

**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CAMPUS DE CASCAVEL
CENTRO DE CIENCIAS EXATAS E TECNOLÓGICAS
PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA**

**SELEÇÃO DE VARIÁVEIS PARA OTIMIZAÇÃO DE CLASSIFICAÇÃO DE
DESMATAMENTO NA PLATAFORMA GOOGLE EARTH ENGINE**

SAMUEL KUHL

**CASCAVEL – PARANÁ
FEVEREIRO – 2024**

SAMUEL KUHL

**SELEÇÃO DE VARIÁVEIS PARA OTIMIZAÇÃO DE CLASSIFICAÇÃO DE
DESMATAMENTO NA PLATAFORMA GOOGLE EARTH ENGINE**

Dissertação de mestrado apresentada ao programa de Pós-Graduação em Engenharia Agrícola, em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, área de concentração em Engenharia de Sistemas Biológicos e Agroindustriais.

Orientador: Prof. Dr. Erivelto Mercante.

**CASCADEL – PARANÁ
FEVEREIRO – 2024**

Ficha de identificação da obra elaborada através do Formulário de Geração Automática do Sistema de Bibliotecas da Unioeste.

Kuhl, Samuel
SELEÇÃO DE VARIÁVEIS PARA OTIMIZAÇÃO DE CLASSIFICAÇÃO DE
DESMATAIMENTO NA PLATAFORMA GOOGLE EARTH ENGINE / Samuel Kuhl;
orientador Erivelto Mercante. -- Cascavel, 2024.
67 p.

Dissertação (Mestrado Acadêmico - Campus de Cascavel) --
Universidade Estadual do Oeste do Paraná, Centro de Ciências
Exatas e Tecnológicas, Programa de Pós-Graduação em
Engenharia Agrícola, 2024.

1. Amazonia. 2. Sensoriamento óptico. 3. Sensoriamento
SAR. 4. Random Forest. I. Mercante, Erivelto, orient. II.
Titulo.



Universidade Estadual do Oeste do Paraná
Reitoria
CNPJ 78.680.337/0001-84
Rua Universitária, 1619, Jardim Universitário
Tel.: (45) 3220-3000 • Fax: (45) 3225-4590 • www.unioeste.br
CEP: 85819-110 • Cx. P.: 701
Cascavel • PARANÁ



SAMUEL KÜHL

SELEÇÃO DE VARIÁVEIS PARA OTIMIZAÇÃO DE CLASSIFICAÇÃO DE DESMATAMENTO NA PLATAFORMA GOOGLE EARTH ENGINE

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, área de concentração Sistemas Biológicos e Agroindustriais, linha de pesquisa Geoprocessamento, Estatística Espacial e Agricultura de Precisão, APROVADO(A) pela seguinte banca examinadora:

Documento assinado digitalmente
ERIVELTO MERCANTE
Data: 22/02/2024 09:07:24-0300
Verifique em <https://validar.it.gov.br>

Orientador(a) - Erivelto Mercante

Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)

Documento assinado digitalmente
VICTOR HUGO ROHDEN PRUDENTE
Data: 22/02/2024 11:58:56-0300
Verifique em <https://validar.it.gov.br>

Victor Hugo Rohden Prudente

University of Michigan - UofM

Documento assinado digitalmente
MARCIO FURLAN MAGGI
Data: 22/02/2024 10:06:09-0300
Verifique em <https://validar.it.gov.br>

Marcio Furlan Maggi

Universidade Estadual do Oeste do Paraná - Campus de Cascavel (UNIOESTE)

Cascavel, 20 de fevereiro de 2024.

BIOGRAFIA

Samuel Kuhl, nascido em 14/08/1998, natural da cidade de Cascavel, Paraná, Brasil, concluiu o curso de graduação em Engenharia Agrícola, pela Universidade Estadual do Oeste do Paraná - UNIOESTE, no ano de 2021. Nos anos de 2017/18 participou do Laboratório de Física do Solo (LAFIS) como bolsista pela Fundação Araucária, no projeto de iniciação Científica (PIBIC) “Modelagem Matemática das Perdas de Água em Latossolo Argiloso Sob Sistema Plantio Direto” sob orientação do professor Dr. Deonir Secco, visando identificar a influência dos parâmetros do solo no escoamento superficial em lavouras sob plantio direto. Do ano de 2018 até 2021 atuou no Laboratório de Topografia e Geoprocessamento (GEOLAB) nos projetos de PIBIC: “Geotecnologias aplicadas à estimativa de produção de biometano e energia elétrica por meio de dejetos provenientes de suinocultura em Cascavel-PR”, sendo bolsista UNIOESTE, “Comparativo de Metodologias de Estimativa de Altura de Culturas Agrícolas por Meio de Imagens de Aeronaves Remotamente Pilotadas”, sendo bolsista Fundação Araucária, e “Correlação entre índices de vegetação do espectro RGB e características físico-químicas das plantas de trigo”, sendo bolsista Parque Tecnológico de Itaipu (PTI), sob orientação do professor Dr. Erivelto Mercante, atuando em aplicações do sensoriamento remoto orbital e aéreo em sistemas agrícolas, por meio de satélites e aeronaves remotamente pilotadas. Em 2021, ingressou no Programa de Pós-Graduação em Engenharia Agrícola – Sistemas Biológicos e Agroindustriais, sob orientação do Prof. Dr. Erivelto Mercante, estudando o tema: Seleção de variáveis para otimização de classificação de desmatamento na plataforma Google Earth Engine, para obtenção do título de Mestre em Engenharia Agrícola.

"A ciência está atrás do que o universo realmente é, não do que nos faz sentir bem".

Carl Sagan

AGRADECIMENTOS

Aos meus pais Alberto e Marivone, por terem sonhado esta etapa junto comigo.

Ao meu amor, Mariel, por dividir o peso dessa jornada, muito obrigado por estar ao meu lado.

À Universidade Estadual do Oeste do Paraná, pelo apoio e infraestrutura.

Ao meu orientador Erivelto Mercante, por apoiar minhas ideias e sanar as dúvidas encontradas pelo caminho.

Ao Programa de Pós-graduação em Engenharia Agrícola - PGEAGRI, e ao seu corpo docente, pelos conhecimentos e ensinamentos compartilhados.

À Fundação Parque Tecnológico de Itaipu, pela concessão de bolsa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes, pela concessão de bolsa.

Ao Laboratório de Topografia e Geoprocessamento – GeoLab e o pessoal que o habita, garantindo boas doses de alegria para continuar a caminhada.

Aos meus amigos e colegas de laboratório, em especial ao Isaque de Souza Mendes, pelo companheirismo e mentoria desde minha iniciação na pesquisa.

KUHL, Samuel. **Seleção de variáveis para otimização de classificação de desmatamento na plataforma Google Earth Engine**. Orientador: Erivelto Mercante; 67 f. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel – Paraná, 2024.

RESUMO

O desmatamento e supressão de biomas nativos, de forma irregular, em prol do desenvolvimento pode trazer consequências em diversas áreas do meio ambiente como, por exemplo, o clima. O sensoriamento remoto aliado à classificação de imagens é uma ferramenta importante para o monitoramento do meio ambiente, porém as informações de entrada no processo devem passar por uma curadoria para reduzir o esforço computacional e obter a maior precisão e acuraria possível. Este trabalho objetivou verificar o custo-benefício de diferentes conjuntos de imagens ópticas e de radar na classificação, por meio do algoritmo Random Forest, de áreas de desmatamento na Amazônia brasileira. A área de estudo localiza-se no município de Portel, no estado do Pará, Brasil. Os conjuntos de imagens derivam das constelações Sentinel 1 (SAR) e Sentinel 2 (MSI), para o ano de 2023. As classificações utilizando todas as imagens por conjunto diferiram visualmente e em suas precisões, o conjunto composto apenas por polarizações SAR obteve exatidão global (EG) de cerca de 92%, conjunto de bandas MSI com EG média de 94%, conjunto de índices com EG de cerca de 94,5% e conjunto Completo obteve EG de 95%. Os recursos computacionais utilizados na plataforma GEE diferiram devido ao uso de imagens SAR ou não, com os conjuntos que continham imagens SAR utilizando maior carga de processamento devido filtros necessários a redução do ruído speckle (Filtro de Frost e Quegan&Yu). O número de imagens influenciou na quantidade de memória utilizada pelo processamento, com o classificador utilizando cerca de 8 vezes mais memória quando se compara o conjunto com menos bandas (Sentinel 1 – 4 imagens) com o conjunto com mais imagens (Completo – 234 imagens). O conjunto que obteve o melhor custo-benefício foi o conjunto Completo 25% (58 imagens), utilizando bandas e índices derivados de ambos os sensores, sua precisão foi elevada e consumo de processamento mediano quando comparado aos demais. Realizou-se uma classificação com esse conjunto para o ano de 2022, ela foi subtraída da classificação de 2023, gerando uma camada de alertas de desmatamento para o ano de 2023, o qual ao se comparar visualmente com os dados oficiais, divulgados pelo PRODES 2023, observou-se concordância de local e forma dos alertas, desempenhando bem a função de classificação com uso otimizado de processamento.

Palavras-chave: Amazonia, Sensoriamento óptico, Sensoriamento SAR, Random Forest.

KUHL, Samuel. **Variable selection for optimization of deforestation classification on the Google Earth Engine platform**. Advisor: Erivelto Mercante; 67 f. Dissertation (Master's in Agricultural Engineering) - Western Paraná State University, Cascavel – Paraná, 2024.

ABSTRACT

The irregular deforestation and suppression of native biomes for the sake of development can have consequences in various areas of the environment, such as the climate. Remote sensing combined with image classification is an important tool for monitoring the environment, but the input information must be curated in order to reduce computational effort and obtain the greatest possible precision and accuracy. This work aimed to verify the cost-effectiveness of different sets of optical and radar images in the classification, using the Random Forest algorithm, of areas of deforestation in the Brazilian Amazon. The study area is located in the municipality of Portel, in the state of Pará, Brazil. The image sets were derived from the Sentinel 1 (SAR) and Sentinel 2 (MSI) constellations for the year 2023. The classifications using all the images per set differed visually and, in their accuracies, with the set composed only of SAR polarizations having an overall accuracy (EG) of around 92%, the set of MSI bands having an average EG of 94%, the set of indices having an EG of around 94.5% and the Complete set having an EG of 95%. The computational resources used on the GEE platform differed due to the use of SAR images or not, with the sets containing SAR images using a greater processing load due to the filters needed to reduce speckle noise (Frost and Quegan&Yu filters). The number of images influenced the amount of memory used for processing, with the classifier using around 8 times more memory when comparing the set with the fewest bands (Sentinel 1 - 4 images) with the set with the most images (Complete - 234 images). The most cost-effective set was the Complete 25% set (58 images), using bands and indices derived from both sensors, with high accuracy and average processing consumption compared to the others. A classification was carried out with this set for the year 2022, which was subtracted from the 2023 classification, generating a layer of deforestation alerts for the year 2023, which when visually compared with the official data released by PRODES 2023, there was agreement in the location and shape of the alerts, performing the classification function well with optimized use of processing.

Keywords: Amazon, Optical sensing, SAR sensing, Random Forest.

SUMÁRIO

LISTA DE FIGURAS	10
LISTA DE TABELAS.....	11
LISTA DE ABREVIações	12
1 INTRODUÇÃO	13
2 OBJETIVOS.....	15
2.1 Objetivo geral	15
2.2 Objetivos específicos.....	15
3 REVISÃO BIBLIOGRÁFICA	16
3.1 Monitoramento de desmatamento	16
3.2 Sensoriamento remoto	17
3.2.1 Sensoriamento remoto orbital para monitoramento ambiental.....	20
3.2.2 Abordagem multisensor: constelação Sentinel	21
3.3 Classificação de imagens	22
3.3.1 Classificadores baseados aprendizado de máquina	24
3.3.2 Otimização de parâmetros em classificadores.....	26
3.3.3 Avaliação de classificação	27
4 MATERIAL E MÉTODOS.....	29
4.1 Área de estudo	29
4.2 Softwares e processamentos	30
4.3 Conjuntos de imagens de entrada.....	31
4.4 Coleta de amostras e classificação	36
4.5 Verificação de acurácia e precisão da classificação	37
4.6 Avaliação do desempenho computacional em nuvem	38
5 RESULTADOS E DISCUSSÃO	40
6 CONCLUSÕES	50
7 CONSIDERAÇÕES FINAIS	51
REFERÊNCIAS.....	52
APÊNDICE.....	61
APÊNDICE A - TABELA DAS IMPORTÂNCIAS	61
APÊNDICE B – MATRIZES DE CONFUSÃO.....	67

LISTA DE FIGURAS

Figura 1 Comportamento espectral de diferentes alvos em cada banda do sensor MSI/Sentinel 2.	20
Figura 2 Frequência de cobertura de nuvens considerando MAIAC e MODIS.	21
Figura 3 Técnicas de aprendizagem de classificação de imagem.	23
Figura 4 Árvore de decisão.	25
Figura 5 Matriz de confusão de uma classificação.	28
Figura 6 Fluxograma dos conjuntos de imagens (a, b, c e d) e processos utilizados no estudo.	29
Figura 7 Mapa de localização da área de estudo.	30
Figura 8 Pontos aleatórios nas classes de floresta e não floresta, utilizados para precisão/acurácia.	37
Figura 9 Classificações dos conjuntos (S1, S2, Índices e Completo) com 100% das bandas para o ano de 2023.	40
Figura 10 Zoom em uma área de diferença de classificação entre S1 e S2.	41
Figura 11 Diferença entre 100% das imagens e apenas o Top10 imagens para cada conjunto (S1, S2, Índices e Completo).	42
Figura 12 Área de maior diferença encontrada no conjunto Índices ao longo da remoção de imagens.	42
Figura 13 Índices EG, AU e AP por conjunto (S1, S2, Índices e Completo).	43
Figura 14 Índices IoU e Dice para as classes floresta e não floresta por cada conjunto (S1, S2, Índices e Completo).	44
Figura 15 Recursos computacionais utilizados – EECU.s no eixo vertical principal e pico de memória no eixo vertical secundário – para cada conjunto (S1, S2 Índices e Completo).	46
Figura 16 Mapa de localização dos alertas emitidos pelo PRODES e derivados da classificação com o conjunto Completo 25%.	48
Figura 17 Mapa de zoom em pontos de grande área de desmatamento comparando PRODES 2023 e Alertas (2023-2022) gerados com o conjunto Completo 25%.	48

LISTA DE TABELAS

Tabela 1 Descrição dos principais comprimentos de onda	19
Tabela 2 Informações espectrais das bandas do sensor MSI	33
Tabela 3 Índices de vegetação para o sensor MSI	34
Tabela 4 Índices de vegetação para o sensor SAR	35
Tabela 5 Conjuntos de imagens e o número de imagens de acordo com a porcentagem....	36
Tabela 6 Índices de avaliação de acurácia e precisão da classificação.	38

LISTA DE ABREVIACOES

AP	Acurcia do Produtor
ARVI	ndice de Vegetao Atmosfericamente Resistente
AU	Acurcia do Usurio
CPU	Unidade Central De Processamento
CR	Cross-Ratio
DETER	Sistema de Deteco de Desmatamento em Tempo Real
EECU	Earth Engine Compute Unit
EG	Exatido Global
ETM+	Enhanced Thematic Mapper Plus
GEE	Google Earth Engine
GNDVI	ndice de Vegetao por Diferena Normalizada Verde
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renovveis
INPE	Instituto Nacional de Pesquisas Espaciais
IoU	Jaccard Index
IW	Interferomtrico de Faixa Larga
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multispectral Imager
MSS	Multi-Spectral Scanner
NDBI	ndice de Construes por Diferena Normalizada
NDRE	Diferena Normalizada da Borda do Vermelho
NDVI	ndice de Vegetao por Diferena Normalizada
NDWI	ndice de gua por Diferena Normalizada
OLI	Operational Land Imager
PMABB	Programa de Monitoramento Ambiental dos Biomas Brasileiros
PRODES	Programa de Monitoramento da Floresta Amaznica Brasileira por Satlite
RENDVI1	NDVI Borda do Vermelho 1
RENDVI2	NDVI Borda do Vermelho 2
RF	Floresta Aleatria
RFE	Recursive Feature Elimination
RVI	Radar Vegetation Index
RVIm	Radar Vegetation Index Modificado
S2REP	Posio da Borda do Vermelho do Sentinel-2
SAR	Radar de Abertura Sinttica
SAVI	ndice de Vegetao Ajustado ao Solo
SVM	Support Vector Machines
SW2_N	Diferena Normalizada Swir2-Nir
TM	Thematic Mapper
VH	Vertical-Horizontal
VV	Vertical-Vertical
WFI	Wide-Field Imager

1. INTRODUÇÃO

A degradação de biomas nativos é um tópico sensível visto que suas consequências são sentidas a longo prazo, e a aceleração da degradação por meio da atividade humana pode criar um prejuízo socioambiental sem capacidade de retorno. O Brasil com sua extensa área territorial e baixo contingente para fiscalização do desmatamento, fica à mercê das consequências que essa condição cria. Aliando tecnologia no combate a destruição do meio ambiente temos, por meio do sensoriamento remoto orbital, a capacidade de identificação remota de áreas desmatadas e em processo de desmatamento, agilizando a captura de informações relevantes para a tomada de ação.

A capacidade do sensoriamento remoto orbital em monitorar uma vasta área territorial associada à identificação do uso e cobertura da terra, por meio da classificação de imagens, é uma ferramenta capaz de gerar informações de maneira rápida e precisa. Atualmente com o desenvolvimento de sensores que atuam em diferentes faixas do espectro eletromagnético, indo das bandas do espectro visível (azul, verde e vermelho) até a região das microondas, com os sensores de radar, que são menos afetados por condições climáticas, a coleta de dados da superfície terrestre gera um volume substancial de dados, o qual demanda para sua análise um poder computacional elevado.

A criação de servidores privados de alto desempenho e com grande capacidade de armazenamento, para o trabalho com essa elevada dimensionalidade dos dados, pode ser substituída pelo uso de plataformas de processamento em nuvem, como o Google Earth Engine (GEE), que buscam oferecer o uso desses servidores sob demanda. A fim de reduzir a carga de processamento necessária para atividades de grande escala, como o monitoramento do Bioma Amazônia ou tarefas que são executadas repetidamente, é necessário que seja feita uma curadoria nos dados de entrada de acordo com o processo, reduzindo o custo operacional envolvido.

Cada técnica de classificação de imagens possui seus desafios e abordagens, o método baseado no algoritmo das Florestas Aleatórias (Random Forest – RF) possui um grande leque de possibilidade de entrada de dados, agrupando dados binários, nominais e contínuos entre outros, fornecendo uma ferramenta de análise do uso de cada dado, ou seja, sua importância para a separabilidade das classes pretendidas. Dessa forma é possível definir um conjunto de dados de entrada que melhor se

adequa à área alvo, porém é necessário garantir que a precisão atinja um nível satisfatório.

Portanto, o presente estudo busca identificar como a redução de dados de entrada em uma classificação supervisionada de uso e cobertura da terra, feita por um algoritmo de RF, implica na redução da demanda computacional necessária para a tarefa e, também, como isso afeta a precisão dos resultados.

2. OBJETIVOS

2.1 Objetivo geral

Analisar a influência de diferentes conjuntos de imagens de sensoriamento remoto orbital multisensor (óptico e radar), em função da acurácia/precisão e do custo computacional, empregados na classificação dos alvos floresta e não floresta, para o Bioma Amazônia, por meio do algoritmo Random Forest implementado na plataforma de processamento em nuvem Google Earth Engine.

