UFSM

Geografia Ensino & Pesquisa

ISSN 2236-4994
ACESSO ABERTO

Geog Ens Pesq, Santa Maria, v. 29, e87967, p. 1-18, 2025 https://doi.org/10.5902/2236499487967
Submissão: 19/06/2024 • Aprovação: 02/12/2024 • Publicação: 10/03/2025

Geoinformação e Sensoriamento Remoto em Geografia

Utilização do classificador *Random Forest* para Classificação do Uso e Cobertura da Terra a partir de Dados Sentinel 1 e 2 em região campestre no bioma Mata Atlântica

Use of the Random Forest classifier to Classify Land Use and Cover using Sentinel 1 and 2 Data in a rural region in the Atlantic Forest biome

Uso del clasificador Random Forest para clasificar el uso y la cobertura del suelo utilizando datos Sentinel 1 y 2 en una región rural del bioma del Bosque Atlántico

Andressa Kossmann Ferla[I], **Tatiana Mora Kuplich**[I], **Igor da Silva Narvaes**[I]

[I]Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brasil

RESUMO

O uso de mapas de uso e cobertura do solo é essencial para o monitoramento ambiental, para isso é necessário a utilização de técnicas de sensoriamento remoto. Pensando nisso, o presente trabalho teve como objetivo a utilização dos atributos: Coeficiente de Retroespalhamento, Decomposição Polarimétrica e Coerência Interferométrica, do sensor Sentinel 1, e as bandas R, G, B, NIR, e índices de vegetação NDVI e SAVI, do sensor Sentinel 2, para identificar a melhor combinação de variáveis de entrada do algoritmo de classificação *Random Forest* (RF) utilizando a acurácia, em uma área nos "Campos de Cima da Serra", pertencente ao bioma Mata Atlântica. O trabalho identificou que a utilização dos três atributos do Sentinel 1, em conjunto com as bandas ópticas do Sentinel 2, teve melhor acurácia (93%), embora a utilização apenas das bandas ópticas obteve 89% de acurácia. Todavia, quando utilizado apenas atributos SAR, obteve a menor acurácia (67%). A elaboração desta metodologia servirá como base para a continuidade da presente pesquisa, utilizando técnicas mais robustas, como análise de séries temporais via SITS (*Satellite Image Time Series Analysis*), com a geração de resultados para o monitoramento da mata atlântica na região sul do País, e de subsídio para testes de monitoramento do bioma pampa, pela sua alta capacidade de análise de séries temporais, a partir de diferentes plataformas, em um pacote de código aberto.

Palavras-chave: Random Forest; Uso e Cobertura do Solo; Classificação de imagem; Radar; Óptico

Artigo publicado por Geografia Ensino & Pesquisa sob uma licença CC BY-NC-SA 4.0.

ABSTRACT

The use of land use and land cover maps is essential for environmental monitoring, and for this purpose, it is necessary to use remote sensing techniques. With this in mind, this study aimed to use the Backscatter Coefficient, Polarimetric Decomposition and Interferometric Coherence attributes of the Sentinel 1 sensor and the R, G, B, NIR bands and NDVI and SAVI vegetation indices of the Sentinel 2 sensor to identify the best combination of input variables for the Random Forest (RF) classification algorithm using accuracy, in an area in the "Campos de Cima da Serra," belonging to the Atlantic Forest biome. The study identified that the use of the three Sentinel 1 attributes together with the optical bands of Sentinel 2 had better accuracy (93%), although the use of only the optical bands obtained 89% accuracy. However, when using only SAR attributes, the lowest accuracy was obtained (67%). The development of this methodology will serve as a basis for the continuation of this research, using more robust techniques such as time series analysis via SITS (Satellite Image Time Series Analysis), with the generation of results for monitoring the Atlantic Forest in the southern region of the country and subsidy for monitoring tests of the pampa biome, due to its high capacity for analyzing time series from different platforms in an open-source package.

Keywords: Random Forest; Land Use and Land Cover; Image Classification; SAR; Optical

RESUMEN

El uso de mapas de uso y cobertura del suelo es fundamental para el monitoreo ambiental, lo que requiere el uso de técnicas de teledetección. Con esto en mente, el presente trabajo tuvo como objetivo utilizar los atributos de Coeficiente de Retrodispersión, Descomposición Polarimétrica y Coherencia Interferométrica del sensor Sentinel 1 y las bandas R, G, B, NIR y los índices de vegetación NDVI y SAVI del sensor Sentinel 2 para identificar el mejor combinación de variables de entrada del algoritmo de clasificación Random Forest (RF) utilizando precisión, en un área de los "Campos de Cima da Serra", perteneciente al bioma de la Mata Atlántica. El trabajo identificó que el uso de los tres atributos de Sentinel 1 en conjunto con las bandas ópticas de Sentinel 2 tuvo mejor precisión (93%), aunque el uso de solo las bandas ópticas obtuvo un 89% de precisión. Sin embargo, al utilizar solo atributos SAR obtuvo la precisión más baja (67%). El desarrollo de esta metodología servirá de base para la continuidad de esta investigación, utilizando técnicas más robustas como el análisis de series de tiempo vía SITS (Satellite Image Time Series Analysis), con la generación de resultados para el monitoreo del bosque atlántico en la región sur. del país y subsidio para pruebas de monitoreo del bioma pampa, debido a su alta capacidad de análisis de series de tiempo desde diferentes plataformas en un paquete de código abierto.

Palabras-clave: Random Forest; Uso y Cobertura del Suelo; Classificación de Imágenes; Radar; Óptico

1 INTRODUÇÃO

A utilização de sensores remotos para o mapeamento de uso e cobertura da terra é essencial, e a base de dados para gestores ambientais, na academia e em

organizações governamentais e não governamentais. Dados de sensoriamento remoto possibilitam o monitoramento ambiental e estratégias de manejo mais efetivas, também facilitando estudos sobre a ocupação e dinâmica da cobertura da terra, entre outras atividades. Cada vez mais, entretanto, tornou-se necessário identificar métodos mais eficientes de utilizar imagens destes sensores, como variáveis dos modelos de classificação e predição de uso e cobertura da terra, utilizando métodos automáticos como *Machine Learning* (ML) e D*eep Learning* (DL).

