

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE GEOCIÊNCIAS - IGEO
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E
EVOLUÇÃO GEOLÓGICA - PPGMEG

DISSERTAÇÃO

MAPEAMENTO DIGITAL DE CLASSES DE SOLO USANDO
DADOS TOPOGRÁFICOS E AERO GEOFÍSICOS NA UNIDADE
HIDROLÓGICA DE PLANEJAMENTO DO RIO GUANDU-MIRIM
E BACIAS LITORÂNEAS, RJ

GUILHERME FREITAS DA SILVA

2024



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE GEOCIÊNCIAS- IGEO
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E
EVOLUÇÃO GEOLÓGICA - PPGMEG

MAPEAMENTO DIGITAL DE CLASSES DE SOLO USANDO DADOS
TOPOGRÁFICOS E AERO GEOFÍSICOS NA UNIDADE
HIDROLÓGICA DE PLANEJAMENTO DO RIO GUANDU-MIRIM E
BACIAS LITORÂNEAS, RJ

GUILHERME FREITAS DA SILVA

Sob a Orientação da Professora
Dra. Helena Saraiva Koenow Pinheiro

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre em geociências**, no Curso de Pós-Graduação em Modelagem e Evolução Geológica, Área de concentração de Modelagem de processos Geológicos e Geodinâmicos, na linha de pesquisa de Processos Exógenos e Geoambientais

Seropédica, RJ
Dezembro de 2024

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

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

S586m

Silva, Guilherme Freitas da, 1990-
MAPEAMENTO DIGITAL DE CLASSES DE SOLO USANDO DADOS
TOPOGRÁFICOS E AERO GEOFÍSICOS NA UNIDADE HIDROLÓGICA
DE PLANEJAMENTO DO RIO GUANDU-MIRIM E BACIAS
LITORÂNEAS, RJ / Guilherme Freitas da Silva. -
Seropédica, 2024.
59 f.: il.

Orientadora: Helena Saraiva Koenow Pinheiro.
Dissertação (Mestrado). -- Universidade Federal Rural
do Rio de Janeiro, Programa de Pós-graduação em
Modelagem e Evolução Ambiental, 2024.

1. Pedometria. 2. Machine Learning. 3. Dados
Gamaespectrométricos. 4. Material de Origem. I.
Saraiva Koenow Pinheiro, Helena, 1985-, orient. II
Universidade Federal Rural do Rio de Janeiro.
Programa de Pós-graduação em Modelagem e Evolução
Ambiental III. Título.



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE GEOCIÊNCIAS**

HOMOLOGAÇÃO DE DISSERTAÇÃO DE MESTRADO Nº 104/2024 - IGEO (11.39.00.34)

Nº do Protocolo: 23083.071626/2024-17

Seropédica-RJ, 20 de dezembro de 2024.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E EVOLUÇÃO**

GEOLOGICA GUILHERME DE FREITAS DA SILVA

Dissertação submetida como requisito parcial para obtenção do grau de Mestre em Geociências, no Programa de Pós-Graduação em Modelagem e Evolução Geológica, área de concentração em MODELAGEM DE PROCESSOS GEOLÓGICOS E GEODINÂMICOS.

DISSERTAÇÃO APROVADA EM 17/12/2024.

Helena Saraiva Koenow Pinheiro. Dra. Departamento de Solos-IA/UFRRJ
(Orientadora/presidente da banca)

Alexis Rosa Nummer. Dr. Departamento de Petrologia e Geotectônica-IGEO/UFRRJ
(Membro Interno)

Mateus Marques Bueno. Dr. IFSULDEMINAS
(Membro Externo)

(Assinado digitalmente em 21/12/2024 09:27)

ALEXIS ROSA NUMMER
PROFESSOR DO MAGISTERIO SUPERIOR
IGEO (11.39.00.34)
Matrícula: ###69#2

(Assinado digitalmente em 20/12/2024 11:49)

HELENA SARAIVA KOENOW PINHEIRO
CHEFE DE DEPARTAMENTO
DeptS (12.28.01.00.00.33)
Matrícula: ###236#8

(Assinado digitalmente em 14/01/2025 13:15)

MATEUS MARQUES BUENO
ASSINANTE EXTERNO
CPF: ###.###.676-##

Visualize o documento original em <https://sipac.ufrj.br/public/documentos/index.jsp> informando seu número: **104**, ano: **2024**, tipo: **HOMOLOGAÇÃO DE DISSERTAÇÃO DE MESTRADO**, data de emissão: **20/12/2024** e o código de verificação: **e4bd2bc77d**

AGRADECIMENTOS

Agradeço a todos que direta ou indiretamente me apoiaram ou não durante essa fase da minha vida. Agradeço aos amigos, professores e familiares, em especial a professora Helena por me ajudar muito e ser paciente comigo.

Agradeço a Universidade Federal Rural do Rio de Janeiro - UFRRJ, em especial ao PPGMEG pelo suporte e estrutura. Agradeço também aos órgãos financiadores FAPERJ, CNPq e CAPES.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 (*This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001*)

RESUMO

SILVA, Guilherme. **“Mapeamento digital de classes de solo usando dados topográficos e aero geofísicos na unidade hidrológica de planejamento do Rio Guandu-Mirim e Bacias litorâneas, RJ”**. 2024. 56 p. Dissertação (Mestrado em Modelagem e Evolução Ambiental). Instituto de Geociências, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ. 2024

O mapeamento de solos é fundamental para o planejamento adequado de atividades relacionadas ao uso do solo, como agricultura, construção de estradas e edificações, ou zoneamento de áreas industriais, entre outras. No entanto, o mapeamento de solos envolve um trabalho extenso e custoso, que muitas vezes resulta em resultados pouco eficientes. Nesse sentido, ferramentas de mapeamento digital de solos têm sido utilizadas para aprimorar a qualidade, mensurar o erro intrínseco nos mapas de solos, e com uma melhor relação custo-benefício. Algumas das principais características e classificações do solo são influenciadas por fatores de sua formação, denominados fatores S.C.O.R.P.A.N., que representam Solo, Clima, Organismos, Relevo, Material Parental, Idade e Posição. Modelos digitais de elevação (MDE) e seus dados derivados são as covariáveis mais utilizadas para representar as condições topográficas, desempenhando um papel importante no mapeamento digital de solos. Feições topográficas como altimetria, declividade, curvatura e outras que representam o relevo de uma área são eficazmente utilizadas para prever características do solo. Como a maioria dos solos existentes são solos minerais, e sua composição mineral reflete suas propriedades e características, justifica-se o uso de covariáveis que reflitam diretamente características do material de origem, como por exemplo, os Dados Aero Geofísicos (DAG), a fim de melhorar a qualidade do mapeamento digital de solos (MDS). O objetivo deste estudo foi utilizar dados topográficos e aero geofísicos para mapear a distribuição espacial das classes de solos na Unidade de Planejamento Hidrológico da Bacia do Rio Guandu-Mirim e Bacias Costeiras, no Rio de Janeiro, Brasil. Os modelos preditivos Decision Tree e Random Forest que obtiveram os melhores resultados estatísticos usaram DAG como covariáveis e obtiveram overall accuracy de 0,87 e 0,97 respectivamente. Dentre as covariáveis provenientes dos DAG, o Urânio apareceu entre as 10 covariáveis mais importantes para todos os modelos que utilizaram DAG. Os resultados obtidos indicam que o uso de DAG como representante do fator de formação material de origem pode contribuir para a obtenção de mapas mais precisos.

Palavras-Chave: pedometria; machine learning; dados gamaespectrométricos; material de origem

ABSTRACT

SILVA, Guilherme. **“Digital soil class mapping using topographic and airborne geophysical data in the hydrological planning unit of Guandu-Mirim river and coastal Basins, RJ.”**. 2024. 59 p. Dissertation (Master in Modeling and Environmental). Instituto de Geociências, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ. 2024

Soil mapping is essential for the appropriate planning of activities related to land use, such as agriculture, road and building construction, or industrial zoning, among others. However, soil mapping involves extensive and costly work, which often results in limited efficiency. In this context, digital soil mapping tools have been employed to enhance map quality, measure intrinsic error, and offer a better cost-benefit operation. Some of the main soil characteristics and classifications are influenced by factors of soil formation, known as S.C.O.R.P.A.N. factors, representing Soil, Climate, Organisms, Relief, Parent Material, Age, and Position. Digital elevation models (DEM) and their derived data are the most commonly used covariates to represent topographic conditions, playing a critical role in digital soil mapping. Topographic features such as elevation, slope, curvature, and others that depict the relief of an area are effectively used to predict soil characteristics. Since most soils are mineral soils, and their mineral composition reflects their properties and characteristics, it is justified to use covariates that directly reflect parent material characteristics, such as Aero Geophysical Data (AGD), to improve the quality of digital soil mapping (DSM). This study aimed to use topographic and aero geophysical data to map the spatial distribution of soil classes in the Guandu-Mirim River Basin and Coastal Basins Hydrological Planning Unit in Rio de Janeiro, Brazil. The predictive models Decision Tree and Random Forest that achieved the best statistical results used AGD as covariates, obtaining overall accuracies of 0.87 and 0.97, respectively. Among the covariates derived from AGD, Uranium ranked among the 10 most important covariates for all models that utilized AGD. The obtained results indicate that using AGD as a representative of the parent material formation factor can contribute to generating more accurate predictive maps.

Keywords: Pedometrics; machine learning; gamma-ray spectrometry data; parental material.

SUMÁRIO

1.INTRODUÇÃO	- 1 -
2. REVISÃO BIBLIOGRÁFICA.....	- 3 -
2.1. Caracterização da Área de Estudo	- 3 -
2.1.1. Características gerais	- 3 -
2.1.2. Relevo	- 4 -
2.1.3. Clima e vegetação	- 4 -
2.1.4. Geologia.....	- 5 -
2.1.5. Pedologia e amostragem de solo.....	- 7 -
2.2. Contextualização do Mapeamento Digital de Solos(MDS).....	- 12 -
2.3. Dados Aero Geofísicos (DAG): um breve resumo	- 15 -
3.ARTIGO: “DIGITAL SOIL CLASS MAPPING USING TOPOGRAPHIC AND AIRBORNE GEOPHYSICAL DATA IN THE HIDROLOGICAL PLANNING UNIT OF GUANDU-MIRIM RIVER AND COASTAL BASINS”	- 19 -
3.1. Introduction	- 19 -
3.2. Material and Methods.....	- 20 -
3.2.1. Study area.....	- 20 -
3.2.2. Pedology and soil sampling	- 21 -
3.2.3. Topographic and airborne geophysical data (AGD)	- 24 -
3.2.4. Predictive models.....	- 27 -
3.3. Result and Discussions	- 28 -
3.3.1. Decision tree (DT) models.....	- 28 -
3.3.2. Random forest (RF) models.....	- 30 -
3.3.3. Models assessment and selection.....	- 33 -
3.4. Conclusions	- 34 -
4. CONSIDERAÇÕES FINAIS	- 36 -
5. REFERÊNCIAS	- 36 -

LISTA DE FIGURAS

Figura 1 - Localização da Unidade Hidrológica de Planejamento Guandu-Mirim e Bacias Litorâneas (AGEVAP, 2018).....	- 3 -
Figura 2- Mapa dos domínios geomorfológicos da área de estudo (IBGE, 2023).....	- 4 -
Figura 3-Mapa Litológico da Unidade Hidrológica de Planejamento do Rio Guandu-Mirim e Bacias Litorâneas - RJ (escala original 1:400.000). Adaptado de Heilbron et. al (2016).....	- 5 -
Figura 4– Mapa de Solos da UHP Guandu-Mirim e Bacias litorâneas, adaptado de Levantamento de reconhecimento de baixa intensidade dos solos do estado do Rio de Janeiro. Rio de Janeiro: Embrapa Solos, por Carvalho Filho, A. et. al – 1:250000 (2003)	- 8 -
Figura 5 – Mapa de uso e ocupação do solo da Unidade Hidrológica de Planejamento do Rio Guandu-Mirim e Bacias Litorâneas (adaptado de MapBiomias, 2017)	- 10 -
Figura 6 – Algumas imagens do trabalho de campo realizado.....	- 11 -
Figura 7 – Preparação das análises Química e Física das amostras de perfis de solo obtidas em campo, realizada no Laboratório de Física dos Solos, Departamentos de Solos – IA – UFRRJ	- 12 -
Figura 8 – Espectros de radiação gama mostrando as posições da janela de energia para cada elemento e da contagem total (Adaptado de: Ribeiro et al, 2015).....	- 16 -
Figura 9 – Mapas gamaespectrométricos utilizados para a modelagem do presente trabalho	- 18 -
Figura 10 – Localization of the Hydrological Planning Unit Guandu-Mirim and Coastal Basins (AGEVAP, 2018) and the soil sampling sites inherited from Pronasolos and the new soil profiles .	- 21 -
Figura 11 - Flow chart summarizing the methodology used in the study	- 23 -
Figura 12 – Covariates Relative Importance for the Decision Tree Model using LOOCV and ROIC, with and with no AGD	- 29 -
Figura 13 – Predictive Soil Class Maps obtained by the Decision Tree Model using LOOCV and ROIC, with and with no AGD	- 30 -
Figura 14 – Covariates Gini and accuracy Importance to the RF model using LOOCV and ROIC, with and with no AGD	- 31 -
Figura 15 – Predictive soil class maps obtained by RF model using LOOCV and ROIC, with and with no AGD	- 32 -
Figura 16 – Predictive soil class map obtained by RF model using ROIC and AGD which showed the best performance at the tests	- 34 -

LISTA DE TABELAS

Tabela 1– Descrição das unidades de mapeamento presentes na área de estudo, reclassificado para o sistema de classificação de solos de 2013, adaptado de Carvalho Filho, A. et. al (2003)	8 -
Tabela 2 – Tabela contendo os valores dos canais e dos picos de energia (em MeV) associados aos canais dos elementos radiométricos e a emissão cósmica. (Adaptado de: Ribeiro et al 2015)	16 -
Tabela 3 – Soil units used, number of soil profiles in each unit and the total number of pixels gattered using ROIC.....	23 -
Tabela 4 - Topographic covariates used in the model construction, along with brief descriptions and references	24 -
Tabela 5 – Gamma Ray covariates, brief descriptions and references	27 -
Tabela 6 – Statistic results of the tested models	33 -

LISTA DE ABREVIACÕES

AD	Árvore de Decisão
AGD	<i>Airborne Geophysics Data</i>
CPRM	Companhia de Pesquisa de Recursos Minerais – Serviço Geológico do Brasil
DAG	Dados Aero Geofísicos
DEM	<i>Digital Elevation Model</i>
DRM	Departamento de Recursos Minerais
DT	<i>Decision Tree</i>
eU	Equivalente em Urânio
eTh	Equivalente em Tório
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
FAO	<i>Food and Agriculture Organization of the United Nations</i>
GIS	<i>geographic information system</i>
GPS	<i>Global Positioning system</i>
HPU	<i>Hydrological Planning Unit</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
INPE	Instituto Nacional de Pesquisas Espaciais
LOOCV	<i>Leave one Out Cross Validation</i> (validação cruzada tirando um item)
MDE	Modelo Digital de Elevação
MDS	Mapeamento Digital de Solos
AO	<i>Overall Accuracy</i> (Acurácia Global)
RF	<i>Random Forest</i>
ROIC	<i>Region of Interest oer Class</i> (Região de Interesse por Classe)
SiBCS	Sistema Brasileiro de Classificação de Solos
SIG	Sistema de informações Geográficas
UHP	Unidade Hidrológica de Planejamento
USGS	<i>United States Geological Survey</i>

1.INTRODUÇÃO

O mapeamento de solos tem como principal função descrever as características físicas, químicas e morfológicas, bem como a distribuição espacial de diferentes classes de solo e seus atributos em uma área específica, fornecendo suporte para um melhor planejamento do uso da terra. No entanto, o mapeamento tradicional é oneroso, demorado e requer mão de obra especializada, somado as dificuldades no compartilhamento de informações sobre solos, como divergência de termos, desatualização, falta de padronização, generalização e imprecisão dos métodos, gerando resultados que muitas vezes carecem de informações importantes para os usuários modernos e tomadores de decisão (Hartemink e McBratney, 2008; Sanchez, 2009; Ten Caten, 2011).

O uso de geotecnologias torna-se altamente relevante nesse campo devido à sua capacidade de lidar com grandes volumes de dados, realizar cálculos estatísticos avançados de forma rápida e eficiente, proporcionando assim a obtenção de produtos com informações mais precisas e de maior qualidade. O mapeamento digital de solos (MDS) surge nesse contexto com o principal objetivo de aprimorar os levantamentos de solos, utilizando dados de sensoriamento remoto, técnicas computacionais como softwares e algoritmos para análise e modelagem de dados espaciais, em conjunto com o mapeamento de campo, a fim de obter informações sobre características importantes da paisagem para delinear a distribuição de classes e propriedades do solo (Lagacherie et al. 2000, 2001, 2006). Esse método confere uma natureza quantitativa aos levantamentos de solos, possibilitando a análise de erros associados e a extrapolação de dados sobre a disposição espacial das classes e propriedades dos solos em modelos matemáticos. Nesse sentido, o MDS torna-se uma ferramenta essencial para o desenvolvimento de levantamentos de solos mais precisos, eficientes, rápidos, com melhor relação custo-benefício e de natureza quantitativa na área de estudo (Chen et al., 2022; Minai et al., 2020; Taghizadeh-Mehrjardi et al., 2021).