2.2 Objetivos específicos

- a) Gerar classificações dos alvos floresta e não floresta para a área de estudo, localizada em Portel (PA), por meio de Random Forest, com diferentes conjuntos de imagens;
- b) Definir o conjunto com melhor custo-benefício computacional, ou seja, melhor relação entre precisão/acurácia e gasto computacional;
- c) Elaborar alertas de potencial desmatamento ocorrido no ano de 2023, para a área em estudo, utilizando o conjunto de melhor custo-benefício;
- d) Comparar os alertas gerados com os alertas de desmatamento provenientes do programa PRODES para o ano de 2023.

3. REVISÃO BIBLIOGRÁFICA

Este capítulo busca explorar o estado da arte acerca dos tópicos que circundam a identificação de desmatamento via sensoriamento remoto orbital multisensor, apresentando informações sobre o processo de monitoramento de desmatamento a nível de país (3.1), como o sensoriamento remoto obtém informação e utiliza para monitorar o desmatamento (3.2) e o processo de classificação de imagens (3.3).

3.1 Monitoramento de desmatamento

O desmatamento é caracterizado pela remoção da cobertura vegetal nativa de um determinado lugar, seu acontecimento está ligado a incidentes naturais, extração de matéria-prima, expansão urbana e atividades como pecuária e agricultura (ALMEIDA et al., 2022; MACHIDA; ANDRADE, 2022). Quando acontece de maneira descontrolada e não planejada pode acarretar negativamente em processos ambientais como ciclo de carbono e água, erosão do solo, redução da biodiversidade e alteração na dinâmica climática (FEARNSIDE, 2006; LAURANCE; SAYER; CASSMAN, 2014; HADDAD, 2015).

O monitoramento do desmatamento é uma necessidade governamental tanto para punir ilegalidades quanto para planejar a expansão de áreas, mitigando os efeitos negativos. O Brasil conta com o Programa de Monitoramento Ambiental dos Biomas Brasileiros (PMABB), como o projeto que engloba todas as políticas de monitoramento ambiental dos seis biomas do país - Amazônia, Caatinga, Cerrado, Mata Atlântica, Pampa e Pantanal (BRASIL, 2017). Devido sua grande área territorial o país mantém programas de monitoramento via satélite, proporcionando agilidade na detecção de focos de desmatamento (SALES; STROBL; ELLIOTT, 2022).

A Amazônia é o maior bioma brasileiro com cerca de 4,2 milhões de km¹, sendo, também, o bioma com maior área desmatada. O monitoramento via satélite do desmatamento nessa região ocorre principalmente por dois programas, o primeiro e mais antigo é o Programa de Monitoramento da Floresta Amazônica Brasileira por Satélite (PRODES) e o segundo é o Sistema de Detecção de Desmatamento em Tempo Real (DETER), ambos pensados e executados pelo Instituto Nacional de

¹ Disponível em:

<https://geoftp.ibge.gov.br/informacoes_ambientais/estudos_ambientais/biomas/mapas/biomas_5000_mil.pdf>

Pesquisas Espaciais (INPE) de acordo com a metodologia descrita por ALMEIDA et al. (2022) e disponibilizados no portal TerraBrasilis.

O PRODES consiste no monitoramento via satélite do desmatamento, por corte raso ou por degradação progressiva na vegetação primária, na Amazônia Legal e gera em um produto anual, desde 1988, que monitora o incremento e a taxa de desmatamento do bioma, que é utilizado como guia das políticas públicas. Sua operação atual conta com diferentes sensores a bordo dos satélites Landsat 8-9, CBERS 4-4a e Sentinel 2. O método de contabilização é incremental, ou seja, são mapeados os novos casos de desmatamento em relação ao ano anterior e a classificação dos polígonos é validada por fotointerpretação de especialistas (ALMEIDA et al., 2022).

Já o DETER, lançado em 2004, foca na detecção diária de alterações da cobertura florestal como desmatamento e degradação, seu produto é a emissão de alertas diários para entidades governamentais como o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) e Secretarias Estaduais de Meio Ambiente, tornando-se ferramenta de suporte à fiscalização. Seus dados também são disponibilizados quinzenalmente para qualquer pessoa no portal TerraBrasilis. Atualmente utiliza sensores a bordo dos satélites Amazônia-1 e CBERS 4-4a (ALMEIDA et al., 2022).

Ambos os programas disponibilizam publicamente os dados de pontos de alteração florestal que possuam área mínima de 6,25 ha, já áreas identificadas menores que 6,25 ha são disponibilizadas apenas para órgãos governamentais (ALMEIDA et al., 2022).

As limitações de uso em larga escala, pela comunidade civil, nesses programas são o tempo análise do PRODES que produz resultados anuais e a utilização exclusivamente de imagens ópticas no atual modelo do DETER, que cria uma barreira na identificação do desmatamento e degradação em áreas com alta taxa de incidência de nuvens. Outro ponto de atenção é a disponibilidade dos polígonos com área mínima de 6,25 ha, limitando a utilização na detecção precoce ou na validação de pontos específicos com área inferior.

3.2 Sensoriamento remoto

O sensoriamento remoto consiste na técnica de obtenção de informações, por meio de sensores, acerca de objetos ou fenômenos, sem contato físico com o alvo em

estudo (JENSEN, 2009). Sua utilização pode ser a nível de solo, a nível aéreo ou orbital, proporcionando a obtenção de dados em diferentes escalas espaciais e de áreas distintas ao mesmo tempo (BEZERRA; SILVA; FERREIRA, 2008).

O sensoriamento remoto usa a interação da energia eletromagnética, discretizada por canais ou faixas do espectro eletromagnético, com o objeto em estudo de forma que a interação pode ser absorção, onde a energia é absorvida e transformada em outra forma de energia como, por exemplo, a absorção da energia nas faixas do azul e vermelho pelas plantas para realização da fotossíntese; transmissão que consiste na passagem da energia pelo objeto, sofrendo ou não mudanças de direção e características; e reflexão que é a capacidade do objeto refletir a energia incidente (JENSEN, 2009).

O sensor que capta a energia pode ser do tipo imageador ou não-imageador, respectivamente, quando o resultado é uma imagem como em uma câmera, e quando não produz imagem como um espectrorradiômetro, que produz um gráfico da área analisada discretizado por comprimento da onda (FORMAGGIO; SANCHES 2017).

As imagens possuem estrutura matricial, onde cada célula possui um valor observado pelo sensor, no qual a capacidade de identificar objetos de interesse e diferenciá-los é determinada por quatro resoluções, descritas por MENESES e ALMEIDA (2012): a) Resolução Espacial: determina o tamanho do menor objeto que pode ser identificado e individualizado na imagem, é dado em função da distância entre sensor e objeto e o ângulo IFOV do sensor, dessa forma uma resolução espacial de 10 m indica que objetos com menos de 10 x 10 m terão seu comportamento dissolvido e atribuído a apenas uma célula da matriz; b) Resolução Radiométrica: capacidade do sensor em medir diferentes intensidades dos níveis de radiância, ou seja, a energia refletida e emitida pelos objetos é detectada pelo sensor e traduzida para um nível de cinza de acordo com sua capacidade. Sua medida é em número de dígitos binários (bits), um sensor com 8 bits de resolução é capaz de diferenciar $2^8 = 256$ valores de radiância, já um sensor com 16 bits é capaz de diferenciar $2^{16} = 65.536$ valores de energia; c) Resolução Temporal: frequência de revisita a área alvo avaliada; e d) Resolução Espectral: envolve três fatores relacionados ao comportamento de onda da radiação eletromagnética: largura da faixa espectral, número de faixas e posição no espectro eletromagnético. A sensibilidade para discretizar alvos está ligada à sua capacidade de captar energia de diferentes partes do espectro eletromagnético, sensores que captam várias partes do espectro, em

comprimentos de onda diferentes e cada faixa com um intervalo estreito do espectro serão eficientes em fornecer informações que diferenciem alvos.

Cada porção do espectro eletromagnético possui um intervalo de comprimento de onda associado, a Tabela 1 apresenta a descrição das principais faixas espectrais (popularmente denominadas de bandas) com seus respectivos valores típicos, onde vê-se que a energia e a frequência são maiores quanto menor o comprimento de onda.

Tabela 1 Descrição dos principais comprimentos de onda

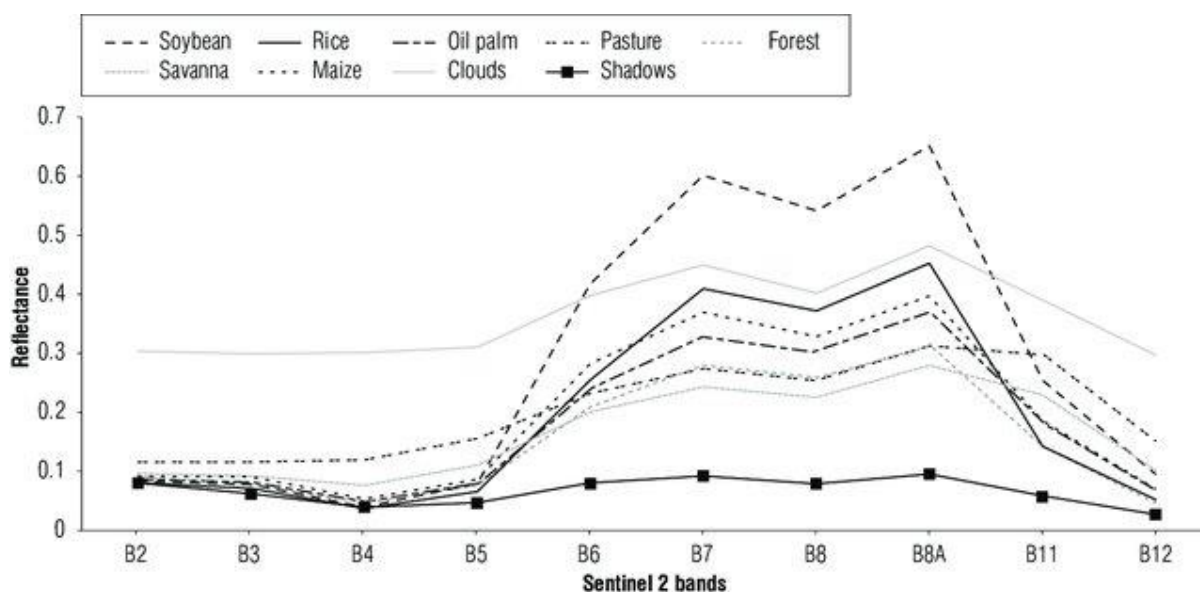
Porção do espectro	Comprimento de onda		Energia	
	Nanômetro (nm)	Frequência (Hzx10 ¹⁴)	Número de onda (Ψ cm ⁻¹)	Elétron-volt (eV) = 1240/ λ
Ultravioleta, onda curta	254	11,82	39.400	4,89
Ultravioleta, onda longa	366	8,19	27.300	3,39
Violeta, limite	400	7,50	25.000	3,10
Azul	450	6,66	22.200	2,75
Verde	500	6,00	20.000	2,48
Verde	550	5,45	18.200	2,25
Amarelo	580	5,17	17.240	2,14
Laranja	600	5,00	16.700	2,06
Vermelho	650	4,62	15.400	1,91
Vermelho, limite	700	4,29	14.300	1,77
Infravermelho próximo	1.000	3,00	10.000	1,24
Infravermelho distante	30.000	0,10	333	0,04

Número de onda (Ψ) é o número de ondas numa unidade de comprimento, usualmente centímetros, logo $\Psi = 1/\lambda$ (cm).

Fonte: adaptado de JENSEN (2009).

A Figura 1 apresenta o comportamento de diferentes alvos (soja, arroz, dendê, pasto, floresta, savana, milho, nuvem e sombra) de acordo com sua reflectância em cada banda captada pelo sensor MSI, que é um sensor a bordo da constelação Sentinel 2, e possui 10 bandas em diferentes locais do espectro com diferentes larguras de banda, permitindo uma maior diferenciação dos alvos. As bandas 2, 3 e 4 são respectivamente na porção azul, verde e vermelho do espectro eletromagnético, já as demais bandas (5 até 12) estão localizadas no infravermelho próximo.

Figura 1 Comportamento espectral de diferentes alvos em cada banda do sensor MSI/Sentinel 2.



Fonte: MARTÍNEZ M. (2017).

3.2.1 Sensoriamento remoto orbital para monitoramento ambiental

O monitoramento ambiental utilizando sensoriamento remoto tende a gerar produtos que carregam informação sobre o uso e cobertura do solo e sua dinâmica de mudanças, mas quando a área monitorada possui grandes dimensões como, por exemplo, um bioma, comumente usam-se sensores montados em satélite devido sua grande cobertura espacial e diversidade de comprimentos de onda disponíveis (TRUJILLO-JIMÉNEZ et al., 2022; WANG, 2023).

A escolha do sensor depende do objetivo do mapeamento e do conjunto de resoluções que mais se adequa, para tal existe uma gama de sensores a bordo de satélites como, por exemplo, o sensor Wide-Field Imager (WFI) a bordo dos satélites CBERS-4, CBERS-4a e Amazonia-1 (OLDONI et al., 2022), os sensores Multi-Spectral Scanner (MSS), Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+) e Operational Land Imager (OLI) a bordo da constelação Landsat (HUSSAIN; MUBEEN; KARUPPANNAN, 2022), o sensor Moderate Resolution Imaging Spectroradiometer (MODIS) a bordo dos satélites Terra e Aqua (ROMÁN et al., 2024), o sensor Multispectral Imager (MSI) a bordo dos satélites Sentinel 2A e 2B (DAMASCENO et al., 2023).

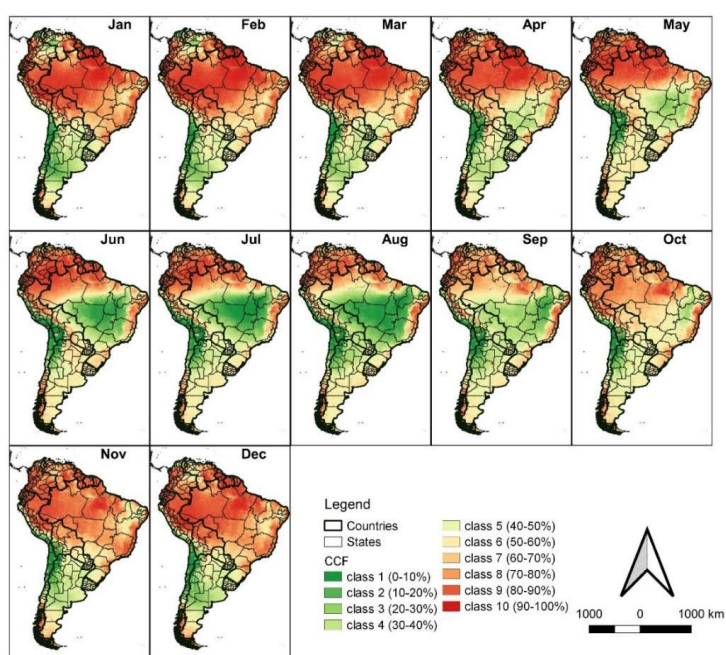
Os sensores acima mencionados possuem suas particularidades, porém todos são sensores ópticos com bandas que possuem comprimento de onda incapazes de transpassar obstáculos como as nuvens, dessa forma em locais com presença de

nuvem o sensor não conseguirá coletar dados da superfície terrestre (ZHANG et al., 2023). Por outro lado, sensores que operam na região das microondas, como os Radares de Abertura Sintética (Synthetic Aperture Radar - SAR), a bordo da constelação de satélites Sentinel-1 por exemplo, são menos afetados por nuvens, possuindo o potencial de imagear a superfície terrestre mesmo em dias nublados. Porém os dados SAR são complexos, exigindo conhecimento especializado para sua manipulação e interpretação (OLDONI et al., 2020; ZHANG et al., 2023).

3.2.2 Abordagem multisensor: constelação Sentinel

Programas como DETER e PRODES, que se baseiam exclusivamente em imagem ópticas, podem ser prejudicados pela redução do número de imagens disponíveis devido à quantidade de nuvens existente na região onde se encontra o bioma Amazônia. Prudente (2021) utilizando produtos do sensor MODIS gerou a frequência de cobertura de nuvens para a América do Sul (Figura 2), nota-se que o melhor período para obtenção de imagens é de junho a setembro, porém para os estados do Amazonas e Pará a cobertura de nuvens ainda se mantém na classe de até 60-70%, dificultando a operação óptica.

Figura 2 Frequência de cobertura de nuvens considerando MAIAC e MODIS



Fonte: PRUDENTE (2021).

A fim de se construir um conjunto de dados robusto para identificação do uso e cobertura do solo torna-se interessante a junção dos sensores ópticos e dos sensores radar, unindo, respectivamente, a capacidade de separação de alvos com a capacidade de observação da superfície mesmo sob condição de nuvem (CHATZIANTONIOU; PSOMÍADIS; PETROPOULOS, 2017). Essa junção de dados já é amplamente utilizada e testada na literatura, principalmente utilizando-se os satélites da família Sentinel, devido ao conjunto de boas resoluções do sensor óptico MSI a bordo do Sentinel 2A e 2B e versatilidade do sensor SAR a bordo do Sentinel 1A e 1B, além da disponibilização gratuita dos dados, como pode ser visto em CHATZIANTONIOU; PSOMÍADIS; PETROPOULOS (2017), TAVARES et al. (2019), ARJASAKUSUMA et al. (2022), MAO; VAN NIEL; MCVICAR (2023) e TORO et al. (2023).

LI et al. (2018) descrevem a missão do Sentinel 2A e 2B, que apresentam 13 bandas do sensor MSI, cobrindo o intervalo entre 433-2286 nm, ou seja, do azul costal até o infravermelho de ondas longas e apresentando resoluções de 10, 20 e 60 m dependendo da banda. Seu tempo de revisita é de 5 dias alternando os satélites. Essa constelação se destaca pela existência das bandas de borda do vermelho, infravermelho próximo e infravermelho de ondas longas, que são utilizadas para vários índices de vegetação.

Já o sensor SAR a bordo dos satélites Sentinel 1A e 1B opera na faixa das microondas utilizando a banda C de frequência 5,405 GHz e resolução espacial de 10 m. Os satélites passavam alternadamente a cada 6 dias no mesmo ponto, porém em dezembro de 2021 o satélite Sentinel 1B sofreu uma falha e foi desativado, com isso a resolução temporal a partir dessa data é de 12 dias. Ele possui 4 modos de operação, mas o mais utilizado em mapeamentos de uso e cobertura é chamado de Interferométrico de faixa larga (Interferometric Wide swath – IW) e fornece polarização dupla – vertical-vertical (VV) e vertical-horizontal (VH) ou horizontal-horizontal (HH) e horizontal-vertical (HV) (TORRES *et al.*, 2012; PRUDENTE, 2021; TORO et al., 2023).