Com maior acesso aos dados, provenientes de diversos sensores orbitais e fontes complementares de sensoriamento remoto, é possível a criação de mapas de uso e cobertura da terra (LULC) mais acurados (Gupta e Shukla, 2023). Os dados SAR (Synthetic Aperture Radar), do Sentinel 1, e os dados ópticos, do Sentinel 2, ambos satélites desenvolvidos pela Agência Espacial Europeia (ESA), estão disponíveis na plataforma *Open Access Hub* do *Copernicus* (OAH) (ESA, 2024) gratuitamente, e são amplamente utilizados com técnicas de inteligência artificial (Adrian; Sagan e Maimaitijiang, 2021; Clerici; Valbuena Calderón e Posada, 2017; Feng *et al*., 2019; Schulz *et al*., 2021).

Os dados ópticos contém informações referentes às características físico-químicas dos objetos, ao passo que os dados SAR incluem informações estruturais, volumétricas e propriedades dielétricas (Al-Ruzouq *et al*., 2019) dos objetos, na superfície terrestre. Desta forma, cada sensor tem habilidades distintas, para classificar certos tipos de cobertura da terra, em relação a outros (Gupta e Shukla, 2023), por isso a utilização desses dados em conjunto pode trazer uma melhora na classificação do LULC. Feng *et al*. (2019) utilizaram dados Sentinel 1 e 2 para mapear a cobertura da terra, e obtiveram maior precisão integrando os dois sensores, enquanto que Chachondhia; Shakya e Kumar (2021) utilizaram dados Sentinel 1 e 2 fundidos, para a classificação de LULC, e obtiveram melhor precisão, quando comparados a dados de sistemas sensores ópticos e SAR em separado.

O algoritmo de classificação de ML intitulado de *Random Forest - RF* (Breiman, 2001) é um dos mais populares na comunidade de Sensoriamento Remoto, apresentando funções de aprendizado em classificações simples e complexas. É um método relativamente fácil de implementar, pois não precisa de ajustes finos de hiperparâmetros. Também pode lidar com tarefas de aprendizagem com uma pequena quantidade de conjunto de dados de treinamento, além de exigir menor complexidade computacional, e ser mais fácil de interpretar, em comparação com modelos de aprendizagem profunda (Sheykhmousa *et al*., 2020).

Até onde sabemos, embora haja vários estudos no monitoramento do uso e cobertura da terra utilizando algoritmos de ML (Wang *et al*., 2022), e DL (Aljebreen *et al*., 2024), [apesar deste último ter como restrição uma demanda de capacidade computacional superior (Southworth *et al*., 2024)], ou aplicados simultaneamente (Naik *et al*., 2024), estes eventualmente efetuados na Mata Atlântica (Bressane *et al*., 2023). Portanto, estudos que preconizam a estruturação metodológica para o monitoramento contínuo de biomas são ainda incipientes (Gu e Zeng, 2023), como verificado em Souza e Azevedo (2017), capitaneado pela iniciativa privada, e em Miranda *et al*. (2024), para o monitoramento contínuo oficial da totalidade dos biomas brasileiros.

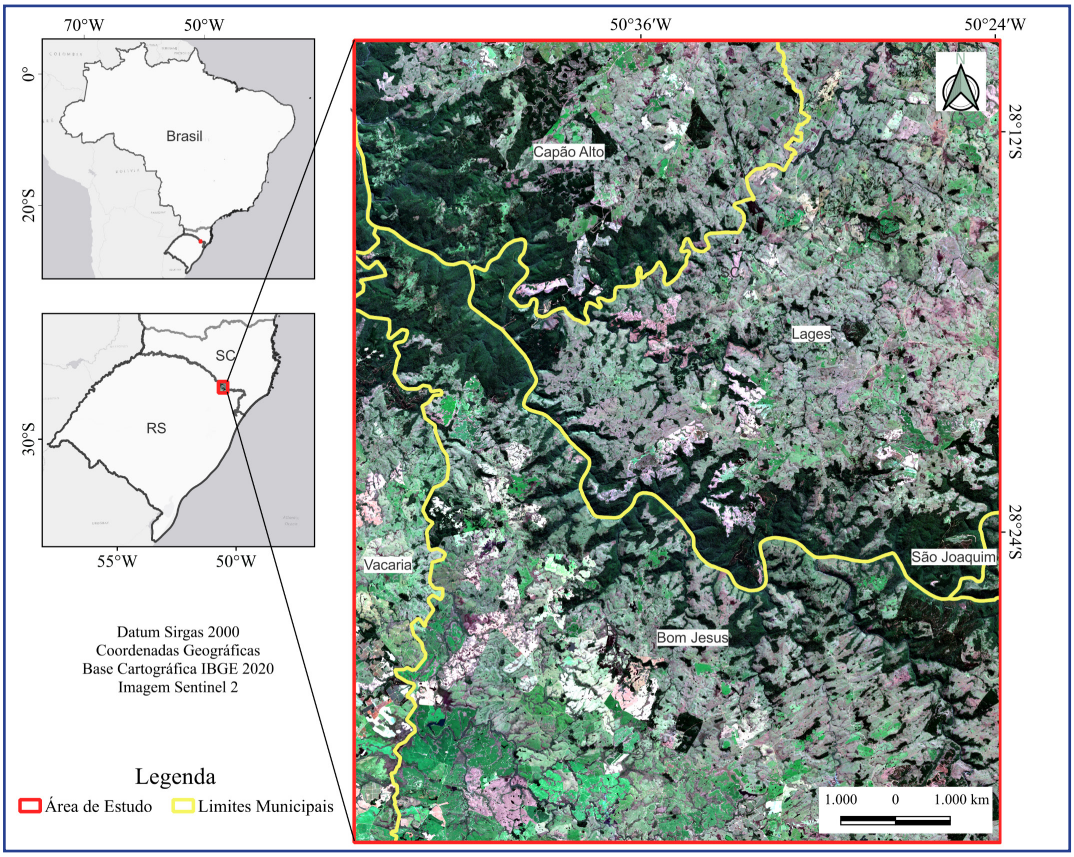
Sendo assim, o objetivo deste trabalho é utilizar o algoritmo RF para classificar o LULC no aprimoramento metodológico do monitoramento da Mata Atlântica, no âmbito do Programa Biomas BR, conduzido pelo INPE. Este Programa está em fase de estruturar, de modo sistemático, a ampliação do monitoramento ambiental para todos os biomas brasileiros, e incorporar novas funcionalidades. Para este fim, nossa análise propõe produzir uma metodologia baseada em ML, em específico no algoritmo de classificação RF, na região dos “Campos de Cima da Serra”, situada no bioma Mata Atlântica. Serão combinadas diferentes bandas dos sensores Sentinel 1 e 2, através de atributos SAR coerentes e incoerentes, bandas e Índices de Vegetação (IVs), do sensor óptico, para que se estabeleça a melhor combinação de variáveis de entrada para a classificação de LULC.

