da área, abrigando a maior parte dos alvos da subárea. A assinatura desse *trend* (Figura 23) é fortemente marcada na gamaespectrometria, com baixos valores de K, comportamento esperado para materiais altamente intemperizados associados à perfis lateríticos bem desenvolvidos, onde o K tende a ser lixiviado devido à sua alta solubilidade. A geomorfologia é marcada por altas cotas (acima de 900 m) de terreno ondulado bastante propício para a formação de bauxita na cabeça dos morros. A área concentra os pontos conhecidos de ocorrência de bauxita, sendo possível sugerir o comportamento da sua distribuição no contexto desse *trend* SW-NE. O trabalho de Iza *et al.* (2020) também destaca essa região, que, além da bauxita, possui ocorrências conhecidas de outros minerais de origem supergênica, como o Mn e de crostas lateríticas ferruginosas. Hayston (2012) sugere que a parte central da subárea 1 tem potencial de ser uma província de bauxita. Outro ponto de destaque para o potencial da região é o contexto climático. Os alvos da área estão geralmente associados com a região de Mata Atlântica, caracterizada por um clima quente e úmido, o que geralmente corresponde a um contexto propício para a ação do intemperismo.

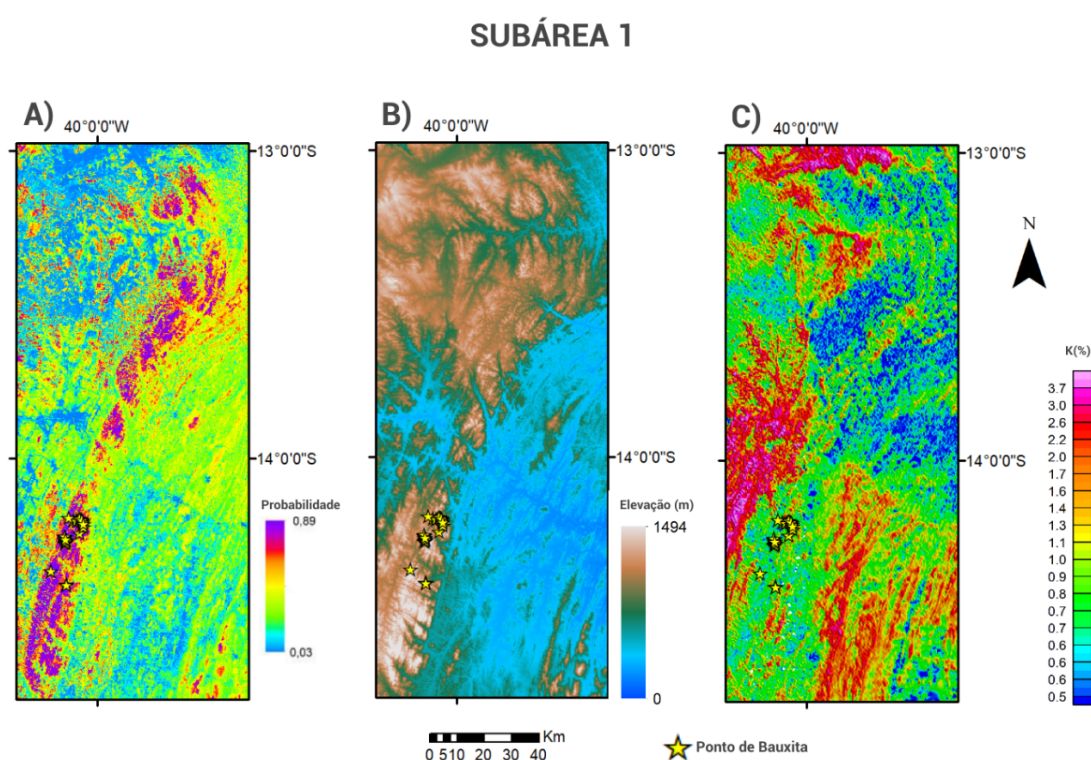


Figura 23 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 1. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.