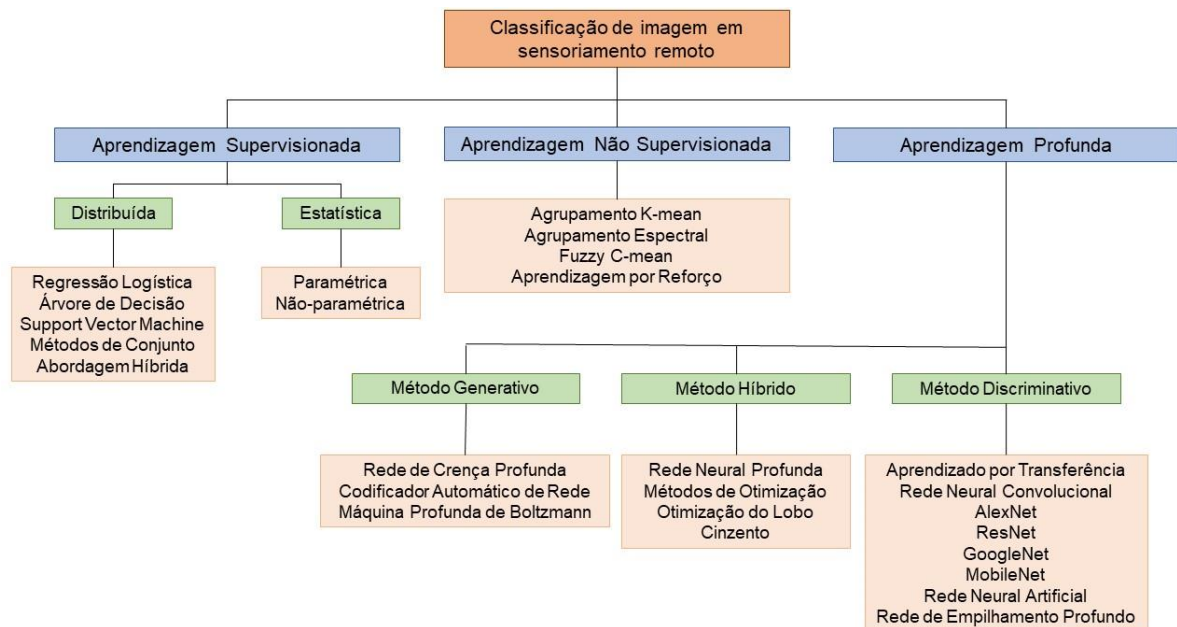
3.3 Classificação de imagens

A tarefa de classificar manualmente elementos em uma imagem é um trabalho demorado e para contornar esse ponto negativo foram propostos vários métodos de automatização (Figura 3) que, por meio de algoritmos, agrupam elementos com

comportamentos similares em classes, com supervisão humana ou de forma automática (MING et al., 2016).

Não existe um classificador ideal e único que se adeque a todos os casos, cada tipo possui especificidades que devem ser atendidas (suposição de normalidade, por exemplo) e um grau de interação humana, dessa forma entende-se que a escolha deve ser guiada pela precisão (quantidade de acertos), quantidade de interação humana (coleta de dados e ajustes de modelo), velocidade computacional e automatização (DEFRIES; CHAN, 2000).

Figura 3 Técnicas de aprendizagem de classificação de imagem



Fonte: adaptado de MEHMOOD et al. (2022).

Classificadores derivados de aprendizado de máquina e aprendizagem profunda estão em evidência na literatura devido maior precisão, robustez e capacidade de lidar com alta quantidade de dados de entrada (dimensionalidade) quando comparados a métodos tradicionais como, por exemplo, a máxima verossimilhança (MING et al., 2016; CHOWDHURY, 2024). O conceito de aprendizado de máquina, cunhado por Arthur Samuel em 1959, é uma junção de conceitos estatísticos e ciência da computação e é associado a algoritmos que permitem que novos dados sejam processados ou classificados com base em informações prévias (dados de treinamento), sendo hoje parte do conjunto denominado inteligência artificial (BANSAL; GOYAL; CHOUDHARY, 2022).

3.3.1 Classificadores baseados em aprendizado de máquina

Os classificadores baseados em aprendizado de máquina apresentam potencial na classificação do uso e ocupação da terra, derivado de imagens orbitais, devido sua capacidade de identificação de padrões não lineares, melhor desempenho com grandes conjuntos de dados quando comparados à modelos tradicionais, simplicidade dos algoritmos de decisão e redução nas situações de super ajustamento aos dados de treinamento (overfitting) (ASADOLLAH; JODAR-ABELLAN; PARDO, 2024). Algoritmo árvore de decisão (3.3.1.1) é a base da construção do RF (3.3.1.2), que possui vasta aplicação na classificação de imagens de sensoriamento remoto, sendo um dos mais estudados para essa função.

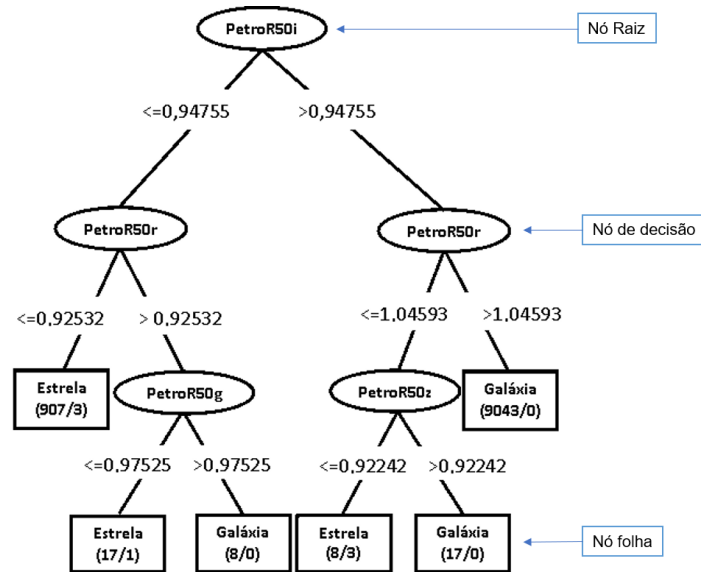
3.3.1.1 Árvore de decisão

Árvore de decisão é um algoritmo de aprendizado de máquina supervisionado (existe um conjunto de dados de treinamento) e hierárquico (segue um caminho de decisão do mais importante para o menos importante) que busca alocar o dado em uma classe pré-determinada (nó folha) e para isso cria um conjunto de decisões (nó de decisão) baseadas nos atributos de treinamento (ABDULRAHMAN et al., 2020; BANSAL; GOYAL; CHOUDHARY, 2022).

Um exemplo, para auxílio na explicação, pode ser visto no trabalho de RUIZ et al. (2009), onde eles utilizaram árvores de decisão para classificar binariamente, como estrela ou galáxia, 10.000 objetos no espaço de acordo com seu brilho em diferentes bandas.

A árvore gerada (Figura 4) contém o nó raiz, que consiste no primeiro nó de decisão, onde o atributo PetroR50i (perfil de brilho superficial Petrosiano) é dividido no ponto que gera maior separação das amostras (maior ganho de informação) - $PetroR50i > 0,94755$ e $PetroR50i \leq 0,94755$ - porém apenas ele como uma decisão não produz informação suficiente para classificação dos objetos, após ele outros dois nós de decisão utilizam diferentes intervalos de valor do atributo PetroR50r para dividir os objetos, agora o caminho da esquerda apresenta um nó folha, indicando que aquele caminho consegue enquadrar 910 objetos na classe “estrela”, de forma que 907 estão corretos e 3 estão classificados de maneira incorreta. O caminho da direita, por sua vez consegue rotular corretamente 9043 objetos como “galáxia”. Por fim há outros dois nós de decisão utilizando os atributos PetroR50g e PetroR50z, de forma que todos os objetos foram alocados em alguma classe.

Figura 4 Árvore de decisão



Fonte: adaptado de RUIZ *et al.* (2009).

Para decidir a ordem de importância dos atributos utilizados nos nós de decisão é empregado o conceito de ganho de informação, que consiste em minimizar a entropia, ou seja, o atributo que divide um maior número de amostras de treinamento de forma correta leva a um maior ganho de informação, como pode ser visto na Figura 4, de forma que quanto mais acima o nó de decisão maior é sua capacidade em alocar as amostras corretamente (ABDULRAHMAN *et al.*, 2020; BANSAL; GOYAL; CHOUDHARY, 2022).

3.3.1.2 Random Forest

Criado por BREIMAN (2001) a técnica de aprendizado de máquina RF é baseada em árvores de decisão e método de ensacamento, que consiste em utilizar diversos classificadores ou diferentes conjuntos de dados de treinamento em um mesmo classificador e atribuir a classe mais votada pelo conjunto. Esse método utiliza o conceito de amostragem bootstrap, que divide aleatoriamente e com reposição o conjunto de dados de treinamento, de forma que cada subconjunto aleatório dará origem a uma árvore de decisão com caminhos e atributos sortidos (RODRIGUEZ-GALIANO *et al.*, 2012).

A totalidade de árvores de decisão (ntree) irá classificar os novos dados de acordo com votação majoritária, ou seja, a classe atribuída ao novo dado levará em

conta todas as árvores (PAL, 2005). Como cada árvore de decisão considera um conjunto aleatório de dados e variáveis preditivas (mtry), há uma redução da correlação entre elas, reduzindo o viés e melhorando desempenho global (BIAU; SCORNET, 2016; MING et al., 2016). Outro ponto positivo é a melhora da precisão global quando a precisão das árvores isoladas (alunos fracos) é maior que 50%, pois a maioria delas estará certa e, calcado na Teoria dos Grandes Números, a resposta tenderá ao acerto (BREIMAN, 2001; RODRIGUEZ-GALIANO et al., 2012).

Atualmente com o número de variáveis possíveis de serem aferidas e armazenadas é comum haver um conjunto de variáveis preditivas que agregam pouco aos classificadores, podendo inserir ruídos na classificação (BREIMAN, 2001), porém o RF ao utilizar o ganho de informação não escolherá essas variáveis, possibilitando a construção de uma métrica chamada “importância”, que diz o quão útil para a classificação a variável é, viabilizando a exclusão das variáveis que são menos utilizadas e com isso diminuindo a dimensionalidade dos dados (RODRIGUEZ-GALIANO et al., 2012).

Para evitar que a floresta fique muito complexa e computacionalmente pesada, pode-se restringir seu crescimento (poda) ao limitar a máxima profundidade da árvore, ou seja, definir um número máximo de nós de decisão ou, também, pode-se utilizar o conceito da importância, de modo que variáveis preditoras que promovam pouco ganho de informação podem ser removidas, ajudando a criar árvores mais simples e interpretáveis (MING et al., 2016).

3.3.2 Otimização de parâmetros em classificadores

A classificação de uso e cobertura utilizando alta dimensionalidade devido a dados variados como, por exemplo, dados ópticos e dados de radar combinados no mesmo banco de dados, consome tempo e poder computacional, por isso deve-se utilizar técnicas de redução da coleção de dados (PRUDENTE, 2021).

Quando são utilizadas séries temporais para classificação de objetos uma opção de redução de dimensionalidade é captar variações do pixel ao longo do tempo por meio de métricas como o mínimo, máximo, média e desvio padrão, reduzindo a série temporal a uma imagem por métrica escolhida (DEFRIES; HANSEN; TOWNSHEND, 1995). Em um caso hipotético de um sensor capturar 4 bandas e a série temporal for composta de 20 imagens, isso resultaria em 80 variáveis preditoras,

porém aplicando a redução pelas métricas acima esse número seria reduzido a 16 variáveis preditoras.

Outra técnica de redução de dados é a eliminação das variáveis preditoras de menor importância, na literatura esse método já foi usado por BADDA *et al.* (2023) que compararam três conjuntos de dados (um índice, quatro índices mais 12 bandas de imagens do satélite Sentinel 2 e, por fim, apenas quatro índices) para detecção de áreas de queimada, de forma que o modelo que melhor se ajustou foi o composto apenas pelos quatro índices espectrais. Proposta por GUYON *et al.* (2002) e chamada de Recursive Feature Elimination (RFE), a automatização da remoção das variáveis de menor importância, realiza primeiramente a modelagem RF seguida da retirada da variável de menor importância, repetindo esse ciclo até o momento que reste apenas uma variável, então é selecionado a iteração que possui o menor número de variáveis com maior precisão (BAZZAN *et al.*, 2023).

3.3.3 Avaliação de classificação

A fim de testar a qualidade das classificações existem dois caminhos, o primeiro visa a verificação global de todos os polígonos, para isso é necessária uma área vetorizada que sirva de verdade de campo e isso a torna o método trabalhoso, essa aplicação pode ser vista nos índices Jaccard index (IoU) (KOENEMAN; CAVANAUGH, 2022) e índice Dice (ALIDOOST *et al.*, 2023):

$$IoU = \frac{TP}{TP + FP + FN} \quad \text{Eq. 01}$$

$$Dice = \frac{2 * TP}{2 * TP + FP + FN} \quad \text{Eq. 02}$$

Em que: TP= verdadeiro positivo (área classificada corretamente como positiva); FP = falso positivo (área de classe negativa classificada como positiva); FN= falso negativo (área de classe positiva classificada como negativa).

Outra forma é a utilização de pontos ou amostras criadas para a validação, as quais são feitas de forma aleatória em cada classe e não são utilizadas para o treinamento do classificador escolhido, ou seja, o modelo não conhece o comportamento dessas amostras. Possuindo a classe real de cada amostra e a classe predita pelo classificador é possível a criação da matriz de confusão (Figura 5), a qual é dividida em acertos (verdadeiro positivo e verdadeiro negativo) e erros (falso positivo

e falso negativo), resultando em um quadro de onde se derivam índices de acurácia e precisão como a acurácia do produtor (AP), acurácia do usuário (AU) e exatidão global (EG) (PHILLIPS et al., 2024).

Figura 5 Matriz de confusão de uma classificação.

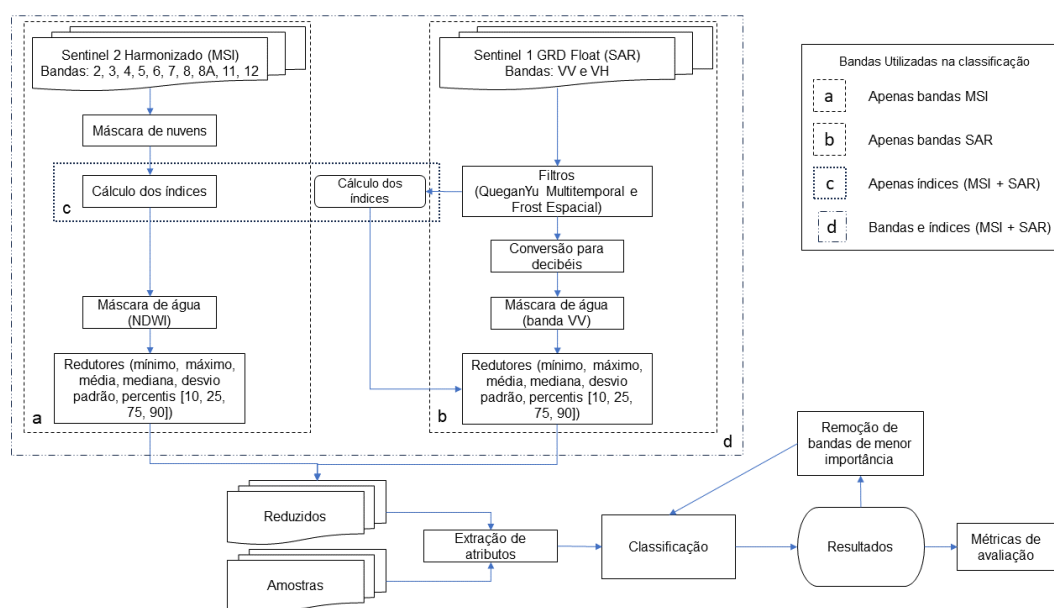
		Predição da Classificação	
		Positivo	Negativo
Verdade de Campo	Positivo	Verdadeiro Positivo	Falso Negativo
	Negativo	Falso Positivo	Verdadeiro Negativo

Fonte: adaptado de Phillips *et al.* (2024).

4. MATERIAL E MÉTODOS

Este estudo foi desenvolvido buscando identificar o comportamento na precisão/acurácia e demanda computacional de um algoritmo de classificação de imagens, baseado em RF, quando exposto a diferentes conjuntos de imagens de entrada, com o objetivo de classificar áreas de floresta e de não floresta (classe mista de alvos como solo exposto, agricultura, pastagem e vegetação rasteira) utilizando imagens de sensoriamento remoto óptico e de radar. A Figura 6 condensa as informações na forma de um fluxograma, apresentando os quatro conjuntos de imagens utilizados e os processamentos realizados, que serão apresentados a seguir.

Figura 6 Fluxograma dos conjuntos de imagens (a, b, c e d) e processos utilizados no estudo



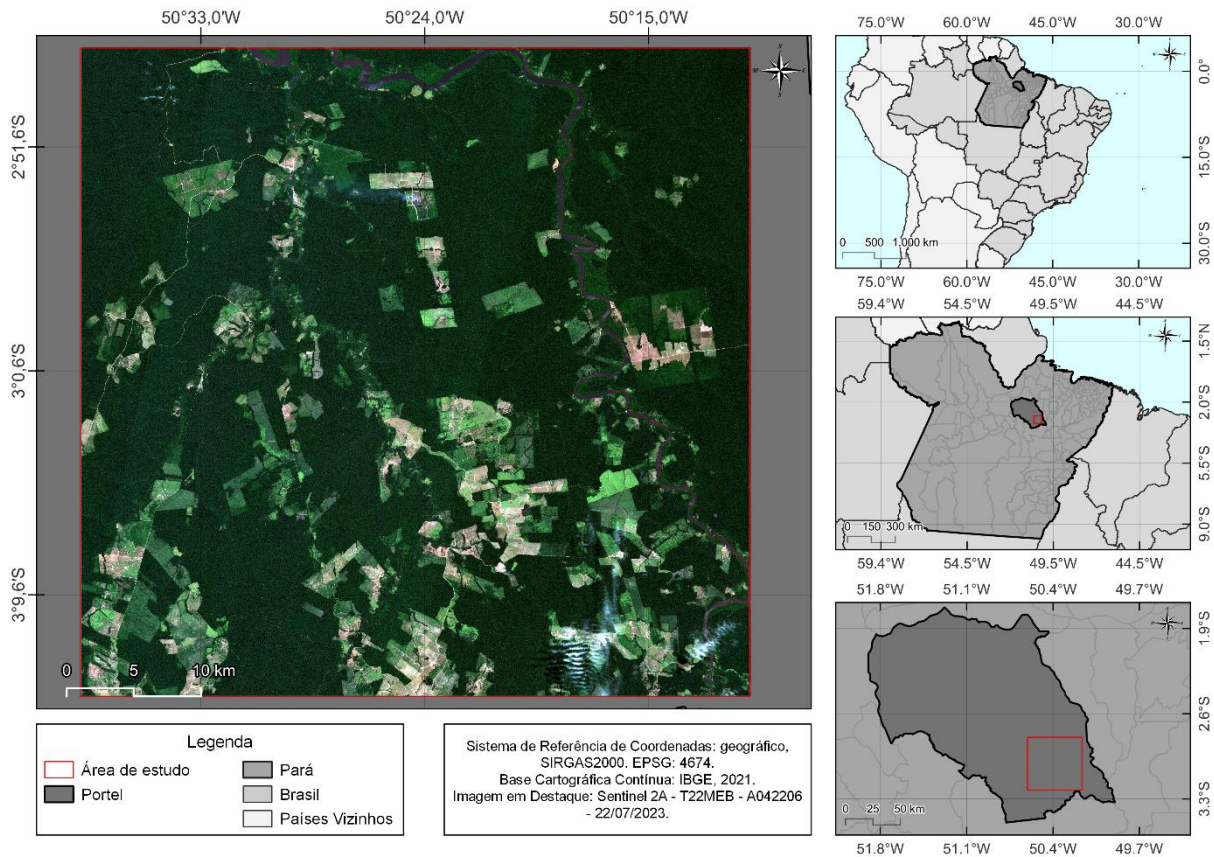
Fonte: O autor (2024).

4.1 Área de estudo

O estudo foi desenvolvido em uma área teste localizada no município de Portel, no estado do Pará, compreendida entre as coordenadas 50° 37' 51,253" O, 2° 47' 37,543" S e 50° 10' 55,850" O, 3° 13' 41,098" S (Figura 7) e possui aproximadamente 240 mil hectares. O clima na região é do tipo Am – clima tropical de monção, segundo a classificação Köppen-Geiger, sua média anual de precipitação é de 1750 – 2500 mm e possui cerca de 8 meses chuvosos e 4 meses secos, sendo alocada na classe de Agressividade Climática nível B1 (baixa agressividade com excesso de umidade) (ANDRADE et al., 2017).