2 MATERIAIS E MÉTODOS

2.1 Área de Estudo

A área de estudo está localizada na região do Planalto Meridional, no extremo nordeste do Rio Grande do Sul, mais precisamente na região chamada “Campos de Cima da Serra”, parte do bioma Mata Atlântica. Situada em altitude média de 684 metros a 1.047 metros, corresponde à Unidade Geomorfológica Planalto dos Campos Gerais (Boldrini, 2009); a precipitação média anual varia de 1.500 a 1.700mm, bem distribuídos; a temperatura média do mês mais quente está abaixo dos 19º C, classificado de acordo com Koppen, como clima tipo Cfb. A área de estudo engloba áreas de campo, florestas, cânions, áreas agrícolas, rios e lago (Figura 1).

Figura 1 – Mapa de localização da área de estudo. É apresentada imagem Sentinel 2 em composição cor verdadeira (4-3-2), de 16/09/2023



Fonte: Autores/as (2024)

2.2. Pré-Processamento

2.2.1 Dados SAR (Sentinel-1)

Foram adquiridos, na plataforma *Open Access Hub* do *Copernicus* (OAH) (ESA, 2024), duas imagens nível 1 SLC (*Single Look Complex)*, modo de faixa Interferométrica Ampla (IW) em dupla polarização (VV + VH), das datas 14/09/2023 e 26/09/2023, como as datas de imageamento mais próximas possíveis para o cálculo do coeficiente de correlação complexa, fundamental para o uso em interferometria (Paradella *et al*., 2021). Todas as operações de pré-processamento para extração dos atributos Coeficiente de Retroespalhamento, Decomposição Polarimétrica Dual de Cloude e Pottier e Coerência Interferométrica foram realizadas no software SNAP (*Sentinel Applications Platform*) (Zuhlke *et al*., 2015).

Na imagem 14/09/2023, foi empregada a operação Split, para dividir as subfaixas e obter apenas a subfaixa IW1, a qual recobre a área de estudo. Para a obtenção dos coeficientes de retroespalhamento (σ^0VH e σ^0VV), foram realizadas as seguintes operações de pré-processamento, a saber: 1 - Aplicação de arquivo de órbita; 2 - Correção do ruído termal, reduzindo a descontinuidade entre as subfaixas; 3 - Calibração, para corrigir a intensidade do sinal; 4 – *Deburst*, para obtenção de uma imagem contínua, removendo as subfaixas; 5 – *Multilooking,* para produzir um elemento de resolução com formato quadrado e atenuar o ruído *speckle*; 6 - Filtro *Speckle Lee* (5x5); 7 - Correção do terreno *Range Doppler*, através do modelo digital de elevação SRTM, com projeção UTM WGS 84; e 8 - Reamostragem para 10 metros, com o intuito de compatibilizar a dimensão dos elementos de resolução das imagens SAR e óptica.

Para a obtenção dos atributos de decomposição polarimétrica Entropia (H) e ângulo alfa (α) a partir da subfaixa IW1, as operações foram: 1 - Aplicação de Arquivo de órbita; 2 - *Deburst*; 3 - *Multilooking*; 4 - Filtro Polarimétrico *Refined Lee Filter* (5x5); 5 - Decomposição H-*Alpha Dual Pol*; 6 - Correção do terreno *Range Doppler*; e 7 - Reamostragem para 10 metros.

Para a extração dos atributos de coerência interferométrica, foram utilizadas duas imagens, de mesma dimensão, com 12 dias de diferença. Após os processos de Split, as seguintes operações foram realizadas: 1 - Aplicação de Arquivo de Órbita; 2 - *Deburst*; 3 - Co-registro (*Coregistration/S-1 Back Geocoding*), correspondência entre pixels; 4 - *Coherence*, coerência interferométrica (grau de correlação complexa entre os mesmos pixels das diferentes imagens); 5 - *Multilooking;* 6 - Filtro *Lee;* 7 - Correção do terreno *Range Doppler*; e 8 - Reamostragem para 10 metros, obtendo as bandas γVH e γVV.

2.2.2 Dados ópticos (Sentinel-2)

Ainda na plataforma (OAH) foi adquirido uma imagem óptica Sentinel 2, sensor MSI nível 1C (Reflectância no topo atmosfera) da data 11/09/2023, incluindo as bandas vermelho (665nm), verde (560nm), azul (490nm) e infravermelho próximo - NIR (842nm) de 10 metros, de resolução espacial e resolução radiométrica de 12 bits, sem a presença de nuvem.

No *software* Sen2Cor foi realizada a correção para Reflectância de Superfície nível 2A. Foram calculados os índices de vegetação – IV - *Soil-adjusted Vegetation Index* – SAVI (Savi e Huete, 1988), utilizando fator L = 0,5. Este fator é um ajuste dos efeitos da cor do solo, e pode variar de 0 (vegetação densa) a 1 (vegetação rala), sendo que o valor aplicado na presente pesquisa, de 0,5, é considerado como um valor intermediário de densidade vegetal (Lima *et al*., 2017). Também foi calculado o *Normalized Difference Vegetation Index* - NDVI (Rouse *et al*.,1973), frequentemente utilizado para medir o vigor e a densidade da vegetação.

