

# USO DE SÉRIES TEMPORAIS SENTINEL-2 PARA MAPEAMENTO DE CLASSES DE COBERTURA VEGETAL DO NORTE DE RONDÔNIA

Filipe Castro Felix<sup>1</sup>, André Silva Tavares<sup>1</sup>, Alexandre Camargo Coutinho<sup>2</sup>, João Francisco Gonçalves Antunes<sup>2</sup>, Júlio César Dalla Mora Esquerdo<sup>2</sup>

<sup>1</sup>Bolsista CNPq do Projeto TerraClass Amazônia, filipe.felix@colaborador.embrapa.br, andre.tavares@colaborador.embrapa.br; <sup>2</sup>Embrapa Agricultura Digital, Av. André Tosello, 209, CEP 13083-886, Campinas, SP, Brasil, alex.coutinho@embrapa.br, joao.antunes@embrapa.br, julio.esquerdo@embrapa.br

## RESUMO

Atualmente, os sistemas de observação da Terra produzem grandes volumes de imagens que permitem o monitoramento de diversos fenômenos espaço-temporais. Neste contexto, este estudo visou explorar o uso de séries temporais do satélite Sentinel-2 e do algoritmo *Random Forest* na classificação supervisionada do uso e cobertura da terra na região do município de Buritis-RO, sudoeste da Amazônia, que é caracterizada pela expansão agrícola acelerada nas últimas décadas. Para isso, foram avaliados dois cenários: (I) estação seca (três meses) e (II) um ano agrícola, a fim de determinar qual o período mais adequado ao mapeamento dessa região. O cenário (II) apresentou a maior acurácia (88,66%), porém nossos resultados demonstraram que a estação seca e o uso das bandas *short wavelength infrared* (SWIR) foram determinantes nos mapeamentos, sendo indicadas para abordagens futuras de mapeamento dessa região.

**Palavras-chave** — desmatamento, *sits*, *Self-organizing Maps*, *Random Forest*, SWIR.

## ABSTRACT

Nowadays, earth observation systems produce large volumes of images that allow the monitoring of several spatiotemporal phenomena. In this context, we aimed to explore the use of satellite image time series of Sentinel-2 and *Random Forest* algorithm to the supervised classification of the land use and land cover (LULC) at the region of Buritis-RO, southwestern of Brazilian Amazon, which represents an area of intense expansion of agricultural frontiers. Then, two scenarios were evaluated: (I) dry season, and (II) one year, aiming to determine which period is most suitable for mapping the region. Scenario II presented the best map, with an accuracy of 88.66%. However, our results showed that the dry season and the use of short wavelength infrared (SWIR) bands were determinants for the mapping. Therefore, we indicate these bands for future approaches that aim to map this region.

**Key words** — deforestation, *sits*, *Self-organizing Maps*, *Random Forest*, SWIR.

## 1. INTRODUÇÃO

A intensificação do desmatamento no bioma Amazônia está relacionada a várias razões, como a transição das áreas de pastagem em plantações de soja e à conversão das áreas naturais em pastagens [1]. Este cenário é justificado por diversos fatores, mas, segundo [2], o crescimento da demanda global por *commodities* de exportação é o principal deles.

Rondônia é o estado mais ameaçado do bioma, com cerca de 61% do seu território afetado por alguma atividade antrópica [3]. Historicamente, a mineração e a construção de estradas eram os vetores de alteração da paisagem no estado [4]. No entanto, com a vasta expansão das áreas para pecuária nas últimas décadas, essa atividade é apontada como a principal contribuinte aos novos desmatamentos e representa o maior uso de florestas tropicais convertidas [5].

Uma alternativa promissora para capturar essas dinâmicas e comparar os cenários em diferentes momentos é o uso de séries temporais de imagens de satélites. Esta abordagem permite a obtenção de dados de maneira consistente e periódica [6], subsidiando a tomada de decisão em diferentes escalas espaço-temporais.

Nesse contexto, este trabalho visou explorar a classificação de séries temporais baseada em aprendizado de máquina por meio do pacote computacional *sits* [7]. O objetivo foi estimar a cobertura vegetal arbustiva/arbórea e herbácea de uma área no sudoeste da Amazônia brasileira, que é caracterizada pela expansão acelerada das fronteiras agrícolas. A hipótese foi a de que a estação seca é o período mais adequado ao mapeamento proposto, já que nessa época as formações herbáceas tendem a apresentar pior condição da biomassa verde em relação às fitofisionomias de estrato arbóreo. Para testar essa hipótese, foram utilizados cubos de dados do Sentinel-2, o algoritmo *Random Forest* e avaliados dois cenários: (I) estação seca / três meses e (II) um ano agrícola.