Esse município foi escolhido como área de estudo por estar entre os 10 municípios que tiveram maior incremento de desmatamento no ano de 2022, com um total de 315,45 km² de área desmatada segundo dados PRODES². Outro ponto de escolha foi sua localização em uma área de extrema condição de nuvens, permanecendo com 60-70% de cobertura de nuvens nos meses secos (junho a setembro) e acima de 80% nos demais meses, condição que limita a disponibilidade de imagens úteis de sensores ópticos, dificultando a operação de classificação (PRUDENTE, 2021).

Figura 7 Mapa de localização da área de estudo.



Fonte: O autor (2024).

4.2 Softwares e processamentos

A plataforma de processamento em nuvem GEE foi utilizada para coleta de amostras, aquisição de coleções de imagem, cálculo dos índices de vegetação, redução temporal das bandas, treinamento do classificador, classificação e exportação dos produtos (classificação em formato raster). Foi utilizado o ambiente

² Disponível em: < <https://www.dpi.inpe.br/prodesdigital/prodesmunicipal.php> >

Spyder v.5.3.3 com a linguagem Python v.3.7.16 para automatização do agrupamento das importâncias, recorte de imagem, vetorização de imagem e geração dos índices IoU e Dice. O software QGIS v.3.22.11 foi utilizado para vetorização da verdade de campo, extração de pontos aleatórios e criação de mapas. O software de planilhas Excel v.2311 foi utilizado para criação das métricas de avaliação e criação dos gráficos e tabelas.

4.3 Conjuntos de imagens de entrada

As imagens foram obtidas entre as datas de 01/01/2023 e 31/12/2023, sendo que as imagens ópticas foram filtradas para utilizar apenas as que apresentarem menos de 75% de cobertura de nuvens. Ao todo foram utilizadas 80 imagens ópticas, sendo 36 do tile T22MEB que abrange toda área de estudo e 44 do tile T22MEC que captura apenas uma fração da área. Para o sensor de radar foram utilizadas 116 imagens, divididas entre imagens da órbita relativa número 68, que cobre totalmente a área e a número 170 que cobre cerca de metade da área.

O conjunto de imagens ópticas é proveniente da coleção Sentinel 2 Harmonizado³, correspondente ao sensor MSI a bordo dos satélites Sentinel 2A e 2B, já as imagens de radar são da coleção Sentinel 1 Ground Range Detection (GRD)⁴ em escala de potência bruta, obtidas pelo sensor SAR a bordo do satélite Sentinel 1A.

O sensor MSI é um sensor óptico que dispõe de 13 bandas (Tabela 2) com resoluções entre 10 e 60 metros, das quais as bandas 01, 09 e 10 são destinadas a correções atmosféricas, devido a isso elas não foram utilizadas como fonte de entrada de dados. Foi realizado reescalonamento para todas as bandas ficarem com resolução espacial de 20 metros. Sua resolução temporal é de 5 dias, alternando entre os satélites.

O sensor SAR possui um canal chamado banda C que opera na frequência de 5,405 GHz, referente às microondas, sua resolução espacial é de 10 m e seu tempo de revisita é de 12 dias para a área em estudo (TORRES *et al.*, 2012). O modo de operação utilizado foi o interferométrico de faixa larga (IW) e as polarizações foram vertical-vertical (transmitida vertical e recebida vertical - VV) e vertical-horizontal

³ Disponível em: <https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED>.

⁴ Disponível em: <https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD>.

(transmitida vertical e recebida horizontal - VH). Também foi realizado o reescalonamento para 20 m de resolução espacial.

Foram aplicados o filtro Quegan & Yu, que é um filtro linear temporal para remoção do ruído inerente ao tipo de coleta de imagens do sensor (speckle), e o filtro Frost que é um filtro espacial baseado em uma janela de pixels e possui o mesmo intuito (FROST et al., 1982; QUEGAN; YU, 2001).

Tabela 2 Informações espectrais das bandas do sensor MSI

Nome da banda	Comprimento de onda central λ_c (nm)	Comprimento de onda mínimo λ_{min} (nm)	Comprimento de onda máximo λ_c (nm)	Largura de banda $\Delta\lambda$ (nm)	Resolução Espacial (m)	Radiância de Referência $L_{ref}(W^*m^{-2}*sr^{-1}*\mu m^{-1})$	Propósito
B01	443	433	452	19	60	129	Correção atmosférica (espalhamento de aerossol).
B02	492	460	524	64	10	128	Sensível à senescência da vegetação, carotenóides, escurecimento e fundo do solo; correção atmosférica (espalhamento de aerossol).
B03	560	543	578	35	10	128	Pico verde, sensível à clorofila total na vegetação.
B04	665	649	680	31	10	108	Absorção máxima de clorofila.
B05	704	697	711	14	20	74.5	Posição da borda vermelha; consolidação da correção atmosférica/linha de base fluorescente.
B06	740	734	747	13	20	68	Posição da borda vermelha, correção atmosférica, recuperação da carga de aerossol
B07	783	773	793	20	20	67	Índice de área foliar (IAF), borda do platô do infravermelho próximo (NIR).
B08	833	782	887	105	10	103	Índice de área foliar (IAF).
B08A	865	854	875	21	20	52.5	Platô NIR, sensível à clorofila total, biomassa, IAF e proteína; referência de absorção de vapor d'água; recuperação de carga e tipo de aerossol.
B09	945	935	955	20	60	9	Absorção de vapor de água, correção atmosférica.
B10	1373	1359	1388	29	60	6	Deteção de Cirrus fina (tipo de nuvem) para correção atmosférica.
B11	1614	1569	1658	89	20	4	Sensível à lignina, amido e biomassa florestal acima do solo, separação neve/gelo/nuvens.
B12	2202	2113	2286	173	20	1.5	Avaliação das condições da vegetação mediterrânica. Distinção de solos argilosos para monitorização da erosão do solo. Distinção entre biomassa viva, biomassa morta e solo, por ex. para mapeamento de cicatrizes de queimaduras.

Fonte: adaptado de LI et al. (2018).

Após aquisição do conjunto de imagens, o conjunto óptico passou por remoção dos pixels que apresentaram mais de 45% de probabilidade de serem nuvens (de acordo com a coleção Sentinel-2: Cloud Probability⁵). Este conjunto filtrado passou pela etapa de cálculo dos índices de vegetação, apresentados na Tabela 3, que consistem em operações matemáticas utilizando as bandas do sensor a fim de destacar informações de acordo com o comportamento do alvo no espectro eletromagnético.

⁵ Disponível em: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY.

Tabela 3 Índices de vegetação para o sensor MSI

Índice	Abreviação em inglês	Fórmula	Referência
Índice de Vegetação Atmosfericamente Resistente	ARVI	$\left(\frac{((NIR - (RED - 1 * (BLUE - RED)))}{(NIR + (RED - 1 * (BLUE - RED)))} \right)$	(KAUFMAN; TANRE, 1992)
Índice de Vegetação por Diferença Normalizada Verde	GNDVI	$\frac{NIR - GREEN}{NIR + GREEN}$	(GITELSON; KAUFMAN; MERZLYAK, 1996)
Índice de Construções por Diferença Normalizada	NDBI	$\frac{SWIR1 - NIR}{SWIR1 + NIR}$	(ZHA; GAO; NI, 2003)
Diferença Normalizada da Borda do Vermelho	NDRE	$\frac{NIR - RE5}{NIR + RE5}$	(BUSCHMANN; NAGEL, 1993)
Índice de Vegetação por Diferença Normalizada	NDVI	$\frac{NIR - RED}{NIR + RED}$	(ROUSE <i>et al.</i> , 1974)
Índice de Água por Diferença Normalizada	NDWI	$\frac{GREEN - NIR}{GREEN + NIR}$	(MCFEETERS, 1996)
NDVI Borda do Vermelho 1	RENDVI1	$\frac{NIR - RE5}{NIR + RE5}$	(FORKUOR <i>et al.</i> , 2018)
NDVI Borda do Vermelho 2	RENDVI2	$\frac{RE6 - NIR}{RE6 + NIR}$	(FORKUOR <i>et al.</i> , 2018)
Posição da Borda do Vermelho do Sentinel-2	S2REP	$705 + 35 * \left(\frac{NIR + RED}{2} - RE5 \right) / (RE6 + RE5)$	(FRAMPTON <i>et al.</i> , 2013)
Índice de Vegetação Ajustado ao Solo	SAVI	$1.5 * \left(\frac{NIR - RED}{NIR + RED + 0.5} \right)$	(HUETE, 1988)
Diferença Normalizada SWIR2-NIR	SW2_N	$\frac{SWIR2 - NIR}{SWIR2 + NIR}$	(AGUILAR <i>et al.</i> , 2016)

Onde: Blue: B02; Green: B03; Red: B04; RE5: B05; RE6: B06; NIR: B08; SWIR1: B11; SWIR2: B12

Após passar pelos filtros de remoção de ruído, as polarizações VV e VH do sensor SAR foram utilizadas para cálculo dos índices, apresentados na Tabela 4, a fim de aumentar a informação disponível. Posteriormente as polarizações foram transformadas para decibéis com intuito de facilitar a interpretação dos alvos.

Tabela 4 Índices de vegetação para o sensor SAR

Índice	Abreviação em inglês	Fórmula	Referência
Radar Vegetation Index	RVI	$\sqrt{\frac{VV}{VV + VH} * \left(\frac{4 * VH}{VV + VH}\right)}$	(KIM; VAN ZYL, 2009)
Radar Vegetation Index modificado	RVI _m	$\left(\frac{4 * VH}{VV + VH}\right)$	(TRUDEL; CHARBONNEAU; LECONTE, 2012)
Cross-Ratio	CR	$VH_{log} - VV_{log}$	(DEN BESTEN et al., 2023)

Após cálculo dos índices, todas as imagens passaram pela remoção de pixels identificados como água, para isso foi utilizado o índice NDWI para as imagens ópticas e a polarização VV em decibéis para as imagens SAR.

A fim de reduzir a quantidade de bandas de entrada foram aplicados redutores temporais, ou seja, cada pixel da imagem foi analisado no tempo e sua variação de valores foi reduzida de acordo com as métricas: valor mínimo, médio, mediano, máximo, desvio padrão e percentis 10, 25, 75 e 90. Esse método reduz a dimensionalidade dos dados ao sintetizar o comportamento temporal do pixel em um conjunto de variáveis, facilitando a diferenciação de alvos no classificador (BECKER et al., 2021).

Os conjuntos de imagens compostos pelas bandas espectrais do sensor MSI, polarizações do sensor SAR e índices de vegetação de ambos os sensores foram divididos em quatro conjuntos de imagens (Tabela 5).

O primeiro foi composto apenas das bandas do sensor SAR aplicando os redutores (conjunto S1), o segundo foi utilizando as polarizações do sensor MSI aplicando os redutores (conjuntos S2), o terceiro emprega apenas os índices (MSI + SAR) aplicando os redutores (conjunto Índices) e o último soma todos os outros, portanto possui bandas e índices de ambos os sensores (conjunto Completo).

Além disso, a partir da primeira rodada de classificação utilizando o máximo de imagens em cada conjunto, procedeu-se a retirada das imagens com menor importância, até que restassem apenas as 10 imagens com maior importância. Após cada remoção de imagens classificou-se novamente a área, com os conjuntos reduzidos, a fim de verificar o comportamento da precisão com a retirada das imagens. Portanto, as classificações utilizaram 100%, 50%, 25% das imagens de cada conjunto e, também, apenas as 10 melhores imagens.

Cada conjunto de dados foi submetido a 5 classificações a fim de reduzir o efeito da aleatoriedade do classificador, possibilitando uma análise mais robusta da importância de cada banda na classificação.

Tabela 5 Conjuntos de imagens e o número de imagens de acordo com a porcentagem

Conjunto	100%	50%	25%	Top 10
S1	18	9	4	10
S2	90	45	22	10
Índices	126	63	31	10
Completo	234	117	58	10

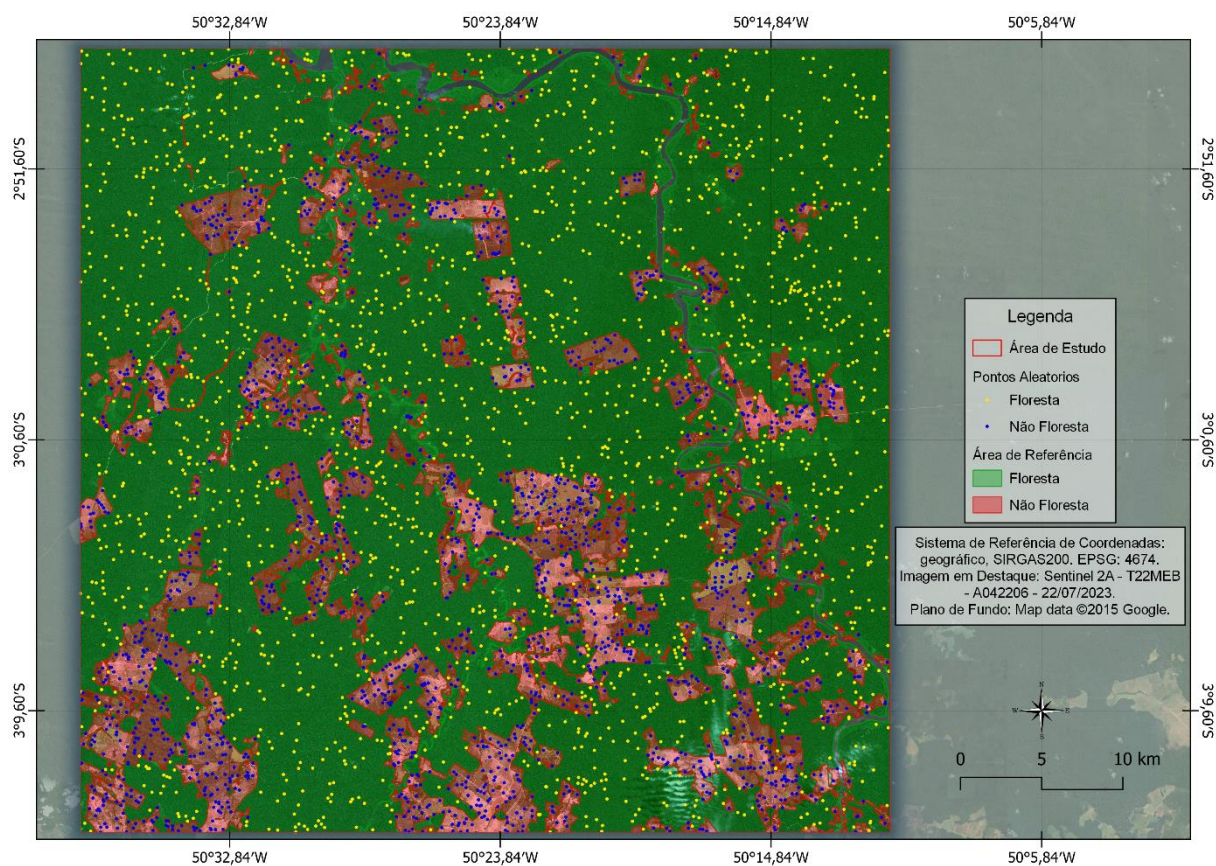
Nota: S1: sensor SAR; S2: sensor MSI; Índice e Completo: multisensor.

4.4 Coleta de amostras e classificação

A amostragem das classes floresta e não floresta foi realizada de forma visual, com base em imagens de cor verdadeira e de falsa cor do sensor MSI, levando em conta o comportamento do ano de 2023, de forma que para se alocar uma amostra, a área alvo deveria pertencer à classe (floresta ou não floresta) durante todo o período. Totalizaram 84 polígonos da classe de floresta e 96 polígonos da classe de não floresta, ao todo foram 26.975 e 6.877 pixels, respectivamente.

As classificações utilizaram o método de RF por meio da função de Classifier.smileRandomForest, presente no GEE, o número de árvores foi fixado em 100, a semente para reduzir a aleatoriedade foi fixada em 1, foram utilizados 70% das amostras para treinamento do algoritmo. A validação do algoritmo e criação da matriz de confusão foi feita com a alocação de 2 mil pontos aleatórios em cada classe (Figura 8).

Figura 8 Pontos aleatórios nas classes de floresta e não floresta, utilizados para precisão/acurácia



Fonte: O autor (2024).

4.5 Verificação de acurácia e precisão da classificação

As métricas utilizadas para a verificação de acurácia e precisão (Tabela 6) derivadas da matriz de confusão foram a Exatidão Global (EG), Acurácia do Produtor (AP), Acurácia do Usuário (AU). Também foram realizados os índices Jaccard index (IoU) e Dice, porém eles necessitaram de uma referência, para isso foi realizada a vetorização manual da área de estudo com base nas imagens do período, sendo assim, a totalidade de pixels da classificação por RF foi confrontada com a vetorização manual de referência para o período de 2023.

Tabela 6 Índices de avaliação de acurácia e precisão da classificação.

Índice	Abreviação	Fórmula	Referência
Acurácia do produtor ou Recall	Ap	$\frac{VP}{VP + FN}$	(CONGALTON, 1991)
Acurácia do usuário ou Precisão	Au	$\frac{VP}{VP + FP}$	(CONGALTON, 1991)
Exatidão Global	EG	$\frac{VP + VN}{VP + VN + FN + FP}$	(CONGALTON, 1991)
Dice ou F1	Dice	$\frac{2 * TP}{2 * TP + FP + FN}$	(POWERS, 2020)
Jaccard index	IoU	$\frac{TP}{TP + FP + FN}$	(POWERS, 2020)

A fim de testar a aplicabilidade da classificação realizada neste estudo, elaborou-se alertas de desmatamento, que consistem na identificação das áreas que eram da classe floresta e passaram para a classe não floresta durante o período do estudo, ou seja, 2023, por fim comparou-se os alertas com o produto do PRODES para o ano de 2023.

Para elaboração dos alertas, procedeu-se a classificação da área de estudo para o período de 2022 (imagens de 01/01/2022 até 31/12/2022) utilizando o conjunto de importâncias elencado como melhor custo-benefício (Completo 25% - 58 imagens), após isso foi realizada a subtração entre a classificação do período de 2023 e a classificação do período de 2022, ambas com o mesmo conjunto, resultando nas áreas potencialmente desmatadas durante o ano de 2023.