5.1.2. Subárea 2

Apesar de demonstrar uma assinatura gamaespectrométrica (Figura 24) similar a que é vista na subárea 1, a subárea 2 se diferencia bastante na componente geomorfológica e climática. Primeiramente, a área está inserida no contexto do bioma Caatinga, que apresenta características climáticas distintas da Mata Atlântica. Uma outra diferença se apresenta na medida que os alvos da subárea 1 estão relacionados espacialmente com relevo tipicamente ondulado, os alvos classificados como de alto potencial mineral na subárea 2 estão associados aos platôs que se formaram nas proximidades do município de Maracás-BA. Os platôs são descritos como ricos em latossolos e crostas lateríticas ferruginosas, além de ocorrências locais de crosta niquilífera associada à alteração intempérica dos corpos ultramáficos encontrados na região (IZA *et al.*, 2022). Nos mesmos platôs, Santos (2010) encontrou perfis lateríticos de até 50 m de espessura através de métodos geoeletricos e descrições de furos de sondagem. Também dentro das áreas prospectivas indicadas pelo modelo na subárea 2, encontra-se o protominério de Ni laterítico de Calembe, descrito por Alves (2018). Apesar de não abrigar ocorrências de bauxita conhecidas, a subárea 2 reúne diversas condições favoráveis para a formação da mineralização, como extensa cobertura laterítica, regolitos espessos, além da ocorrência de outros minerais supergênicos.

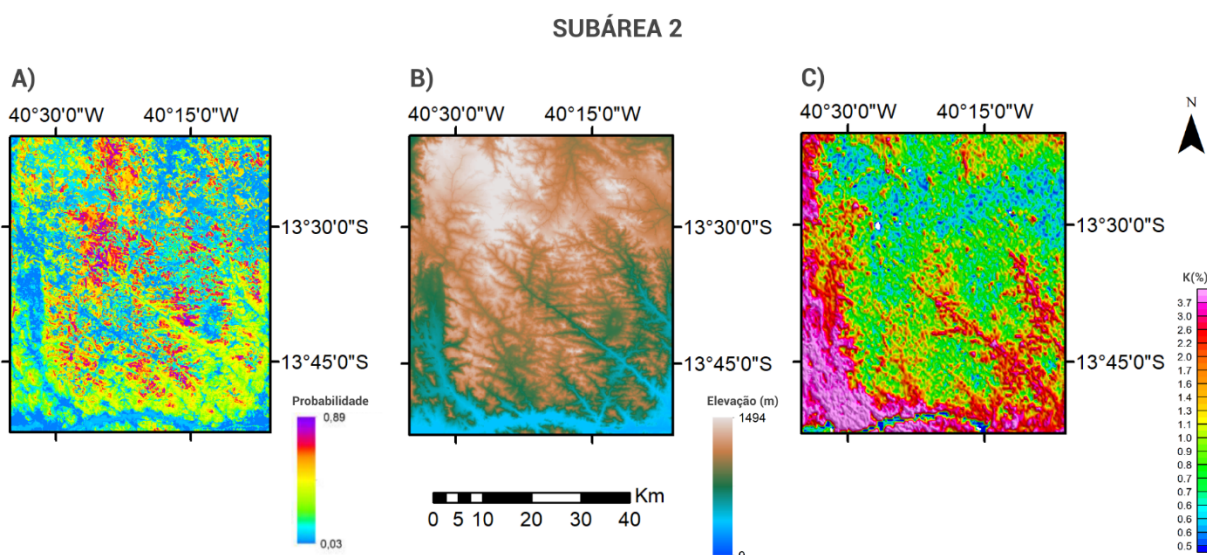


Figura 24 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 2. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.

5.1.3. Subárea 3

A assinatura (Figura 25) gamaespectrométrica relacionada aos alvos prospectivos da subárea 3 seguem o comportamento já visto nas áreas analisadas anteriormente e discutidos na literatura, com padrão de alto Th, que tende a se acumular em condições intempéricas e baixo K, que é muito solúvel e tende a ser lixiviado. Quase toda a subárea é caracterizada pela feição geomorfológica associada ao planalto de Vitória da Conquista, sendo que os alvos prospectivos estão localizados nas bordas desse planalto. A área possui ocorrências conhecidas de crosta laterítica ferruginosa e uma ocorrência de crosta manganésifera (IZA; SANTOS; FILHO, 2020). Como indicado no modelo e pelas ocorrências estudadas, a subárea reúne algumas condições que dão indícios de seu potencial mineral, embora com menos embasamento do que as subáreas anteriores.