O MDS baseia-se na hipótese da existência de uma correlação entre as feições da paisagem (covariáveis) e os fatores de formação do solo, fundamentada no modelo de formação do solo de Jenny (Jenny, 1941), posteriormente atualizado para o chamado modelo SCORPAN (McBratney et al., 2003). Chen et al. (2022) e Wilson (2019) apontam que relevo, organismos e clima são as três covariáveis ambientais mais frequentemente

utilizadas, enquanto outros fatores, como material parental, recebem menos atenção na literatura. Estudos recentes (Bastos et al., 2023; Mello et al., 2021, 2022; Maino et al., 2022; Loiseau et al., 2020, 2021) mostraram que dados geofísicos podem ser utilizados para representar a covariável material parental do modelo SCORPAN, adicionando informações valiosas para os modelos preditivos, melhorando sua capacidade de predição. A gamaespectrometria é um ramo da geofísica que mede a radiação natural emitida pelas rochas e pelo solo. Esse método permite medir as concentrações de K (potássio), U (urânio) e Th (tório) a partir de suas séries de decaimento radioativo (^{40}K , ^{238}U e ^{232}Th), geralmente utilizando detectores de NaI(Tl) em levantamentos in situ ou em levantamentos aéreos (Minty, 1997). A hipótese em foco, é de que a utilização de Dados Aero Geofísicos (DAG) como variável nos modelos preditivos, representando o material de origem dos solos, pode melhorar a capacidade preditiva dos modelos acerca da distribuição espacial das classes de solo.

O objetivo deste estudo foi a calibração e avaliação de modelos preditivos para representar a variação espacial das classes de solo com base nas características da bacia hidrográfica do Rio Guandu-Mirim, utilizando dados topográficos e dados aero geofísicos (DAG) como covariáveis.

A dissertação foi estruturada da seguinte forma: inicialmente, foi realizada uma revisão bibliográfica sobre os dados aerogeofísicos e exemplos de produtos associados; em seguida, foi feita uma revisão da área de estudo, incluindo caracterização geral da área, clima e vegetação, contexto geológico, geomorfológico e pedológico; seguindo com a apresentação do artigo submetido para publicação: “Digital Soil Class Mapping using topographic and airborne geophysical data on the Hydrological Planning Unit of Guandu-Mirim River and Coastal Basins” que inclui descrição metodológica, apresentação dos resultados, discussão e conclusão.

2. REVISÃO BIBLIOGRÁFICA

2.1. Caracterização da Área de Estudo

2.1.1. Características gerais

A Bacia Hidrográfica do Rio Guandu-Mirim (Figura 1) está localizada no oeste do estado do Rio de Janeiro e cobre aproximadamente 260 km², drenando a partir do Complexo Alcalino do Mendanha, conhecido como Guandu-do-Sena, formado por várias nascentes e mudando de nome diversas vezes antes de desaguar na Baía de Sepetiba. A área das bacias costeiras abrange cerca de 210 km², consistindo em uma série de pequenos rios e canais que eventualmente também deságuam na Baía de Sepetiba. A Unidade Hidrológica de Planejamento do Rio Guandu-Mirim e Bacias Litorâneas engloba parte dos municípios de Itaguaí, Rio de Janeiro, Seropédica e Nova Iguaçu (AGEVAP, 2018). Dados oficiais fornecidos pela plataforma online SIGA Web - Guandu, do Comitê de Bacia Hidrográfica do Guandu (disponível em: www.sigaguandu.org.br), foram utilizados para representar limites políticos, estradas de acesso, regiões hidrográficas e o limite da bacia da unidade hidrológica de planejamento (UHP).

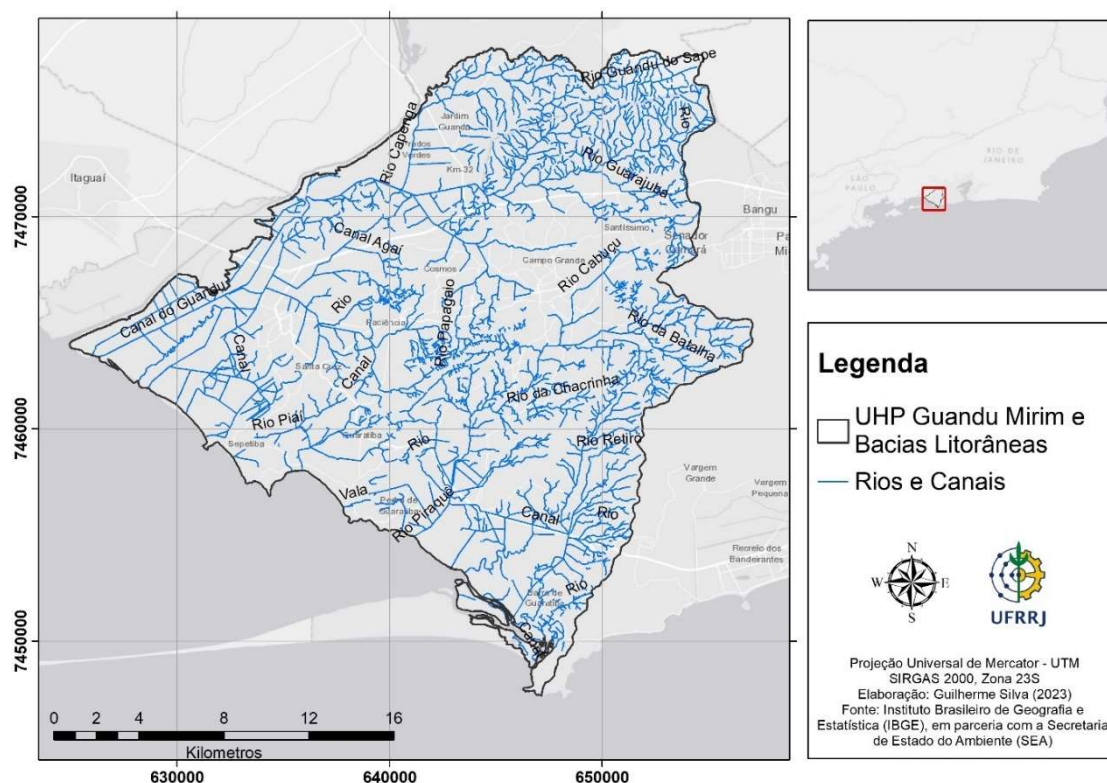


Figura 1 - Localização da Unidade Hidrológica de Planejamento Guandu-Mirim e Bacias Litorâneas (AGEVAP, 2018)

2.1.2. Relevo

O terreno da região é caracterizado principalmente por áreas de encostas e escarpas da porção oceânica da Serra do Mar (domínio das Terras Altas) nas regiões sudeste e nordeste da bacia, com altitudes que atingem até 1000 m, e pela extensa planície flúvio-marinha (domínio das Terras Baixas) que forma as Baixadas de Sepetiba (Figura 2). Destacam-se na área de terras baixas as colinas baixas e a extensão da planície fluvial dos rios Guandu, Guandu-Mirim e Guarda, com altitude média abaixo de 60 m (ANA, 2007).

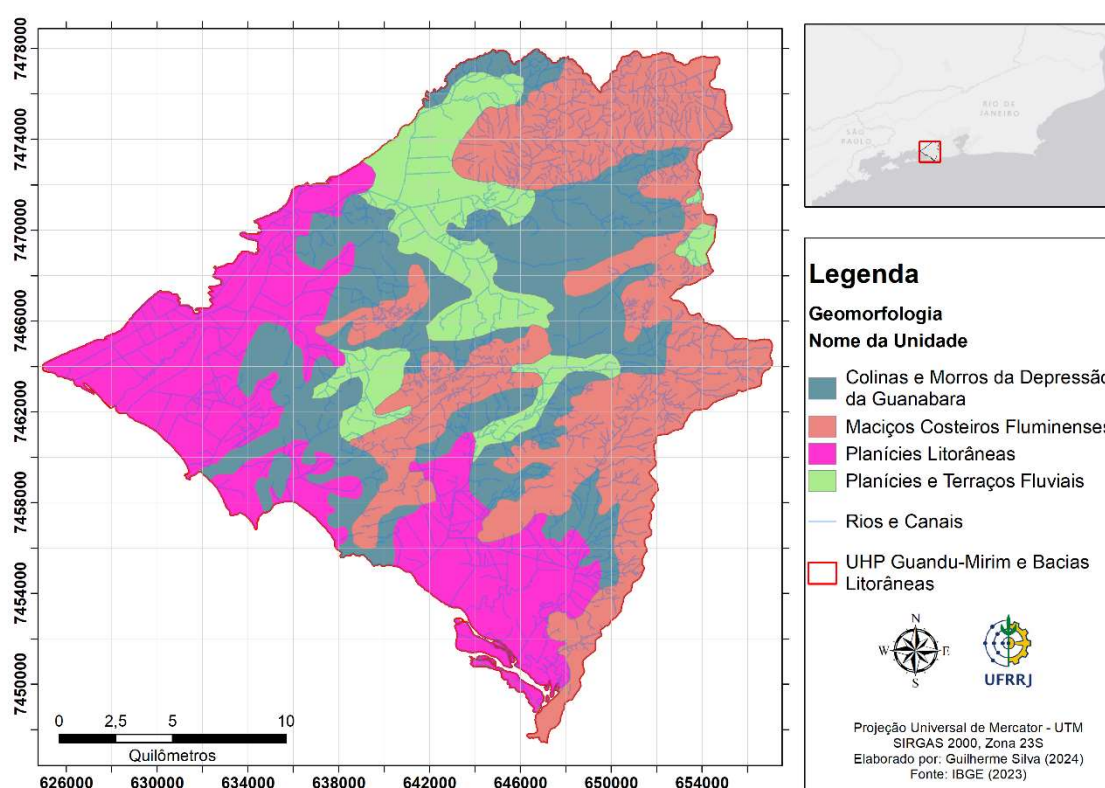


Figura 2- Mapa dos domínios geomorfológicos da área de estudo (IBGE, 2023)

2.1.3. Clima e vegetação

O clima é classificado como tropical úmido (Af, de acordo com o sistema de classificação de Köppen), com temperatura média anual variando entre 20°C e 27°C, e precipitação anual média entre 1000 e 2300 mm. As temperaturas e a precipitação são mais altas nas planícies e encostas da Serra do Mar, enquanto temperaturas mais baixas e períodos secos mais longos são observados nas áreas de bacia hidrográfica e no lado

oposto das cristas (área do reservatório de Lajes, Rio Claro e Pirai). A vegetação natural da região é composta por floresta ombrófila densa, vegetação arbustiva e formações de mangue, remanescentes do bioma Mata Atlântica (AGEVAP, 2018).

2.1.4. Geologia

O contexto geológico da área de estudo envolve a interseção entre dois grandes grupos, o primeiro sendo composto pelas rochas metamórficas e ígneas da porção oeste da Serra do Mar, com idades desde o Paleoproterozóico ao Cambriano superior. O segundo grupo é composto pelo mar de morros e coberturas sedimentares cenozóicas da baixada na zona oeste do Rio de Janeiro, como é possível observar no mapa geológico abaixo (Figura 3).

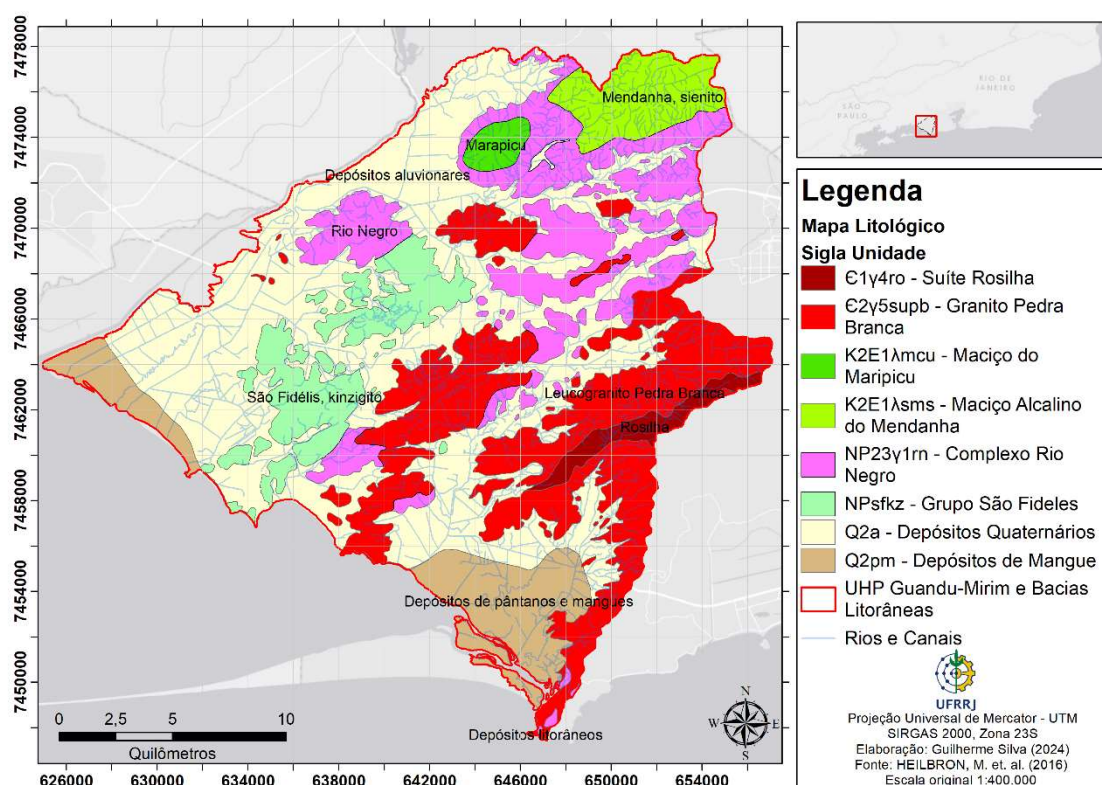


Figura 3-Mapa Litológico da Unidade Hidrológica de Planejamento do Rio Guandu-Mirim e Bacias Litorâneas - RJ (escala original 1:400.000). Adaptado de Heilbron et. al (2016)

As formações geológicas presentes na área são:

NPsfkz – Grupo São Fidélis, Gnaiss Kinzigítico

É uma unidade metassedimentar neoproterozóica da faixa ribeira no RJ, constitui a unidade basal do grupo e é composta predominantemente por um gnaiss kinzigítico, sendo um (sillimanita)-biotita-gnaiss granatífero, localmente com cordierita e grafita. (Heilbron et al., 2012)

NP23γ1rn - Complexo Rio Negro

Definido por Tupinambá et al. (1996) como um conjunto ortoderivado representante de um arco magmático neoproterozóico, com idade entre 630 e 600 Ma, constituído por ortognaisses e granitoides - diorito, tonalito, gnaiss, leucogranito e gnaiss porfíroide (Fernandes, 2001, Tupinambá et al., 2012).

€1γ4ro – Suíte Rosilha

A Suíte Rosilha é um conjunto de rochas granitoides que representa uma série de magmatismo sin a tardi colisional neoproterozóico. compreendendo um grupo de litotipos que variam composicionalmente desde dioritos (gabros) até granodioritos (granitos), sendo que os tipos tonalíticos/ granodioríticos são predominantes. Afloram, principalmente, nas zonas oeste e norte da cidade do Rio de Janeiro (Folha Baía da Guanabara 1:100.000), intrudindo as rochas do Domínio Costeiro, no Terreno Oriental (Esteves, 2004).

€2γ5supb – Granito Pedra Branca

Faz parte da suíte Suruí, sendo proveniente de um evento de magmatismo tardi a pós colisional. O Granito Pedra Branca é tipicamente representado pela fácies porfírica com megacristais (2 a 4 cm) de microclina em matriz monzogranítica de granulação média, onde o mineral máfico é a biotita, eventualmente hornblenda. Minerais acessórios incluem apatita, titanita, allanita, zircão e minerais opacos (Porto Jr., 1994, 2004)

K2E1λsms – Mendanha

O maciço é proveniente do magmatismo alcalino do neocretáceo-Eoceno, sendo formado por sienitos na sua porção central com traquitos na área ao redor (Mota 2012). Nas áreas periféricas ocorrem ainda brechas intrusivas. Brechas piroclásticas ocorrem distribuídas irregularmente ao longo do maciço, assim como brechas oligomíticas com clastos de rochas do embasamento (Klein, 1993; Valente et al., 2005)

K2E1λmcu - Marapicu

Também proveniente do magmatismo alcalino do neocretáceo-Eoceno o maciço do Marapicu é formado predominantemente de fácies sieníticas e nefelina sieníticas incluindo fácies tinguaíticas e traquitos. Os sienitos são equigranulares com textura variando de fina a grossa ocorrendo também variações porfiríticas com matriz de granulometria fina e com fenocristais de álcali feldspato e hornblenda (CPRM, 2016).

Q2a - Depósito Aluvionares Recentes

São os depósitos nas margens, fundos de canal, barras arenosas e planícies de inundação de rios, além de areias, cascalheiras, siltes e argilas provenientes dos processos de erosão/transporte e deposição a partir de áreas fontes diversas.

Q2pm – Depósitos de pântanos e mangues

Os depósitos de pântanos e mangues observados no estado do Rio de Janeiro são geralmente associados a sistemas lagunares ou estuarinos, como aqueles no fundo das baías da Guanabara, Sepetiba e da Ilha Grande. Nestes ambientes são também observadas planícies de marés que interagem com as porções da desembocadura dos baixos cursos dos canais fluviais.

2.1.5. Pedologia e amostragem de solos

Atualmente, o mapa pedológico (Carvalho Filho et al., 2003) que representa os solos da área de estudo está em uma escala de baixo detalhamento (1:250.000), o que é inadequado para o planejamento local. As unidades de levantamento de solos são compostas por associações e complexos de solos. Seguindo as montanhas e áreas de maior altitude nas porções leste e nordeste da área de estudo, as associações e complexos são predominantemente formados por Argissolos Vermelho-Amarelos e Argissolos Amarelos, com pequenas áreas de Cambissolos, Chernossolos, Latossolos, Luvisolos, Neossolos Litólico e Flúvico. As porções central e oeste da área de estudo apresentam complexos compostos principalmente por Gleissolos e Planossolos nas áreas mais baixas, com Argissolos presentes nas áreas de maior elevação. Na porção sudoeste, nas terras baixas costeiras, os complexos e associações são predominantemente compostos por

Gleissolos, frequentemente intercalados com Planossolos, Espodossolos, Organossolos e pequenas áreas de mangue (Figura 4).