4.6 Avaliação do desempenho computacional em nuvem

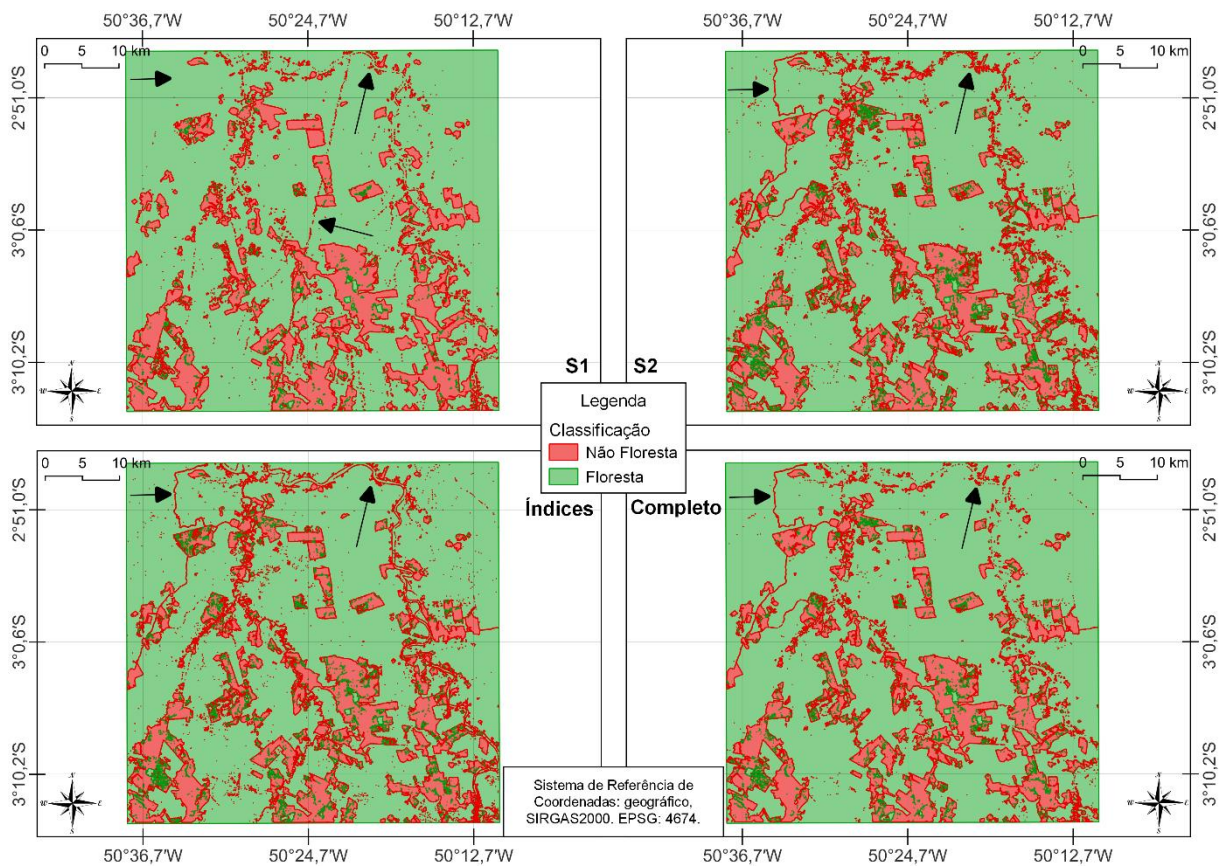
A plataforma GEE, por meio da função de executar com criador de perfil, exibe informações sobre consumo de memória e tempo de CPU (Unidade Central de Processamento) para cada função executada no script, seja ela nativa ou criada pelo usuário. A medida de uso de CPU é dada por uma representação chamada Earth Engine Compute Unit (EECU), que consiste na abstração dos diferentes núcleos de processamento e arquiteturas utilizadas pela plataforma, sua unidade é dada por EECU ao longo do tempo, nesse caso EECU-segundos (GOOGLE, 2024). Já a medida da memória é dada pelo valor de pico, ou seja, o valor máximo de memória utilizada pelo processo avaliado.

Para avaliar o desempenho do uso computacional foi utilizado o consumo total médio de EECU-segundo e de pico de memória das cinco classificações de cada conjunto de imagens.

5. RESULTADOS E DISCUSSÃO

O resultado das classificações iniciais usando 100% das bandas de cada conjunto, apresentado na Figura 9, pode-se identificar diferenças principalmente ao leito do rio que segue do norte da imagem para o sudeste, sendo que a classificação que utilizou o conjunto S2 (MSI) obteve uma detecção de não floresta maior nessa região, seguido pelo conjunto Índices (multisensor) que conseguiu delimitar as margens do rio (ressaltando que a água foi mascarada e não entrou na análise). Já o conjunto Completo (multisensor) e conjunto S1 (SAR) tiveram detecção reduzida e este último, por sua vez, não detectou a estrada na parte noroeste da imagem. Outro ponto de ruído dos dados S1 é a linha de junção das diferentes faixas de imageamento do sensor, presente como uma diagonal no centro da área de estudo.

Figura 9 Classificações dos conjuntos (S1, S2, Índices e Completo) com 100% das bandas para o ano de 2023



Fonte: O autor (2024).

Também, observa-se que há uma tendência de homogeneidade nos alvos na classificação com o conjunto S1, sendo mais sensível à vegetação parcialmente

removida, enquanto o conjunto S2 ainda identifica como floresta locais que possuem uma vegetação mais baixa e menos densa (Figura 10).

Figura 10 Zoom em uma área de diferença de classificação entre S1 e S2

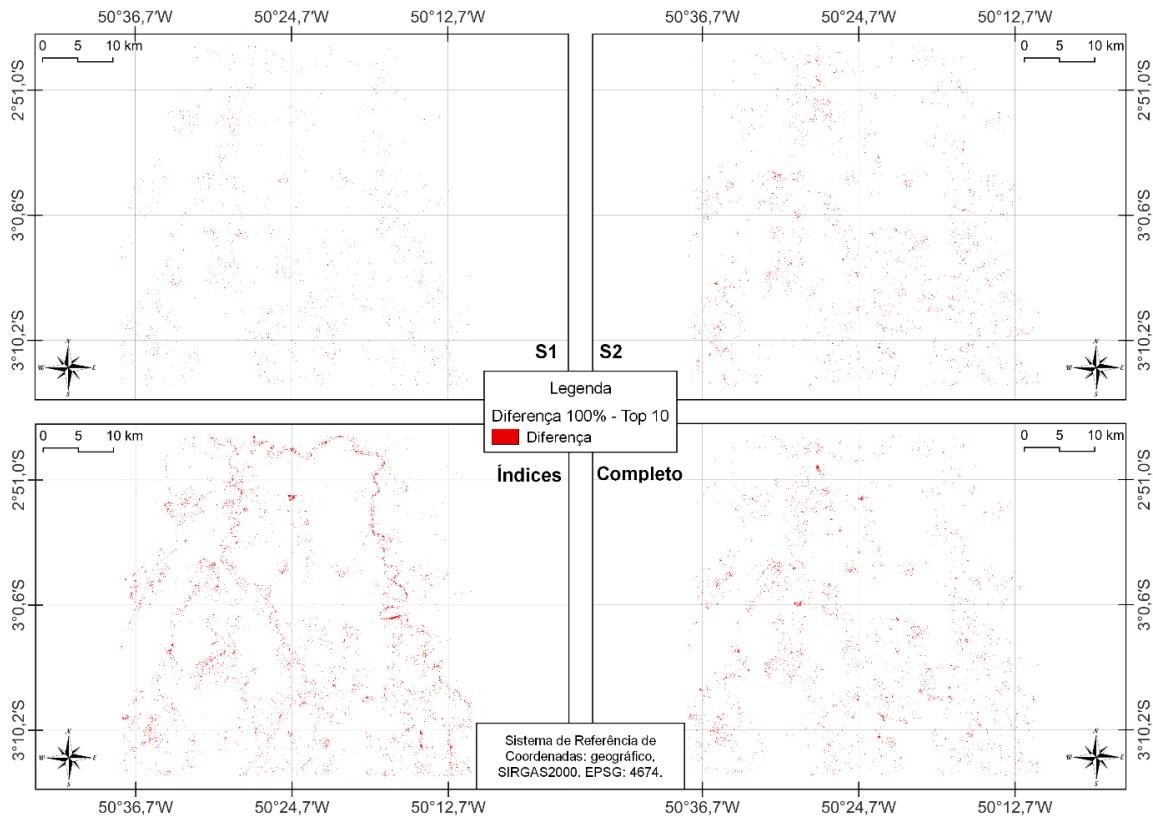


Fonte: O autor (2024).

Analisando visualmente a diferença entre as classificações com 100% das imagens e com apenas o Top10 imagens do mesmo conjunto de imagens - S1, S2, Índices e Completo (Figura 11), vê-se que no conjunto Índices apresenta a maior diferença, seguido pelo conjunto Completo. Em termos de área, os conjuntos S1, S2, Índices e Completo possuíram, respectivamente, 827, 1591, 3528 e 2230 ha de diferença na classificação. O agrupamento de maior tamanho de diferença na classificação foi encontrado no conjunto Índices, totalizando cerca de 46 ha (Figura 12), onde a área classificada como floresta é gradualmente trocada para não floresta com a redução do número das imagens de entrada.

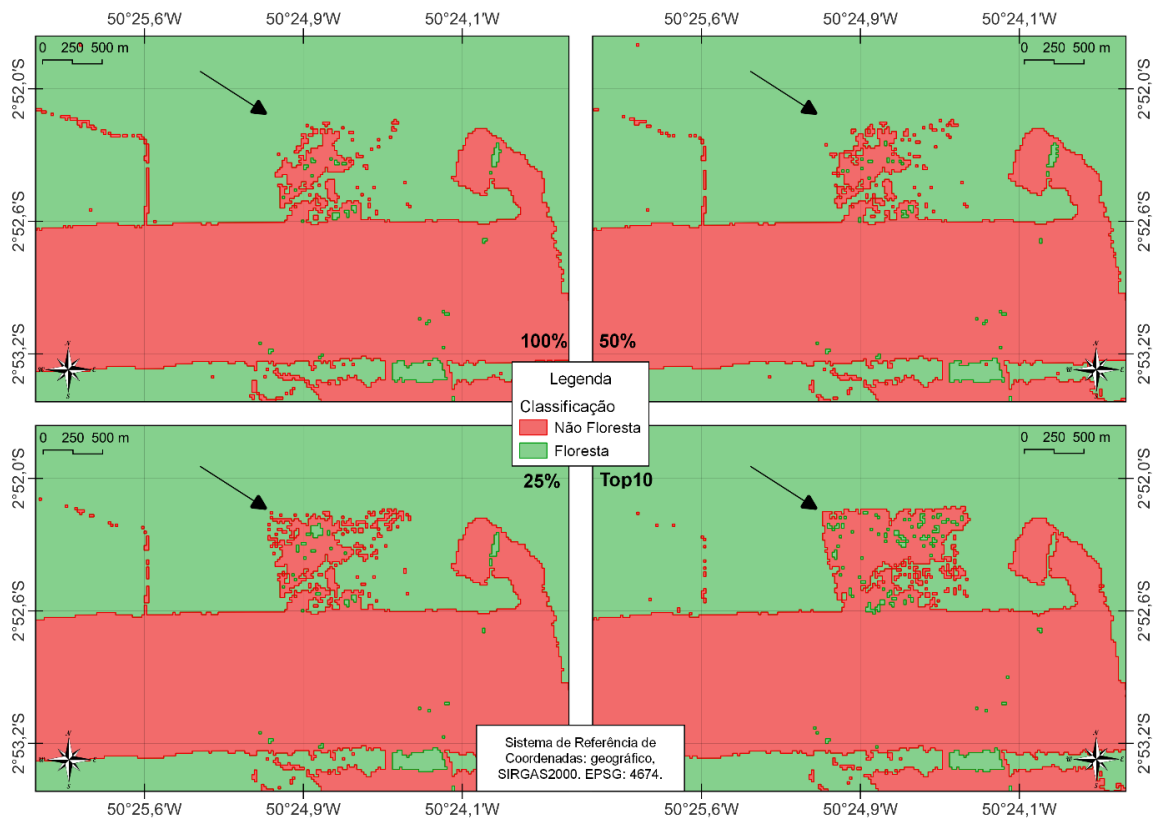
As importâncias utilizadas como guia de corte para 100%, 50%, 25% e Top10 estão apresentadas no APÊNDICE A – Tabela das importâncias, como dados complementares de identificação individual de todas as imagens utilizadas.

Figura 11 Diferença entre 100% das imagens e apenas o Top10 imagens para cada conjunto (S1, S2, Índices e Completo)



Fonte: O autor (2024).

Figura 12 Área de maior diferença encontrada no conjunto Índices ao longo da remoção de imagens

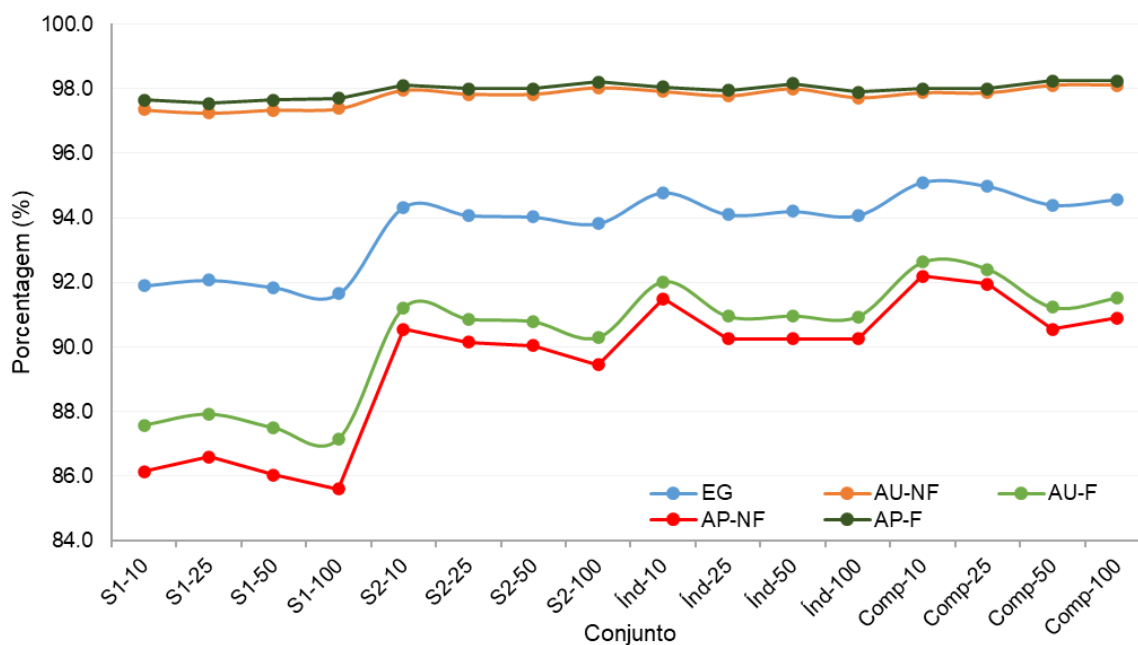


Fonte: O autor (2024).

Derivando as métricas de precisão e acurácia das matrizes de confusão (APÊNDICE B – Matrizes de confusão), por conjunto, deu-se origem ao gráficos de comportamento dos índices EG, AU, AP (Figura 13) e confrontando as classificações com a vetorização manual do ano de 2023, criou-se o gráfico dos índices IoU floresta e não floresta, Dice floresta e não floresta (Figura 14).

É possível identificar uma tendência inversa de qualidade de classificação em relação ao número de imagens, ou seja, os conjuntos performaram melhor a cada retirada de variáveis, com os picos de performance entre o Top 10 e 25% do conjunto. Outro ponto a salientar é a dominância dos conjuntos que utilizaram dados de ambos os sensores, com destaque ao conjunto Completo, que possui as maiores pontuações em todos os índices de acurácia, e seu ponto ótimo é entre 10 e 58 imagens (Top 10 e 25%). O conjunto S1, que utilizou apenas bandas de radar, obteve o pior desempenho. Apesar disso, todas as classificações figuram com EG maior que 91%, chegando a 95,1% no conjunto Completo com as imagens Top 10.

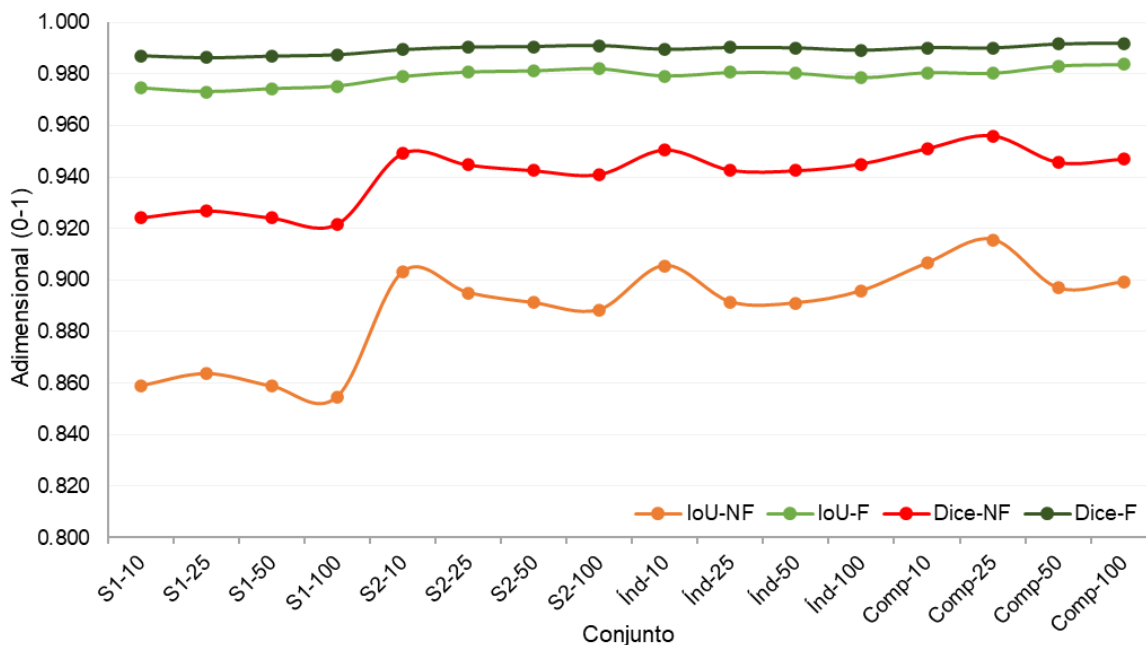
Figura 13 Índices EG, AU e AP por conjunto (S1, S2, Índices e Completo)



Nota: EG: Exatidão Global; AU-NF: Acurácia do Usuário da classe não floresta; AU-F: Acurácia do Usuário da classe floresta; AP-NF: Acurácia do Produtor da classe não floresta; AP-F: Acurácia do Produtor da classe floresta.

Fonte: O autor (2024).

Figura 14 Índices IoU e Dice para as classes floresta e não floresta por cada conjunto (S1, S2, Índices e Completo)



Nota: IoU-NF: índice de Jaccard da classe não floresta; IoU-F: índice de Jaccard da classe floresta; Dice-NF: índice Dice da classe não floresta; Dice-F: índice Dice da classe floresta.

Fonte: O autor (2024).

Observando-se a diferença entre o índice AP das duas classes, vê-se que a classe não floresta teve áreas marcadas erroneamente como floresta, ou seja, a classificação tende a classificar como floresta os pixels que deveriam ser não floresta, dessa forma ela comete erros de omissão, no mesmo sentido, verificando o índice AU das duas classes, vê-se que existem erros de comissão na classe de floresta, ou seja, os pontos de floresta estão sendo alocados onde não deveriam ser. Este comportamento pode estar atrelado ao número de pontos utilizados para treinamento de cada classe, já que a classe floresta possui quase 4 vezes mais pontos para treinamento. Outro fator que também pode ter contribuído é a mistura de alvos presente na classe não floresta, visto que ela engloba todos os alvos que não são floresta como, por exemplo, agricultura, estrada, solo exposto, vegetação rasteira e solo recentemente queimado ou revolvido, podendo ter dificultado a identificação de padrões pelo classificador.

Esses resultados corroboram com o trabalho de TAVARES *et al.* (2019) que utilizando um classificador RF para mapeamento multiclasse, em Belém (PA), obtiveram a melhor acurácia global para a integração dos dados Sentinel-1 e Sentinel-2 com 91,07%, seguida por dados apenas Sentinel-2 com 89,53%, em terceiro lugar

ficaram dados Sentinel-2 com índices radiométricos com 89,45% e o pior resultado foi apenas para os dados Sentinel-1 com 56,01%.

Os achados de CHATZIANTONIOU, PSOMIADIS e PETROPOULOS (2017) também indicam melhora na precisão global (93,85%) e no coeficiente Kappa (0,894) quando se adicionam índices e bandas Sentinel 1 ao conjunto Sentinel 2 em uma classificação de zona húmida utilizando o classificador Support Vector Machines (SVM) no norte da Grécia.

Mapeando a região do baixo Magdalena na Colômbia CLERICI, VALBUENA CALDERÓN e POSADA (2017), tiveram a precisão de classificação aumentada quando utilizaram um empilhamento de bandas Sentinel 1, Sentinel 2 e índices do infravermelho, seu trabalho comparou três classificadores SVM, RF e k-Nearest Neighbor, suas respectivas acurácias globais máximas são 88,75%, 55,50% e 39,37%, e os Kappas são 0,86, 0,49 e 0,26.