A partir da imagem óptica Sentinel 2, composição RGB (4-3-2), foram coletadas amostras das classes Floresta, Água, Solo Exposto, Campo e Agricultura, após uma segmentação de imagem, em que foram definidos os parâmetros: algoritmo de segmentação = *meanshift, Spatial radius* = 5, realizada na ferramenta Orfeu Toolbox, no *software* Qgis 3.22.10. Para cada classe, por interpretação visual, foi selecionado

cuidadosamente o segmento, seguindo seus limites e, quando necessário, editando o polígono, de forma a evitar o efeito de borda.

Foram coletados 61 polígonos da classe Floresta, contendo 644.511 pixels; 37 polígonos da classe Água, contendo 71.741 pixels; 47 polígonos da classe Solo Exposto, contendo 241.653 pixels; 64 polígonos da classe Campo, contendo 658.140 pixels; e 57 polígonos da classe Agricultura, contendo 404.757 pixels. As amostras foram separadas em 70% para treinamento e 30% para teste (validação).

2.3. Processamento e Avaliação da Precisão

Foram realizadas nove classificações, cada uma com uma composição de banda e polarizações diferentes, ordenadas de acordo com a ordem decrescente de sua acurácia (Tabela 1). O processamento dos dados foi realizado em linguagem Python, onde o classificador RF foi empregado a partir da biblioteca *Scikit-Learning,* utilizando o pacote RandomForestClassifier, descrito em Pedregosa *et al.* (2011). Os parâmetros definidos foram de acordo com o padrão do modelo, a saber: *n_estimators* = 500 (*ntree*) e *max_features* = raiz quadrada do número de variáveis de cada classificação (*mtry*).

A avaliação da precisão foi feita comparando as classes de LULC resultantes da fase de treinamento com os dados produzidos pela fase de teste (validação). Com base nas matrizes de confusão, foram calculadas métricas de acurácia: exatidão total, Coeficiente Kappa e F1 Score. Através do teste estatístico Z, de Congalton e Mead (1983), foi identificado se houve diferença estatisticamente significativa entre as classificações, o qual compara os coeficientes Kappa entre duas classificações diferentes. Se o valor Z calculado para o teste for maior que o valor Z tabulado, o resultado é dito significativo, e a hipótese nula é rejeitada (Ho:K1=K2), ou seja, os dois classificadores são estatisticamente diferentes. Para o cálculo do teste Z foi utilizada a planilha de cálculo proposta por Rennó e Silva (2023), que segue as etapas e procedimentos definidos por Hudson (1987) e Congalton (1999).

3 RESULTADOS E DISCUSSÃO

O modelo testado apenas com bandas SAR obteve uma classificação de 67% de acurácia, sendo a menor acurácia entre as classificações (Tabela 2), consistente com a representação incorreta da forma dos objetos, além de não ter apresentado minimização do efeito do ruído *speckle* (Figura 2 f), resultado inferior ao encontrado em Diniz, Gama e Adami (2022), onde classificaram um trecho da Floresta Amazônica utilizando dados SAR. O algoritmo de classificação RF aplicado em Mishra *et al*. (2023) utilizando alguns índices de vegetação desenvolvidos exclusivamente para imagens SAR obteve valores superiores de acurácia, enquanto para Wang *et al*. (2021) a utilização de dados SAR multisensores com diferentes bandas também obtiveram acurácia superior ao nosso estudo para as bandas X, C e L, o que era esperado, devido aos diferentes comprimentos de onda utilizados.

Ao utilizar apenas dados do Sentinel 2, o modelo obteve 89% de acurácia, similar aos estudos de Adebalu *et al*. (2013), utilizando imagens do RapidEye. Em Trindade *et al.* (2023), para a classificação de uma área declivosa no bioma Pampa, obtiveram os melhores resultados por meio da combinação de bandas, IV e modelo digital de elevação.

Tabela 2 – Acurácia das classificações com diferentes composições de bandas, utilizando classificador RF em ordem decrescente de significância estatística, de acordo com Teste Z