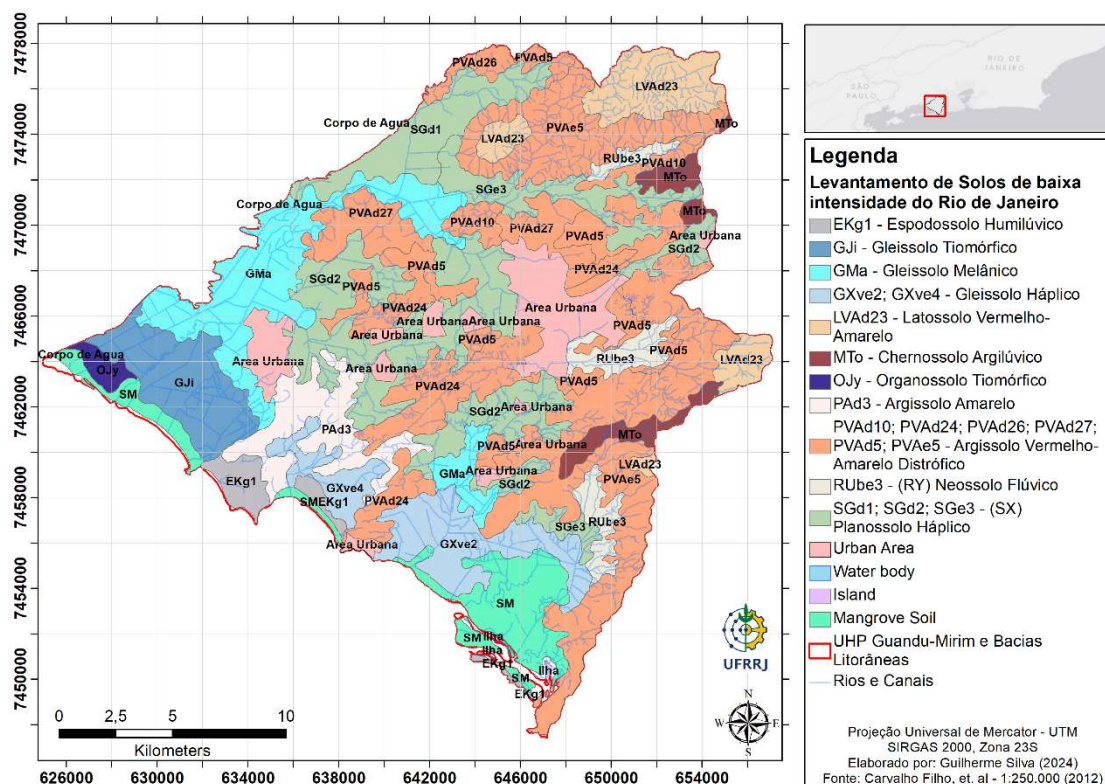


Figura 4– Mapa de Solos da UHP Guandu-Mirim e Bacias litorâneas, adaptado de Levantamento de reconhecimento de baixa intensidade dos solos do estado do Rio de Janeiro: Embrapa Solos, por Carvalho Filho, et. al – 1:250.000 (2003)

Tabela 1– Descrição das unidades de mapeamento presentes na área de estudo, reclassificado para o Sistema Brasileiro de Classificação de Solos de 2018(Santos et al, 2018), adaptado de Carvalho Filho, A. et. al (2003)

Legenda	Classificação SiBCS (2013)
EKg1	ESPODOSSOLO HUMILÚVICO + ESPODOSSOLO FERRI-HUMILÚVICO, ambos textura arenosa, A moderado ou A proeminente + NEOSSOLO QUARTZARÊNICO, A moderado ou A fraco, todos fase campo de restinga arbóreo-arbustiva, relevo plano (50-30-20%)
GJi	GLEISSOLO TIOMÓRFICO, A moderado, textura argilosa ou muito argilosa, fase campo halófilo de várzea, relevo plano
GMa	GLEISSOLO MELÂNICO, A húmico ou A proeminente + GLEISSOLO TIOMÓRFICO, ambos textura argilosa ou muito argilosa + GLEISSOLOS HÁPLICO, ambos textura argilosa, A moderado, todos fase campo tropical higrófilo de várzea, relevo plano (40-20-40%)
GXve2	GLEISSOLO HÁPLICO, textura argilosa ou média A moderado + GLEISSOLO MELÂNICO, textura argilosa ou muito argilosa, A húmico ou A chernozêmico, ambos fase campo tropical hidrófilo de várzea + GLEISSOLO TIOMÓRFICO, textura argilosa ou muito argilosa ou muito argilosa, A húmico ou A proeminente, fase campo halófilo de várzea, todos relevo plano (40-40-20%)
GXve4	GLEISSOLO HÁPLICO, textura argilosa fase campo tropical hidrófilo de várzea + PLANOSSOLO HÁPLICO, textura arenosa/argilosa fase floresta tropical subcaducifólia de várzea, ambos A moderado, fase relevo plano (50-50%)

LVAd23	LATOSSOLO VERMELHO-AMARELO + LATOSSOLO AMARELO, ambos textura argilosa ou muito argilosa + CAMBISSOLO HÁPLICO, textura argilosa ou média cascalhenta, todos A moderado, fase floresta tropical subperenifólia, relevo montanhoso e forte ondulado (40-30-30%)
MTo	CHERNOSSOLO ARGILÚVICO, fase pedregosa + LUVISSOLO CRÔMICO + ARGISSOLO VERMELHO-AMARELO, ambos A moderado, todos textura média/argilosa, fase floresta tropical subcaducifólia, relevo montanhoso e forte ondulado (50-30-20%)
OJy	ORGANOSSOLO TIOMÓRFICO + ORGANOSSOLO HÁPLICO + GLEISSOLO TIOMÓRFICO, textura argilosa ou média, todos fase campo tropical hidrófilo de várzea relevo plano (40-30-30%)
PAd3	ARGISSOLO AMARELO, textura média/argilosa ou média/muito argilosa + ARGISSOLO VERMELHO-AMARELO, textura média/argilosa, ambos A moderado, fase floresta tropical subcaducifólia, relevo suave ondulado e ondulado (60-40%)
PVAd5	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura média/argilosa ou média/muito argilosa, A moderado, fase floresta tropical subcaducifólia relevo suave ondulado e ondulado (50-50%)
PVAd10	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura média/argilosa ou média/muito argilosa, A moderado, fase floresta tropical subcaducifólia, relevo ondulado e suave ondulado (60-40%)
PVAd24	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura média/argilosa + NEOSSOLO LITÓLICO, textura média, todos A moderado, fase floresta tropical subcaducifólia, relevo forte ondulado e montanhoso (40-40-20%)
PVAd26	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura média/argilosa, fase relevo suave ondulado e ondulado + PLANOSSOLO HÁPLICO, textura arenosa/argilosa, fase relevo plano, todos A moderado, fase floresta tropical subcaducifólia (40-30-30%)
PVAd27	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura arenosa/argilosa ou média/argilosa + PLANOSSOLO HÁPLICO, textura arenosa/argilosa ou arenosa/muito argilosa, todos A moderado, fase floresta tropical subcaducifólia, relevo suave ondulado e plano (40-30-30%)
PVAe5	ARGISSOLO VERMELHO-AMARELO + ARGISSOLO AMARELO, ambos textura média/argilosa, A moderado, fase relevo montanhoso e forte ondulado + NEOSSOLO LITÓLICO, textura média, A moderado ou A chernozêmico, fase relevo montanhoso, todos fase não rochosa ou rochosa, floresta tropical subcaducifólia (40-40-20%)
RUbe3	NEOSSOLO FLÚVICO, textura média ou argilosa, fase floresta tropical subperenifólia de várzea + GLEISSOLOS HÁPLICO, textura argilosa ou muito argilosa, fase campo tropical higrófilo de várzea e floresta tropical subperenifólia de várzea, todos A moderado, fase relevo plano (60-40%)
SGd1	PLANOSSOLO HÁPLICO, textura arenosa/argilosa ou média/argilosa, fase floresta tropical subcaducifólia de várzea + GLEISSOLO HÁPLICO, textura argilosa ou muito argilosa, fase campo tropical higrófilo de várzea, ambos A moderado, fase relevo plano (70-30%)
SGd2	PLANOSSOLO HÁPLICO, textura arenosa/argilosa ou arenosa/média, A moderado, fase floresta tropical subcaducifólia de várzea, relevo plano
SGe3	PLANOSSOLO HÁPLICO, textura média/argilosa ou arenosa/argilosa, fase floresta tropical subcaducifólia de várzea + GLEISSOLO HÁPLICO, textura argilosa, fase campo tropical higrófilo de várzea, todos A moderado, fase relevo plano (80-20%)

Para análises iniciais e definição das áreas-alvo para campanhas de campo, foram utilizados dados de 32 perfis de solo do banco de dados legados do PronaSolos, e os mapas temáticos de geologia, geomorfologia, uso e cobertura dos solos e mapa de solos existente, seguindo a ideia de obter amostras de perfis de solo nas áreas que não possuíam amostras mantendo a importância da representatividade das amostras para a região como um todo (Carvalho Junior, 2014). Como é possível observar no mapa de uso e ocupação do solo (Figura 5) a área de estudo é intensamente povoada, sendo a atividade antrópica, a ocupação urbana, além da existência das áreas de risco na zona oeste do Rio de Janeiro, uma dificuldade a mais na obtenção de amostras em algumas áreas.

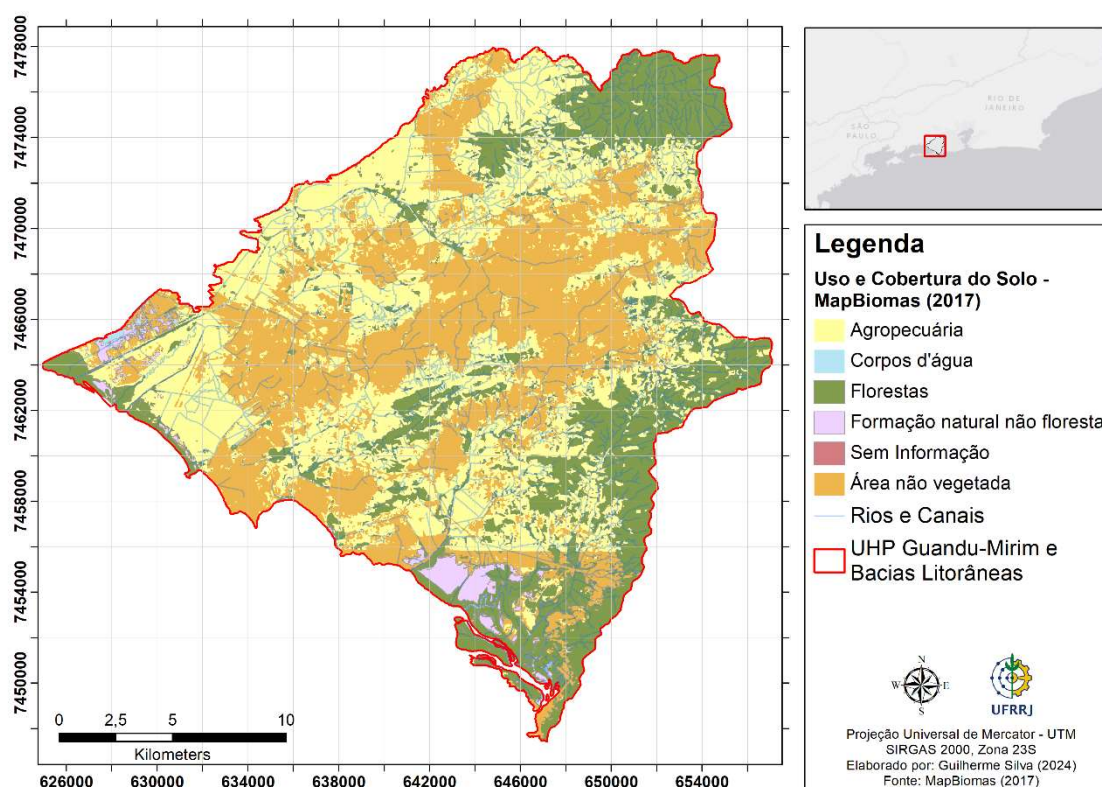


Figura 5 – Mapa de uso e ocupação do solo da Unidade Hidrológica de Planejamento do Rio Guandu-Mirim e Bacias Litorâneas (adaptado de MapBiomas, 2017)

Após a análise inicial, organização e tratamento dos dados legados, campanhas de campo (Figura 6) foram realizadas por uma equipe multidisciplinar entre os dias 13 e 28 de abril de 2023 e dados de 28 perfis de solo foram georreferenciados para integração ao banco de dados geral, totalizando 60 perfis de solo para a área de estudo.



Figura 6 – Algumas imagens do trabalho de campo realizado

Os exames de perfis de solo e as coletas de amostras foram realizados de acordo com Santos et al. (2005). Todas as amostras de solo foram preparadas e analisadas no Laboratório de Física do Solo do Departamento de Solos da UFRRJ (Figura 7), utilizando os métodos descritos por Embrapa (1997). As análises físicas e químicas incluíram: distribuição do tamanho das partículas em areia total, fina e grossa, silte e argila, pH em H₂O e KCl 1N, bases trocáveis (Ca^{2+} , Mg^{2+} , K^{+} , Na^{+}), soma de bases (valor S), saturação de bases (valor V), acidez extraível, alumínio trocável (Al^{3+}), capacidade de troca catiônica (valor T), saturação por alumínio e sódio, fósforo assimilável e carbono orgânico.



Figura 7 – Preparação das análises Química e Física das amostras de perfis de solo obtidas em campo, realizada no Laboratório de Física dos Solos, Departamentos de Solos – IA – UFRRJ

2.2. Contextualização do Mapeamento Digital de Solos (MDS)

A maior parte do território brasileiro possui apenas dados sobre os solos em nível exploratório (1:1 000 000 a 1:250 000) que não são suficientes para o planejamento local de atividades ligadas ao uso dos solos. Essa carência de mapas em escala adequada para subsidiar o planejamento, o zoneamento ambiental, a ocupação ordenada dos espaços, dentre outras atividades com uso do solo, gera uma lacuna de planejamento e gestão que tem levado ao surgimento de uma série de problemas ambientais nas áreas com maior pressão de uso e ocupação. São vários os motivos que dificultam a execução de um levantamento de solos em detalhe eficiente no país, mas se destacam as restrições orçamentárias e logísticas, a escassez de profissionais qualificados, a grande área e a própria variabilidade intrínseca dos ambientes de estudo, além do desmantelamento e fusão dos órgãos e instituições responsáveis pelo levantamento de solo no Brasil com outras instituições de atribuições generalizadas. O planejamento e a gestão de atividades ligadas ao uso dos solos, os projetos ambientais e os projetos agrícolas, necessitam de levantamentos de solo em escala de maior detalhe (Chagas, 2006, 2011; Menezes et al. 2011, 2013; Pinheiro, 2015; Cunha, 2013).

A recente valorização do conhecimento e da pesquisa sobre os solos vem da racionalização deste como recurso, mediante um sistema com crescente aumento da população global e necessidade de expansão das áreas agropecuárias, segundo Hartemink & McBratney (2008). Surge então a preocupação quanto ao uso racional dos recursos naturais, onde o mapeamento digital de solos tem se mostrado de extrema importância.

O uso das geotecnologias, associadas ao mapeamento digital de solos, passa a ser de grande importância para as pesquisas envolvendo os solos, devido a capacidade dessas tecnologias em lidar de maneira mais precisa e rápida com grandes quantidades de dados, fornecendo como produto, informações mais precisas e de melhor qualidade. O Mapeamento Digital de Solos (MDS) se utiliza de dados de sensores remotos, técnicas computacionais, como softwares e algoritmos para modelagem e análise espacial de dados, em conjunto com os trabalhos de um levantamento de solos tradicional, para melhorar os produtos das pesquisas acerca de solos, obtendo informações importantes sobre a paisagem para estimar a distribuição espacial das classes e atributos dos solos. (Crivellini, 2019, Lagacherie and McBratney, 2006; Minasny et al., 2003; Zhu, 2001;).

O uso dessa metodologia, confere uma natureza quantitativa aos produtos dos levantamentos de solos, permitindo a análise de erros associados e a extrapolação das correlações identificadas através de modelos matemáticos para inferir sobre a disposição espacial das classes e propriedades dos solos. O MDS torna-se então, uma ferramenta fundamental para o desenvolvimento de levantamentos de solos mais precisos, eficientes, rápidos, com melhor relação custo-benefício, que possuam natureza quantitativa permitindo a mensuração de erros associados (Chen et al., 2022; Minai et al., 2020; Taghizadeh-Mehrjardi et al., 2021).

As inovações tecnológicas das últimas décadas, como o uso do GPS, os Sistemas de Informação Geográfica – SIG, o advento dos sensores proximais e remotos, além dos dados fornecidos por Modelos Digitais de Elevação (MDE) (Odeha, 1994), trouxeram além de uma quantidade infinitamente maior de dados, também o desafio de analisar e interpretar todos esses dados (McBratney et. al, 2003; Minasny et al. 2002). Diversos autores sugerem que o levantamento de solos, passe por mudanças, incorporando algoritmos preditivos, integração de sistemas de informações geográficas e ferramentas de geoestatística, inovações tecnológicas, modelagem global e dinâmica da gênese de solos e uso de imagens de alta resolução, para a expansão dos horizontes de possibilidades da pesquisa de solos e pedometria.

A pedometria e o MDS ganham notoriedade neste contexto de carência de informações em nível de detalhe sobre os solos, pois possuem um caráter prático e eficiente para dar suporte as decisões sobre uso da terra, além da capacidade de lidar com grandes volumes de dados provenientes do grande avanço em tecnologia e poder de computação, sendo os novos desafios desenvolver e entender novas ferramentas e técnicas de Data Mining e Machine Learning para lidar com esses grandes volumes de dados e integra-los de maneira eficiente (Hastie et al., 2009; Minasny et al., 2003, 2013).