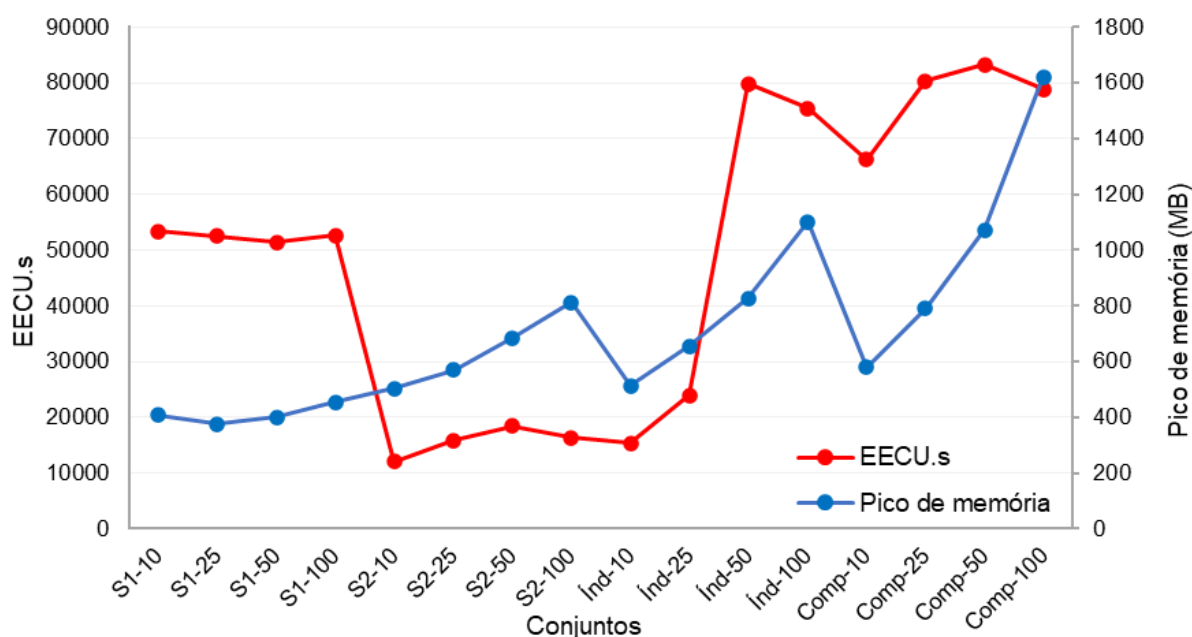
Utilizando Sentinel 1, modelo de terreno de SRTM e imagens multiespectrais RapidEye, BRAUN e HOCHSCHILD (2015) obtiveram uma acurácia global de 83,9% no seu mapeamento de uso e cobertura multiclasse com classificador RF. BADDA *et al.* (2023) ao mapear três áreas queimadas no Marrocos, obtiveram melhora da precisão com a redução de bandas de entrada no classificador RF. ARJASAKUSUMA *et al.* (2022) também tiveram melhora no desempenho de um classificador RF quando utilizaram um conjunto Sentinel 1 e Sentinel 2, chegando a uma precisão de 95,8% na detecção de áreas queimadas na Indonésia.

Na Figura 15, um gráfico de consumo de processamento (EECU.s no eixo vertical principal e memória de pico no eixo vertical secundário) foi condensado em função dos conjuntos de dados, e pode-se identificar que a coleção apenas de bandas SAR teve o menor custo de memória, porém seu processamento foi médio. A hipótese para isso ter ocorrido é o uso excessivo da função imagem *reduceNeighborhood*, que foi a maior contribuidora para o gasto de processamento, sendo aplicada no filtro de Frost para remoção do ruído speckle, de forma que em conjuntos sem imagens SAR (conjunto S2 e conjunto Índices com Top 10 e 25% de imagens) foram os que menos fizeram uso do processador.

A memória por outro lado apresenta maior consumo conforme o número de imagens aumenta, sendo a função que mais utilizou memória não possui descrição e a segunda função é de tráfego de dados. O treinamento do classificador gastou 200 kb de forma constante entre os conjuntos, devido os parâmetros de entrada serem

iguais e inertes, já a classificação no cenário com menos bandas (S1 25% - 4 imagens) usou 600 kb, mas para o conjunto de maior número de bandas (Completo 100% - 234 imagens) utilizou 4.915 kb, ou seja, houve um aumento de 8 vezes no consumo de memória.

Figura 15 Recursos computacionais utilizados – EECU.s no eixo vertical principal e pico de memória no eixo vertical secundário – para cada conjunto (S1, S2 Índices e Completo)



Fonte: O autor (2024).

A classificação deste estudo foi realizada em uma área de aproximadamente 240 mil ha, ao se estender os limites da classificação para um bioma inteiro, como o caso do Bioma Amazônia que possui cerca de 4,2 milhões de km, ou seja, 420 milhões de hectares, essa tarefa se torna computacionalmente desafiadora, ainda mais se o objetivo necessitar de classificações recorrentes. Nesse contexto todo acréscimo necessário de processamento em uma tarefa é multiplicado por diversas vezes. Essa métrica de recursos gastos é importante tanto a nível de processamento na nuvem, com planos do GEE variando entre U\$ 500,00 a U\$2.000,00 mensais⁶ com adicional dependendo do volume de uso, mas, também, para dimensionamento de máquinas de processamento local devido ao custo de peças e de manutenção.

Com base nas informações adquiridas, elegeu-se o conjunto Completo 25%, composto por 58 imagens multisensor (MSI + SAR), como melhor custo-benefício com

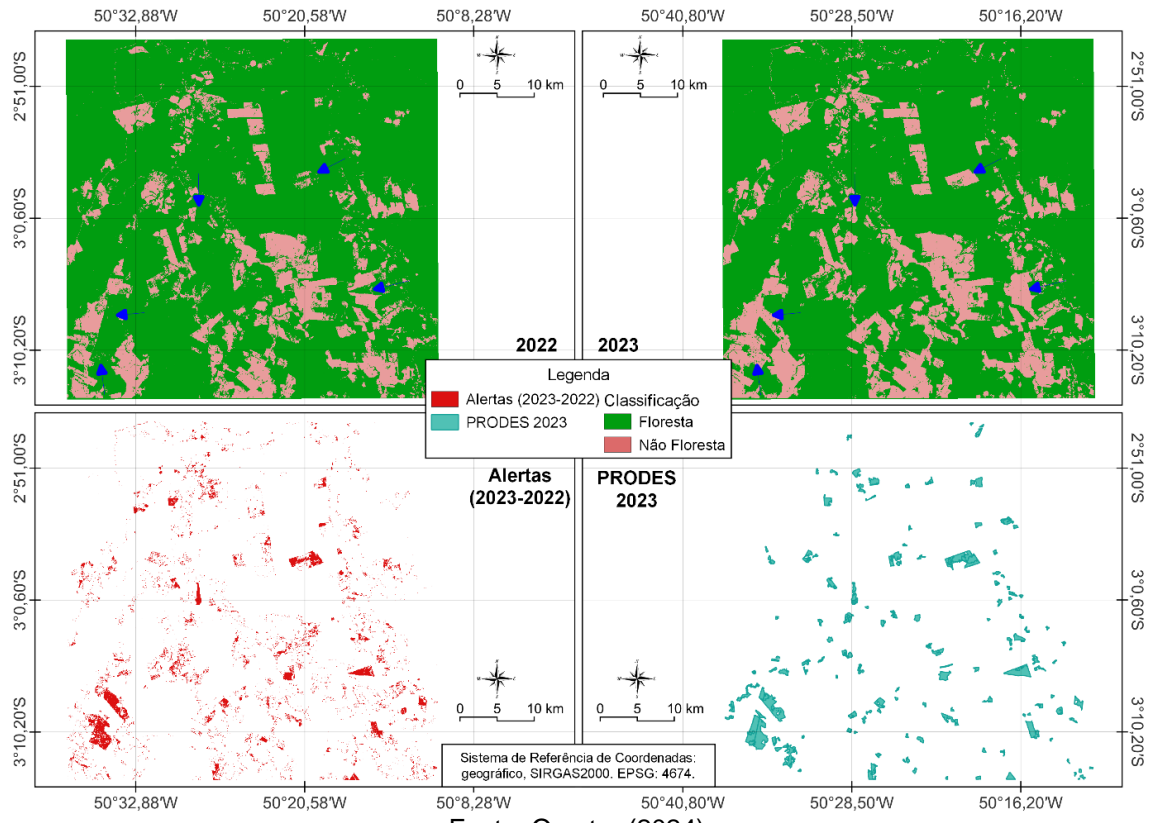
⁶ Disponível em: <<https://cloud.google.com/earth-engine/pricing?hl=pt-br>> . Acesso em: 14 jun. 2024.

requisitos computacionais medianos e acurácia muito alta. A fim de se gerar as áreas de possível desmatamento para o ano de 2023, ou seja, todos os pixels que mudaram da classe floresta para a classe não floresta durante o ano de 2023, para isso utilizou-se esse conjunto de imagens para classificar o período de 2022 (01/01/2022 até 31/12/2022), realizou-se a subtração entre as duas classificações (2023 menos 2022) obtendo uma camada de alertas de desmatamento.

Inicialmente a camada de alertas de desmatamento foi comparada ao PRODES 2023 (Figura 16) derivado de cenas prioritárias, portanto não é a classificação final do PRODES para o ano. Observa-se a presença de ruídos na camada de alertas gerada, então foi empregado um filtro espacial simples de janela 3x3 para remoção, resultando na Figura 17, onde a comparação visual com a camada do PRODES indica que há grande similaridade nos focos de desmatamento.

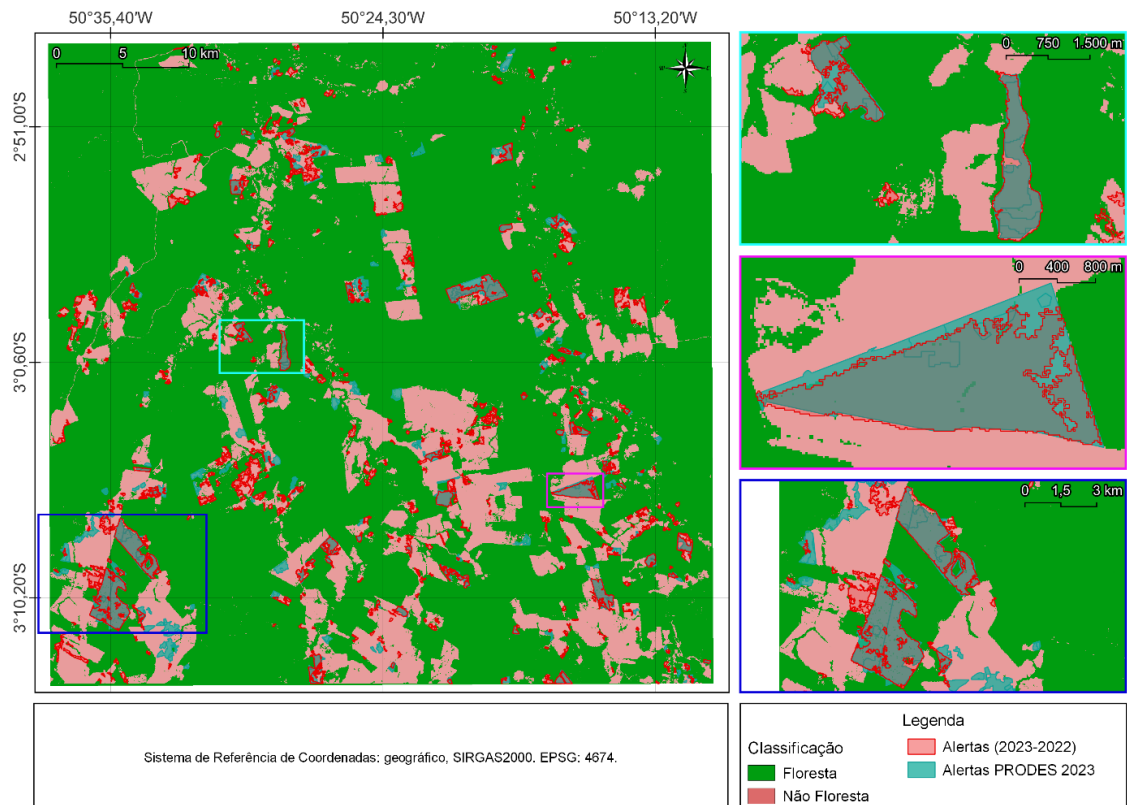
A área divulgada pelo PRODES foi de aproximadamente 6634 ha com 504 polígonos, já a área classificada como alertas de desmatamento neste trabalho foi de cerca de 7106 ha com 3778 polígonos, a diferença entre as duas pode ser devido à disponibilização de polígonos acima de 6,25 ha pelo PRODES, enquanto a classificação deste trabalho possui área mínima de 0,04 ha. Confirmando a expectativa de que as bandas desempenhariam bem a proposta de classificação para identificação de desmatamento.

Figura 16 Mapa de localização dos alertas emitidos pelo PRODES e derivados da classificação com o conjunto Completo 25%.



Fonte: O autor (2024).

Figura 17 Mapa de zoom em pontos de grande área de desmatamento comparando PRODES 2023 e Alertas (2023-2022) gerados com o conjunto Completo 25%



Fonte: O autor (2024).

Vale ressaltar que o desmatamento emitido pelo PRODES passa por conferência por meio da fotointerpretação de imagens de satélites por especialistas treinados, proporcionando confiabilidade dos dados, além disso os polígonos são disponibilizados somente a partir de 6,25 ha (ALMEIDA et al., 2022).

A partir disso, pode-se determinar que a camada de alertas de possíveis desmatamentos, feita pela subtração das classificações de 2023 e 2022 com base no conjunto de imagens Completo 25%, totalizando 58 imagens, apresenta um grande potencial para classificação de áreas desmatadas, sendo uma alternativa econômica para o monitoramento de áreas de interesse particular, abrindo um leque de possibilidade de determinação de datas arbitrárias e obtenção de polígonos de desmatamento sem área mínima pré-estabelecida, sofrendo menos interferência climática devido ao uso da abordagem multisensor.

6. CONCLUSÕES

As classificações geradas com diferentes conjuntos de imagens de entrada resultaram em produtos com diferenças perceptíveis, além disso suas precisões e acurácias diferiram entre cada conjunto e entre o próprio conjunto com quantidades diferentes de bandas. A tendência foi de elevar a precisão e acurácia, reduzindo o consumo de memória, de acordo com a redução de bandas do conjunto.

O conjunto eleito como melhor custo-benefício foi o Completo 25%, que possui 58 bandas, mesclando bandas e índices derivados dos dois sensores, aliando a diferenciação de alvos obtida com o sensor óptico MSI e a potencialidade de trabalho com condições de nuvem do sensor radar SAR. A comparação da banda de alertas derivada de classificações usando o conjunto Completo 25% em 2022 e 2023 teve boa correlação de forma e localização com os dados oficiais para o desmatamento de 2023 divulgados pelo PRODES.

7. CONSIDERAÇÕES FINAIS

Diante da insuficiência de memória e processamento disponibilizado pela plataforma GEE não foi possível realizar a remoção automática das bandas de baixa importância pelo método RFE, dessa forma se faz necessário recorrer a outras linguagens de programação e ao processamento local para testes mais criteriosos acerca do número exato de bandas ótimas, além de sua identificação.

REFERÊNCIAS

- ABDULRAHMAN, S. A.; KHALIFA, W.; ROUSHDY, M.; SALEM, A.-B. M. Comparative Study for 8 Computational Intelligence Algorithms for Human Identification. **Computer science review**, v. 36, p. 100237, 2020. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S1574013719303740>>. Acesso em: 2 fev. 2024.
- AGUILAR, M.; NEMMAOUI, A.; NOVELLI, A.; AGUILAR, F.; GARCÍA LORCA, A. Object-Based Greenhouse Mapping Using Very High Resolution Satellite Data and Landsat 8 Time Series. **Remote sensing**, v. 8, n. 6, p. 513, 2016. Disponível em: <<https://www.mdpi.com/2072-4292/8/6/513>>. Acesso em: 7 fev. 2024.
- ALIDOOST, M.; GHODRATI, V.; AHMADIAN, A.; SHAFIEE, A.; HASSANI, C. H.; BEDAYAT, A.; WILSON, J. L. Model Utility of a Deep Learning-Based Segmentation Is Not Dice Coefficient Dependent: A Case Study in Volumetric Brain Blood Vessel Segmentation. **Intelligence-based medicine**, v. 7, p. 100092, 2023. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2666521223000066>>. Acesso em: 3 fev. 2024.
- ALMEIDA, C. A. de; MAURANO, L. E. P.; VALERIANO, D. de M.; CÂMARA, G.; VINHAS, L.; MOTTA, M. da; GOMES, A. R.; MONTEIRO, A. M. V.; SOUZA, A. A. de A.; MESSIAS, C. G.; RENNÓ, C. D.; MARCOS, M. S. A.; SOLER, L. de S.; AMARAL, S. **METODOLOGIA UTILIZADA NOS SISTEMAS PRODES E DETER - 2a EDIÇÃO (ATUALIZADA)**. São José dos Campos: INPE, 2022. Disponível em: <<http://mtc-m21d.sid.inpe.br/col/sid.inpe.br/mtc-m21d/2022/08.25.11.46/doc/thisInformationItemHomePage.html>>. Acesso em: 30 jan. 2024.
- ANDRADE, V. M. S. de; CORDEIRO, I. M. C. C.; SCHWARTZ, G.; RANGEL-VASCONCELOS, L. G. T.; OLIVEIRA, F. de A. Considerações sobre clima e aspectos edafoclimáticos da mesorregião Nordeste paraense. *Em*: CORDEIRO, I. M. C. C.; RANGEL-VASCONCELOS, L. G. T.; SCHWARTZ, G.; OLIVEIRA, F. DE A. **Nordeste Paraense: panorama geral e uso sustentável das florestas secundárias**. Belém: EDUFRA, 2017. p. 59–96.
- ARJASAKUSUMA, S.; KUSUMA, S. S.; VETRITA, Y.; PRASASTI, I.; ARIEF, R. Monthly Burned-Area Mapping Using Multi-Sensor Integration of Sentinel-1 and Sentinel-2 and Machine Learning: Case Study of 2019's Fire Events in South Sumatra Province, Indonesia. **Remote Sensing Applications Society and Environment**, v. 27, p. 100790, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2352938522000982>>. Acesso em: 31 jan. 2024.
- ASADOLLAH, S. B. H. S.; JODAR-ABELLAN, A.; PARDO, M. Á. Optimizing Machine Learning for Agricultural Productivity: A Novel Approach with RScv and Remote Sensing Data over Europe. **Agricultural systems**, v. 218, n. 103955, p. 103955, 2024. Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S0308521X24001057>>. Acesso em: 14 jun. 2024.

BADDA, H.; CHERIF, E. K.; BOULAASSAL, H.; WAHBI, M.; YAZIDI ALAOU, O.; MAATOUK, M.; BERNARDINO, A.; COREN, F.; EL KHARKI, O. Improving the Accuracy of Random Forest Classifier for Identifying Burned Areas in the Tangier-Tetouan-Al Hoceima Region Using Google Earth Engine. **Remote sensing**, v. 15, n. 17, p. 4226, 2023. Disponível em: <<https://www.mdpi.com/2072-4292/15/17/4226>>. Acesso em: 3 fev. 2024.

BANSAL, M.; GOYAL, A.; CHOUDHARY, A. A Comparative Analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory Algorithms in Machine Learning. **Decision Analytics Journal**, v. 3, p. 100071, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2772662222000261>>. Acesso em: 2 fev. 2024.