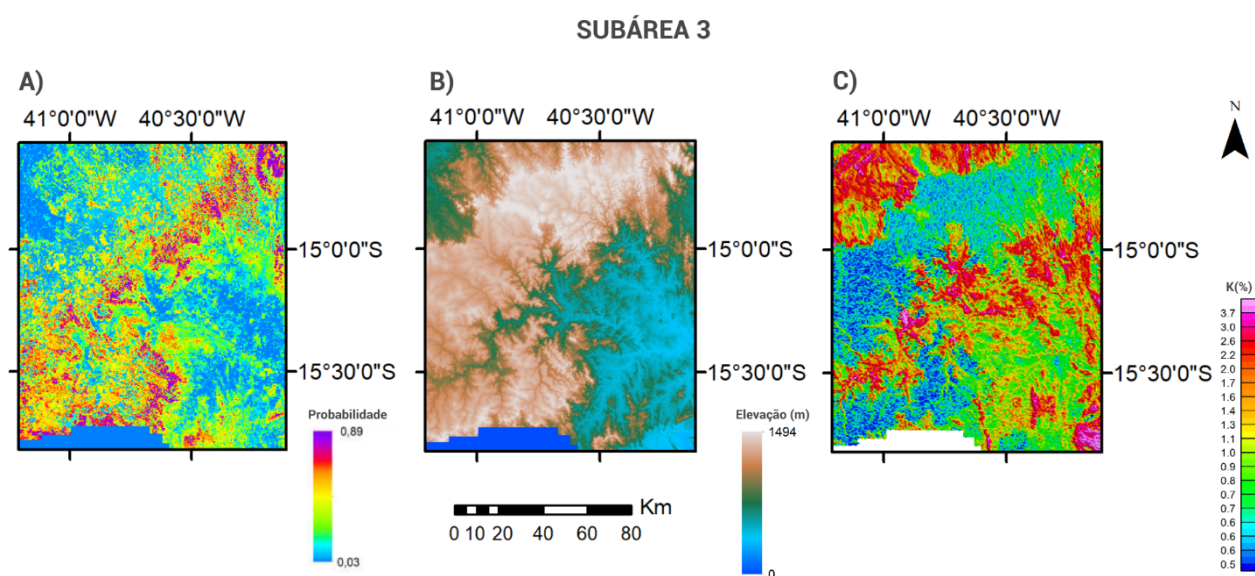


Figura 25 – Mapas em detalhe para visualização de assinaturas relacionadas a subárea 3. A) Modelo de Potencial Mineral; B) SRTM; C) Potássio.

5.2. Importância Relativa das Variáveis

A importância relativa das variáveis foi calculada com o intuito de verificar se os conceitos construídos na literatura se aplicam para o problema em questão, fornecer *insights* acerca dos mapas de evidências utilizados e possivelmente servir como embasamento na escolha dos mapas ou até mesmo no descarte de dados.

Foram utilizados 12 mapas de evidências na construção do modelo do presente trabalho. Cada mapa recebeu um valor de importância de acordo com a média do quanto essa variável aumentou a precisão das árvores de decisão em que foi utilizada. As variáveis que receberam os maiores valores de importância foram o SRTM (17,8), o urânio (10,3) e o potássio (9,8). O valor mais baixo (3,5) foi atribuído à banda 6 do *Landsat 8* (SWIR-1). A partir dessas informações algumas questões podem ser analisadas.

O alto valor de importância atribuído ao SRTM está dentro das expectativas geradas a partir da base conceitual. A componente geomorfológica é um fator de grande peso para a formação do regolito e de seus materiais associados, com os modelos digitais de elevação sendo utilizados largamente em estudos da área.

Os valores altos atribuídos ao potássio e ao urânio também estão de acordo com o que é discutido na literatura. Os radioelementos possuem comportamentos típicos relacionados com o desenvolvimento de perfis lateríticos, sendo bons marcadores no mapeamento de regiões muito intemperizadas com potencial de formação de depósitos de bauxita.

Apesar dos mapas de evidências possuírem importâncias maiores ou menores na melhora da precisão do modelo, tanto o *F-score* quanto o coeficiente *kappa* foram mais altos quando utilizamos todo o conjunto de mapas (ver tabela 3) se comparado ao uso de cada tema individualmente, indicando que todas as variáveis contribuíram para a construção de um modelo mais preciso e confiável.

6. CONCLUSÕES

Os resultados do estudo demonstram que é possível construir uma metodologia de modelagem de potencial mineral utilizando dados geoespaciais de diversas fontes e que pode ser implementada a partir de uma plataforma gratuita e acessível a todos. O uso do *Google Earth Engine* foi essencial para cumprir o objetivo proposto, visto que se mostrou uma ferramenta robusta e versátil ao lidar com dados de diversas naturezas de forma eficiente.

Com o universo de possibilidades que estão disponíveis quando se trata de dados geoespaciais, a modelagem conceitual realizada na parte inicial do estudo foi essencial para a escolha de mapas de evidências que representassem os fatores críticos na formação de depósitos lateríticos de bauxita. Esse modelo conceitual do sistema mineral pode servir como base na construção de outros modelos conceituais para mineralizações formadas a partir de processos semelhantes.