Conforme resumido por Lagacherie e McBratney (2006), "o Mapeamento Digital de Solos (DSM) é a criação e a população de sistemas espaciais de informação sobre solos, utilizando modelos numéricos que inferem as variações espaciais e temporais de tipos e propriedades do solo a partir de observações de solo e conhecimento de variáveis ambientais relacionadas". Essa abordagem numérica ou quantitativa foi testada nos anos 1990 por McKenzie e Austin (1993), que propuseram uma abordagem de correlação ambiental. Hudson (1992) argumenta que os levantamentos de solos são uma estratégia científica baseada no conceito dos fatores de formação do solo, aliados às relações solo-paisagem. O paradigma solo-paisagem tem sua origem na equação dos fatores de formação do solo (Eq.1) delineada por Dokuchaev (Glinka, 1927) e Hilgard (Jenny, 1941).

$$S = f(cl, o, r, p, t) \quad (Eq.1)$$

Essa conhecida equação identifica os cinco fatores de formação do solo. O solo é caracterizado como sendo o resultado em função do clima (cl), organismos (o), relevo (r), material de origem (p) e tempo (t). A equação implica que, ao observar mudanças em um ou mais desses fatores à medida que a paisagem é percorrida, é possível localizar com precisão os limites entre diferentes corpos de solo (Hudson, 1992).

O conceito de DSM foi posteriormente formalizado como uma estrutura para a criação de mapas digitais de solo, com base nos cinco fatores de formação do solo de Jenny, mas modificados na estrutura do Scorpan-SSPFe (Soil Spatial prediction function with spatially correlated erros ou Função de Predição Espacial do Solo com Erros espacialmente autocorrelacionados) (Eq. 2), por McBratney et al. (2003) para uma predição espacial quantitativa.

$$Sa \text{ ou } Sc = f(s, c, o, r, p, a, n) + e \quad (\text{Eq.2})$$

Nessa equação, Sa e Sc representam os Atributos e Classes de solo, respectivamente. O termo (s) refere-se à informação já existente do solo; (c) refere-se ao clima; (o) a organismos, vegetação, fauna ou atividade humana; (r) ao relevo; (p) ao material de origem; (a) ao fator idade ou tempo; (n) à posição espacial ou geográfica; e (e) aos erros residuais espacialmente correlacionados (Chen, et. al, 2022).

2.3. Dados Aero Geofísicos (DAG): Um Breve Resumo

A gamaespectrometria é o ramo da geofísica que estuda os fenômenos ligados a emissão natural de radiação através de raios gama por elementos radioativos. Existem mais de 50 isótopos radioativos diferentes na natureza, mas a maior parte é muito rara ou fracamente radioativa. Sendo assim, as principais fontes de radiação gama detectadas na superfície terrestre são provenientes da desintegração natural do potássio (^{40}K), e dos elementos das séries do Urânio (^{238}U) e do Tório (^{232}Th), presentes na composição da maioria das rochas. As primeiras aquisições radiométricas aéreas foram realizadas nos EUA, Canada e na antiga URSS em 1947, impulsionadas pela exploração de urânio, e na década de 60 começa a ser amplamente utilizada na exploração mineral e monitoramento ambiental (Ribeiro et.al, 2015).

Os dados Aero Gamaespectrométricos (DAG) são coletados por meio de sobrevoo da área utilizando um helicóptero ou um avião normalmente entre 60 e 100 metros de altitude do solo. A resolução dos dados Aero Gamaespectrométricos está diretamente ligada ao espaçamento das linhas de voo, que é definido normalmente levando-se em consideração um balanço entre o custo da aquisição e a resolução a ser obtida. Levantamentos regionais normalmente usam espaçamentos de linha de voo entre 200 e 400 metros, enquanto levantamentos mais detalhados costumam ter espaçamento entre 50 e 100 metros (Wilford, 2002)

A contagem da radiação é feita através de um equipamento chamado espectrômetro, que mede todos os raios gama dentro de uma janela energética que vai de 0,41 a 2,81 MeV. Dentro desse intervalo, cada elemento é associado a um canal, um

pedaço desse espectro, onde suas energias estão concentradas. Os raios gama emitidos pelo potássio (^{40}K), por exemplo, estão centradas no pico de energia de 1,46 MeV (Figura 8 – Tabela 2). Os isótopos de ^{238}U e ^{232}Th não emitem radiação gama, mas os isótopos produto dos seus decaimentos radioativos (^{214}Bi e ^{208}Tl , respectivamente) emitem radiação gama com picos de energia centrados em 1,76 e 2,61 MeV (Kaplan, 1964). Por isso as estimativas de Urânio e Tório são referidas na literatura como Urânio Equivalente e Tório equivalente. (Ribeiro et al, 2015)

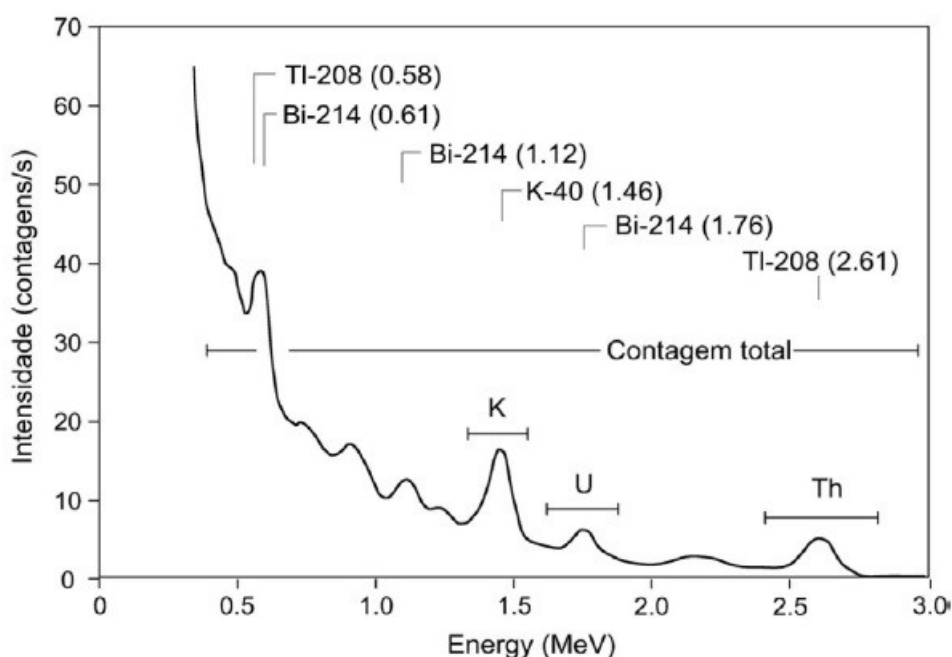


Figura 8 – Espectros de radiação gama mostrando as posições da janela de energia para cada elemento e da contagem total (Adaptado de: Ribeiro et al, 2015)

Tabela 2 – Tabela contendo os valores dos canais e dos picos de energia (em MeV) associados aos canais dos elementos radiométricos e a emissão cósmica. (Adaptado de: Ribeiro et al 2015)

Canal	Faixa de Energia (MeV)	Pico (MeV)
^{40}K	1,37 - 1,57	1,46
^{238}U	1,66 - 1,86	1,76
^{232}Th	2,41 - 2,81	2,61
CT	0,41 - 2,81	-
Cósmico	3,0 - 8,0	-

A radiação gama emitida por esses isótopos e seus equivalentes, tem a capacidade de penetração de cerca de 30 cm no solo ou rocha, ou de várias centenas de metros na atmosfera, por isso são a opção disponível para sensoriamento remoto (aéreo), mas vale a pena ressaltar que representam apenas uma camada superficial de cerca de 30 cm de solo (Minty, 1997). Os raios gama emitidos pela superfície terrestre estão diretamente ligados a mineralogia e a geoquímica do substrato rochoso subjacente, e também ao material fruto do intemperismo que é remobilizado através dos processos intempéricos (como por exemplo: solos, saprólitos, sedimentos aluviais e coluviais, etc.). Os processos naturais de intemperismo e erosão do solo, com o tempo, modificam a concentração dos radioelementos em comparação com o substrato rochoso original, resultando em diferentes comportamentos ambientais e levando a processos de redistribuição em várias escalas espaciais. As medições de raios gama, portanto, podem ajudar a diferenciar entre rocha matriz e solo, além de detectar a intensidade do intemperismo, as propriedades texturais e o estado nutricional do solo (Cook et al., 1996). O entendimento da resposta das rochas, regolitos e solos aos processos intempéricos e a remobilização desses radioelementos tem sem provado de muito valor tanto para o mapeamento de material regolítico, quanto para o entendimento dos processos geomorfológicos (Wilford, et. al, 1995, 1997).

O potássio, de maneira geral, tem uma diminuição da concentração com a evolução do intemperismo, devido em grande parte ao fato de o potássio ser altamente solúvel sob diferentes condições de intemperismo, sendo rapidamente lixiviado do perfil regolítico. A exceção a isso é quando o potássio é incorporado a argilas potássicas, como a illita. Em contraste, o Urânio e o Tório, estão associados a minerais resistentes e são retidos por óxidos de ferro no perfil de intemperismo. Portanto, a concentração de Urânio e Tório tende a permanecer a mesma ou aumenta preferencialmente nos materiais do regolito à medida que outros minerais mais solúveis são lixiviados (Wilford, 2002)

A concentração média na crosta terrestre de Potássio é de cerca de 2,3% e ele está presente em diversos minerais formadores de rocha, como os K-feldspatos, micas e argilas. Urânio e Tório possuem uma concentração média muito menor, cerca de 3 ppm e 12 ppm, respectivamente. Normalmente o Urânio e o Tório são associados a minerais acessórios e resistentes, como zircão, titanita, apatita, monazita e epidoto (Wilford, 1992). As concentrações de Potássio, Urânio e Tório nas rochas apresentam uma tendência a estarem correlacionados ao teor de sílica, quanto maior o teor de sílica maior a

concentração desses elementos. Essa correlação está refletida na tendência de se encontrar concentrações maiores desses elementos em rochas ácidas do que em rochas básicas. Altos teores de Potássio são normalmente associados a rochas ígneas ácidas, como granitos, riolitos e pegmatitos. O Urânio é frequentemente associado a pegmatitos, sienitos, granitos radioativos e alguns xistos. Enquanto o Tório é associado granitos, pegmatitos e gnaisses (Wilford, et. al, 2002).

Na Figura 9 são exibidos os mapas aero gamaespectrométricos utilizados para a modelagem, os dados foram obtidos a partir do projeto Aero Geofísica 1117 - Rio de Janeiro (CPRM, 2012) disponibilizados online de forma gratuita (<http://geosgb.cprm.gov.br/>) e processados no software Oasis Montaj v2012.1.

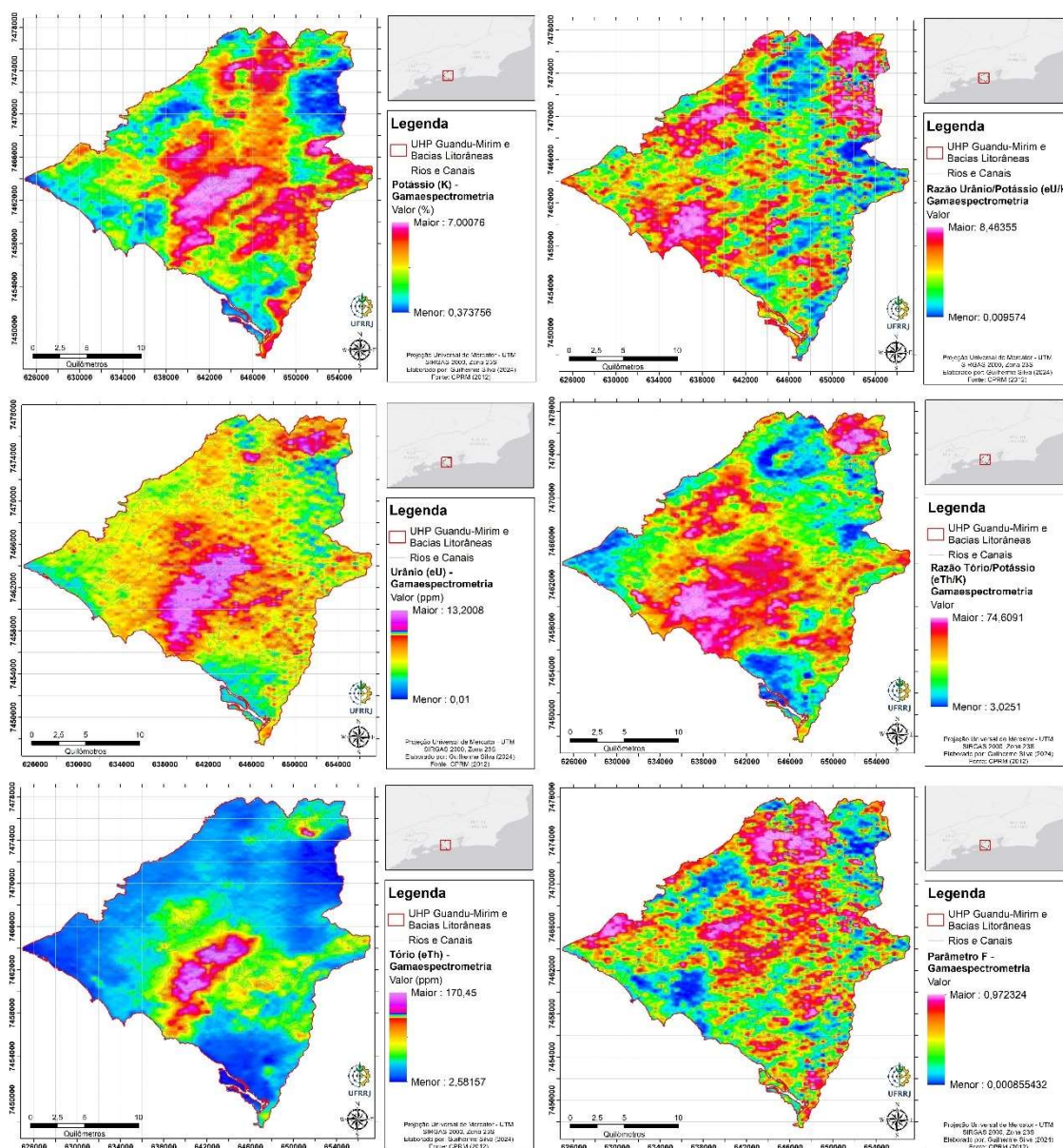


Figura 9 – Mapas gamaespectrométricos utilizados para a modelagem do presente trabalho

3.ARTIGO: “DIGITAL SOIL CLASS MAPPING USING TOPOGRAPHIC AND AIRBORNE GEOPHYSICAL DATA IN THE HIDROLOGICAL PLANNING UNIT OF GUANDU-MIRIN RIVER AND COASTAL BASINS, RJ.”

Este tópico se refere ao artigo de título “Digital Soil Class Mapping using topographic and airborne geophysical data in the Hidrological Planning Unit of Guandu-Mirim River and Coastal Basins” (Silva et al., 2024), fruto desta dissertação de mestrado e submetido à publicação na Revista Brasileira de Ciência dos Solos – RBCS, com coautoria de Helena S. K. Pinheiro, Waldir de Carvalho Junior, Blenda P. Bastos, Pedro A.M. Xavier e Joyce F. Alves.

3.1. Introduction

Soil mapping has the main function of describing the physical, chemical, and morphological characteristics, as well as the spatial distribution of different soil classes and their attributes within a specific area, providing support for better land-use planning. However, this activity is costly, time-consuming, and requires specialized labor, often resulting in outcomes lacking important information for modern users, such as the accuracy of maps, for example.

The use of geotechnologies has become highly important due to their ability to handle large volumes of data, thereby providing more precise and higher-quality information. Digital soil mapping (DSM) emerges in this context with the main purpose of enhancing soil surveys, utilizing remote sensing images, computational techniques such as software and algorithms for spatial data analysis and modeling, in conjunction with field mapping, to obtain information on important landscape characteristics for delineating soil classes and properties distribution (Lagacherie and McBratney, 2007). This method lends a quantitative nature to these soil surveys, enabling the analysis of associated errors and extrapolation of data on the spatial arrangement of soil classes and properties into mathematical models. In this sense, DSM becomes a key tool for the development of a more precise, efficient, rapid, better cost-benefit, and quantitative soil survey within the study area. (Chen et al., 2022; Minai et al., 2020; Taghizadeh-Mehrjardi et al., 2021).

DSM relies on the hypothesis of an existing correlation between landscape features(covariates) and soil formation factors based on jenny's soil formation factors model (Jenny, 1941), which then were upgraded to the so called SCORPAN model (McBratney et al. 2003). Chen et al. (2022) points out that relief, organisms and climate are the three most frequently used environmental covariates, while other factors such as parental material receive less attention in the literature. Recent studies (Bastos et al, 2023; Mello et al, 2021, 2022; Maino et al, 2022; Loiseau et al, 2020, 2021) have shown that geophysical data can be used to represent the parental material covariate from the SCORPAN model. So, the idea was to use airborne geophysical data (AGD) to represent this covariate, enhancing the quality of DSM providing further information about the soil's characteristics.

The goal of the study was the calibration and assessment of predictive models to represent the spatial variation of soil classes based on the characteristics of the Guandu-Mirim River watershed, by using topographic data and airborne geophysical data (AGD) as covariates.

3.2. Materials and Methods

3.2.1. Study area

The Guandu-Mirim River Watershed (Figure 10) is located on Rio de Janeiro state, and covers approximately 260 km², draining from the Mendanha Alkaline Complex, known as Guandu-do-Sena, formed by several springs, and changing names several times before flowing into Sepetiba Bay. The coastal basin area spans around 210 km², consisting of a series of small rivers and channels that eventually flow into Sepetiba Bay (AGEVAP, 2018). Official data provided by the Guandu Watershed Committee's online platform SIGA Web - Guandu (available at: www.sigaguandu.org.br) were used for representing political boundaries, access roads, hydrographic regions, watershed boundary of the hydrological planning unit (UHP).

The region's terrain is primarily characterized by hillslope areas and escarpments from the oceanic portion of the Serra do Mar (Highland's domain) in the southeast and northeast regions of the basin, with altitudes reaching up to 1000 m, and by the extensive

fluvial-marine plain (Lowland's domain) forming the Sepetiba Lowlands. Noteworthy in the lowlands area are the low hills and the extension of the fluvial plain of the Guandu River and Guarda River, with an average altitude below 60m (ANA, 2007).