BAZZAN, T.; RENNÓ, C. D.; LIMA, D. L. C.; RECKZIEGEL, E. W. INTEGRAÇÃO DA INCERTEZA NA AMOSTRAGEM E CLASSIFICAÇÃO RANDOM FOREST UTILIZANDO BANDAS E ÍNDICES ESPECTRAIS PARA O MAPEAMENTO DE INUNDAÇÃO. **Geociências**, v. 41, n. 4, p. 905–925, 2023. Disponível em: <<https://www.periodicos.rc.biblioteca.unesp.br/index.php/geociencias/article/view/16802>>. Acesso em: 3 fev. 2024.

BECKER, W. R.; LÓ, T. B.; JOHANN, J. A.; MERCANTE, E. Statistical Features for Land Use and Land Cover Classification in Google Earth Engine. **Remote Sensing Applications Society and Environment**, v. 21, n. 100459, p. 100459, 2021. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2352938520306340>>. Acesso em: 4 maio. 2024.

BEZERRA, B. G.; SILVA, B. B. da; FERREIRA, N. J. Estimativa da evapotranspiração real diária utilizando-se imagens digitais TM - Landsat 5. **Revista Brasileira de Meteorologia**, [S.L.], v. 23, n. 3, p. 305-317, set. 2008. Disponível em: https://www.scielo.br/scielo.php?pid=s010277862008000300005&script=sci_abstract&tlng=p. Acesso em: 04 fev. 2021

BIAU, G.; SCORNET, E. A Random Forest Guided Tour. **Test (Madrid, Spain)**, v. 25, n. 2, p. 197–227, 2016. Disponível em: <<https://link-springer-com.ez89.periodicos.capes.gov.br/article/10.1007/s11749-016-0481-7#ref-CR21>>. Acesso em: 3 fev. 2024.

BRASIL, Ministério do Meio Ambiente. **ESTRATÉGIA DO PROGRAMA DE MONITORAMENTO AMBIENTAL DOS BIOMAS BRASILEIROS**. 2. ed. Brasília, DF: MMA, 2017. 51 p.

BRAUN, A.; HOCHSCHILD, V. Combining SAR and Optical Data for Environmental Assessments around Refugee Camps. **GI Forum-Journal for Geographic Information Science**, v. 1, p. 424–433, 2015. Disponível em: <https://www.austriaca.at/buecher/Open%20Access/GI_Forum/GI_Forum_2015___8>

211__Geospatial_Minds_for_Society/3_07_10114_FP_Braun_10.pdf?ts=1544310483161>. Acesso em: 8 fev. 2024.

BREIMAN, L. Random Forests. **Machine learning**, v. 45, n. 1, p. 5–32, 2001. Disponível em: <<https://link.springer.com/article/10.1023/A:1010933404324>>. Acesso em: 2 fev. 2024.

BUSCHMANN, C.; NAGEL, E. In Vivo Spectroscopy and Internal Optics of Leaves as Basis for Remote Sensing of Vegetation. **International journal of remote sensing**, v. 14, n. 4, p. 711–722, 1993. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01431169308904370>>. Acesso em: 7 fev. 2024.

CHATZIANTONIOU, A.; PSOMIADIS, E.; PETROPOULOS, G. Co-Orbital Sentinel 1 and 2 for LULC Mapping with Emphasis on Wetlands in a Mediterranean Setting Based on Machine Learning. **Remote sensing**, v. 9, n. 12, p. 1259, 2017. Disponível em: <<https://www.mdpi.com/2072-4292/9/12/1259>>. Acesso em: 31 jan. 2024.

CHOWDHURY, M. S. Comparison of Accuracy and Reliability of Random Forest, Support Vector Machine, Artificial Neural Network and Maximum Likelihood Method in Land Use/Cover Classification of Urban Setting. **Environmental Challenges**, v. 14, p. 100800, 2024. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2667010023001233>>. Acesso em: 1 fev. 2024.

CLERICI, N.; VALBUENA CALDERÓN, C. A.; POSADA, J. M. Fusion of Sentinel-1A and Sentinel-2A Data for Land Cover Mapping: A Case Study in the Lower Magdalena Region, Colombia. **Journal of maps**, v. 13, n. 2, p. 718–726, 2017. Disponível em: <<https://www-tandfonline.ez89.periodicos.capes.gov.br/doi/full/10.1080/17445647.2017.1372316>>. Acesso em: 8 fev. 2024.

CONGALTON, R. G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. **Remote sensing of environment**, v. 37, n. 1, p. 35–46, 1991. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/003442579190048B>>. Acesso em: 7 fev. 2024.

DAMASCENO, A. S. da S.; BOECHAT, C. L.; SOUZA, H. A. de; CAPRISTO-SILVA, G. F.; MENDES, W. de S.; TEODORO, P. E.; MORAIS, P. G. C.; OLIVEIRA, R. I. de; DELLA-SILVA, J. L.; SOUZA, I. A. M. de; SILVA JUNIOR, C. A. da. Nutritional Monitoring of Boron in Eucalyptus Spp. in the Brazilian Cerrado by Multispectral Bands of the MSI Sensor (Sentinel-2). **Remote Sensing Applications Society and Environment**, v. 29, p. 100913, 2023. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S235293852200221X>>. Acesso em: 31 jan. 2024.

DEFRIES, R. S.; CHAN, J. C.-W. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. **Remote sensing of environment**, v. 74, n. 3, p. 503–515, 2000. Disponível em: <<https://www->

sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0034425700001425#BIB22>. Acesso em: 1 fev. 2024.

DEFRIES, R.; HANSEN, M.; TOWNSHEND, J. Global Discrimination of Land Cover Types from Metrics Derived from AVHRR Pathfinder Data. **Remote sensing of environment**, v. 54, n. 3, p. 209–222, 1995. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/0034425795001425?via%3Dihub>>. Acesso em: 3 fev. 2024.

DEN BESTEN, N.; STEELE DUNNE, S.; MAHMUD, A.; JACKSON, D.; AOUIZERATS, B.; DE JEU, R.; BURGER, R.; HOUBORG, R.; MCGLINCHEY, M.; VAN DER ZAAG, P. Understanding Sentinel-1 Backscatter Response to Sugarcane Yield Variability and Waterlogging. **Remote sensing of environment**, v. 290, n. 113555, p. 113555, 2023. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425723001062>>. Acesso em: 7 fev. 2024.

FEARNSIDE, P. M. Desmatamento na Amazônia: dinâmica, impactos e controle. **Acta amazonica**, v. 36, n. 3, p. 395–400, 2006. Disponível em: <<https://www.scielo.br/j/aa/a/97R3WH4L93x74Sr5q7X7BVy/?lang=pt>>. Acesso em: 29 jan. 2024.

FORKUOR, G.; DIMOBE, K.; SERME, I.; TONDOH, J. E. Landsat-8 vs. Sentinel-2: Examining the Added Value of Sentinel-2's Red-Edge Bands to Land-Use and Land-Cover Mapping in Burkina Faso. **GIScience & remote sensing**, v. 55, n. 3, p. 331–354, 2018. Disponível em: <<https://www-tandfonline.ez89.periodicos.capes.gov.br/doi/full/10.1080/15481603.2017.1370169>>. Acesso em: 7 fev. 2024.

FORMAGGIO, A. R.; SANCHES, I. D. **Sensoriamento remoto em Agricultura**. São Paulo, SP: Oficina de Textos, 2017. 288 p.

FRAMPTON, W. J.; DASH, J.; WATMOUGH, G.; MILTON, E. J. Evaluating the Capabilities of Sentinel-2 for Quantitative Estimation of Biophysical Variables in Vegetation. **ISPRS journal of photogrammetry and remote sensing: official publication of the International Society for Photogrammetry and Remote Sensing (ISPRS)**, v. 82, p. 83–92, 2013. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S092427161300107X>>. Acesso em: 7 fev. 2024.

FROST, V. S.; STILES, J. A.; SHANMUGAN, K. S.; HOLTZMAN, J. C. A model for radar images and its application to adaptive digital filtering of multiplicative noise. **IEEE transactions on pattern analysis and machine intelligence**, v. PAMI-4, n. 2, p. 157–166, 1982. Disponível em: <<https://ieeexplore-ieee-org.ez89.periodicos.capes.gov.br/document/4767223>>. Acesso em: 7 fev. 2024.

GITELSON, A. A.; KAUFMAN, Y. J.; MERZLYAK, M. N. Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS. **Remote sensing of environment**, v. 58, n. 3, p. 289–298, 1996. Disponível em:

<<https://www.sciencedirect.com/science/article/abs/pii/S0034425796000727>>. Acesso em: 7 fev. 2024.

GOOGLE. **Computation overview**. Disponível em: <https://developers.google.com/earth-engine/guides/computation_overview>. Acesso em: 7 fev. 2024.

GUYON, I.; WESTON, J.; BARNHILL, S.; VAPNIK, V. Gene Selection for Cancer Classification using Support Vector Machines. **Machine learning**, v. 46, p. 389–422, 2002. Disponível em: <<https://link.springer.com/article/10.1023/A:1012487302797>>. Acesso em: 3 fev. 2024.

HADDAD, N. M.; BRUDVIG, L. A.; CLOBERT, J.; DAVIES, K. F.; GONZALEZ, A.; HOLT, R. D.; LOVEJOY, T. E.; SEXTON, J. O.; AUSTIN, M. P.; COLLINS, C. D.; COOK, W. M.; DAMSCHEN, E. I.; EWERS, R. M.; FOSTER, B. L.; JENKINS, C. N.; KING, A. J.; LAURANCE, W. F.; LEVEY, D. J.; MARGULES, C. R.; MELBOURNE, B. A.; NICHOLLS, A. O.; ORROCK, J. L.; SONG, D.-X.; TOWNSHEND, J. R. Habitat Fragmentation and Its Lasting Impact on Earth's Ecosystems. **Science advances**, v. 1, n. 2, p. e1500052, 2015. Disponível em: <<https://www.science.org/doi/10.1126/sciadv.1500052>>.

HUETE, A. R. A Soil-Adjusted Vegetation Index (SAVI). **Remote sensing of environment**, v. 25, n. 3, p. 295–309, 1988. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/003442578890106X>>. Acesso em: 7 fev. 2024.

HUSSAIN, S.; MUBEEN, M.; KARUPPANNAN, S. Land Use and Land Cover (LULC) Change Analysis Using TM, ETM+ and OLI Landsat Images in District of Okara, Punjab, Pakistan. **Physics and chemistry of the Earth, Parts A/B/C**, v. 126, p. 103117, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S1474706522000122>>. Acesso em: 31 jan. 2024.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres**. 2. ed. São José dos Campos: Parênteses, 2009. 2009

KAUFMAN, Y. J.; TANRE, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. **IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society**, v. 30, n. 2, p. 261–270, 1992. Disponível em: <<https://ieeexplore-ieee-org.ez89.periodicos.capes.gov.br/document/134076>>. Acesso em: 7 fev. 2024.

KIM, Y.; VAN ZYL, J. J. A time-series approach to estimate soil moisture using polarimetric radar data. **IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society**, v. 47, n. 8, p. 2519–2527, 2009. Disponível em: <<https://ieeexplore-ieee-org.ez89.periodicos.capes.gov.br/document/4810143>>. Acesso em: 7 fev. 2024.

KOENEMAN, S. H.; CAVANAUGH, J. E. An Improved Asymptotic Test for the Jaccard Similarity Index for Binary Data. **Statistics & probability letters**, v. 184, p.

109375, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0167715222000074>>. Acesso em: 3 fev. 2024.

LAURANCE, W. F.; SAYER, J.; CASSMAN, K. G. Agricultural Expansion and Its Impacts on Tropical Nature. **Trends in ecology & evolution**, v. 29, n. 2, p. 107–116, 2014. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0169534713002929?via%3Dihub>>.

LI, Y.; CHEN, J.; MA, Q.; ZHANG, H. K.; LIU, J. Evaluation of sentinel-2A surface reflectance derived using Sen2Cor in north America. **IEEE journal of selected topics in applied earth observations and remote sensing**, v. 11, n. 6, p. 1997–2021, 2018. Disponível em: <<https://ieeexplore-ieee-org.ez89.periodicos.capes.gov.br/document/8386426>>. Acesso em: 31 jan. 2024.

MACHIDA, F.; ANDRADE, E. Modeling and Analysis of Deforestation Prevention by Uncrewed Aerial Vehicles-Based Monitoring Systems. **Environmental modelling & software: with environment data news**, v. 158, n. 105540, p. 105540, 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1364815222002407>>. Acesso em: 13 jun. 2024.

MAO, Y.; VAN NIEL, T. G.; MCVICAR, T. R. Reconstructing Cloud-Contaminated NDVI Images with SAR-Optical Fusion Using Spatio-Temporal Partitioning and Multiple Linear Regression. **ISPRS journal of photogrammetry and remote sensing**, v. 198, p. 115–139, 2023. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0924271623000618>>. Acesso em: 31 jan. 2024.

MARTÍNEZ M., L. J. Relationship between crop nutritional status, spectral measurements and Sentinel 2 images. **Agromonia colombiana**, v. 35, n. 2, p. 205–215, 2017. Disponível em: <<https://revistas.unal.edu.co/index.php/agrocol/article/view/62875>>. Acesso em: 13 jun. 2024.

MCFEETERS, S. K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. **International journal of remote sensing**, v. 17, n. 7, p. 1425–1432, 1996. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01431169608948714>>. Acesso em: 7 fev. 2024.

MEHMOOD, M.; SHAHZAD, A.; ZAFAR, B.; SHABBIR, A.; ALI, N. Remote Sensing Image Classification: A Comprehensive Review and Applications. **Mathematical problems in engineering**, v. 2022, p. 1–24, 2022. Disponível em: <<https://www.hindawi.com/journals/mpe/2022/5880959/>>. Acesso em: 1 fev. 2024.

MENESES, P. R.; ALMEIDA, T. de. **Introdução ao Processamento de Imagens de Sensoriamento Remoto**. Brasília: [s. n.], 2012. 2012.

MING, D.; ZHOU, T.; WANG, M.; TAN, T. Land Cover Classification Using Random Forest with Genetic Algorithm-Based Parameter Optimization. **Journal of applied remote sensing**, v. 10, n. 3, p. 035021, 2016. Disponível em: <<https://www.spiedigitallibrary.org/journals/journal-of-applied-remote-sensing/volume-10/issue-3/035021/Land-cover-classification-using-random-forest-with-genetic-algorithm-based/10.1117/1.JRS.10.035021.short>>. Acesso em: 31 jan. 2024.

OLDONI, L. V.; DEL'ARCO SANCHES, I.; PICOLI, M. C. A.; ROHDEN PRUDENTE, V. H.; ADAMI, M. Geometric Accuracy Assessment and a Framework for Automatic Sub-Pixel Registration of WFI Images from CBERS-4, CBERS-4A, and Amazonia-1 Satellites over Brazil. **Remote Sensing Applications Society and Environment**, v. 28, p. 100844, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2352938522001525>>. Acesso em: 31 jan. 2024.

OLDONI, L. V.; PRUDENTE, V. H. R.; DINIZ, J. M. F. S.; WIEDERKEHR, N. C.; SANCHES, I. D.; GAMA, F. F. Polarimetric Sar Data from Sentinel-1a Applied to Early Crop Classification. **ISPRS - International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences**, v. XLIII-B3-2020, p. 1039–1046, 2020. Disponível em: <<https://isprs-archives.copernicus.org/articles/XLIII-B3-2020/1039/2020/isprs-archives-XLIII-B3-2020-1039-2020.html>>. Acesso em: 31 jan. 2024.

PAL, M. Random Forest Classifier for Remote Sensing Classification. **International journal of remote sensing**, v. 26, n. 1, p. 217–222, 2005. Disponível em: <<https://www-tandfonline.ez89.periodicos.capes.gov.br/doi/full/10.1080/01431160412331269698>>. Acesso em: 3 fev. 2024.

PHILLIPS, G.; TEIXEIRA, H.; KELLY, M. G.; SALAS HERRERO, F.; VÁRBÍRÓ, G.; LYCHE SOLHEIM, A.; KOLADA, A.; FREE, G.; POIKANE, S. Setting Nutrient Boundaries to Protect Aquatic Communities: The Importance of Comparing Observed and Predicted Classifications Using Measures Derived from a Confusion Matrix. **The Science of the total environment**, v. 912, p. 168872, 2024. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0048969723075010>>. Acesso em: 3 fev. 2024.

POWERS, D. M. W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. **International Journal of Machine Learning Technology**, v. 2, n. 1, p. 37–63, 2020. Disponível em: <<https://arxiv.org/ftp/arxiv/papers/2010/2010.16061.pdf>>. Acesso em: 7 fev. 2024.

PRUDENTE, V. H. R. **Multisensor optical-sar approach to land use and land cover characterization in Roraima**. 2021. Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, 2021. Disponível em: <<http://mtc-m21d.sid.inpe.br/col/sid.inpe.br/mtc-m21d/2021/08.12.23.51/doc/publicacao.pdf>>. Acesso em: 31 jan. 2024.

QUEGAN, S.; YU, J. J. Filtering of multichannel SAR images. **IEEE transactions on geoscience and remote sensing**, v. 39, n. 11, p. 2373–2379, 2001. Disponível em: <<https://ieeexplore-ieee-org.ez89.periodicos.capes.gov.br/document/964973>>. Acesso em: 7 fev. 2024.

RODRIGUEZ-GALIANO, V. F.; GHIMIRE, B.; ROGAN, J.; CHICA-OLMO, M.; RIGOL-SANCHEZ, J. P. An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. **ISPRS journal of photogrammetry and remote sensing**, v. 67, p. 93–104, 2012. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0924271611001304>>. Acesso em: 3 fev. 2024.

ROMÁN, M. O.; JUSTICE, C.; PAYNTER, I.; BOUCHER, P. B.; DEVADIGA, S.; ENDSLEY, A.; ERB, A.; FRIEDL, M.; GAO, H.; GIGLIO, L.; GRAY, J. M.; HALL, D.; HULLEY, G.; KIMBALL, J.; KNYAZIKHIN, Y.; LYAPUSTIN, A.; MYNENI, R. B.; NOOJIPADY, P.; PU, J.; RIGGS, G.; SARKAR, S.; SCHAAF, C.; SHAH, D.; TRAN, K. H.; VERMOTE, E.; WANG, D.; WANG, Z.; WU, A.; YE, Y.; SHEN, Y.; ZHANG, S.; ZHANG, S.; ZHANG, X.; ZHAO, M.; DAVIDSON, C.; WOLFE, R. Continuity between NASA MODIS Collection 6.1 and VIIRS Collection 2 Land Products. **Remote sensing of environment**, v. 302, n. 113963, p. 113963, 2024. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0034425723005151>>. Acesso em: 31 jan. 2024.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the Great Plains with ERTS. (S. C. Freden, E. P. Mercanti, M. A. Becker) Em: **Third Earth Resources Technology Satellite-1 Symposium**, 1974, Washington. [...]. Washington: NASA, 1974. p. 309–317.

RUIZ, R. S. R.; DE CAMPOS VELHO, H. F.; SANTOS, R. D. C.; TREVISAN, M. Árvores de Decisão na Classificação de Dados Astronômicos. **TEMA. Tendências em Matemática Aplicada e Computacional**, v. 10, n. 1, p. 75–86, 2009. Disponível em: <https://www.researchgate.net/publication/272800585_Arvores_de_Decisao_na_Classificacao_de_Dados_Astronomicos>. Acesso em: 2 fev. 2024.

SALES, V. G.; STROBL, E.; ELLIOTT, R. J. R. Cloud Cover and Its Impact on Brazil's Deforestation Satellite Monitoring Program: Evidence from the Cerrado Biome of the Brazilian Legal Amazon. **Applied geography (Sevenoaks, England)**, v. 140, n. 102651, p. 102651, 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0143622822000224>>. Acesso em: 13 jun. 2024.