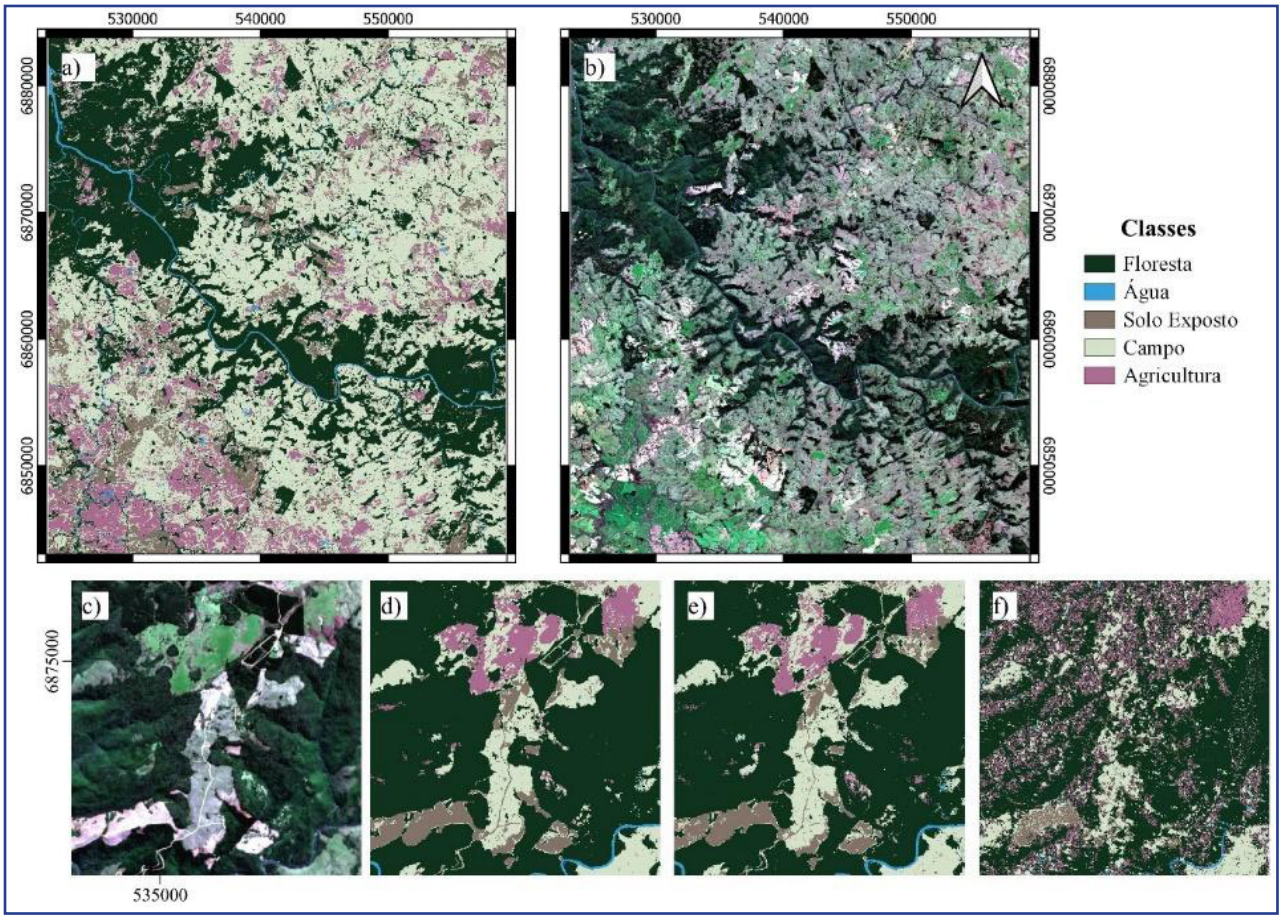
Classificação	Composição de Bandas	Acurácia	Kappa
1	H e α, σ^0VH e σ^0VV, γVH e γVV, RGB, NIR, NDVI, SAVI	0,93	0,90
9	σ^0VH e σ^0VV, γVH e γVV, RGB, NIR, NDVI, SAVI	0,92	0,90
5	σ^0VH e σ^0VV, RGB, NIR, NDVI, SAVI	0,92	0,89
7	H e α, γVH e γVV, RGB, NIR, NDVI, SAVI	0,92	0,89
4	H e α, RGB, NIR, NDVI, SAVI	0,91	0,88
6*	γVH e γVV , RGB, NIR, NDVI, SAVI	0,91	0,88
8*	H e α, σ^0VH e σ^0VV, RGB, NIR, NDVI, SAVI	0,91	0,88
2	RGB, NIR, NDVI, SAVI	0,89	0,86
3	H e α, σ^0VH e σ^0VV, γVH e γVV	0,67	0,54

Fonte: Autores/as (2024)

Em * não há diferença estatisticamente significativa entre classificações

A classificação com os dados em conjunto do Sentinel 1 e 2 (classificação 1) (Figura 2 a), obteve 93% de acurácia. Isso demonstra que os dados SAR utilizados juntamente com ópticos podem aumentar a acurácia do algoritmo de classificação (Monsalve-Tellez; Torres-León; Garcés-Gómez, 2022). Esta tendência de aumento da precisão, quando da utilização conjunta de dados Sentinel 1 e 2, também verificada por Feng *et al*. (2019), contribuiu em sua análise, para um aumento de 3,24% na precisão da classificação. Todavia, a classificação 2 (Figura 2 e) nas áreas onde havia alta declividade e efeito de sombra das copas, verificou-se o efeito sal e pimenta, também visto em Raczko e Zagajewski (2017).

Figura 2 – Mapa de uso e cobertura do solo, mostrando o resultado das classificações geradas com classificador *Random Forest*



Fonte: Autores/as (2024)

a) Classificação 1; b) Imagem Sentinel 2 composição 4-3-2; c) extrato de imagem Sentinel 2 composição 4-3-2; d) extrato da classificação 1; e) extrato da classificação 2; f) extrato da classificação 3

As demais classificações (4, 5, 6, 7, 8 e 9), onde foram alternadas as composições entre as bandas SAR+Ópticas, apresentaram diferenças mínimas (1%) na acurácia entre as diferentes composições, tendência decorrente da relevância das variáveis ópticas destas classificações na composição do modelo, calculada pelo índice de Gini (Breiman, 2001), podendo estar associada a multicolinearidade (Alin, 2010). Belgiu e Drăguţ (2016). Entretanto, afirmam que este classificador pode lidar com sucesso com alta dimensionalidade e multicolinearidade de dados, sendo rápido e insensível ao sobreajuste (*overfitting*).

Ao realizar o teste Z, identificamos que a classificação 9 é estatisticamente superior à classificação 7 e 5, as quais possuem as segundas maiores acurácias dentre as nove classificações. Neste caso, as bandas σ^0VH e σ^0VV, γVH e γVV se mostraram mais relevantes quando utilizadas em conjunto com bandas ópticas e IVs. Isto foi também comprovado por Khalil e Saad-Ul-Haque (2018), que conseguiram melhor distinção às feições do terreno, ao contrário de usar apenas estes dados de forma individual, obtendo uma acurácia maior, ao mapear as classes de água, areia, edificação e vegetação. Em adição, Diniz, Gama e Adami (2022) também comprovaram que a utilização dos atributos de coerência interferométrica e coeficiente de retroespalhamento apresentaram grande importância para a classificação, em que obtiveram os melhores resultados, ao separar classes florestais na Amazônia.

Nas matrizes de confusão das classificações 1 e 2 (Tabelas 3 e 4, respectivamente), observa-se que os erros de classificação ocorreram principalmente entre as categorias de agricultura e campo, bem como entre as categorias de solo exposto e campo. Neste caso, nestas classificações a grande relevância das variáveis ópticas no modelo de classificação do RF é devido à semelhança das características espectrais entre esses tipos de cobertura da terra, o que também foi verificado em Tavares *et al*. (2019), aliado a importância complementar dos dados SAR na melhora do desempenho do classificador, comprovado pela menor confusão (maiores valores de F1-score) para estas classes, com a inserção das variáveis SAR (coerência interferométrica,

decomposição dual de Cloude-Pottier e retroespalhamento em σ^0) (teste Z = 61,55; Valor p= 0,0000; α = 0,05), tendência encontrada na maioria das pesquisas sobre o tema (Solórzano *et al*., 2023).