O modelo de potencial mineral de bauxita gerado sugere novos alvos para prospecção, além de reforçar o potencial de áreas com ocorrências conhecidas, como o caso da parte central da subárea 1. Como outros materiais associados ao regolito possuem assinaturas semelhantes em alguns mapas de evidências, a área de estudo também demonstra potencial para descoberta de outros minerais de origem supergênica. Isso é reforçado pelas diversas ocorrências já encontradas na área, como o Níquel de Calembe, crostas ferruginosas e manganesíferas, ETR, entre outras.

O modelo *Random Forest* confirmou o indicado pela literatura, sendo uma metodologia de implantação relativamente rápida, que gerou um bom resultado, atingindo *F-score* de 0,81 e coeficiente *kappa* de 0,83. Como os scripts utilizados na geração do modelo estão disponibilizados (Anexo I) e o *Random Forest* possui a característica de ser um classificador com poucos parâmetros para ajuste, acredito que seu uso contribuiu para o objetivo proposto de construir uma metodologia que possa ser replicável por outros pesquisadores.

Os resultados acerca da Importância Relativa das Variáveis nos fornece *insights* acerca da relação das variáveis utilizadas no modelo e a mineralização estudada, reforçando conceitos previamente estabelecidos na literatura e sugerindo temas para futuros trabalhos. Entre esses temas sugeridos, destaca-se o estudo de possíveis correlações entre a mineralização e a vegetação através de índices espectrais, estudos de *follow-up* nas áreas sugeridas como de alto potencial para geração de mapas de prospectividade de mais detalhe utilizando técnicas de espectrorradiometria, entre outros.

O resultado do modelo diminui de forma expressiva a área de estudo, indicando que a utilização de modelos de potencial mineral podem reduzir o custo, tempo e riscos envolvidos na prospecção mineral, além de constituírem uma importante estratégia de fomento da indústria mineral pelos órgãos governamentais.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ABDELKAREEM, M. I. *et al.* Using of remote sensing and aeromagnetic data for predicting potential areas of hydrothermal mineral deposits in the Central Eastern Desert of Egypt. **Remote Sensing**, v. 7, n. 1, 2018.
- ALMEIDA, F. F. M. O Cráton do São Francisco. **Revista Brasileira de Geociências**, v. 7, p. 349–364, 1977.
- ALVES, N. S. **Petrografia e litogeoquímica do protominério de níquel laterítico de Calembe, Lajedo do Tabocal**. Monografia—[s.l.] Universidade Federal da Bahia, 2018.
- ANAND, R. R. The terminology and classification of the deeply weathered regolith. p. 29, 1988.
- ANAND, R. R. *et al.* **Laterite geochemistry for mineral exploration : chapter 3 : laterite types and associated ferruginous materials, Yilgarn Block, WA : terminology, classification, and atlas / R. R. Anand ... [and others]**. [s.l.] Australian Mineral Industries Research Association, 1989.
- ANAND, R. R.; PAINE, M. Regolith geology of the Yilgarn Craton, Western Australia: Implications for exploration. **Australian Journal of Earth Sciences**, v. 49, n. 1, p. 3–162, 2002.
- ARHIN, E. *et al.* Regolith mapping of deeply weathered terrain in savannah regions of the Birimian Lawra Greenstone Belt, Ghana. **Journal of Geochemical Exploration**, v. 159, p. 194–207, 2015.
- ARHIN, E.; NUDE, P. M. Significance of regolith mapping and its implication for gold exploration in northern Ghana: a case study at Tinga and Kunche. **Geochemistry: Exploration, Environment, Analysis**, v. 9, n. 1, p. 63–69, 2009.
- BAGHDADI, N. *et al.* The contribution of radar satellite imagery to geological exploration in and areas. **Comptes Rendus Geoscience**, v. 337, n. 8, p. 719–728, 2005.
- BÁRDOSSY, G.; ALEVA, G. J. J. **Lateritic Bauxites**. Amsterdam: Elsevier Science, 1990.
- BLATT, H. **Sedimentary Petrology**. Subsequent edição ed. New York, NY: W H Freeman & Co, 1992.
- BOGATYREV, B. A.; ZHUKOV, V. V.; TSEKHOVSKY, YU. G. Formation conditions and regularities of the distribution of large and superlarge bauxite deposits. **Lithology and Mineral Resources**, v. 44, n. 2, p. 135–151, 2009.
- BONHAM-CARTER, G. F.; BONHAM-CARTER, G. **Geographic Information Systems for Geoscientists: Modelling with GIS**. [s.l.] Elsevier, 1994.
- BOOTHROYD, R. J. *et al.* Applications of Google Earth Engine in fluvial geomorphology for detecting river channel change. **WIREs Water**, v. 8, n. 1, p. e21496, 2021.

- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- BRIGGS, I. Machine Contouring Using Minimum Curvature. **Geophysics**, v. 39, 1974.
- CAMPOS, L. D. *et al.* Predictive Mapping of Prospectivity in the Gurupi Orogenic Gold Belt, North–Northeast Brazil: An Example of District-Scale Mineral System Approach to Exploration Targeting. **Natural Resources Research**, v. 26, n. 4, p. 509–534, 2017.
- CARRANZA, E. J. Logistic Regression for Geologically Constrained Mapping of Gold Potential, Baguio District, Philippines. **Exploration and Mining Geology - EXPLORATION MINING GEOLOGY**, v. 10, p. 165–175, 2001.
- CARRANZA, E. J. M. Controls on mineral deposit occurrence inferred from analysis of their spatial pattern and spatial association with geological features. **Ore Geology Reviews**, v. 35, n. 3, p. 383–400, 2009.
- CARRANZA, E. J. M.; LABORTE, A. G. Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. **Ore Geology Reviews**, v. 71, p. 777–787, 2015a.
- CARRANZA, E. J. M.; LABORTE, A. G. Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). **Computers & Geosciences**, v. 74, p. 60–70, 2015b.
- CARRINO, T. A. *et al.* Detecção de ocorrências de coberturas supergênicas a partir de imagens altimétricas e gamaespectrométricas: os alvos do extremo leste do Amazonas. **Simpósio Brasileiro de Sensoriamento Remoto**, n. 14, 2009.
- CARRINO, T. A. *et al.* Discriminação de áreas de espesso regolito do leste do Estado do Amazonas usando estatística multivariada, algoritmo hiperespectral e modelagem de dados espaciais. **Revista Brasileira de Geofísica**, v. 29, p. 155–172, 2011.
- CBPM - COMPANHIA BAIANA DE PESQUISA MINERAL. **Projeto levantamento aerogeofísico da Área Ruy Barbosa / Vitória da Conquista: relatório final de aquisição e processamento de dados**. [s.l.] Lasa Engenharia e Prospecções S.A, 2007.
- CBPM - COMPANHIA BAIANA DE PESQUISA MINERAL. **Projeto levantamento aerogeofísico da Área Cândido Sales / Mascote: relatório final de aquisição e processamento de dados**. [s.l.] Lasa Engenharia e Prospecções S.A, 2009.
- CBPM - COMPANHIA BAIANA DE PESQUISA MINERAL; CPRM - SERVIÇO GEOLÓGICO DO BRASIL. **Projeto levantamento aerogeofísico da Área Ipirá-Ilhéus: relatório final de levantamento e processamento dos dados magnetométricos e gamaespectrométricos**. Rio de Janeiro: Microsurvey, 2011.
- CHAVES, J.; ROCHA, W. **Geotecnologias trilhando novos caminhos nas geociências**. [s.l.: s.n.].
- COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37–46, 1960.

COSTA, I. S. L.; TAVARES, F. M.; OLIVEIRA, J. K. M. DE. Predictive lithological mapping through machine learning methods: a case study in the Cinzento Lineament, Carajás Province, Brazil. **Journal of the Geological Survey of Brazil**, v. 2, n. 1, p. 26–36, 2019.

COSTA, M. Introdução ao intemperismo laterítico e à lateritização. **Prospecção geoquímica**, p. 299–344, 2007.

CRAIG, M. A. *et al.* Regolith-landform mapping in the Yilgarn Craton, Western Australia : towards a standardized approach. **CSIRO Australia, Division of Exploration Geoscience Report**, 1 fev. 1993.

DAMASCENO, H.; BASSINI, A. **CienTec-USP - O Espectro eletromagnético**. Disponível em:

<<https://www.parquecientec.usp.br/passeio-virtual/tudo-sao-ondas/o-espectro-eletromagnetico>>. Acesso em: 3 jun. 2022.

DAUTH, C. Airborne magnetic, radiometric and satellite imagery for regolith mapping in the Yilgarn Craton of Western Australia. **Exploration Geophysics**, v. 28, n. 2, p. 199–203, 1997.

DAVIES, R. S. *et al.* Towards producing mineral resource-potential maps within a mineral systems framework, with emphasis on Australian orogenic gold systems. **Ore Geology Reviews**, v. 119, p. 103369, 2020.

DICKSON, B. L.; SCOTT, K. M. Interpretation of aerial gamma-ray surveys - adding the geochemical factors. **AGSO Journal of Australian Geology and Geophysics**, v. 17, n. 2, p. 187–200, 1997.

DONG, J. *et al.* Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. **Remote Sensing of Environment**, Landsat 8 Science Results. v. 185, p. 142–154, 2016.

