

MAPEAMENTO DA COBERTURA FLORESTAL EM LAGOINHA (SP) POR MEIO DE IMAGENS SENTINEL-2 E ALGORITMO RANDOM FOREST

Gustavo Bayma¹, Taya Cristo Parreiras², Danielle Elis Garcia Furuya³, Édson Luis Bolfe^{2,3}, Edson Eyji Sano⁴, Victoria Beatriz Soares⁵, Victória Hellena Matusevicius e de Castro²

¹Embrapa Meio Ambiente, Jaguariúna, SP, gustavo.bayma@embrapa; ²Universidade Estadual de Campinas, tayacristo1@gmail.com, v245192@dac.unicamp.br; ³Embrapa Agricultura Digital, Campinas, SP, danielle.furuya@colaborador.embrapa.br; edson.bolfe@embrapa.br; ⁴Embrapa Cerrados, Planaltina, DF, edson.sano@embrapa, e ⁵Pontifícia Universidade Católica de Campinas, Campinas, SP, victoria.leandro@colaborador.embrapa.br

RESUMO

Este artigo apresenta um método para mapear a cobertura florestal no município de Lagoinha (SP) utilizando o algoritmo Random Forest disponível no Google Earth Engine (GEE). A classificação foi realizada com bandas multiespectrais e índices de vegetação por diferença normalizada (NDVI) e índice de água por diferença normalizada (NDWI), derivadas do Sentinel-2 MSI, entre 2021 e 2023. No Nível 1, foram discriminadas áreas de Floresta e Não-Floresta, e no Nível 2, as áreas de Floresta foram discriminadas entre Floresta Natural e Silvicultura. A exatidão global foi de 91,0% no Nível 1 e de 93,8% no Nível 2. A área total de Floresta mapeada foi de 8.586 hectares, com 7.918 hectares de Floresta Natural e 668 hectares de Silvicultura. A área de Floresta Natural em Áreas de Preservação Permanente foi de 777 hectares. O método oferece uma abordagem rápida e eficiente para a análise da cobertura florestal em nível municipal.

Palavras-chave — classificação, uso da terra, Mata Atlântica, vegetação natural, silvicultura

ABSTRACT

This article presents an analysis of the compliance of Permanent Preservation Areas through forest cover mapping in the municipality of Lagoinha (SP), using the Random Forest algorithm available in Google Earth Engine (GEE). The classification was performed using multispectral bands and NDVI and NDWI vegetation indices derived from Sentinel-2 MSI images between 2021 and 2023. In Level 1, areas of Forest and Non-Forest were distinguished, while in Level 2, Forests were discriminated between Natural Forest and Silviculture. The overall accuracy was 91.0% for Level 1 and 93.8% for Level 2. The total mapped forest area was 8,586 hectares, comprising 7,918 hectares of Natural Forest and 668 hectares of Forestry. The area of Natural Forest within Permanent Preservation Areas was 777 hectares (72.5% of the total).

Key words — classification, land cover, Atlantic Forest, natural vegetation, forestry.

1. INTRODUÇÃO

O bioma Mata Atlântica tem passado por intensas mudanças de uso e cobertura da terra, desde o período colonial no século XVI. Atualmente, encontra-se bastante fragmentado devido à expansão de extensas áreas agropecuárias e centros urbanos, restando apenas 11,7% da sua cobertura vegetal original [1] [2]. A consequente fragmentação da paisagem acaba por afetar a estrutura, disposição, composição e biodiversidade, dificultando o fluxo de genes e indivíduos entre os fragmentos remanescentes. O isolamento de populações causadas pela fragmentação antrópica pode intensificar a competição, resultando em uma redução da qualidade do habitat [3].

Dada a importância dos serviços ecossistêmicos florestais, os avanços nas tecnologias de sensoriamento remoto por satélite e o rápido aumento da disponibilidade de dados, sistemas de monitoramento da perda florestal têm sido desenvolvidos utilizando dados de sensoriamento remoto por satélites com séries temporais consistentes e resoluções espaciais que geralmente variam entre 10 e 50 m [4]. Um exemplo deste avanço tecnológico é a investigação da conformidade do uso das terras dentro dos limites das áreas legalmente protegidas [5], [6], [7], [8].

O Código Florestal (Lei Federal nº 12.651 de 25 de maio de 2012) [9] estabelece que as áreas situadas ao longo de cursos d'água são áreas de proteção permanente. A largura da faixa varia de acordo com o tamanho do curso d'água e a vegetação ao longo destas áreas deve ser mantida pelo proprietário. Ainda, a Lei Estadual (SP) nº 15.684 de 14 de janeiro de 2015 [10] indica que espécies como Pinus e Eucalipto não podem ser utilizadas na recomposição de áreas de preservação permanente (APP). Neste contexto, o objetivo deste trabalho foi realizar o mapeamento da cobertura florestal natural (FN) e de silvicultura (SILV) no município de Lagoinha (SP), utilizando o algoritmo *Random Forest* aplicado a dados do satélite Sentinel-2. A análise teve como foco o diagnóstico das áreas de FN dentro das áreas de APP relacionadas aos cursos d'água, excluindo, assim, da análise, as áreas de topo de morro e com declividade superior a 45°.