Tabela 3 – Matriz de confusão da classificação 1

Classe	Floresta	Água	Solo Exposto	Campo	Agricultura	F1 Score
Floresta	182797	324	472	4366	591	0,96
Água	270	19203	122	168	3	0,97
Solo Exposto	1207	66	59680	8015	455	0,88
Campo	5015	188	4836	177584	4689	0,91
Agricultura	1723	42	517	9853	105564	0,92
Acurácia (Exatidão total)	**0,93**					
Kappa	**0,90**					

Tabela 4 – Matriz de confusão da classificação 2

Classe	Floresta	Água	Solo Exposto	Campo	Agricultura	F1 Score
Floresta	180376	594	66	5754	1117	0,95
Água	799	18350	148	326	210	0,93
Solo Exposto	1452	109	55579	11772	758	0,83
Campo	5903	304	7301	171846	7375	0,87
Agricultura	2336	70	756	14149	100427	0,88
Acurácia (Exatidão total)	**0,89**					
Kappa	**0,86**					

As classes Floresta e Água foram classificadas com melhor precisão em ambas classificações, por serem áreas uniformes e distintas, além de possuírem coeficiente de variação muito baixo (Talukdar *et al*., 2020) e resposta espectral distinta.

A discriminação da classe Solo Exposto, embora sendo a única com valor de F-score inferior a 90% [neste caso 0,83 (classificação 2)], apresentou alta confusão com a classe Campo, decorrente da grande semelhança espectral em áreas de baixo teor de umidade, além da classe Campo ser influenciada pelos efeitos de fundo do solo (Chaves *et al*., 2023).

4 CONCLUSÕES

Os resultados deste estudo sugerem que a integração de dados de sensores remotos SAR e ópticos do Sentinel 1 e 2, respectivamente, como variáveis de entrada do classificador *Random Forest*, aliado a um método eficaz de coleta de amostras, é uma abordagem robusta para o mapeamento preciso do uso e cobertura da terra em áreas de interesse na Mata Atlântica. As variáveis testadas conjuntamente, e que apresentaram as maiores acurácias nas classificações testadas, foram: σ^0VH e σ^0VV, γVH e γVV, RGB, NIR, NDVI e SAVI (classificação 9).

Essa metodologia tem potencial para auxiliar gestores ambientais, pesquisadores e organizações governamentais no monitoramento e manejo sustentável dessas áreas, por intermédio da consolidação de uma metodologia semiautomática. Podem ser incorporadas, com algumas modificações, para o monitoramento dos biomas Mata Atlântica e Pampa, no âmbito do Programa BiomasBR, proposto pelo Ministério do Meio Ambiente, e capitaneado pelo Instituto Nacional de Pesquisas Espaciais (INPE).

Esta metodologia proposta se mostrou eficaz na escolha das melhores combinações de variáveis para a classificação. Apesar disto, torna-se necessário adequações, utilizando o método proposto para outras regiões e/ou biomas, dada as variações das características fito fisionômicas, apenas trocando as variáveis de entrada do modelo.

AGRADECIMENTOS

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pela bolsa (Processo: 301336/2024-5), do Programa de Capacitação Institucional (PCI) (projeto 400077/2022-1 - Pesquisa e Desenvolvimento em Ciências e Tecnologias Espaciais e suas Aplicações), e ao CNPQ (processo 409336/2021-1).

REFERÊNCIAS

ADRIAN, J.; SAGAN, V.; MAIMAITIJIANG, M. Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. **ISPRS Journal of Photogrammetry and Remote Sensing**, *[S. l.]*, v. 175, p. 215–235, mai. 2021. DOI: 10.1016/j.isprsjprs.2021.02.018.

ALIN, A. Multicollinearity. **WIREs Interdisciplinary Reviews: Computational Statistics**, *[S. l.]*, v. 2, n. 3, p. 370–374, 25 mai. 2010.

ALJEBREEN, M *et al*. Land Use and Land Cover Classification Using River Formation Dynamics Algorithm With Deep Learning on Remote Sensing Images. **IEEE Access**, *[S. l.]*, v. 12, n. 99, p. 11147-11156, 2024. DOI: 10.1109/ACCESS.2023.3349285.

AL-RUZOUQ, R. *et al*. Multi-scale correlation-based feature selection and random forest classification for LULC mapping from the integration of SAR and optical Sentinel images. *In*: Remote Sensing Technologies and Applications in Urban Environments, 4., 2019, Strasbourg. **Proceedings...** Strasbourg: SPIE, 2019. p. 1. DOI: 10.1117/12.2533123.

BELGIU, M.; DRĂGUŢ, L. Random forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, *[S. l.]*, v. 114, p. 24–31, abr. 2016. DOI: 10.1016/j.isprsjprs.2016.01.011.

BOLDRINI, I. L. D. **Biodiversidade Dos Campos Do Planalto Das Araucárias.** Brasília: Ministério do Meio Ambiente, 2009.

BREIMAN, L. Random Forests. **Machine Learning**, *[S. l.]*, v. 45, p. 5–32, 2001. DOI: 10.1023/A:1010933404324.

BRESSANE, A.; GOMES, I. G.; DA ROSA, G. C. S.; BRANDELIK, C. C. M.; SILVA, M. B.; SIMINSKI, A.; NEGRI, R. G. Computer-aided classification of successional stage in subtropical Atlantic Forest: a proposal based on fuzzy artificial intelligence. **Environ Monit Assess**, *[S. l.]*, v. 195, n. 1, p. 184, 2023. DOI: 10.1007/s10661-022-10799-x.

CHACHONDHIA, P.; SHAKYA, A.; KUMAR, G. Performance evaluation of machine learning algorithms using optical and microwave data for LULC classification. **Remote Sensing Applications: Society and Environment**, *[S. l.]*, v. 23, p. 100599, ago. 2021. DOI: 10.1016/j.rsase.2021.100599.