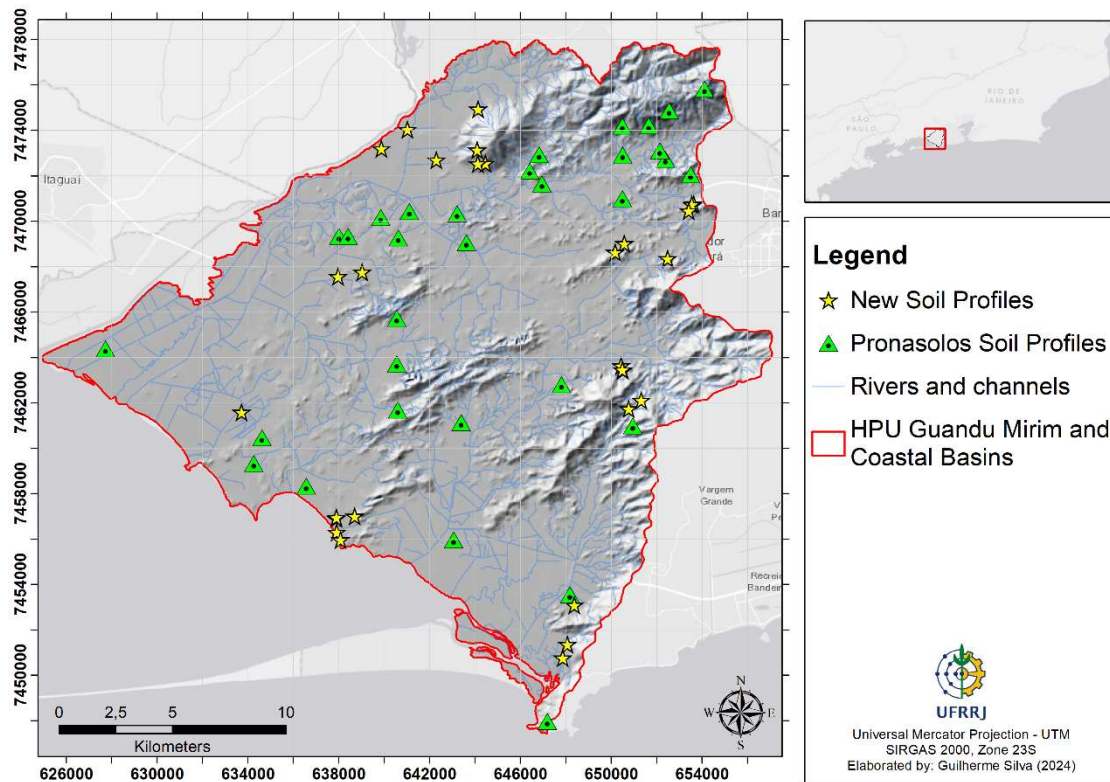


Figure 10 – Localization of the Hydrological Planning Unit Guandu-Mirim and Coastal Basins (AGEVAP, 2018) and the soil sampling sites inherited from Pronasolos and the new soil profiles

3.2.2. Pedology and soil sampling

Currently, the pedological map (Carvalho Filho et al., 2003) representing the soils of the study area is at a low-detail scale (1:250.000), which is inadequate for local planning. The soil survey units are composed of soil associations and complexes. Following the mountains and higher elevation areas in the eastern and northeastern portions of the study area, the associations and complexes are predominantly formed by acrisols, with small areas of cambisols, chernozems, ferrasols, luvisols, leptosols and fluvisols. The central and western portions of the study area present complexes primarily composed of gleyssols and planosols in the lower areas, with acrisols present in the higher elevations. In the southwestern portion, in the coastal lowlands, the complexes and

associations are predominantly composed of gleysols, which are frequently interspersed with planosols, podzols, histosols and small mangrove areas.

For initial analyses and to define the target areas for field campaigns, data from 32 soil profiles were utilized from the legacy database of PronaSolos, with the profile locations highlighted at Figure 10. The dataset includes morphological descriptions, physical and chemical analytic data, and classification according to the Brazilian Soil Classification System- SiBCS (Santos et al., 2018), which were transformed to World Reference Base for Soil Resources – WRB (IUSS Working Group WRB. 2022). Following initial analysis, organization, and treatment of the legacy data, field campaigns were conducted and data from 28 soil profiles were geo-referenced for integration with the general database, totaling 60 soil profiles.

Soil profile examinations and sampling were conducted following Santos et al. (2005). All soil samples were prepared and analyzed using methods described in Embrapa (1997). Physical and chemical analyses included: particle size distribution, pH in H₂O and KCl 1N, exchangeable bases (Ca²⁺, Mg²⁺, K⁺, Na⁺), sum of bases (S value), percentage base saturation (V value), extractable acidity, exchangeable aluminum (Al³⁺), cation exchange capacity (T value), percentage aluminum and sodium saturation, assimilable phosphorus and organic carbon.

A sample recollection technique called Region of Interest per Class – ROIC (Carvalho Junior, et. al, 2020) was adopted to amplify and balance the number of soil samples used to calibrate the algorithms. The method consists of the creation of irregular polygons around the sampling site, to represent the soil class. The placement of those polygons is made carefully respecting the main trend on the covariate's values, aiming to grasp the occurrence of different patterns for different soil classes, delineating areas of homogeneous pixel tonality around the sampling site, through visual inspection of each covariate layer, since the soil classification usually does not vary around the original sample (Figure 11). The covariates were analyzed one by one for the selection of each pixel. The 60 soil samples turned into 100 polygons, covering an area of approximately 1,3 Km², totalizing 1527 pixels being used, 1018 pixels being used to train the models and 509 pixels utilized to test the models (Table 3).

Table 3 – Soil units used, number of soil profiles in each unit and the total number of pixels gathered using ROIC

Soil Unit	Number of soil profiles	Number of Pixels
CX - Cambisols and Regosols	11	210
ES - Podzols	3	180
GX - Gleysols	5	198
GZ - Solonchaks + Gleysols +Histosols	7	180
LA - Ferrasols	3	180
PA - Acrisols + Planosols	21	213
PVA - Acrisols	9	186
RY - Fluvisols	1	180
TOTAL	60	1527

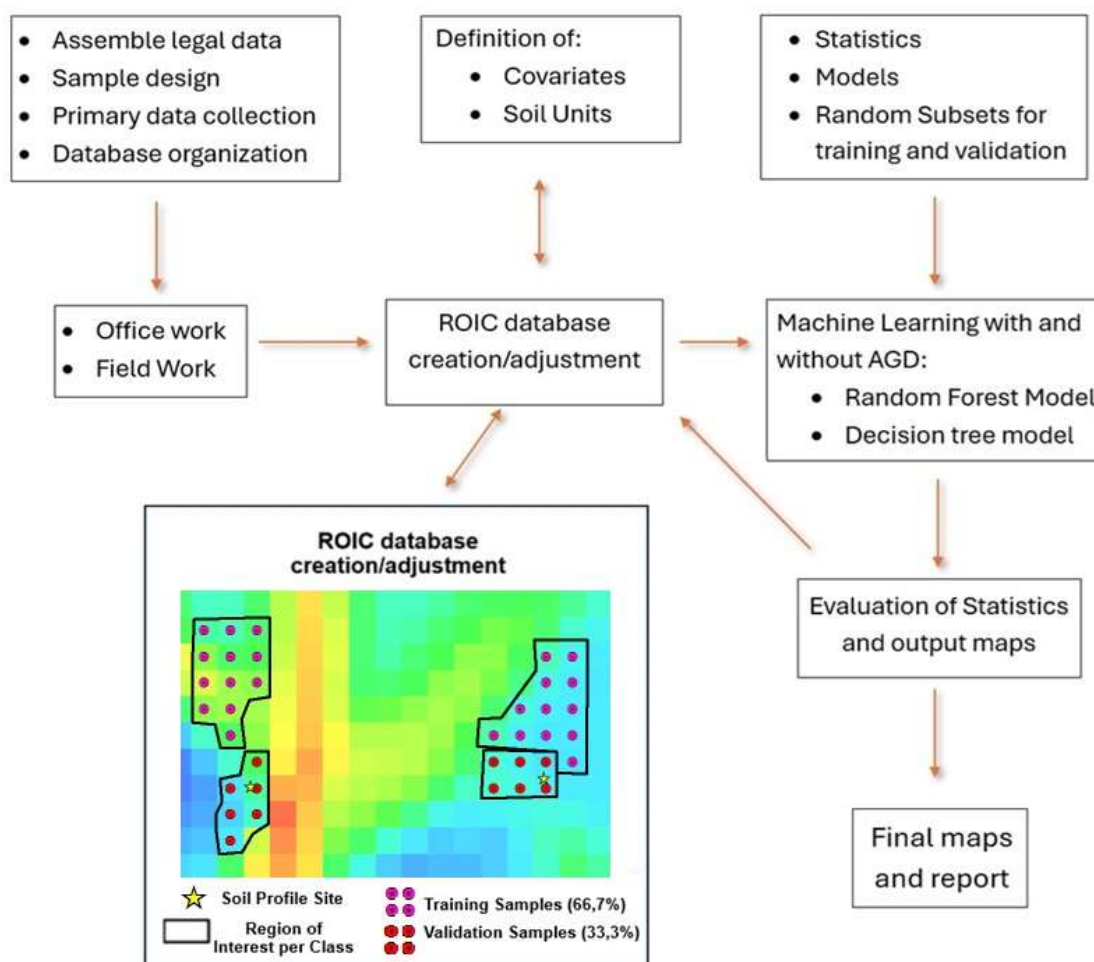


Figure 11 - Flow chart summarizing the methodology used in the study

3.2.3. Topographic and airborne geophysical data (AGD)

The initial step was to acquire and organize a database containing information from online platforms (IBGE, CPRM, BDGEx, SIGAWeb Guandu, Embrapa Solos - Geoinfo, for instance). Subsequently, a digital elevation model (DEM) was generated to derive topographic covariates used in the modeling procedures. Elevation data provided by the Brazilian Institute of Geography and Statistics (IBGE), in collaboration with the State Environmental Secretariat (SEA), at a 1:25.000 scale, comprising contour lines at 10m equidistance and elevation points, which combined with hydrography and the watershed boundary through TopotoRaster interpolator in ArcGIS Desktop v. 10.8. The DEM was generated with 30m spatial resolution, followed by a correction procedure aiming to fill the spurious depressions from the interpolation process. This operation aimed to obtain a hydrologically consistent DEM for deriving primary and secondary topographic attributes (Boehner et al., 2006; Hall and Olson, 1991; Moore et al., 1991, 1993), which was performed using the SAGA-GIS software (available in: www.saga-gis.org).

Using the hydrologically consistent DEM, over 60 topographic covariates were derived through SAGA-GIS terrain analysis tools, and then on a first selection some were excluded through visual inspection, mainly due to low variability in the study area. The 25 remaining covariates are listed in Table 4 with a brief description and a reference.

Table 4 - Topographic covariates used in the model construction, along with brief descriptions and references

Topographic Covariates	Description	Reference
Relative slope position	Evaluates the elevation of a point relative to the general slope of the surrounding terrain	Conrad (2015)
Slope	Quantifies the degree of slope at a specific point	Florinsky (2009); Olaya (2006)
Valley Depth	The difference between the elevation and an interpolated ridge level	Conrad (2012)
Slope Heigh	Height difference between the central point and neighboring points on a slope	Boehner. and Selige (2006)
Standard Heigh	Height difference between the central point and the average height of neighboring points in a specific area	Boehner and Selige (2006)

Normalized Height	Normalized index representing the difference between a point and its neighbors	Boehner and Selige (2006)
Plan Curvature	Rate of change of the slope of the terrain relative to the horizontal plane	Olaya (2006); Florinsky (2009)
Tangential Curvature	Rate of change of the slope of the terrain along a line tangent to the surface	Heerdegen and Beran (1982)
Upslope Curvature	Evaluates the curvature of the terrain surface relative to the upslope direction	Zevenbergen and Thorne (1987)
Downslope Curvature	Evaluates the curvature of the terrain surface relative to the downslope direction	Freeman (1991)
Profile Curvature	Measures the curvature of the terrain surface along the topographic profile	Evans (1979)
Local Curvature	Calculates the local curvature of a cell as sum of the gradients to its neighbor cells	Freeman (1991)
Longitudinal Curvature	Represents the curvature of the terrain surface along contour lines	Zevenbergen and Thorne (1987)
General Curvature	A general assessment of the curvature of the terrain surface in all directions	Bauer et al. (1985):
Channel Network Distance	indicates the proximity of each point on the terrain surface to the channel or drainage system	Conrad (2015)
Multiresolution Index of Valley Bottom Flatness (MRVBF)	Characterizes the flatness and concavity of the lower parts of valleys at different resolutions of analysis	Gallant and Dowling (2003)
Multiresolution index of the ridge top flatness (MRRTF)	Evaluates the flatness of ridge tops and hilltops at various scales of analysis	Gallant and Dowling (2003)
Vector Ruggedness Measure (VRM)	Quantifies terrain roughness in a landscape based on variations in slope orientations	Sappington et al. (2007)
Terrain Ruggedness Index (TRI)	Quantifies the roughness and vertical variation of the terrain in a given area	Riley et al. (1999)
Real Surface Area	Calculates the actual total area of a cell or three-dimensional object,	Olaya (2004)
LS Factor	Calculates the slope length factor (LS), as used in the Universal Soil Loss Equation (USLE), based on slope and specific catchment area	Desmet and Govers (1996)
Multi-Scale Topographic Position Index (TPI)	Difference between the elevation of a central point and the average elevations of surrounding points within a given window or radius, at various scales	Guisan et al. (1999)

Topographic Position Index (TPI)	An index that quantifies the elevation of a point relative to the average elevations of its neighborhood	Guisan et al. (1999)
Topographic Wetness Index (TWI)	Evaluates the terrain moisture in a given area based on topographic characteristics	Boehner and Selige (2006)
SAGA Wetness Index	Measures the relative moisture of the terrain based on topographic characteristics, using a modified catchment area calculation	Boehner et al. (2002)
Elevation	The vertical distance from the sea level	Conrad (2015)
Terrain Classification Index for Lowlands (TCI Low).	Classifies lowland areas based on their topographic characteristics, such as elevation, slope, and curvature of the terrain	Bock et al. (2007)

The radiometric data were provided by the Geological Survey of Brazil (CPRM), originating from the Aero geophysical Project 1117 – Rio de Janeiro (CPRM 2012, available at <http://geosgb.cprm.gov.br/>). According to the project's final report, the flight line spacing was 500m (N-S direction), control lines were spaced 10,000m apart (E-W direction), and the flight altitude was 100m. The data processing was conducted using Oasis Montaj (Geosoft) v.2021.2. First, minimum curvature (Briggs, 1974) interpolation was performed to generate primary gamma-ray variables: total count ($\mu\text{R/h}$), K (potassium), eU (uranium equivalent), and eTh (thorium equivalent). The ratios between elements, Factor F or F-parameter, were generated from those variables. Through visual inspection some were discarded, mainly due to low variability in the area, and the remaining used covariates are presented in Table 5. After processing, all geophysical raster images were resampled from 100 m to 30 m of spatial resolution in the Oasis Montaj software, using the Regrid tool to adapt them to the finer resolution of the DEM. These input data were used as covariates in the models to predict the spatial distribution of soil classes in the study area. Geophysical covariates have a stronger affinity with geology and relief factors, but they also relate to soil formation as they indicate variations in the texture of the relief, and the mineralogical and chemical characteristics of the rock or exposed soil at that point (Bastos et al., 2023; Reinhardt et al., 2019; Guimarães et al., 2009; Souza, 2005; Foote, 1968). By integrating these covariates with robust mathematical models, we can identify complex patterns and non-linear relationships between geophysics and the spatial distribution of soil classes.

Table 5 – Gamma Ray covariates, brief descriptions and references

Gamma-ray covariates	Description	Reference
Uranium equivalent (eU)	Measurement of gamma radiation emitted by Uranium	Ulbrich et al. (2009)
Thorium equivalent (eTh)	Measurement of gamma radiation emitted by Thorium	Ulbrich et al. (2009)
Potassium (percentage)	Measurement of gamma radiation emitted by Potassium	Ulbrich et al. (2009)
eTh/Potassium	Ratio between Thorium and Potassium data	Wilford et al. (1997)
eU/Potassium	Ratio between Uranium and Potassium data	Iza et al. (2018)
F Parameter or Factor	Ratio between potassium and uranium and thorium highlighting high potassium concentrations	Gnojek and Prichystal; (1985)

3.2.4. Predictive models

Predictive models, statistical analyses, and graphs were developed in R and R Studio (R Core Team, 2022). The predictive models were tested using the Classification Tree (CART), also called Decision Tree (DT), through the 'Rpart' package (Therneau et al., 2015), and Random Forest (RF) using the 'Random Forest' package (Liaw and Wiener, 2002).

For the DT and RF models using only the database with 60 soil samples available, with and without Airborne Geophysical Data (AGD) due to the small number of samples for the total area, it was used the method of Leave one out cross validation – LOOCV (Bishop et al., 2015; Carvalho Junior et. al, 2020), and a seed number (499) was defined randomly for later reproduction of the model. For the models using the ROIC technique, the samples were divided into three parts, so two thirds (66,67%) of the samples were used to train the model and one third (33,33%) were used to validate it.

For the DT-based method, for all 4 models tested, with and without the use of ROIC technique and with and without AGD, the complexity parameter adopted was (cp) = 0.006 for model calibration, the "class" method was used, and the deviance function was applied.

For the RF method, on the model training using LOOCV and AGD the parameters used for model training were mtry 8 and ntree 500. For the model using LOOCV and no

AGD, and the models using ROIC, with and without AGD, the parameters used for model training were mtry = 6 and ntree = 500.

3.3. Result and Discussions

3.3.1. Decision tree (DT) models

The DT model, using LOOCV without AGD, obtained an Overall Accuracy (OA) of 0,41 and Kappa 0,20. These low values indicate that this model's predictive power is weak and that the model was not able to learn meaningful patterns, probably because of the small number of soil profiles samples available to train the model. The 10 most important covariates were: Vector Terrain Measure, Slope Percentage, Elevation, LS Factor, Terrain Ruggedness Index, Surface Area, MRRTF, General Curvature, MRVBF and Upslope Curvature (Figure 12).