## 2. MATERIAL E MÉTODOS

### 2.1. Área de estudo

O estudo foi desenvolvido na região do município de Buritis, localizado no Norte de Rondônia (Figura 1).

Segundo a classificação de Köppen, o clima é tropical úmido (Am), que é caracterizado por uma estação seca bem definida entre junho e agosto/setembro e precipitação média anual de 2.600 mm [8]. Já a vegetação predominante é a Floresta Ombrófila Aberta, que é um tipo de vegetação que marca a transição da Floresta Amazônica nessa região [9].

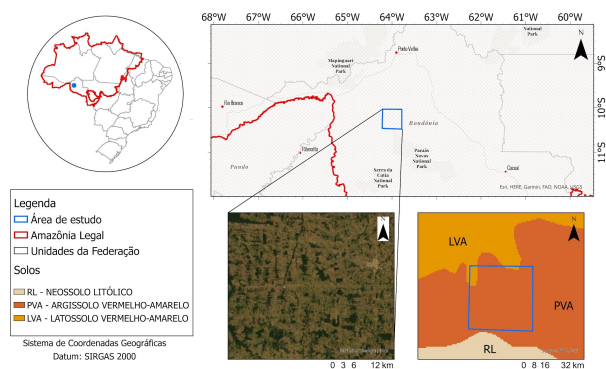


Figura 1. Localização geográfica e caracterização do solo da área de estudo no Norte de Rondônia.

## 2.2. Processamento

### 2.2.1. Cubos Sentinel-2

Os cubos de dados Sentinel-2/MSI (“S2-SEN2COR\_10\_16D\_STK-1”) foram gerados a partir das coleções do *Brazil Data Cube* (BDC). Assim, foram estabelecidos dois períodos de estudo: (I) estação seca / 3 meses (11 de julho – 29 setembro de 2020); (II) um ano agrícola (27 de julho de 2019 – 26 de julho de 2020). Nos dois cenários avaliados foram utilizadas 14 variáveis, sendo compostas por 10 espectrais do Sentinel-2 (B01, B02, B03, B04, B05, B06, B07, B08A, B09, B11, B12), os índices de vegetação *Normalized Difference Vegetation Index* (NDVI) e *Enhance Vegetation Index* (EVI) e o produto CLOUD de máscara de nuvem.

### 2.2.2. Coleta e filtragem das amostras

Para a coleta das amostras de treinamento foi utilizado o mosaico anual de imagens Sentinel-2 sem nuvens, referente ao ano de 2020, disponibilizado gratuitamente pela Comissão Europeia (<https://maps.eox.at>) e integrado ao software QGIS 3.22.9 por meio de um serviço WMS. Assim, foram coletadas 170 amostras das classes temáticas vegetação natural (50), pastagem arbustiva (50), pastagem herbácea (50) e água (20).

As amostras de vegetação natural consideraram as áreas de vegetação arbórea e dossel contínuo. Já as amostras de pastagem arbustiva foram baseadas nas áreas de vegetação arbustiva esparsa e com presença de

indivíduos arbóreos, representando os diversos estágios de degradação e regeneração observados na área. No caso da pastagem herbácea, consideramos as áreas em processo produtivo ou com baixa infestação de invasoras arbustivas, assim como a ausência de indivíduos arbóreos.

Posteriormente, avaliamos a qualidade do conjunto amostral a partir do *Self-organizing Maps* (SOM) [10], que produz agrupamentos de séries temporais e inferência Bayesiana para remoção de amostras com baixo poder discriminatório.

### 2.2.3. Treinamento e classificação com Random Forest

As amostras selecionadas foram submetidas ao treinamento do algoritmo *Random Forest* (RF). Embora o *sits* também disponibilize outros algoritmos de aprendizado de máquina, como o *Extreme Gradient Boosting* e o *lightTAE*, a escolha do RF foi baseada nos diversos resultados já reportados na literatura voltados à classificação do uso e cobertura da terra. A parametrização de treinamento utilizada foi número de

árvores (“*n<sub>tree</sub>*”) = 500, número de variáveis preditivas (“*mtry*”) = *default* do pacote *sits*.