EBELE, J. E.; OFOEGBU, C. O.; NUR, A. Interpretation of high-resolution aeromagnetic and radiometric data for delineation of mineral potential zones over Abuja and Environs, North-Central Nigeria. **Arabian Journal of Geosciences**, v. 14, n. 18, p. 1947, 2021.

EGGLETON, R. A. **The regolith glossary : surficial geology, soils and landscape**. [s.l.] Cooperative Research Centre for Landscape Evolution and Mineral Exploration, 2001.

FARR, T. G. *et al.* The Shuttle Radar Topography Mission. **Reviews of Geophysics**, v. 45, n. 2, 2007.

FERNANDES, P. C. D. Projeto Terras Raras no Complexo Jequié, Bahia. 1996.

FITZ, P. R. **Geoprocessamento sem complicação**. [s.l.] Oficina de Textos, 2008.

FORD, A. Practical Implementation of Random Forest-Based Mineral Potential Mapping for Porphyry Cu–Au Mineralization in the Eastern Lachlan Orogen, NSW, Australia. **Natural Resources Research**, v. 29, n. 1, p. 267–283, 2020.

FREYSSINET, P. *et al.* Ore-Forming Processes Related to Lateritic Weathering. 2005.

FURQUIM, C. **10 Algoritmos de Aprendizagem de Máquinas (Machine Learning) que você precisa saber.** Medium, 2018. Disponível em: <<https://medium.com/@cristianofurquim/10-algoritmos-de-aprendizagem-de-m%C3%A1quinas-machine-learning-que-voc%C3%AA-precisa-saber-c49f9eefe319>>. Acesso em: 6 jun. 2022

GEOLOGICAL SURVEY OF WESTERN AUSTRALIA. **Revised classification system for regolith in Western Australia, and the recommended approach to regolith mapping.** Perth: Geological Survey of Western Australia, 2013.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.** [s.l.] O'Reilly Media, Inc., 2019.

GORELICK, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, Big Remotely Sensed Data: tools, applications and experiences. v. 202, p. 18–27, 2017.

GOZZARD, J. R.; TAPLEY, I. J. **Improved regolith-landform mapping using Landsat TM imagery as an aid to mineral exploration in the Lawlers District, North-eastern Goldfields Region, Western Australia.** . Em: AUSTRALASIAN REMOTE SENSING CONFERENCE. Melbourne: 1994.

GUHA, A. *et al.* Analysis of ASTER data for mapping bauxite rich pockets within high altitude lateritic bauxite, Jharkhand, India. **International Journal of Applied Earth Observation and Geoinformation**, v. 21, p. 184–194, 2013.

HARIHARAN, S. *et al.* Random Forest-Based Prospectivity Modelling of Greenfield Terrains Using Sparse Deposit Data: An Example from the Tanami Region, Western Australia. **Natural Resources Research**, v. 26, n. 4, p. 489–507, 2017.

HARRIS, D.; PAN, G. Mineral Favorability Mapping: A Comparison of Artificial Neural Networks, Logistic Regression, and Discriminant Analysis. **Natural Resources Research**, v. 8, n. 2, p. 93–109, 1999.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning – Data Mining, Inference, and Prediction.** , 2009.

HAYSTON, P. **A descoberta da província costeira de bauxita na Bahia.** . Em: SIMEXMIN 2012. Ouro Preto - MG: 2012. Disponível em: <http://www.adimb.com.br/simexmin2012/wp-content/themes/simexmin/palestras/10%20-%20Casos%20de%20Sucesso/XI_5_Hayston.pdf>

HERRERA, I. L. I. E. Utilização de dados altimétricos, geomorfológicos e gamaespectrométricos para a identificação de crostas lateríticas em uma área da porção norte do estado de Rondônia. 2016.

HORBE, A. M. C.; ANAND, R. R. Bauxite on igneous rocks from Amazonia and Southwestern of Australia: Implication for weathering process. **Journal of Geochemical Exploration**, v. 111, n. 1, p. 1–12, 2011.

HRONSKY, J. M. A.; GROVES, D. I. Science of targeting: definition, strategies, targeting and performance measurement. **Australian Journal of Earth Sciences**, v. 55, n. 1, p. 3–12, 2008.

IAEA. **Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data**. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2003.

IZA, E. R. H. DE F. *et al.* **Mapa de regolito: folha SD.24-V-D-I Maracás**, 2022. Accepted: 2022-03-25T13:03:48Z.