The DT model using LOOCV, with the AGD obtained OA of 0,36 and Kappa 0,18, values slightly lower than the model with no AGD, but statistically pretty much the same. These low values indicate that this model's predictive power is weak and that the model was not able to learn meaningful patterns, probably because of the small number of soil profile samples available to train the model. The 10 most important covariates to the model were elevation, vector terrain measure, MRRTF, LS factor, normalized heigh, Slope Percentage, Topographic Wetness, MRVBF, Uranium equivalent and Slope Heigh (Figure 12). Only 1 of the AGD covariates was amongst the 10 most important covariates to the model.

The DT model using ROIC and with no AGD obtained OA of 0,93 and Kappa 0,92 on the training of the model, suggesting that the model is very well-fitted to the training data and effectively capturing the complexity of the soil classes spatial distribution. The validation obtained OA of 0,76 and Kappa 0,73 to the model, values relatively high, but the drop in the values indicate that model is less accurate when predicting on new data and that the model may not generalize as well as expected.

But in general, this model using ROIC showed far better results than the ones using LOOCV. The 10 most important covariates to the model were: Standardized Heigh, Normalized Heigh, MRVBF, Elevation, Topographic Wetness, Valley Depth, Channel

Network Distance, SAGA Wetness, TCI Low Terrain Classification and MRRTF (Figure 12).

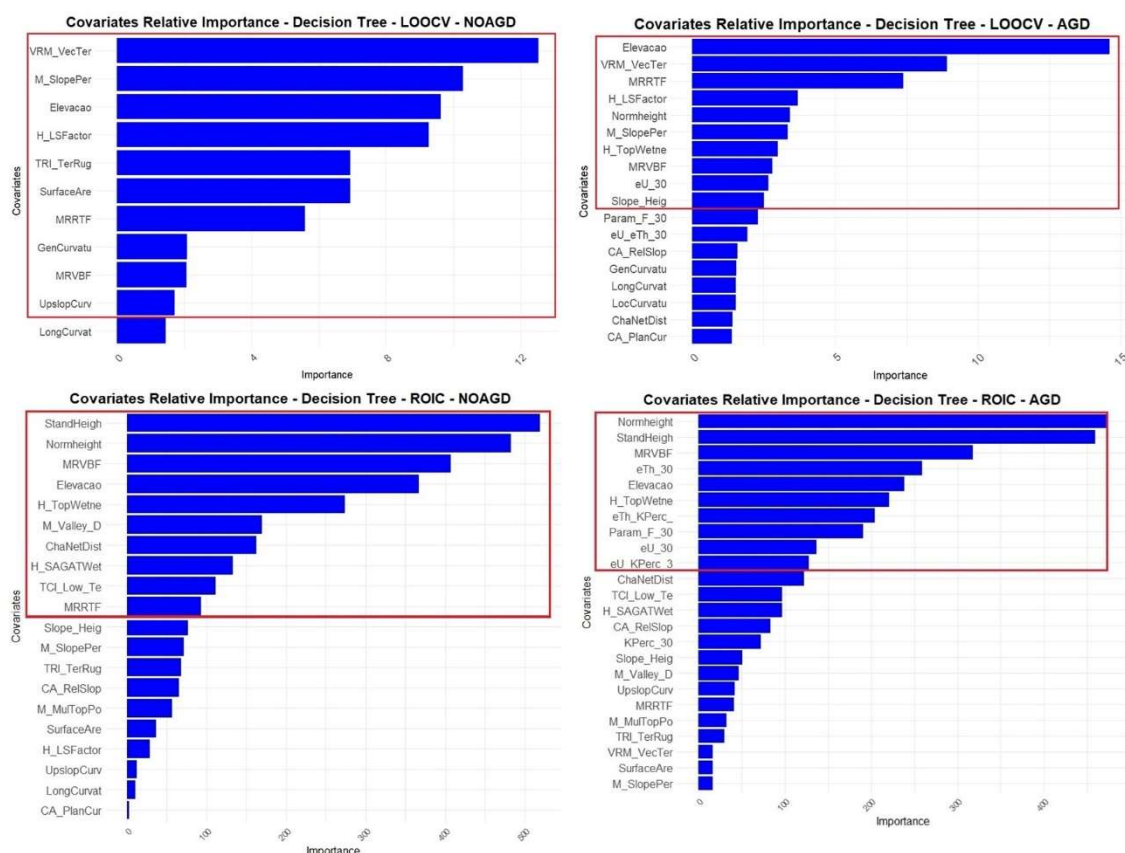


Figure 12 – Covariates Relative Importance for the Decision Tree Model using LOOCV and ROIC, with and with no AGD

The DT model using ROIC and AGD presented the highest values of OA and kappa, among the DT models. The OA value of 0,95 for training and 0,87 for validation, and Kappa of 0,94 for training and 0,85 for validation, suggest that the model has effectively captured the relationships between the covariates and the soil classes spatial distribution. The relatively small drop in performance from training to validation (OA drops by 0,08; and Kappa by 0,09) indicates that the model is likely well fitted and has learned general patterns rather than just memorizing the training data. This model seems to be robust and reliable when applied to unseen data, handling the complexity of the soil class spatial distribution well. The 10 most important covariates to the model were the following: Normalized Heigh, Standardized Heigh, MRVBF, Thorium, Elevation, Topographic Wetness, Thorium Potassium ratio, F Parameter, Uranium and Uranium Potassium Ratio (Figure 12). Five of the six covariates from AGD used were ranked amongst the 10 most important covariates to the model, only Potassium didn't make it,

but it showed some importance to the model. The result predictive soil class maps are presented in Figure 13.

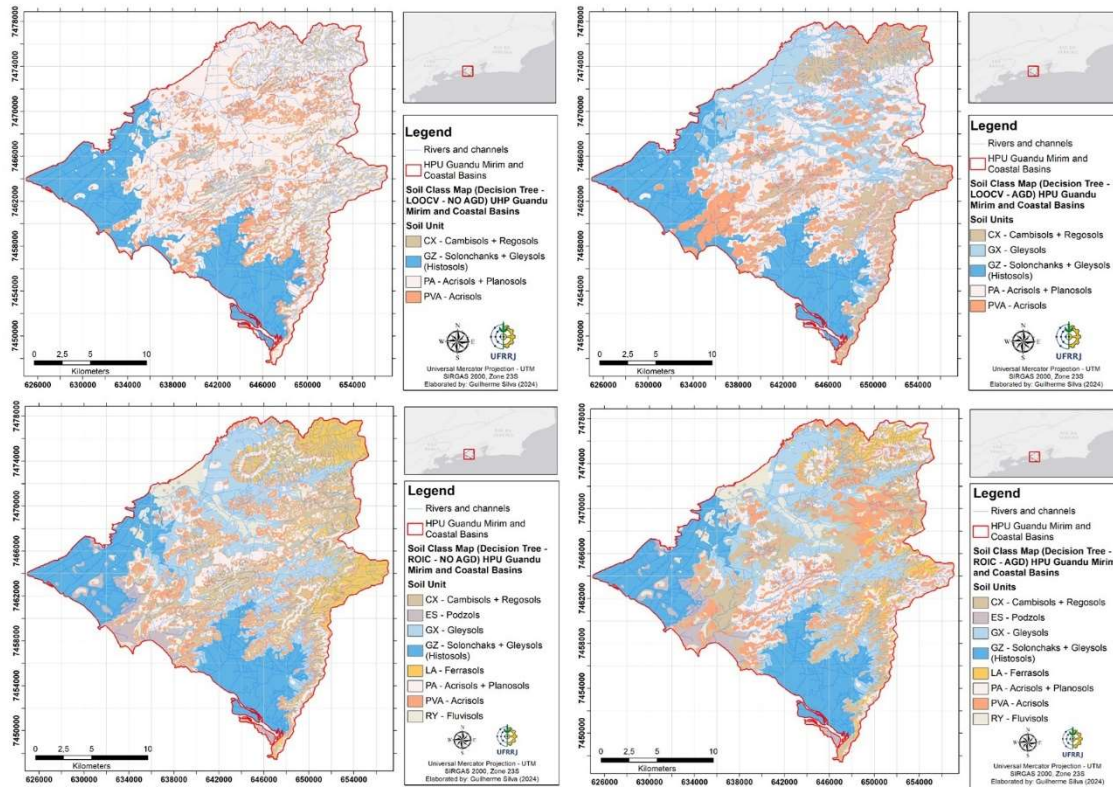


Figure 13 – Predictive Soil Class Maps obtained by the Decision Tree Model using LOOCV and ROIC, with and with no AGD

3.3.2. Random forest (RF) models

The RF model using LOOCV and no AGD, obtained OA of 0,36 and Kappa of 0,12. These low results show that the model is not well fitted, not predicting efficiently soil class spatial distribution. The covariates with the 10 highest Gini Importance were: Elevation, Standardized Heigh, Vector Terrain Measure, Upslope Curvature, MRRTF, LS Factor, Topographic Position Index, General Curvature, Topographic Wetness and Normalized Heigh (Figure 14). Since some soil classes (Ferrasol and Fluvisol) were represented by only one soil sample, even using the LOOCV the model was not capable of calculating the accuracy importance, only the Gini importance, which measures the purity of the node split created in the model construction process (Breiman, L. et. al., 1984), thus giving an idea of the covariate importance to the model.

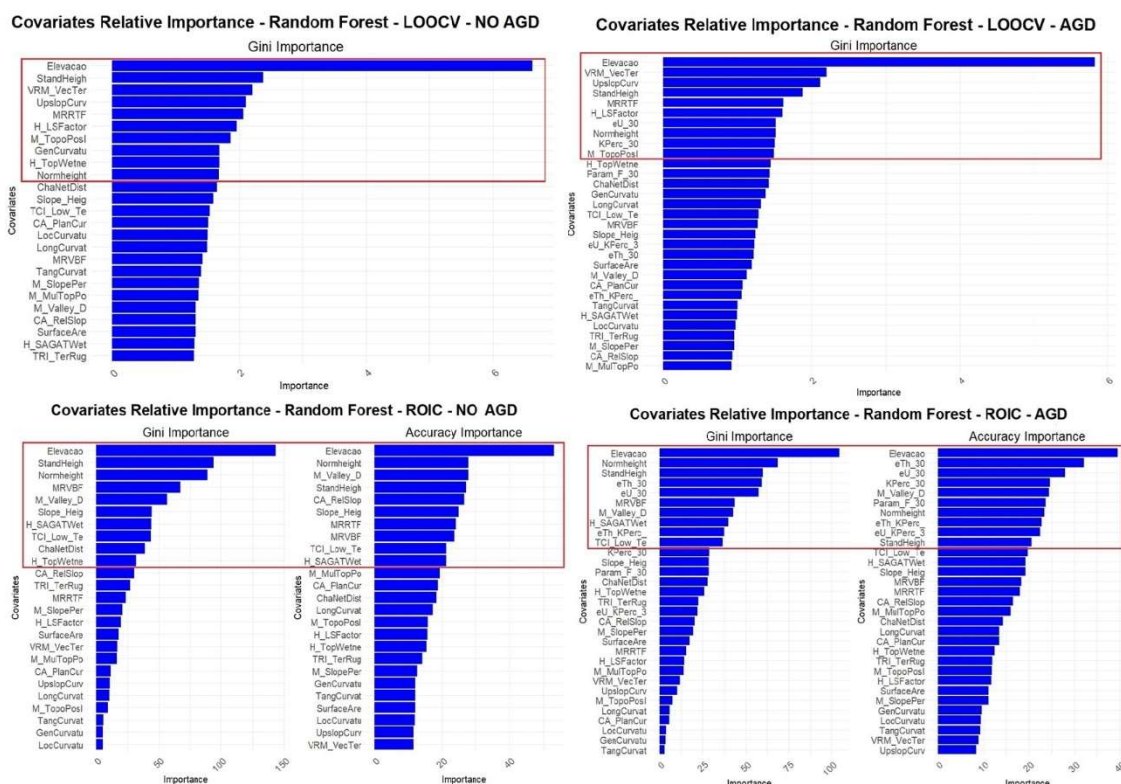


Figure 14 – Covariates Gini and accuracy Importance to the RF model using LOOCV and ROIC, with and with no AGD

The RF model using LOOCV and AGD obtained an OA value of 0,36 and Kappa of 0,11. Also low values suggesting that the model is not well fitted, not predicting efficiently the soil class spatial distribution. The covariates with the 10 highest Gini importance were Elevation, Vector Ruggedness Measure, Upslope Curvature, Standardized Heigh, MRRTF, LS Factor, Uranium, Normalized Heigh, Potassium, Uranium and Topographic Position Index (Figure 14). Since some soil classes were represented by only soil sample (Ferrasol and Fluvisol), even using the LOOCV the model was not capable of calculating accuracy importance, only the Gini Importance. However, the fact that Elevation had the highest Gini importance corroborates the importance of this covariate to the models. Two covariates from AGD (Uranium and Potassium) were ranked among the 10 highest Gini importance. The result of the predicted soil class maps are presented in Figure 15.

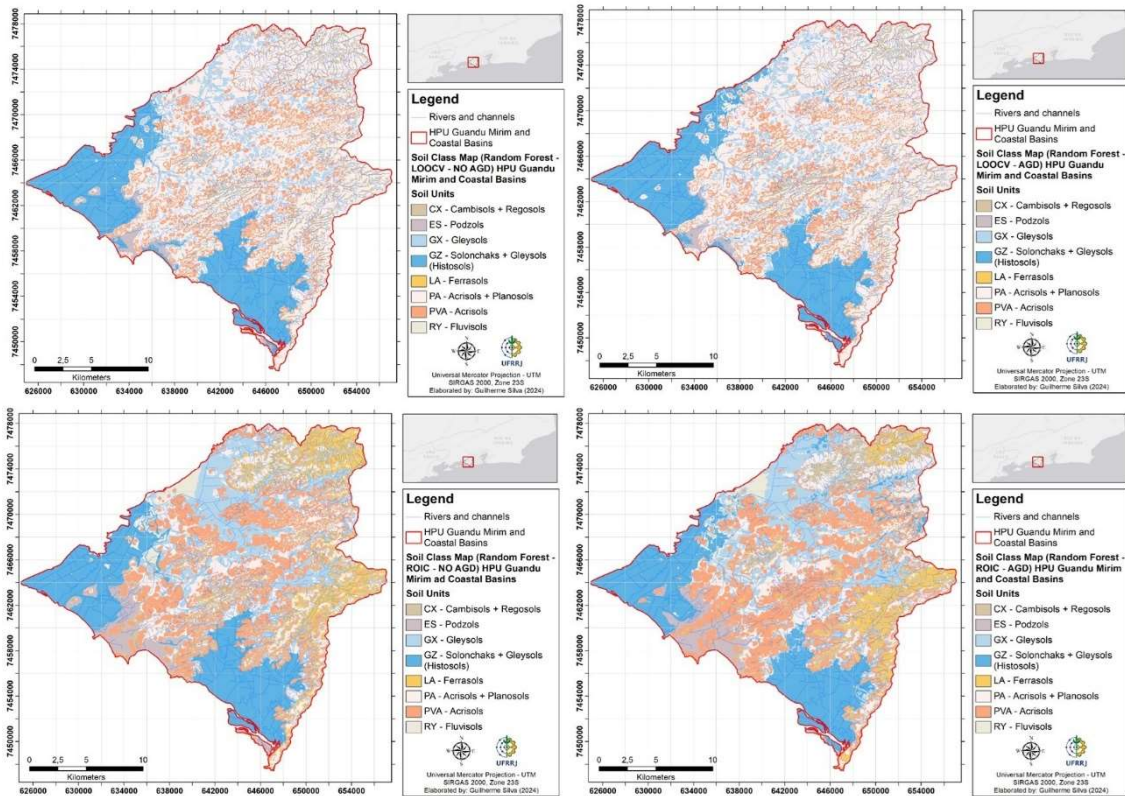


Figure 15 – Predictive soil class maps obtained by RF model using LOOCV and ROIC, with and without AGD

The RF model using ROIC and no AGD obtained OA of 0,93 for training and 0,76 for validation and Kappa of 0,92 for training and 0,73 for validation. Overall, the model performed well, especially on the training data. The small drop in performance from training to validation indicates that the model has some difficulty with new data, not generalizing very well, but still with a relatively high level of predictive capability on unseen soil data, however, there may be room for improvement. The 10 most important covariates to the model accuracy were: Elevation, Normalized Height, Valley Depth, Standardized Height, Relative Slope Position, Slope Height, MRRTF, MRVBF, Low Terrain Classification Index and SAGA Wetness (Figure 14). The covariates with the top 10 Gini importances were: Elevation Standardized Height, Normalized Height, MRVBF, Valley Depth, Slope Height, SAGA Wetness, Low Terrain Classification Index, Channel Network Distance and Topographic Wetness. The performance of this model was significantly higher than the models with LOOCV.

The RF model using the ROIC and AGD was the most robust and accurate model for soil class spatial distribution modeling between the tested models. The OA obtained for training was 0,99 and for validation 0,97. The Kappa for training was 0,99 and for

validation 0,97. These results indicate that the model is well fitted and that the model has learned the underlying patterns between the covariates and the spatial distribution of soil classes. The slight drop in validation performance shows that the model generalizes extremely well to new soil data.

The 10 most important covariates to the model accuracy were: Elevation, Thorium, Uranium, Potassium Valley Depth, F Parameter, Normalized Heigh, Thorium Potassium Ratio, Uranium Potassium Ratio and Standardized Heigh (Figure 15). All the six covariates from AGD used were ranked between the 10 highest accuracy importances, showing that the use of AGD can be used to enhance DSM quality and accuracy. The covariates with the 10 higher Gini Importance were: Elevation, Normalized Heigh, Standardized Heigh, Thorium, Uranium, MRVBF, Valley Depth, SAGA Wetness, Thorium Potassium Ratio and TCI Low Terrain Classification. The presence of all the six covariates from AGD in the 10 highest Gini Importance also shows how AGD can aid to reduce errors by increasing the purity of node splits. The result predictive map is presented, previously, in Figure 15; and the selected map to represent the variability of soil units, due the best performance, is presented in Figure 16.