## 3. RESULTADOS E DISCUSSÃO

O cenário (I) do estudo foi composto por seis imagens e o processo de filtragem das amostras pelo SOM permitiu a remoção de 41 amostras. Assim, foram utilizadas 129 amostras para o treinamento do RF: vegetação natural (41), pastagem arbustiva (32), pastagem herbácea (41) e água (15). A acurácia deste modelo para o mapeamento foi de 86,05%. Em contrapartida, o cenário (II) foi composto por 24 imagens e foram removidas 73 amostras. Portanto, o conjunto final de treinamento foi de 97 amostras: vegetação natural (35), pastagem arbustiva (17), pastagem herbácea (34) e água (11). A acurácia deste modelo foi de 88,66% e o resultado das duas classificações é apresentado na Figura 2.

Além disso, o uso do RF possibilitou avaliar a importância das variáveis utilizadas como preditoras à separabilidade das classes a partir da diminuição média do índice de desigualdade de Gini (DMG) e da diminuição média da precisão (DMP). A Figura 3 apresenta os resultados desses indicadores para os dois cenários. Os gráficos demonstraram que as bandas *short wavelength infrared* (SWIR) 11 e 12 foram as variáveis de maior importância preditiva nos modelos das classificações, sobretudo durante o período de 27 de julho a 28 de agosto.

Os resultados das classificações não confirmaram a hipótese do estudo, em vista que a acurácia do cenário (II) foi superior. Na estação seca, ao contrário do que era

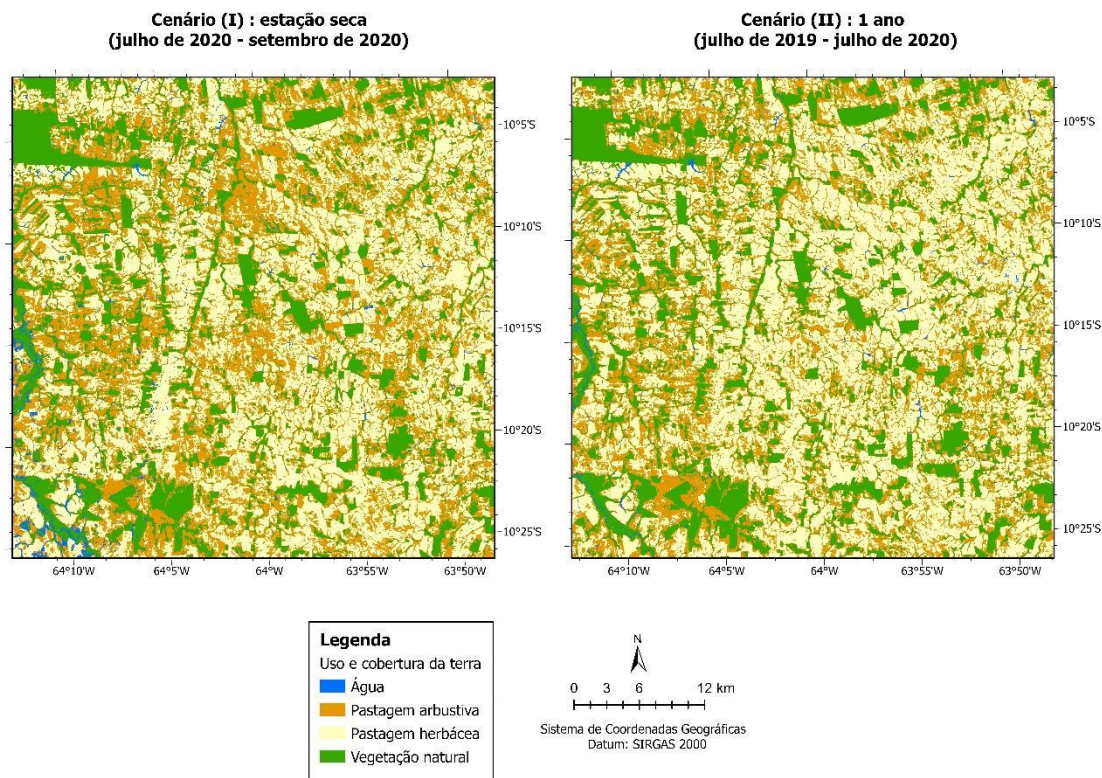


Figura 2. Mapas de uso e cobertura da terra da região de Buritis, Norte de Rondônia.