CHAVES, M. E. D.; SOARES, A. R.; MATAVELI, G. A. V; SANCHEZ, A. H.; SANCHES, I. D. A Semi-Automated Workflow for LULC Mapping via Sentinel-2 Data Cubes and Spectral Indices. **Automation**, *[S. l.]*, v. 4, n. 1, p. 94–109, 23 fev. 2023. DOI: 10.3390/automation4010007.

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**, *[S. l.]*, v. 13, n. 2, p. 718–726, 30 nov. 2017. DOI 10.1080/17445647.2017.1372316.

CONGALTON, R. G. **Assessing the accuracy of Remotely Sensed data**: principles and practices. Boca Raton: Lewis, 1999. p. 34-46.

CONGALTON, R. G.; MEAD, R. A. A quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric Engineering and Remote Sensing**, *[S. l.]*, v. 49, n. 1, p. 69-74, 1983.

DINIZ, J. M. F. DE S.; GAMA, F. F.; ADAMI, M. Evaluation of polarimetry and interferometry of sentinel-1A SAR data for land use and land cover of the Brazilian Amazon Region. **Geocarto International**, *[S. l.]*, v. 37, n. 5, p. 1482–1500, 4 mar. 2022. DOI: 10.1080/10106049.2020.1773544.

FENG, Q.; YANG, J.; ZHU, D.; LIU, J.; GUO, H.; BAYARTUNGALAG, B.; LI, B. Integrating Multitemporal Sentinel-1/2 Data for Coastal Land Cover Classification Using a Multibranch Convolutional Neural Network: A Case of the Yellow River Delta. **Remote Sensing**, *[S. l.]*, v. 11, n. 9, p. 1006, 28 abr. 2019. DOI: 10.3390/rs11091006.

GU, Z.; ZENG, M. The use of artificial intelligence and satellite remote sensing in land cover change detection: review and perspectives. **Sustainability**, *[S. l.]*, v. 16, n. 1, p. 274, 2023. DOI: 10.3390/su16010274.

GUPTA, P.; SHUKLA, D. P. Combined Optical and SAR remote sensing for LULC mapping of Imphal valley using Machine Learning Algorithm. *In*: 2023 INTERNATIONAL CONFERENCE ON MACHINE INTELLIGENCE FOR GEOANALYTICS AND REMOTE SENSING (MIGARS), 2023, *[S. l.]*. **Proceedings...** *[S. l.]*: IEEE, 27 jan. 2023. DOI: 10.1109/MIGARS57353.2023.10064582.

ESA. Copernicus Open Access Hub. 2024. Disponível em: https://scihub.copernicus.eu/. Acesso em: 29 nov. 2024.

HUDSON, W. D. Correct Formulation of the Kappa Coefficient of Agreement. **Photogrammetric Engineering & Remote Sensing**, *[S. l.]*, v. 53, n. 4, p. 421-422, 1987.

HUETE, A. R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, *[S. l.]*, v. 25, n. 3, p. 295-309, 1988. DOI: 10.1016/0034-4257(88)90106-X.

KHALIL, R. Z.; SAAD-UL-HAQUE. InSAR coherence-based land cover classification of Okara, Pakistan. **The Egyptian Journal of Remote Sensing and Space Science**, *[S. l.]*, v. 21, n. supl. 1, p. S23–S28, jul. 2018. DOI 10.1016/j.ejrs.2017.08.005.

LIMA, D. R. M.; DLUGOSZ, F. L.; IURK, M. C.; PESCK, V. A. Uso de NDVI e SAVI para Caracterização da Cobertura da Terra e Análise Temporal em Imagens RapidEye. **Revista Espacios**, online, v. 38, n. 36, p. 7-21, 2017.

MIRANDA, M. S.; JÚNIOR V. A. S.; KÖRTING T. S.; MONTEIRO E. C. S.; E SILVA J. Q. AI4LUC: deep learning and automated mask labelling to support land use and land cover mapping in the *Cerrado* biome, **Remote Sensing Letters**, *[S. l.]*, v. 15, n 8, p. 850–860, ago. 2024. DOI: 10.1080/2150704X.2024.2382845.

MISHRA, D.; PATHAK, G.; SINGH, B. P.; MOHIT; SIHAG, P.; RAJEEV; SINGH, K.; SINGH, S. Crop classification by using dual-pol SAR vegetation indices derived from Sentinel-1 SAR-C data. **Environmental Monitoring and Assessment**, *[S. l.]*, v. 195, n. 1, p. 115, 2023. DOI: 10.1007/s10661-022-10591-x.

MONSALVE-TELLEZ, J. M.; TORRES-LEÓN, J. L.; GARCÉS-GÓMEZ, Y. A. Evaluation of SAR and Optical Image Fusion Methods in Oil Palm Crop Cover Classification Using the Random Forest Algorithm. **Agriculture**, *[S. l.]*, v. 12, n. 7, p. 955, 1 jul. 2022. DOI: 10.3390/agriculture12070955.

NAIK, N.; CHANDRASEKARAN, K.; MEENAKSHI SUNDARAM, V.; PANNEER, P. Assessment of land use and land cover change detection and prediction using deep learning techniques for the southwestern coastal region, Goa, India. **Environ Monit Assess**, *[S. l.]*, v. 196, n. 6, p. 1-34, 2024. DOI: 10.1007/s10661-024-12598-y.

PARADELLA, W. R.; MURA, J. C.; GAMA, F. F. **Monitoramento DInSAR para mineração e geotecnia**. 1. ed. São Paulo: Oficina de Textos, 2021.

PEDREGOSA *et al*. Scikit-learn: Machine Learning in Python. **Journal of machine learning research**, *[S. l.]*, v. 12, n. 85, p. 2825-2830, 2011.

RACZKO, E.; ZAGAJEWSKI, B. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. **European Journal of Remote Sensing**, *[S. l.]*, v. 50, n. 1, p. 144–154, 9 jan. 2017. DOI: 10.1080/22797254.2017.1299557.

RENNÓ, C. D.; SILVA, T. S. F. **Planilha**: Índice Kappa. 2023. Disponível em: https://www.dpi.inpe.br/~camilo/estatistica/xls/kappa10cond.xlsx. Acesso em: 24 outubro 2023.