IZA, E. R. H. DE F.; SANTOS, R. S. V. DOS; FILHO, B. E. DA C. Integration of multisource data to support the identification of lateritic regolith in Eastern - Bahia, northeastern Brazil. **Journal of the Geological Survey of Brazil**, v. 3, n. 1, p. 1–24, 2020.

IZA, E. R. H. F. **Coberturas lateríticas do SW do Cráton Amazônico: aspectos geofísicos e geoquímicos**. Doutorado em Geologia—[s.l.] Universidade de Brasília, 2017.

IZA, E. R. H. F. *et al.* Integration of Geochemical and Geophysical Data to Characterize and Map Lateritic Regolith: An Example in the Brazilian Amazon. **Geochemistry, Geophysics, Geosystems**, v. 19, n. 9, p. 3254–3271, 2018.

KEAREY, P.; BROOKS, M.; HILL, I. **An Introduction to Geophysical Exploration**. [s.l.] John Wiley & Sons, 2002.

LAWLEY, C. J. M. *et al.* Prospectivity modelling of Canadian magmatic Ni (\pm Cu \pm Co \pm PGE) sulphide mineral systems. **Ore Geology Reviews**, v. 132, p. 103985, 2021.

LICHT, O. A. B. **Prospecção geoquímica : depósitos minerais metálicos, não-metálicos, óleo e gás /**.

LIN, N. *et al.* A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity. **Minerals**, v. 11, n. 2, p. 159, 2021.

MARTINS, A. A. M.; SANTOS, R. A. DOS. Ibicarai: folha SD.24-Y-B-V: estado da Bahia. **CPRM**, 1997.

MCCUAIG, T. C.; BERESFORD, S.; HRONSKY, J. Translating the mineral systems approach into an effective exploration targeting system. **Ore Geology Reviews**, Mineral Prospectivity Analysis and Quantitative Resource Estimation. v. 38, n. 3, p. 128–138, 2010.

MCKAY, G.; HARRIS, J. R. Comparison of the Data-Driven Random Forests Model and a Knowledge-Driven Method for Mineral Prospectivity Mapping: A Case Study for Gold Deposits Around the Huritz Group and Nuelin Suite, Nunavut, Canada. **Natural Resources Research**, v. 25, n. 2, p. 125–143, 2016.

MCQUEEN, K. G.; CRAIG, M. A. Developments and new approaches in regolith mapping. **Centre for Australian Regolith Studies Occasional Publication**, n. 3, p. 80, 1995.

MEDEIRO, F. S. L. **CLASSIFICAÇÃO DO USO E COBERTURA DAS TERRAS E QUANTIFICAÇÃO DOS PADRÕES ESPACIAIS DE ÁREAS DE CAATINGA EM SÉRIES TEMPORAIS**. [s.l.] Universidade Estadual de Feira de Santana, 2018.

METELKA, V. *et al.* Automated regolith landform mapping using airborne geophysics and remote sensing data, Burkina Faso, West Africa. **Remote Sensing of Environment**, v. 204, p. 964–978, 2018.

MINTY, B. Short note: on the use of radioelement ratios to enhance gamma-ray spectrometric data. **Exploration Geophysics**, v. 42, n. 1, p. 116–120, 2011.

NABIGHIAN, M. N. The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation. **GEOPHYSICS**, v. 37, n. 3, p. 507–517, 1972.

OLIVEIRA, L. P. **Uso e cobertura das terras no entorno da Baía de Todos os Santos, Bahia, Brasil: mudanças dos padrões espaciais em séries temporais**. [s.l.] Universidade Estadual de Feira de Santana, 2016.

PARSA, M.; MAGHSOUDI, A. Assessing the effects of mineral systems-derived exploration targeting criteria for random Forests-based predictive mapping of mineral prospectivity in Ahar-Arasbaran area, Iran. **Ore Geology Reviews**, v. 138, p. 104399, 2021.

PATEL, N. N. *et al.* Multitemporal settlement and population mapping from Landsat using Google Earth Engine. **International Journal of Applied Earth Observation and Geoinformation**, v. 35, p. 199–208, mar. 2015.

PORWAL, A. *et al.* Fuzzy inference systems for prospectivity modeling of mineral systems and a case-study for prospectivity mapping of surficial Uranium in Yeelirrie Area, Western Australia. **Ore Geology Reviews**, v. 71, p. 839–852, 2015.

PORWAL, A.; CARRANZA, E. J. M.; HALE, M. Artificial Neural Networks for Mineral-Potential Mapping: A Case Study from Aravalli Province, Western India. **Natural Resources Research**, v. 12, n. 3, p. 155–171, 2003.