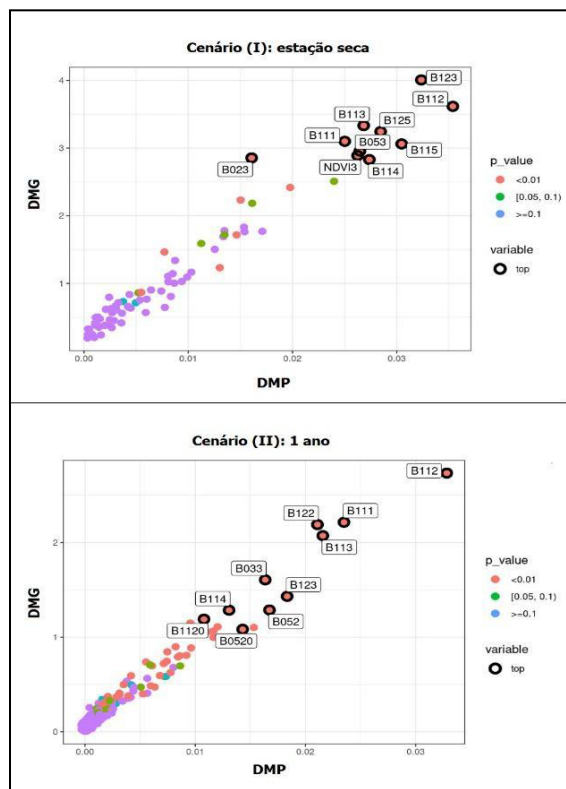
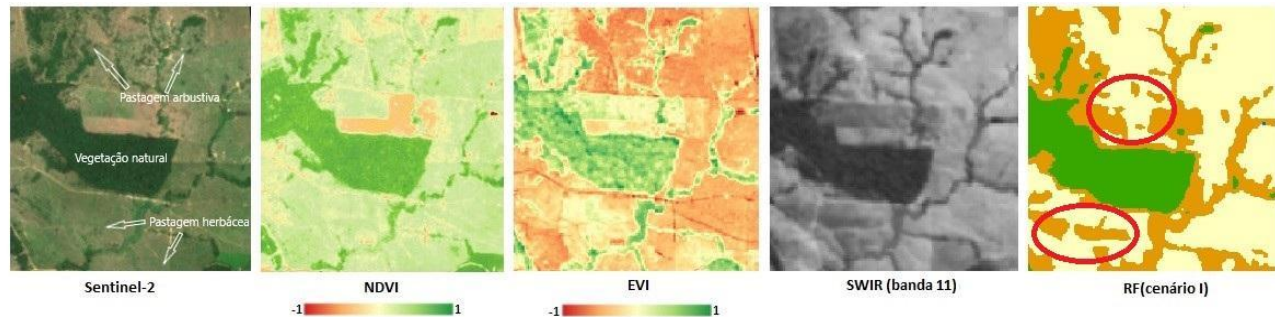


Figura 3. Importância das variáveis derivadas do modelo *Random Forest* em cada cenário de estudo.

esperado, as confusões entre as classes de pastagens foram elevadas, sendo observada principalmente a inclusão da classe arbustiva em áreas com padrão herbáceo bem definido (Figura 4).

Desta maneira, os resultados indicaram que, embora as diferenças na refletância das pastagens seja mais intensa na estação seca, o uso das bandas do visível e, também, dos índices de vegetação NDVI e EVI é insuficiente para o mapeamento dessa região, que é predominantemente composta por remanescentes florestais e pastagens em diferentes estágios de desenvolvimento. Tal fato foi confirmado pela Figura 3, em que a estação seca e as bandas do SWIR se mostraram determinantes nas classificações das duas séries temporais. Portanto, sugere-se que essas faixas espectrais possam ser exploradas em abordagens futuras, com o objetivo de mapear esta região da Amazônia ou outras áreas de contexto semelhante.

Em contrapartida, [11] empregou o algoritmo RF e séries temporais de imagens Landsat (1985-2017) no mapeamento de pastagens no Brasil e confirmaram as potencialidades de uso do NDVI na distinção das classes. A acurácia do estudo (87%) foi similar à obtida nos cenários (I) e (II), o que reflete os desafios de mapeamento dos tipos de pastagem em diferentes escalas e sensores.



**Figura 4.** Exemplo de uma imagem de julho de 2020 e a sua representação com o NDVI, EVI, a banda 11 (SWIR) e o mapa gerado com o modelo RF do cenário (I); com destaque às áreas em vermelho que demonstram regiões de confusão do classificador.