3.3.3. Models' assessment and selection

In total 8 models were tested, 4 using DT and 4 using RF, each using and not using ROIC and Airborne Geophysical Data (AGD). The outcome results of the Overall Accuracy (OA) and Kappa Index of the models are shown below in Table 6.

Table 6 – Statistic results of the tested models

No Airborne Geophysical Data (AGD)						
LOOCV			ROIC			
			Training		Validation	
Modelo	Overall Accuracy	Kappa	Overall Accuracy	Kappa	Overall Accuracy	Kappa
DT	0,41	0,20	0,93	0,92	0,76	0,73
RF	0,36	0,12	0,98	0,98	0,90	0,89

With Airborne Geophysical Data (AGD)						
LOOCV			ROIC			
			Training		Validation	
Modelo	Overall Accuracy	Kappa	Overall Accuracy	Kappa	Overall Accuracy	Kappa
DT	0,36	0,18	0,95	0,94	0,87	0,85
RF	0,36	0,11	0,99	0,99	0,97	0,97

The selected model to represent the soil classes in the study area was the Random Forest model using Region of Interest per Class (ROIC) and Airborne Geophysical Data (AGD) (Figure 16) since it showed the best performance among the tested models.

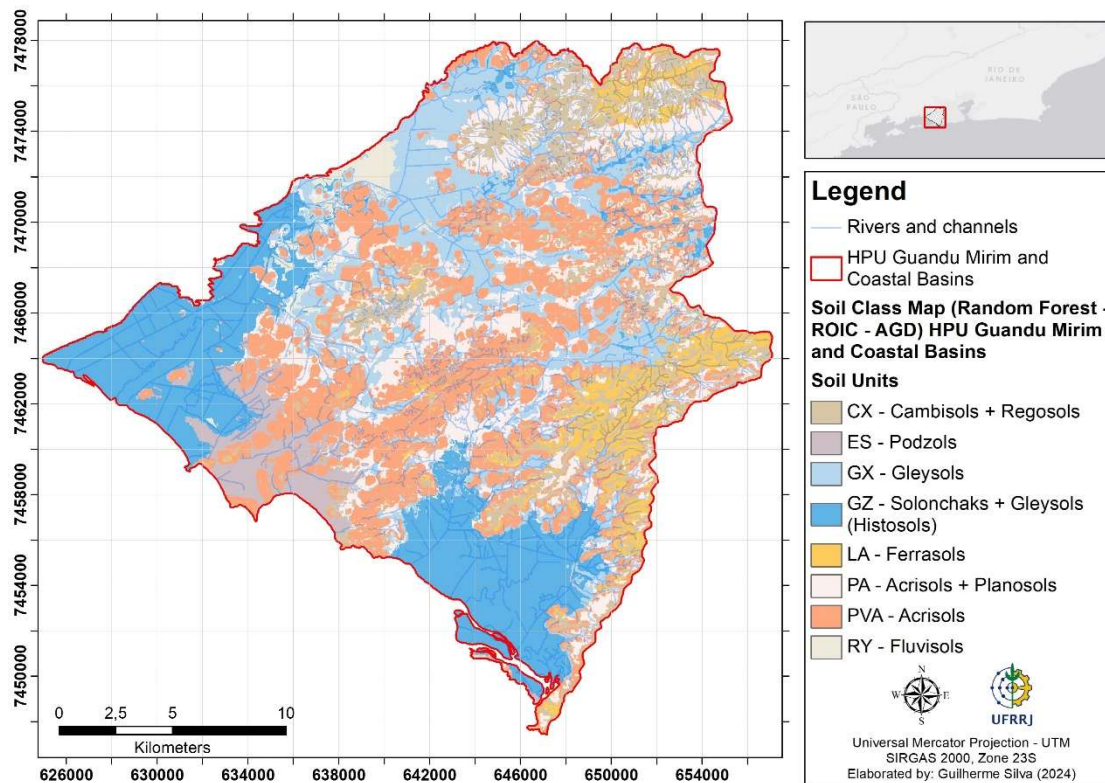


Figure 16 – Predictive soil class map obtained by RF model using ROIC and AGD which showed the best performance at the tests

3.4. Conclusions

The models using ROIC and Airborne Geophysical Data (AGD) presented the best results for both RF and DT models, and the RF models got a slightly better performance than the DT model. AGD were more efficient in improving the map quality when combined with the use of ROIC. This must be linked to the fact that a larger number of samples provide statistical analyses and more reliable interpolations, resulting in more precise predictions.

Elevation was amongst the 5 most important covariates to all the 8 models tested, highlighting the importance of an accurate digital elevation model (DEM) to Digital Soil Mapping.

For the DT models, with AGD, Uranium was ranked among the 10 highest importance to both models, using LOOCV and ROIC. Furthermore, on the DT model using ROIC, all 6 covariates from AGD shown importance to the model, reassuring its potential to improve digital soil mapping adding valuable information for spatial distribution of soil class prediction as a parental material covariate.

For the RF models, besides Elevation, Normalized Height and Standardized Height were present as some of the 10 most important covariates to all 4 models tested. Uranium was also present as one of the 10 most important covariates on both models (LOOCV and ROIC) using AGD, in fact Uranium was ranked amongst the 10 most important covariates to all 4 models tested with AGD (DT and RF). On the RF model using ROIC all 6 covariates from AGD made to the top 10 most important covariates to the model also reassuring its potential to improve digital soil mapping quality adding valuable information for soil class prediction as parental material covariate.

The combination of ROIC and AGD appears to provide very strong predictive power to the models, which demonstrates extremely high accuracy and reliability both during training and validation. Its minimal performance drop between training and validation is a sign of robust model training and exceptional generalization.

An accurate digital soil mapping survey is influenced by many factors, like the type and amount of data available, the scale, type and quality of the covariates obtained, time, budget, amongst others, but these results show that is possible to achieve a faster, more efficient, better cost-benefit, reproducible process, and which contains an error measurement, which is not possible on traditional soil mapping, since the model is formed in the specialist mind based on its knowledge. The soil specialist's knowledge will always be necessary to understand every step of the process, and to detect anomalies, errors, associations between covariates and to evaluate if the final maps generated are consistent.

Therefore, it is highly recommended that digital soil mapping techniques should be incorporated in every soil mapping campaign aiming to reduce time length, improve quality, provide error accountancy and more accurate, detailed and better cost-benefit soil maps suitable for local activities planning.

4. CONSIDERAÇÕES FINAIS

O presente estudo ressalta a contribuição positiva dos dados aero geofísicos para o MDS, em específico dos dados gamaespectrométricos. Esses dados desempenharam um papel complementar em relação às variáveis ambientais (topográficas) tradicionalmente utilizadas. Esses dados oferecem uma perspectiva mais abrangente ao considerar covariáveis representativas do material de origem, fator de formação do solo menos frequentemente utilizado no MDS. A realização de estudos adicionais se faz necessária, a fim de promover discussões mais aprofundadas, o que por sua vez contribuirá para uma compreensão mais aprimorada dessa abordagem e uma apreciação mais completa de suas limitações, especialmente no que se refere à correlação desses dados com as propriedades dos solos e distribuição espacial das classes de solo.

5. REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS – ANA (2007) Plano estratégico de recursos hídricos das bacias hidrográficas dos Rios Guandu, da Guarda e Guandu Mirim: relatório gerencial. Superintendência de Planejamento de Recursos Hídricos; elaboração, Sondotécnica Engenharia de Solos S. A. Brasília,DF 63 p. <https://www.ana.gov.br/arquivos/institucional/sge/CEDOC/Catalogo/2007/PlanoEstrategicoRHGuandu.pdf>

AGEVAP (2018) Relatório final do Plano Estratégico de Recursos Hídricos da Bacia do Rio Guandu, da Guarda e Guandu Mirim, PERH-Guandu, Comite Guandu, Apoio AGEVAP e Profill Engenharia e Ambiente S.A. <https://www.agevap.org.br/downloads/relatorios/execucao/guandu/2018/relatorio-de-execucao-2018.pdf>.

BASTOS, B.P.; PINHEIRO, H.S.K.; FERREIRA, F.J.F.; CARVALHO JUNIOR, W.d.; dos ANJOS, L.H.C. (2023) Could Airborne Geophysical Data Be Used to Improve Predictive Modeling of Agronomic Soil Properties in Tropical Hillslope Area? Remote Sens. 15, 3719. <https://doi.org/10.3390/rs15153719>.

BISHOP, T. F. A.; HORTA, A.; KARUNARATNE, S. B. (2015). Validation of digital soil maps at different spatial supports. *Geoderma*, 241, 238-249. <https://doi.org/10.1016/j.geoderma.2014.11.026>.

BOCK, M.; BOEHNER, J.; CONRAD, O.; KOETHE, R.; RINGELER, A. (2007): Methods for creating Functional Soil Databases and applying Digital Soil Mapping with SAGA GIS. In: Hengl, T., Panagos, P., Jones, A., Toth, G. [Eds.]: Status and prospect of soil information in south-eastern Europe: soil databases, projects and applications. EUR 22646 EN Scientific and Technical Research series, Office for Official Publications of the European Communities, Luxemburg, p.149-162. http://eusoils.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/esb_rr/EUR22646EN.pdf.

BÖEHNER, J.; MCCLOY, K.R.; STROBL, J. [eds.] (2006): SAGA - Analysis and Modelling Applications. Göttinger Geographische Abhandlungen, Vol.115, 130pp <https://doi.org/10.5194/gmd-8-1991-2015>.

BOEHNER, J.; SELIGE, T. (2006): Spatial Prediction of Soil Attributes Using Terrain Analysis and Climate Regionalisation. In: BOEHNER, J.; MCCLOY, K.R.; STROBL, J.: 'SAGA - Analysis and Modelling Applications', Goettinger Geographische Abhandlungen, Vol.115, p.13-27. <https://gmd.copernicus.org/articles/8/1991/2015/>.

BRIGGS, I.C. (1974) Machine contouring using minimum curvature. *Geophysics*, 39, 39–48. <http://dx.doi.org/10.1190/1.1440410>.

BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R.A.; & STONE, C.J. (1984). Classification and Regression Trees (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315139470>.

CARVALHO JUNIOR, W., CHAGAS, C. S., MUSELLI, A., PINHEIRO, H. S. K., PEREIRA, N. R., BHERING, S. B. (2014) Método do hipercubo latino condicionado para a amostragem de solos na presença de covariáveis ambientais visando o mapeamento digital de solos. *Revista Brasileira de Ciência do Solo*, 38(2), 386-396. <https://doi.org/10.1590/S0100-06832014000200003>.

CARVALHO JUNIOR, W.; LAGACHERIE, P.; CHAGAS, C. S.; CALDERANO FILHO, B.; BHERING, S. B. (2014) A regional-scale assessment of digital mapping of soil attributes in a tropical hillslope environment. *Geoderma* (Amsterdam), v. 232-234, p. 479-486.

https://ui.adsabs.harvard.edu/link_gateway/2014Geode.232..479D/doi:10.1016/j.geoderma.2014.06.007.

CARVALHO JUNIOR, W.; PEREIRA, N.R.; FERNANDES FILHO E.I.; CALDERANO FILHO, B.; PINHEIRO, H.S.K.; CHAGAS, C.S.; BHERING, S.B.; PEREIRA, V.R.; LAWAL, S. (2020) Sample design effects on soil unit prediction with machine: randomness, uncertainty, and majority map. *Rev Bras Cienc Solo* <https://doi.org/10.36783/18069657rbcs20190120>.

CARVALHO FILHO, A. de; LUMBRERAS. J. F.; WITTERN. K. P.; LEMOS, A. L.; SANTOS, R. D. dos; CALDERANO FILHO. B.; MOTHCI, E. P.; ITURRI LARACH, J. O.; CONCEIÇÃO, M. da; TAVARES. N. P.; SANTOS, H. G. dos; GOMES. J. B. V.; CALDERANO, S. B.; GONÇALVES, A. O.; MARTORANO, L. G.; SANTOS, L. C. de O; BARRETO, W. de O.; CLAESSEN, M. E. C.; PAULA, J. L. de; SOUZA, J. L. R. de; LIMA, T. da C.; ANTONELLO, L. L.; LIMA, P. C. de; OLIVEIRA, R. P. de; AGLIO, M. L. D. (2003) Levantamento de reconhecimento de baixa intensidade dos solos do estado do Rio de Janeiro. Rio de Janeiro: Embrapa Solos (Rio de Janeiro, RJ). Escala 1:250.000. (Embrapa Solos. Boletim de Pesquisa e Desenvolvimento. 32). <http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1037806>.

CHAGAS, C. DA S. (2006) Mapeamento digital de solos por correlação ambiental e redes neurais em uma bacia hidrográfica de domínio de mar de morros. Tese de Doutorado. UFV, MG. 239p. <https://locus.ufv.br/items/aa6f09e9-654d-4ce8-8c32-e70d832c8f26>.

CHAGAS, C. S., CARVALHO JÚNIOR, W., BHERING, S. B. (2011) Integração de dados do Quickbird e atributos do terreno no mapeamento digital de solos por redes neurais artificiais. *R. Bras. Ci. Solo*, v.35, p.693-704. <https://doi.org/10.1590/S0100-06832011000300004>.

CHEN, S.; ARROUAYS, D.; MULDER, V.L.; POGGIO L.; MINASNY, B.; ROUDIER, P.; LIBOHOVA, Z.; LAGACHERIE, P.; SHI Z.; HANNAM, J.; MEERSMANS, J.; RICHER-DE-FORGES, A. C; WALTER, C. (2022) Digital mapping of GlobalSoilMap soil properties at a broad scale: A review, *Geoderma*, Volume 409, 115567, ISSN 0016-7061, <https://doi.org/10.1016/j.geoderma.2021.115567>.

CONRAD, O.; BECHTEL, B.; BOCK, M.; DIETRICH, H.; FISCHER, E.; GERLITZ, L.; WEHBERG, J.; WICHMANN, V.; BÖEHNER, J. (2015): System for Automated

GEOSCIENTIFIC ANALYSES – SAGA (2015) v. 2.1.4, Geosci. Model Dev., 8, 1991-2007, <https://doi.org/10.5194/gmd-8-1991-2015>.

COMPANHIA DE PESQUISA DE RECURSOS MINERAIS – CPRM – Serviço Geológico do Brasil (2001) Geologia do Estado do Rio de Janeiro: texto explicativo e mapa geológico do Estado do Rio de Janeiro. Departamento de Recursos Minerais – DRM. Brasília; 2 ed. Disponível em: https://rigeo.sgb.gov.br/jspui/bitstream/doc/18458/1/mapa_geologico_rio_janeiro.pdf

COMPANHIA DE PESQUISA DE RECURSOS MINERAIS - CPRM – Serviço Geológico do Brasil (2016) - Mapa Geológico e de recursos minerais do Estado do Rio de Janeiro. Escala 1:400.000. Disponível em: <http://rigeo.cprm.gov.br/jspui/handle/doc/18458>.

COOK, S. E.; CORNER, R. J.; GROVES, P. R.; GREALISH, G. J. (1996) Use of airborne gamma radiometric data for soil mapping. Aust. J. Soil. Res. 34, 183–194. <https://doi.org/10.1071/SR9960183>

CRIVELENTI, R. C.; COELHO, R. M.; ADAMI, S. F. & OLIVEIRA, S. R. de M. (2009) Mineração de dados para a inferência de relações solo-paisagem em mapeamentos digitais de solo. Rev. Agro. Bras. 44:1707-1715. <https://doi.org/10.1590/S0100-204X2009001200021>.

CUNHA, A. de M. (2013) Seleção de Variáveis Ambientais e de Algoritmos de Classificação para Mapeamento Digital de Solos. Tese de Doutorado. UFV. 132p. <https://locus.ufv.br/items/e2433be9-45a8-4189-8f5c-4b6dab9c9dd7>.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. EMBRAPA. (1995) Centro Nacional de Pesquisa de Solos. Procedimentos normativos de levantamentos pedológicos. Brasília, DF: EMBRAPA – SPI, 101 p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. EMBRAPA (1997) Centro Nacional de Pesquisa de Solos. Manual de métodos de análise de solo. Rio de Janeiro: Centro Nacional de Pesquisa de Solos-CNPS. 212p. (Embrapa-CNPS. Documentos. 1).

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. EMBRAPA (1999) Centro Nacional de Pesquisa de Solos. Sistema Brasileiro de Classificação de Solos. Rio

de Janeiro. 412p. <https://www.agroapi.cnptia.embrapa.br/portal/assets/docs/SiBCS-2018-ISBN-9788570358004.pdf>.

ESTEVEZ, A.A. (2004) Evolução tectônica do Domínio Costeiro no Maciço da Pedra Branca, Cidade do Rio de Janeiro, RJ. Dissertação de Mestrado, UERJ Rio de Janeiro, 2004. 92p.

FLORINSKY, I.V. (2009): Computation of the third order partial derivatives from a digital elevation model. *International Journal of Geographical Information Science*, 23(2), p.213-231. <https://doi.org/10.1080/13658810802527499>.

FOOTE, R.S. (1968) Improvement in airborne gamma radiation data analyses for anomalous radiation changes. In: *Simp. Sobre uso de Técnicas Nucleares na prospecção e desenvolvimento de Recursos Minerais. Anais...Buenos Aires*. https://inis.iaea.org/collection/NCLCollectionStore/_Public/44/058/44058725.pdf?r=1

FREEMAN, G.T. (1991): Calculating catchment area with divergent flow based on a regular grid. *Computers and Geosciences*, 17:413-22. *ScienceDirect*. [https://doi.org/10.1016/0098-3004\(91\)90048-I](https://doi.org/10.1016/0098-3004(91)90048-I).

GALLANT, J.C.; DOWLING, T.I. (2003): A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39/12:1347-1359. Wiley. <https://doi.org/10.1029/2002WR001426>.