ROUSE JR, J. W.; HASS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the Great Plains with ERTS. *In*: PROCEEDINGS OF THE THIRD EARTH RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM, 3., 1973, Washington D.C. **Anais...** Washington D.C.: NASA. p. 309.

SOLÓRZANO, J. V.; MAS, J. F.; GALLARDO-CRUZ, J. A.; GAO, Y.; DE OCA, A. F. M. Deforestation detection using a spatio-temporal deep learning approach with synthetic aperture radar and multispectral images. **ISPRS Journal of Photogrammetry and Remote Sensing**, *[S. l.]*, v. 199, p. 87-101, 2023. DOI: 10.1016/j.isprsjprs.2023.03.017.

SOUTHWORTH, J.; SMITH, A. C.; SAFAEI, M.; RAHAMAN, M.; ALRUZUQ, A.; TEFERA, B. B.; ... & HERRERO, H. V. Machine learning versus deep learning in land system science: a decision-making framework for effective land classification. **Frontiers in Remote Sensing**, *[S. l.]*, v. 5, p. 1-25, 2024. DOI: 10.3389/frsen.2024.1374862.

SOUZA, C.; AZEVEDO, T. **MapBiomas general "Handbook"**. São Paulo: MapBiomas, 2017.

SHEYKHMOUSA, M.; MAHDIANPARI, M.; GHANBARI, H.; MOHAMMADIMANESH, F.; GHAMISI, P.; HOMAYOUNI, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, *[S. l.]*, v. 13, p. 6308–6325, 2020. DOI: 10.1109/JSTARS.2020.3026724.

SCHULZ, D.; YIN, H.; TISCHBEIN, B.; VERLEYSDONK, S.; ADAMOU, R.; KUMAR, N. Land use mapping using Sentinel-1 and Sentinel-2 time series in a heterogeneous landscape in Niger, Sahel. **ISPRS Journal of Photogrammetry and Remote Sensing**, *[S. l.]*, v. 178, p. 97–111, ago. 2021. DOI 10.1016/j.isprsjprs.2021.06.005.

TALUKDAR, S.; SINGHA, P.; MAHATO, S.; PAL, S.; LIOU, Y. A.; RAHMAN, A. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations— A Review. **Remote Sensing**, *[S. l.]*, v. 12, n. 7, p. 1135, 2 abr. 2020. DOI: 10.3390/rs12071135.

TAVARES, P.; BELTRÃO, N. E. S.; GUIMARÃES, U. S.; TEODORO, A. C. Integration of Sentinel-1 and Sentinel-2 for Classification and LULC Mapping in the Urban Area of Belém, Eastern Brazilian Amazon. **Sensors**, *[S. l.]*, v. 19, n. 5, p. 1140, 6 mar. 2019. DOI: 10.3390/s19051140.

TRINDADE, P. M. P.; PEIXOTO, D. W. B.; SILVEIRA, G. V.; KUPLICH, T. M.; NARVAES, I. S. Mapeamento de cobertura e uso da terra no bioma Pampa utilizando diferentes sensores orbitais e classificador *Random Forest.* **GEOgrafias**, *[S. l.]*, v. 19, n. 2, jul./dez. 2023. DOI: 10.35699/2237-549X.2023.46915.

WANG, W.; GADE, M.; STELZER, K.; KOHLUS, J.; ZHAO, X.; FU, K. A Classification Scheme for Sediments and Habitats on Exposed Intertidal Flats with Multi-Frequency Polarimetric SAR. **Remote Sensing**, [S. l.], v. 13, n. 3, p. 360, 2021. DOI: 10.3390/rs13030360.

WANG, J.; BRETZ, M.; DEWAN, M. A. A.; & DELAVAR, M. A. Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. **Science of The Total Environment**, v. 822, p. 153559, 2022. DOI: 10.1016/j.scitotenv.2022.153559.

ZUHLKE, M.; FOMFERRA, N., BROCKMANN, C., PETERS, M., VECI, L., MALIK, J., & REGNER, P. SNAP (sentinel application platform) and the ESA sentinel 3 toolbox. *In*: SENTINEL-3 FOR SCIENCE WORKSHOP, 3., 2015, Venice. **Proceedings...** Venice: ESA, 2015. p. 21.

Contribuições de autoria

1 – Andressa Kossmann Ferla

Instituto Nacional de Pesquisas Espaciais, Mestre em Ciência e Tecnologia Ambiental e Engenheira Sanitarista e Ambiental pela Universidade Federal de Santa Maria, UFSM.
https://orcid.org/0000-0003-1801-3355 • andressaferla94@gmail.com
Contribuição: Análise Formal, Investigação, Metodologia, Software, Visualização, Escrita – primeira redação, Escrita – revisão e edição

2 – Tatiana Mora Kuplich

Instituto Nacional de Pesquisas Espaciais, Mestre em Sensoriamento Remoto pelo Instituto Nacional de Pesquisas Espaciais (INPE, São José dos Campos) e PhD pela School of Geography da University of Southampton na Inglaterra.
https://orcid.org/0000-0003-0657-4024 • tkuplich@gmail.com
Contribuição: Administração do projeto, Supervisão, Escrita – revisão e edição

3 – Igor da Silva Narvaes

Instituto Nacional de Pesquisas Espaciais, Doutor em Sensoriamento Remoto pelo Instituto Nacional de Pesquisas Espaciais (2010) e Pesquisador Sênior da Coordenação Espacial Sul (COESU) do Instituto Nacional de Pesquisas Espaciais (INPE).

https://orcid.org/0000-0002-9950-895X • igor.narvaes@inpe.br

Contribuição: Conceituação, Metodologia, Administração do projeto, Supervisão, Escrita – revisão e edição

Como citar este artigo

FERLA, A. K.; NARVAES, I. S.; KUPLICH, T. M. Utilização do classificador Random Forest para Classificação do Uso e Cobertura da Terra a partir de Dados Sentinel 1 e 2 em região campestre no bioma Mata Atlântica. **Geografia Ensino & Pesquisa**, Santa Maria, v. 29, e 87967, 2025. Disponível em: 10.5902/2236499487967. Acesso em: dia mês abreviado ano.