Ainda, cumpre ressaltar que a área de estudo apresenta uma baixa fertilidade natural dos solos e sofre com a prática recorrente das queimadas, que inviabilizam grande parte das novas áreas de pastagens implantadas. Neste sentido, ao considerarmos que esta região da Amazônia concentra 77% do desmatamento nos últimos cinco anos, identificar e monitorar a regeneração dessas áreas torna-se imprescindível à compreensão das dinâmicas locais a longo prazo e os nossos resultados podem direcionar a formulação de metodologias para tais finalidades.

### 5. Conclusões

As diferenças na refletância da cobertura vegetal arbustiva/arbórea e herbácea são mais intensas na estação seca e as bandas do SWIR são determinantes na distinção dessas classes. O uso das séries temporais de imagens Sentinel-2 é promissor para a compreensão das dinâmicas da paisagem e pode subsidiar políticas públicas regionais para a Amazônia brasileira.

### Agradecimentos

Ao Centro Gestor e Operacional do Sistema de Proteção da Amazônia (Censipam) pelos recursos financeiros utilizados nesta pesquisa (Projeto TerraClass Amazônia/TED Censipam-CNPq); ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio; ao projeto Brazil Data Cube (INPE) por ceder os cubos de imagens Sentinel-2 utilizados neste trabalho; e aos desenvolvedores do pacote computacional sits.

### 7. Referências

[1] E. Barona, N. Ramankutty, G. Hyman and O. T. Coomes. The role of pasture and soybean in deforestation of the Brazilian Amazon, *Environmental Research Letters*, 5: 9pp, 2010.  
[2] P. M. Fearnside, A. M. R. Figueiredo and S. C. M. Bonjour. Amazonian forest loss and the long reach of China's influence, *Environment, Development and Sustainability*, 15 (2): 13pp, 2013.

[3] C. R. Sanquetta, A. S. Bastos, M. N. I. Sanquetta, I. M. Barberena, A. P. D. Corte, A. Queiroz and L. P. U. Almeida. Assessing the carbon stock of cultivated pastures in Rondônia, southwestern Brazilian Amazon, *Anais da Academia Brasileira de Ciências*, 94 (4): e20210262, 2022.  
[4] L. S. Guild, J. B. Kauffman, W. B. Cohen, C. A. Hlavka and D. E. Ward. Modeling biomass burning emission for Amazon forest and pastures in Rondônia, Brazil, *Ecological Applications*, 14: 14pp, 2004.  
[5] J. F. L. de Moraes, B. Volkoff, C. C. Cerri and M. Bernoux. Soil properties under Amazon forest and changes due pasture installation in Rondônia, Brazil, *Geoderma*, 70: 18pp, 1996.  
[6] R. Simões, G. Câmara, G. Queiroz, F. Souza, P. R. Andrade, L. Santos, A. Carvalho and K. Ferreira. Satellite image time series analysis for big earth observation data, *Remote Sensing*, 13 (13): 2428, 2021.  
[7] G. Câmara, R. Simões, F. Souza, C. Pelletier, A. Sanchez, P. R. Andrade, K. Ferreira and G. Queiroz. *Satellite image times series analysis on earth observation data cubes*. Online < disponível em: <https://e-sensing.github.io/sitsbook/index.html>>, 2021.  
[8] C. A. Alvares, J. L. Stape, P. C. Sentelhas, J. L. M. Gonçalves and G. Sparovek. Köppen's climate classification map for Brazil, *Meteorologische Zeitschrift*, 22(6), 17pp. 2013.  
[9] IBGE (Instituto Brasileiro de Geografia e Estatística). *Manual técnico da vegetação brasileira*, IBGE, Rio de Janeiro, 2012.  
[10] L. Santos, K. R. Ferreira, G. Câmara, M. C. A. Picoli and R. E. Simões. Quality control and class noise reduction of satellite image time series, *ISPRS Journal of Photogrammetry and Remote Sensing*, 177, 13 pp. 2021.  
[11] L. Parente, L. Ferreira, A. Faria, S. Nogueira, F. Araújo, L. Teixeira and S. Hagen. Monitoring the Brazilian pasturelands: A new mapping approach based on the Landsat 8 spectral and temporal domains. *International Journal of Applied Earth Observation and Geoinformation*, 62, 8pp. 2017.