GLINKA, K.D. (1927) The great soil groups of the world and their development. (transl. from german by C.F. Marbut.) Edwards Bros., Ann Arbor, MI. <https://reader.library.cornell.edu/docviewer/digital?id=chla3055800#mode/1up>

GNOJEK, I., PRICHYSTAL, A. (1985) A new zinc mineralization detected by airborne gamma-ray spectrometry in Northern Moravia (Czechoslovakia). *Geoexploration*, 23 (4), 491-502. [https://doi.org/10.1016/0016-7142\(85\)90076-6](https://doi.org/10.1016/0016-7142(85)90076-6).

GUIMARÃES, G. B.; FRUCHTING, A. (2009) Levantamentos gamaespectrométricos em granitos diferenciados. I: Revisão da metodologia e do comportamento geoquímico dos elementos K, Th e U. *Geologia USP. Série Científica*, 9(1), 33-53. <https://doi.org/10.5327/Z1519-874X2009000100003>.

GUISAN, A.; WEISS, S.B.; WEISS, A.D. (1999): GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143: 107-122. <https://doi.org/10.1023/A:1009841519580>.

HARTEMINK, A.E.; MCBRATNEY, A.B. (2008) A soil science renaissance. *Geoderma*, v.148, n.2, p.123-129. Disponível em: www.sciencedirect.com/science/article/pii/S0016706108002802.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J.; (2009) The elements of statistical learning: data mining, inference and prediction. Springer Series in Statistics. Springer-Verlag, New York. <https://www.sas.upenn.edu/~fdiebold/NoHesitations/BookAdvanced.pdf>

HEERDEGEN, R.G.; BERAN, M.A. (1982): Quantifying source areas through land surface curvature. *Journal of Hydrology*, Vol.57. [https://doi.org/10.1016/0022-1694\(82\)90155-X](https://doi.org/10.1016/0022-1694(82)90155-X).

HEILBRON, M.; EIRADO, L. G.; ALMEIDA, J. (2016) Mapa geológico e de recursos minerais do estado do Rio de Janeiro. Belo Horizonte: CPRM. Escala 1:400.000

HUDSON, B.D. (1992) The soil survey as paradigm-based science. *Soil Science Society of America Journal* 56, 836–841. <https://doi.org/10.2136/sssaj1992.03615995005600030027x>

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE (2002). Mapa de Clima do Brasil [pdf]. 1:5.000.000. http://geoftp.ibge.gov.br/informacoes_ambientais/climatologia/mapas/brasil/Map_BR_clima_2002.pdf

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE (2009) Manual Técnico de Geomorfologia. Coordenação de Recursos Naturais e Estudos Ambientais. Rio de Janeiro, RJ: IBGE. <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?id=281612&view=detalhes>

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE (2023) Base Vetorial de Geomorfologia do Estado do Rio de Janeiro 1:250.000. Coordenação de Recursos Naturais e Estudos Ambientais. Rio de Janeiro, RJ: IBGE. https://geoftp.ibge.gov.br/informacoes_ambientais/geomorfologia/vetores/escala_250_mil/versao_2023/

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE (2018). Base Cartográfica Vetorial Contínua do Estado do Rio de Janeiro [shp] 1:25.000. Consultado em 07/2023. <https://www.ibge.gov.br/geociencias/downloads-geociencias.html>

INSTITUTO ESTADUAL DO AMBIENTE – INEA. (2012) Bacia Hidrográfica dos Rios Guandu, da Guarda e Guandu-Mirim/Comitê da Bacia Hidrográfica Guandu/ organizado por Décio Tubbs Filho, Julio Cesar Oliveira Antunes, Janaina Silva Vettorazzi. - Rio de Janeiro: INEA. comiteguandu.org.br/conteudo/livroguandu2013.pdf.

IUSS WORKING GROUP WRB (2022) World Reference Base for Soil Resources. International soil classification system for naming soils and creating legends for soil maps. 4th edition. International Union of Soil Sciences (IUSS), Vienna, Austria. [WRB_fourth_edition_2022-12-18.pdf](http://www.isric.org/WRB_fourth_edition_2022-12-18.pdf) (isric.org).

IZA, E.R.H.F.; HORBE, A.M.C.; CASTRO, C.C.; HERRERA, I.L.I.E. (2018) Integration of Geochemical and Geophysical Data to Characterize and Map Lateritic Regolith: An Example in the Brazilian Amazon. *Geochem. Geophys. Geosyst.* 19, 3254–3271. <https://doi.org/10.1029/2017GC007352>.

JENNY, H. (1994) *Factors of Soil Formation: A System of Quantitative Pedology*; Dover: New York, NY, USA, ISBN 978-0-486-68128-3.

LAGACHERIE, P.; McBRATNEY, A.B. (2006) Chapter 1 Spatial Soil Information Systems and Spatial Soil Inference Systems: Perspectives for Digital Soil Mapping. In *Developments in Soil Science*; Elsevier: Amsterdam, The Netherlands, Volume 31, pp. 3–22, ISBN. 978-0-444-52958-9. [http://dx.doi.org/10.1016/S0166-2481\(06\)31001-X](http://dx.doi.org/10.1016/S0166-2481(06)31001-X).

LAGACHERIE, P.; VOLTZ, M. (2000) Predicting soil properties over a region using sample information from a mapped reference area and digital elevation data: a conditional probability approach. *Geoderma*, v.97. p.187-208. [https://doi.org/10.1016/S0016-7061\(00\)00038-0](https://doi.org/10.1016/S0016-7061(00)00038-0).

LAGACHERIE, P., ROBBEZ-MASSON, J. M., NGUYEN-THE, N., & BARTHÈS, J. P. (2001) Mapping of reference area representativity using a mathematical soilscape distance. *Geoderma*, Amsterdam, v.101, p.105-118. [https://doi.org/10.1016/S0016-7061\(00\)00101-4](https://doi.org/10.1016/S0016-7061(00)00101-4).

LIAW A., WIENER, M. (2002) Classification and regression by randomForest. *R News*. 2(3):18–22. <https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf>

LOISEAU, T.; RICHER-DE-FORGES, A.C.; MARTELET, G.; BIALKOWSKI, A.; NEHLIG, P.; ARROUAYS, D. (2020) Could Airborne Gamma-Spectrometric Data Replace Lithological Maps as Co-Variates for Digital Soil Mapping of Topsoil Particle-Size Distribution? A Case Study in Western France. *Geoderma Reg.* 22, e00295. <https://doi.org/10.1016/j.geodrs.2020.e00295>.

LOISEAU, T.; ARROUAYS, D.; RICHER-DE-FORGES, A.C.; LAGACHIERE, P.; DUCOMMUN, C.; MINASNY, B. (2021) Density of Soil Observations in Digital Soil Mapping: A Study in the Mayenne Region, France. *Geoderma Reg.* 2021, 24, e00358. <https://doi.org/10.1016/j.geodrs.2021.e00358>

MAINO, A.; ALBERI, M.; ANCESCHI, E.; CHIARELLI, E.; CICALA, L.; COLONA, T.; DE CESARE, M.; GUASTALDI, E.; LOPANE, N.; MANTOVANI, F.; et al. (2022) Airborne Radiometric Surveys and Machine Learning Algorithms for Revealing Soil Texture. *Remote Sens*, 14, 3814. <https://doi.org/10.3390/rs14153814>.

MCBRATNEY, A. B.; ODEH, I. (2000) An overview of pedometric techniques for use in soil survey. *Geoderma*, v.97. p.293-327. [doi.org/10.1016/S0016-7061\(00\)00043-4](https://doi.org/10.1016/S0016-7061(00)00043-4).

MCBRATNEY, A.B.; MENDONÇA-SANTOS, M.L.; MINASNY, B. (2003) On digital soil mapping. *Geoderma*, v.117, p.3-52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4).

MCKENZIE, N. J.; RYAN, P. J. (1999) Spatial prediction of soil properties using environmental correlation. *Geoderma*, v.89. p.67-94. [Spatial_Prediction_of_Soil_Properties_Us20161113-24950-6u95b5-libre.pdf](https://doi.org/10.1016/S0016-7061(99)00043-4) (d1wqtxts1xzle7.cloudfront.net).

MELLO, D.C.D.; DEMATTÊ, J.A.M.; MELLO, F.A.D.O.; POPPIEL, R.; SILVERO, N.E.Q.; SAFANELLI, J.L.; SOUZA, A.B. et al. (2021). Applied Gamma-Ray Spectrometry for Evaluating Tropical Soil Processes and Attributes. *Geoderma* 2021, 381, 114736. <https://doi.org/10.1016/j.geoderma.2020.114736>.

MELLO, D.C.D.; VELOSO, G.V.; LANA, M.G.D.; MELLO, F.A.O.; POPPIEL, R.R.; CABRERO, D.R.O.; DR-RAIMO, L.A.D.L.; SCHAEFER, C.E.G.R.; FILHO, E.I.F.; LEITE, E.P.; et al. (2022) A New Methodological Framework by Geophysical Sensors Combinations Associated with Machine Learning Algorithms to Understand Soil

Attributes. *Earth Space Sci. Inform.* 15, 1219–1246. <https://doi.org/10.5194/gmd-15-1219-2022>.

MENEZES, M.D. (2011) Levantamento pedológico de hortos florestais e mapeamento digital de atributos físicos do solo para estudos hidrológicos. Tese de Doutorado. Lavras: UFLA, 224p. <http://repositorio.ufla.br/jspui/handle/1/3575>.

MENEZES, M. D. de; SILVA, S. H. G.; OWENS, P. R.; CURI, N. (2013) Digital soil mapping approach based on fuzzy logic and field expert knowledge. *Ciênc. Agrotec.* [online]., vol.37, n.4, pp. 287-298. <https://doi.org/10.1590/S1413-70542013000400001>.

MINAI, J.; LIBOHOVA, Z.; SCHULZE, D. G. (2020) Disaggregation of the 1:100,000 Reconnaissance soil map of the Busia Area, Kenya using a soil landscape rule-based approach, *CATENA*, Volume 195, 104806, ISSN 0341-8162, <https://doi.org/10.1016/j.catena.2020.104806>.

MINASNY, B. & McBRATNEY, A.B. (2002) Uncertainty analysis for pedotransfer functions. *Eur. J. Soil Sci.*, 53:417-430. <https://doi.org/10.1046/j.1365-2389.2002.00452.x>.

MINASNY, B.; MCBRATNEY, A. B.; SANTOS, M. de L. M.; SANTOS H. G. dos. (2003) Revisão sobre Funções de Pedotransferência (PTFs) e Novos Métodos de Predição de Classes de Solos e Atributos do Solo. Rio de Janeiro: Embrapa Solos, 50 p. (Embrapa Solos. Documentos n. 45). <https://ainfo.cnptia.embrapa.br/digital/bitstream/CNPS/11589/1/pedotransferencia.pdf>.

MINASNY, B., MCBRATNEY, A. B., MALONE, B. P., WHEELER, I. (2013) Digital mapping of soil carbon. *Adv. Agron.* 118: 1-47. <https://doi.org/10.1016/B978-0-12-405942-9.00001-3>.

MINTY, B. R. S. (1997). Fundamentals of airborne gamma-ray spectrometry. *AGSO Journal of Australian Geology & Geophysics*, 17(2), 39-50. https://researchgate.net/publication/231111111_Fundamentals_of_airborne_gamma-ray_spectrometry.

MOORE, I.D.; GRAYSON, R.B.; LADSON, A.R. (1991) Digital terrain modelling: a review of hydrological, geomorphological and biological application. *Hydrology Processes*, v.5. p.3- 30. <https://doi.org/10.1002/hyp.3360050103>.

MOORE, I.D.; GESSLER, P.E.; NIELSEN, G.A.; PETERSON, G.A. (1993) Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, v.57. p.443-452. <https://doi.org/10.2136/sssaj1993.03615995005700020026x>.

ODEHA, I.O.A.; MCBRATNEY, A. B.; CHITTLEBOROUGH, D. J. (1994) Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, v. 63. p.197-214. [https://doi.org/10.1016/0016-7061\(94\)90063-9](https://doi.org/10.1016/0016-7061(94)90063-9).

OLAYA, V. (2006): Basic Land-Surface Parameters. In: Hengl, T., Reuter, H.I. [Eds.]: *Geomorphometry: Concepts, Software, Applications. Developments in Soil Science*, Elsevier, Vol.33, 141-169. [https://doi.org/10.1016/S0166-2481\(08\)00006-8](https://doi.org/10.1016/S0166-2481(08)00006-8).

PINHEIRO, H.S.K. (2015) Métodos de mapeamento digital aplicados na predição de classes e atributos dos solos da Bacia Hidrográfica do Rio Guapi Macacu, RJ. 178p. [Tese Doutorado]. Universidade Federal do Rio de Janeiro, Rio de Janeiro. <https://rima.ufrj.br/jspui/handle/20.500.14407/9043>.

PROJETO MAPBIOMAS – Coleção 2017 da Série Anual de Mapas de Cobertura e Uso da Terra do Brasil, acesso através do link: <https://brasil.mapbiomas.org/>

REINHARDT, N.; HERRMANN, L. (2019) Gamma-ray spectrometry as versatile tool in soil science: A critical review. *Journal of Plant Nutrition and Soil Science*, v. 182, n. 1, p. 9–27. <https://doi.org/10.1002/jpln.201700447>

RIBEIRO, V. B., MANTOVANI, M.; LOURO, V. H. A. (2015). Aerogamaespectrometria e suas aplicações no mapeamento geológico. *Terrae Didactica*, 10(1), 29–51. <https://doi.org/10.20396/td.v10i1.8637386>

RILEY, S.J.; DE GLORIA, S.D.; ELLIOT, R. (1999): A Terrain Ruggedness that Quantifies Topographic Heterogeneity. *Intermountain Journal of Science*, Vol.5, No.1-4, pp.23-27.

online.https://researchgate.net/A_Terrain_Ruggedness_Index_that_Quantifies_Topographic_Heterogeneity.

SANCHEZ, P.A. et al. (2009) Digital soil map of the world. *Science*, v.325, p.680-681. Disponível em: <<http://www.sciencemag.org/content/325/5941/680.full.pdf>> doi: 10.1126/science.1175084.

SANTOS, M. DE L. M., SANTOS, H. G. dos (2003) Mapeamento digital de classes e atributos de solos. Embrapa solos. Documentos. 55p. <https://embrapa.br/mapeamentodigitaldeclasseseatributosdesolo.pdf>.

SANTOS, R. D. dos; LEMOS, R. C. de; SANTOS, H. G. dos; KER, J. C.; ANJOS, L. H. C. dos (2005). Manual de descrição e coleta de solo no campo. 5ª ed. Revista e ampliada. SBCS. Viçosa. 100p. 00064310.pdf (embrapa.br).

SANTOS, H. G., JACOMINE, P. K. T.; ANJOS, L. H. C.; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; CUNHA, T. J. F.; OLIVEIRA, J. B. (2013) Sistema Brasileiro de Classificação de Solos. 3ª ed. Rio de Janeiro. Embrapa Solos, Rio de Janeiro, Brasil. <http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1094003>.

SANTOS, H. G., JACOMINE, P. K. T.; ANJOS, L. H. C.; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; ARAÚJO FILHO, J. C. de; CUNHA, T. J. F.; OLIVEIRA, J. B. (2018) Sistema Brasileiro de Classificação de Solos. 5ª ed. Brasília, DF : Embrapa. <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/199517/1/SiBCS-2018-ISBN-9788570358004.pdf>

SOUZA, J. L. D.; FERREIRA, F. J. F. (2005) Anomalias aerogamaespectrométricas (K, eU e eTh) da quadrícula de Araras (SP) e suas relações com processos pedogenéticos e fertilizantes fosfatos. Revista Brasileira de Geofísica, 23, 251-274. <https://doi.org/10.1590/S0102-261X2005000300005>.

TAGHIZADEH-MEHRJARDI, R.; HAMZEHPUR, N.; HASSANZADEH, M.; HEUNG, B.; GOYDARAGH M. G.; SCHMIDT, K.; SCHOLTEN, T. (2021) Enhancing the accuracy of machine learning models using the super learner technique in digital soil mapping, Geoderma, Volume 399, 115108, ISSN 0016-7061, <https://doi.org/10.1016/j.geoderma.2021.115108>.

TEN CATEN, A. (2011) Mapeamento digital de solos: metodologias para atender a demanda por informação espacial em solos. Tese de Doutorado. Universidade Federal de Santa Maria. 108p. <http://repositorio.ufsm.br/handle/1/3326>.

THERNEAU, T.; ATKINSON, B.; RIPLEY, B. (2015) RPART: Recursive Partitioning and Regression Trees. (Author of initial R port).

WILFORD, J. R. (1992) Regolith mapping using integrated Landsat TM imagery and high-resolution gamma-ray spectrometric imagery – Cape York Peninsula. Bureau of Mineral Resources Record 1992/78: 35 pp.
<https://ecat.ga.gov.au/geonetwork/srv/api/records/a05f7892-74d0-7506-e044-00144fdd4fa6>

WILFORD, J. R. (1995) Airborne gamma-ray spectrometry as a tool for assessing relative landscape activity and weathering development of regolith, including soils. AGSO Res. News, v. 22, p. 12-14.
https://d28rz98at9flks.cloudfront.net/90442/ResearchNews_22_p012.pdf.

WILFORD, J. R.; BIERWIRTH, P. N.; CRAIG, M. A. (1997). 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. 210-216.
<https://inis.iaea.org/search/28049087>.

WILFORD, J.R. (2002). Airborne gamma-ray spectrometry. Geophysical and Remote Sensing Methods for Regolith Exploration, CRCLEME Open File Report. 144.
https://www.researchgate.net/publication/228716956_Airborne_gamma-ray_spectrometry

WILSON, M.J. (2019) The Importance of Parent Material in Soil Classification: A Review in a Historical Context. Catena, 182, 104131.
<https://doi.org/10.1016/j.catena.2019.104131>.

ZEVENBERGEN, L.W.; THORNE, C.R. (1987): Quantitative analysis of land surface topography. Earth Surface Processes and Landforms, 12: 47-56.
<https://doi.org/10.1002/esp.3290120107>.

ZHU, A. X. (2001) Soil mapping using GIS, expert knowledge, and fuzzy logic. Soil Science Society of American Journal, v. 65, p.1463-1472.
<https://doi.org/10.2136/sssaj2001.6551463x>