PRADO, E. M. G. *et al.* Modeling of Cu-Au prospectivity in the Carajás mineral province (Brazil) through machine learning: Dealing with imbalanced training data. **Ore Geology Reviews**, v. 124, p. 103611, 2020.

REDDY, R. K. T.; BONHAM-CARTER, G. F. A Decision-Tree Approach to Mineral Potential Mapping in Snow Lake Area, Manitoba. **Canadian Journal of Remote Sensing**, v. 17, n. 2, p. 191–200, 1991.

RODRIGUEZ-GALIANO, V. *et al.* Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. **Ore Geology Reviews**, v. 71, p. 804–818, 2015.

RODRIGUEZ-GALIANO, V. F.; CHICA-OLMO, M.; CHICA-RIVAS, M. Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain. **International Journal of Geographical Information Science**, v. 28, n. 7, p. 1336–1354, 2014.

SAMUEL, A. Some Studies in Machine Learning Using the Game of Checkers. **IBM J. Res. Dev.**, 1959.

SANTOS, C. B. DOS. Avaliação hidrogeológica do alto da bacia do rio Jiquiriçá, BA. 2010.

SCHODDE, R. C. **Challenges and opportunities for under-cover exploration in Australia**. Em: UNCOVER SUMMIT, 2014. Disponível em: <<https://minexconsulting.com/challenges-and-opportunities-for-under-cover-exploration-in-australia/>>

SOUZA, C. M. *et al.* Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. **Remote Sensing**, v. 12, n. 17, p. 2735, 2020.

SOUZA, J. D. DE; MELO, R. C. DE; KOSIN, M. (COORDS). Mapa geológico do estado da Bahia. 2003.

SRUTHI, E. R. **Random Forest | Introduction to Random Forest Algorithm**. **Analytics Vidhya**, 2021. Disponível em: <<https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>>. Acesso em: 6 jun. 2022

SUN, L. *et al.* Geologic characteristics and potential of bauxite in China. **Ore Geology Reviews**, v. 120, p. 103278, 2020.

TAYLOR, G.; EGGLETON, R. A. **Regolith Geology and Geomorphology**. [s.l.] John Wiley & Sons, 2001.

THOMAS, M. F. **Tropical geomorphology: a study of weathering and landform development in warm climates** /.

TRIPATHI, M. K.; GOVIL, H. Regolith mapping and geochemistry of hydrothermally altered, weathered and clay minerals, Western Jahajpur belt, Bhilwara, India. **Geocarto International**, v. 37, n. 3, p. 879–895, 2020.

VASCONCELOS, C. H.; NOVO, E. M. L. DE M. Mapping land use cover using segmentation and classification of fraction images, soil, shade and vegetation, derived from a linear mixing model applied to Landsat-5 TM data, Tucuruí reservoir region - Pará. **Acta Amazonica**, v. 34, p. 487–493, 2004.

WANG, L. *et al.* A summary of the special issue on remote sensing of land change science with Google earth engine. **Remote Sensing of Environment**, v. 248, p. 112002, 2020.

WANG, Y. *et al.* Ediacaran carbon cycling and Shuram excursion recorded in the Tarim Block, northwestern China. **Precambrian Research**, v. 377, p. 106694, 1 ago. 2022.

WEDEPOHL, K. H.; CORRENS, C. W., Shaw, Denis Martin, Turekian, Karl K. **Handbook of geochemistry**. Berlin; Heidelberg; New York: Springer-Verlag, 1969.

WILFORD, J. A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis. **Geoderma**, v. 183–184, p. 124–142, 2012.

WILFORD, J. R. **New regolith mapping approaches for old Australian landscapes.** Thesis—[s.l: s.n.].

WILFORD, J. R.; BIERWIRTH, P. N.; CRAIG, M. A. Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology. **AGSO Journal of Australian Geology and Geophysics**, v. 17, n. 2, p. 201–216, 1997.

WYBORN, L. A. I.; HEINRICH, C. A.; JAQUES, A. Australian Proterozoic mineral systems: essential ingredients and mappable criteria. **Publication Series of the Australasian Institute of Mining and Metallurgy**, v. 5, p. 109–115, 1994.

ZHANG, Z.; ZUO, R.; XIONG, Y. A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. **Science China Earth Sciences**, v. 59, n. 3, p. 556–572, 2016.

ANEXO 1 – Links para acesso aos *scripts* no *Google Earth Engine*

- Link para o *script* para geração do mosaico *Landsat 8* com filtro de nuvem:

<https://code.earthengine.google.com/6a411300c0a6032c9be70d5546648366>

- Link para o *script* utilizado para modelagem do potencial mineral de bauxita:

<https://code.earthengine.google.com/7f8ebdfd82874061fdace017df999ba0>