TAVARES, P.; BELTRÃO, N.; GUIMARÃES, U.; TEODORO, A. Integration of Sentinel-1 and Sentinel-2 for Classification and LULC Mapping in the Urban Area of Belém, Eastern Brazilian Amazon. **Sensors**, v. 19, n. 5, p. 1140, 2019. Disponível em: <<https://www.mdpi.com/1424-8220/19/5/1140>>. Acesso em: 31 jan. 2024.

TORRES, R.; SNOEIJ, P.; GEUDTNER, D.; BIBBY, D.; DAVIDSON, M.; ATTEMA, E.; POTIN, P.; ROMMEN, B.; FLOURY, N.; BROWN, M.; TRAVER, I. N.; DEGHAJE,

P.; DUESMANN, B.; ROSICH, B.; MIRANDA, N.; BRUNO, C.; L'ABBATE, M.; CROCI, R.; PIETROPAOLO, A.; HUCHLER, M.; ROSTAN, F. GMES Sentinel-1 Mission. **Remote sensing of environment**, v. 120, p. 9–24, 2012. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0034425712000600?via%3Dihub>>. Acesso em: 4 maio. 2024.

TORO, A. P. S. G. D. D.; BUENO, I. T.; WERNER, J. P. S.; ANTUNES, J. F. G.; LAMPARELLI, R. A. C.; COUTINHO, A. C.; ESQUERDO, J. C. D. M.; MAGALHÃES, P. S. G.; FIGUEIREDO, G. K. D. A. SAR and Optical Data Applied to Early-Season Mapping of Integrated Crop–Livestock Systems Using Deep and Machine Learning Algorithms. **Remote sensing**, v. 15, n. 4, p. 1130, 2023. Disponível em: <<https://www.mdpi.com/2072-4292/15/4/1130>>. Acesso em: 31 jan. 2024.

TRUDEL, M.; CHARBONNEAU, F.; LECONTE, R. Using RADARSAT-2 Polarimetric and ENVISAT-ASAR Dual-Polarization Data for Estimating Soil Moisture over Agricultural Fields. **Canadian Journal of Remote Sensing**, v. 38, n. 4, p. 514–527, 2012. Disponível em: <<https://www-tandfonline.ez89.periodicos.capes.gov.br/doi/full/10.5589/m12-043>>. Acesso em: 7 fev. 2024.

TRUJILLO-JIMÉNEZ, M. A.; LIBEROFF, A. L.; PESSACG, N.; PACHECO, C.; DÍAZ, L.; FLAHERTY, S. SatRed: New Classification Land Use/Land Cover Model Based on Multi-Spectral Satellite Images and Neural Networks Applied to a Semiarid Valley of Patagonia. **Remote Sensing Applications Society and Environment**, v. 26, p. 100703, 2022. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S2352938522000118>>. Acesso em: 31 jan. 2024.

WANG, Y.; SUN, Y.; CAO, X.; WANG, Y.; ZHANG, W.; CHENG, X. A Review of Regional and Global Scale Land Use/Land Cover (LULC) Mapping Products Generated from Satellite Remote Sensing. **ISPRS journal of photogrammetry and remote sensing: official publication of the International Society for Photogrammetry and Remote Sensing (ISPRS)**, v. 206, p. 311–334, 2023. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0924271623003209>>.

ZHA, Y.; GAO, J.; NI, S. Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery. **International journal of remote sensing**, v. 24, n. 3, p. 583–594, 2003. Disponível em: <<https://www-tandfonline.ez89.periodicos.capes.gov.br/doi/abs/10.1080/01431160304987>>. Acesso em: 7 fev. 2024.

ZHANG, S.; LI, X.; ZHOU, X.; WANG, Y.; HU, Y. Cloud Removal Using SAR and Optical Images via Attention Mechanism-Based GAN. **Pattern recognition letters**, v. 175, p. 8–15, 2023. Disponível em: <<https://www-sciencedirect.ez89.periodicos.capes.gov.br/science/article/pii/S0167865523002623>>. Acesso em: 31 jan. 2024.

APÊNDICE

APÊNDICE A - Tabela das importâncias

A seguir está apresentada a Tabela referente às importâncias médias obtidas com o uso de todas as bandas para cada conjunto, servindo de guia para o corte das bandas de acordo com as porcentagens definidas. A cor azul refere-se ao Top 10 bandas, a cor verde ao corte de 25% de mais importantes, o amarelo ao corte de 50% de maiores importâncias e o cinza representa 100% das bandas disponíveis, de forma que cada cor deve ser somada às bandas da cor anterior.

S1		S2		Índices		Completo	
Bandas	Média 100% (18 bandas)	Bandas	Média 100% (90 bandas)	Bandas	Média 100% (126 bandas)	Bandas	Média 100% (234 bandas)
vhDB_stdDev	60.35	B11_mean	4.91	NDBI_p90	8.18	vhDB_stdDev	3.25
vvDB_stdDev	39.45	B11_median	4.64	NDBI_max	6.02	B12_median	3.02
vhDB_min	27.00	B11_p75	4.24	SW2_N_mean	5.09	B11_p75	2.85
vhDB_p10	25.69	B12_mean	3.49	SAVI_max	4.98	B11_median	2.50
vhDB_max	23.52	B12_median	3.35	SAVI_p75	4.97	B5_median	1.99
vvDB_p10	22.55	B5_mean	2.84	SAVI_p90	4.60	NDBI_p90	1.74
vvDB_max	22.29	B5_median	2.32	SW2_N_p90	4.43	NDBI_max	1.62
vvDB_min	21.85	B12_p90	2.06	NDBI_mean	4.23	SW2_N_mean	1.60
vvDB_median	20.53	B11_p90	1.90	SAVI_mean	3.27	B11_mean	1.53
vvDB_mean	19.46	B12_p75	1.81	NDBI_p75	3.15	B12_mean	1.45
vvDB_p90	18.62	B5_p75	1.68	SW2_N_median	3.11	B5_mean	1.40
vhDB_median	18.55	B5_p25	1.61	RENDVI1_p25	3.10	B12_p75	1.28
vvDB_p25	18.37	B3_median	1.47	SW2_N_p75	3.08	B5_p75	1.28
vhDB_p25	18.02	B8A_p90	1.40	NDRE_p10	2.97	vvDB_stdDev	1.26
vvDB_p75	17.28	B12_stdDev	1.35	NDBI_stdDev	2.88	ARVI_p25	1.20
vhDB_p90	16.99	B12_p25	1.33	NDBI_median	2.56	SW2_N_p75	0.99
vhDB_p75	16.31	B12_max	1.33	NDRE_p25	2.56	B3_median	0.92
vhDB_mean	15.67	B4_median	1.29	SAVI_median	2.45	SW2_N_p90	0.89
		B7_p75	1.23	SAVI_p25	2.35	B11_p90	0.85
		B4_mean	1.13	ARVI_mean	2.33	B4_p75	0.82
		B11_max	1.09	RENDVI1_mean	2.26	B6_p90	0.73
		B8_p90	1.05	SW2_N_stdDev	2.12	B8A_p90	0.69
		B4_p25	1.03	RENDVI2_max	2.08	B8A_p75	0.67
		B6_p75	1.01	SAVI_p10	2.07	B4_p25	0.64
		B11_p25	0.88	RENDVI1_p10	2.02	B5_p25	0.62
		B2_mean	0.88	NDWI_stdDev	1.98	SW2_N_median	0.61
		B4_p75	0.86	RENDVI2_p75	1.75	ARVI_mean	0.60
		B8_stdDev	0.83	NDRE_mean	1.74	NDRE_stdDev	0.59
		B8A_p10	0.81	ARVI_min	1.68	NDBI_median	0.58
		B8_max	0.78	SW2_N_max	1.65	NDBI_p75	0.58

S1	S2	Índices	Completo
	B3_p25 0.77	ARVI_stdDev 1.61	NDRE_p75 0.57
	B7_stdDev 0.74	SAVI_min 1.60	NDBI_stdDev 0.56
	B3_mean 0.74	ARVI_p75 1.60	RENDVI2_median 0.55
	B4_max 0.73	NDVI_median 1.60	B12_p25 0.54
	B8A_min 0.71	NDVI_mean 1.50	NDVI_p10 0.54
	B6_stdDev 0.71	SAVI_stdDev 1.48	vvDB_min 0.53
	B6_p25 0.71	RENDVI2_p90 1.47	NDVI_mean 0.53
	B6_median 0.69	ARVI_p25 1.44	SW2_N_stdDev 0.48
	B11_min 0.66	GNDVI_mean 1.42	NDRE_p25 0.46
	B6_mean 0.66	NDVI_min 1.40	B4_median 0.45
	B7_max 0.64	RENDVI2_stdDev 1.40	NDBI_mean 0.42
	B5_p90 0.64	NDVI_p10 1.39	B8A_median 0.42
	B8A_p75 0.63	RENDVI1_stdDev 1.38	NDBI_p25 0.40
	B3_p75 0.63	ARVI_median 1.37	B11_p25 0.38
	B4_p90 0.63	GNDVI_p75 1.31	SAVI_p25 0.38
	B8A_stdDev 0.59	NDRE_p75 1.30	B12_max 0.36
	B4_stdDev 0.58	NDWI_p90 1.24	NDRE_mean 0.35
	B3_stdDev 0.57	ARVI_p10 1.22	B12_p90 0.34
	B8_p10 0.56	RENDVI1_median 1.17	B6_median 0.33
	B6_p90 0.55	NDVI_stdDev 1.14	B7_max 0.33
	B8_p75 0.54	SW2_N_p10 1.10	RENDVI1_p25 0.33
	B7_p90 0.53	NDBI_p10 1.10	NDVI_p75 0.33
	B8_median 0.48	NDRE_stdDev 1.10	B8_median 0.32
	B3_p10 0.48	NDBI_min 1.09	B4_mean 0.31
	B8_min 0.48	RENDVI1_min 1.07	NDRE_median 0.31
	B8A_max 0.48	S2REP_stdDev 1.06	ARVI_stdDev 0.31
	B2_p90 0.47	NDRE_p90 1.05	GNDVI_stdDev 0.30
	B5_p10 0.44	NDVI_p25 1.04	RENDVI1_p10 0.30
	B6_max 0.43	RENDVI2_min 1.04	B8A_mean 0.30
	B3_max 0.43	RVIm_min 1.04	B5_p90 0.29
	B7_min 0.41	CR_mean 0.98	ARVI_median 0.29
	B11_p10 0.40	GNDVI_stdDev 0.96	NDVI_median 0.28
	B8_mean 0.40	RVI_stdDev 0.95	vhDB_min 0.28
	B8A_p25 0.40	NDBI_p25 0.93	B5_stdDev 0.28
	B2_p25 0.39	RVIm_stdDev 0.93	vhDB_p10 0.27
	B2_p10 0.39	NDRE_min 0.93	B6_p75 0.26
	B8_p25 0.38	ARVI_p90 0.89	RENDVI2_min 0.26
	B2_max 0.38	RENDVI2_median 0.87	B8_p90 0.25
	B2_p75 0.37	SW2_N_p25 0.85	GNDVI_p90 0.24
	B3_p90 0.37	ARVI_max 0.85	SAVI_p75 0.24
	B5_max 0.34	RVIm_max 0.84	ARVI_p90 0.24
	B7_median 0.32	NDRE_median 0.83	NDRE_p90 0.24
	B7_mean 0.30	RENDVI1_max 0.82	B3_p25 0.23
	B7_p10 0.28	NDVI_p90 0.82	SAVI_min 0.22
	B4_min 0.28	NDWI_median 0.80	B7_median 0.22
	B8A_mean 0.26	GNDVI_min 0.79	RENDVI2_p25 0.22
	B8A_median 0.26	NDWI_p75 0.79	B7_stdDev 0.22
	B2_median 0.24	GNDVI_p10 0.73	B8A_p10 0.22

S1	S2	Índices		Completo	
	B2_min 0.24	RVI_p75	0.71	B2_stdDev	0.22
	B12_min 0.23	SW2_N_min	0.70	SAVI_p90	0.22
	B4_p10 0.21	NDWI_min	0.69	NDRE_p10	0.21
	B6_min 0.20	NDVI_p75	0.69	B2_p75	0.21
	B6_p10 0.18	RVIm_mean	0.69	B12_stdDev	0.20
	B5_min 0.16	NDWI_p10	0.68	B4_p10	0.20
	B2_stdDev 0.16	NDWI_mean	0.66	B6_max	0.20
	B7_p25 0.15	GNDVI_max	0.65	B7_p10	0.20
	B11_stdDev 0.11	NDWI_max	0.64	vhDB_p90	0.19
	B5_stdDev 0.04	RVI_p10	0.63	B4_max	0.19
	B12_p10 0.01	RVI_p90	0.62	B11_stdDev	0.19
	B3_min 0.00	RENDVI2_p25	0.60	RVI_p10	0.18
		CR_p90	0.59	NDWI_p25	0.18
		NDWI_p25	0.58	B8_p75	0.18
		RENDVI2_mean	0.57	ARVI_p75	0.18
		GNDVI_median	0.57	RVI_p25	0.18
		RENDVI1_p90	0.57	NDBI_p10	0.17
		GNDVI_p25	0.54	NDRE_min	0.17
		RVIm_p90	0.53	B8A_max	0.16
		RVI_median	0.53	RENDVI2_p75	0.16
		NDRE_max	0.52	RENDVI1_median	0.16
		RVI_p25	0.52	S2REP_stdDev	0.15
		RENDVI1_p75	0.52	B2_p10	0.14
		RENDVI2_p10	0.51	ARVI_p10	0.14
		RVIm_p75	0.51	B7_min	0.13
		NDVI_max	0.50	B8_stdDev	0.13
		RVIm_p10	0.45	B8_max	0.13
		CR_p10	0.42	B6_mean	0.12
		RVI_min	0.40	GNDVI_p25	0.12
		RVIm_p25	0.39	RENDVI1_p75	0.12
		CR_median	0.35	NDWI_max	0.12
		CR_max	0.34	B11_max	0.12
		CR_p25	0.29	GNDVI_p75	0.12
		GNDVI_p90	0.24	SW2_N_p25	0.11
		RVI_mean	0.23	B3_p10	0.11
		RVIm_median	0.21	vhDB_max	0.11
		CR_min	0.19	RENDVI2_p90	0.11
		CR_stdDev	0.19	SAVI_max	0.11
		CR_p75	0.12	RVI_min	0.11
		RVI_max	0.02	RVI_max	0.11
		S2REP_max	0.00	SAVI_median	0.11
		S2REP_mean	0.00	NDWI_p75	0.11
		S2REP_median	0.00	ARVI_min	0.11
		S2REP_min	0.00	GNDVI_mean	0.10
		S2REP_p10	0.00	B3_stdDev	0.10
		S2REP_p25	0.00	RVIm_mean	0.10
		S2REP_p75	0.00	SW2_N_min	0.10
		S2REP_p90	0.00	NDVI_p90	0.10

S1	S2	Índices	Completo
			B6_stdDev 0.09
			NDWI_min 0.09
			vvDB_p75 0.09
			B6_min 0.09
			RENDVI2_stdDev 0.09
			RVI_m_p90 0.09
			vhDB_mean 0.09
			NDRE_max 0.09
			RENDVI1_mean 0.09
			B7_p75 0.08
			SAVI_stdDev 0.08
			NDVI_min 0.08
			NDBI_min 0.08
			CR_p10 0.08
			B7_p25 0.08
			B8_mean 0.07
			B5_max 0.07
			vvDB_p10 0.07
			B2_max 0.06
			NDVI_p25 0.06
			CR_stdDev 0.06
			NDWI_p90 0.06
			GNDVI_max 0.06
			ARVI_max 0.06
			RENDVI2_mean 0.05
			B2_p90 0.05
			B3_min 0.05
			vvDB_p90 0.04
			CR_min 0.04
			RENDVI1_p90 0.04
			RENDVI2_max 0.04
			B3_mean 0.04
			B8A_stdDev 0.04
			B6_p10 0.04
			SW2_N_max 0.04
			B12_p10 0.03
			GNDVI_p10 0.03
			B4_stdDev 0.03
			B8A_p25 0.03
			RVI_p90 0.03
			B4_p90 0.03
			B2_mean 0.03
			vvDB_mean 0.02
			GNDVI_min 0.02
			B6_p25 0.02
			B5_p10 0.02
			B8_min 0.02
			B2_median 0.02

S1	S2	Índices	Completo
			SAVI_p10 0.01
			B2_p25 0.01
			B8_p25 0.01
			B11_min 0.01
			RVI_m_median 0.01
			B7_p90 0.01
			RVI_m_max 0.01
			RVI_m_stdDev 0.01
			NDWI_stdDev 0.01
			B11_p10 0.01
			B3_p75 0.01
			NDVI_stdDev 0.00
			CR_p90 0.00
			RVI_m_p75 0.00
			B7_mean 0.00
			B8_p10 0.00
			B4_min 0.00
			B2_min 0.00
			B5_min 0.00
			RENDVI1_stdDev 0.00
			B3_p90 0.00
			RENDVI1_min 0.00
			NDWI_median 0.00
			B8A_min 0.00
			vvDB_p25 0.00
			CR_max 0.00
			vhDB_p25 0.00
			B12_min 0.00
			B3_max 0.00
			CR_mean 0.00
			CR_median 0.00
			CR_p25 0.00
			CR_p75 0.00
			GNDVI_median 0.00
			NDVI_max 0.00
			NDWI_mean 0.00
			NDWI_p10 0.00
			RENDVI1_max 0.00
			RENDVI2_p10 0.00
			RVI_mean 0.00
			RVI_median 0.00
			RVI_p75 0.00
			RVI_stdDev 0.00
			RVI_m_min 0.00
			RVI_m_p10 0.00
			RVI_m_p25 0.00
			S2REP_max 0.00
			S2REP_mean 0.00

S1	S2	Índices	Completo
			S2REP_median 0.00
			S2REP_min 0.00
			S2REP_p10 0.00
			S2REP_p25 0.00
			S2REP_p75 0.00
			S2REP_p90 0.00
			SAVI_mean 0.00
			SW2_N_p10 0.00
			vhDB_median 0.00
			vhDB_p75 0.00
			vvDB_max 0.00
			vvDB_median 0.00

APÊNDICE B – Matrizes de confusão

Abaixo são apresentadas as matrizes de confusão de todos os conjuntos de dados.

Completo 100%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1818	182	2000
	0	35	1965	2000
	TOTAL	1853	2147	4000

Completo 50%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1811	189	2000
	0	35	1965	2000
	TOTAL	1846	2154	4000

Completo 25%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1839	161	2000
	0	40	1960	2000
	TOTAL	1879	2121	4000

Completo TOP 10				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1844	156	2000
	0	40	1960	2000
	TOTAL	1884	2116	4000

S1 100%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1712	288	2000
	0	46	1954	2000
	TOTAL	1758	2242	4000

S1 50%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1721	279	2000
	0	47	1953	2000
	TOTAL	1768	2232	4000

S1 25%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1732	268	2000
	0	49	1951	2000
	TOTAL	1781	2219	4000

S1 TOP 10				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1723	277	2000
	0	47	1953	2000
	TOTAL	1770	2230	4000

Índices 100%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1805	195	2000
	0	42	1958	2000
	TOTAL	1847	2153	4000

Índices 50%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1805	195	2000
	0	37	1963	2000
	TOTAL	1842	2158	4000

Índices 25%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1805	195	2000
	0	41	1959	2000
	TOTAL	1846	2154	4000

Índices TOP 10				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1830	170	2000
	0	39	1961	2000
	TOTAL	1869	2131	4000

S2 100%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1789	211	2000
	0	36	1964	2000
	TOTAL	1825	2175	4000

S2 50%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1801	199	2000
	0	40	1960	2000
	TOTAL	1841	2159	4000

S2 25%				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1803	197	2000
	0	40	1960	2000
	TOTAL	1843	2157	4000

S2 TOP 10				
Matriz de Erros				
	Classes	1	0	TOTAL
Verdade de campo	1	1811	189	2000
	0	38	1962	2000
	TOTAL	1849	2151	4000