2. MATERIAL E MÉTODOS

2.1 Área de estudo

A área de estudo é o município de Lagoinha (SP) localizado no trecho paulista do Vale do Paraíba, latitude 23°05' S; longitude 45°12' O (Figura 1). Este município é um dos dez municípios que compõem o Centro de Ciência para o Desenvolvimento da Agricultura Digital – SemeAR Digital [11]. Este projeto é liderado pela Embrapa, em parceria com diversas instituições de pesquisa e tecnologia, possui foco no pequeno e médio produtor e visa desenvolver e difundir tecnologias voltadas à produção agrícola considerando eficiência do uso da terra e dos recursos naturais.

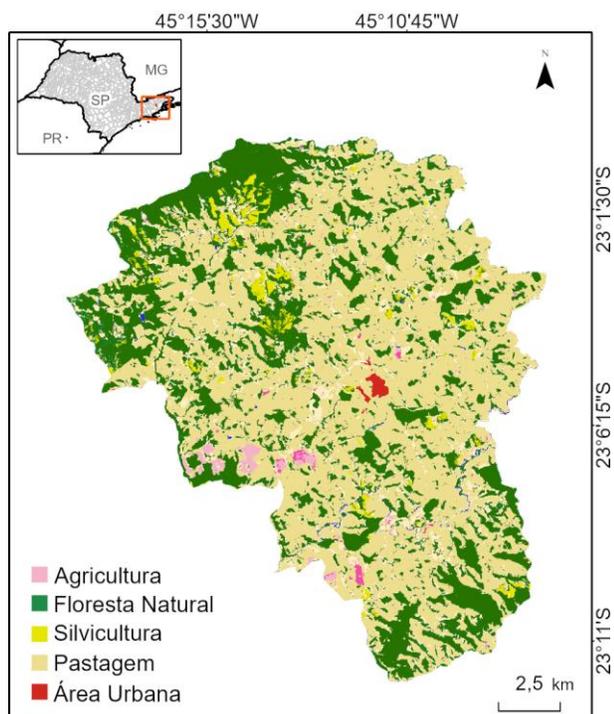


Figura 1. Principais classes de uso e cobertura do solo do município de Lagoinha (SP), no ano de 2023. Fonte: Souza et al. (2020).

Lagoinha (SP) está localizada no bioma Mata Atlântica e abrange aproximadamente 25,5 mil hectares. Em 2023, 19,3 mil hectares correspondiam às classes de uso agropecuário, sendo 13,4 mil hectares de áreas de pastagens. As áreas de floresta natural totalizaram cerca de 6,1 mil hectares, enquanto as áreas de silvicultura alcançaram 827 hectares no mesmo ano [2]. Apresenta dois tipos predominantes de solo: Latossolos Vermelho-Amarelo, concentrados na parte sul, e Argissolos Vermelho-Amarelos, presentes ao norte. Conforme a classificação climática de Köppen, os climas predominantes são o subtropical de altitude (Cwb) e o oceânico temperado (Cfb), ambos característicos de regiões temperadas [12]. Em relação ao relevo, a classe montanhosa,

com declividade entre 45% e 75%, cobre mais de 48% da área de estudo.

2.2 Material e Métodos

No Sistema Nacional de Cadastro Ambiental Rural - SICAR [13] (MMA, 2024) foram obtidos os arquivos, no formato *shapefile*, correspondentes às APPs, Área Rural Consolidada (ARC), Área do Imóvel (AI), Reserva Legal (RL) e Vegetação Nativa (VN). Em cada *shapefile* adquirido, foi utilizada a ferramenta “*Dissolve*” para mesclar os polígonos que se sobrepunham, de forma que a mesma área fosse contabilizada apenas uma única vez.

Para o mapeamento, a etapa inicial consistiu na geração dos pontos amostrais, realizada em dois níveis de análise e de forma hierárquica. No Nível 1, foram consideradas as áreas de cobertura florestal (F) e não florestal (NF). Para a geração aleatória dos pontos amostrais de F, foi selecionado o *shapefile* das áreas de RL, seguido pelo cálculo de um *buffer* interno de 50 metros para mitigar o efeito de borda. Em seguida, foram geradas aleatoriamente 225 pontos amostrais, com uma distância mínima de 150 metros entre cada ponto amostral. Para os pontos amostrais de NF, o processo incluiu a seleção do *shapefile* de AI, subtraindo-se as áreas de APP, RL e VN, e o cálculo de um *buffer* interno de 50 metros para reduzir o efeito de borda. Foram gerados aleatoriamente 250 pontos amostrais, também assegurando uma distância mínima de 150 metros entre os pontos amostrais. Ambos os conjuntos amostrais foram submetidos a uma inspeção visual no *Google Earth Pro*, com o objetivo de garantir que as amostras não estivessem associadas a classes distintas.

No Nível 2, para a geração dos 138 pontos amostrais de FN, foi criado um *buffer* interno de 50 metros nos polígonos com área superior a 2 hectares do mapa de F gerado no nível 1. Em seguida, realizou-se uma inspeção visual no *Google Earth Pro* para evitar a seleção de amostras em áreas de SILV. Os 46 pontos amostrais de SILV foram selecionados por meio da inspeção visual das imagens de alta resolução do ano de 2023 disponíveis no *Google Earth Pro*. Cabe ressaltar que a inspeção visual considerou informações como cor, textura, forma e dados de séries temporais na interpretação de cada ponto amostral [14].

O mapeamento foi realizado por meio de classificação supervisionada, *pixel-a-pixel*, no *Google Earth Engine* (GEE), a partir das bandas multiespectrais e de índices de vegetação por diferença normalizada (NDVI) [15] e índice de água por diferença normalizada (NDWI) [16] do Sentinel-2 *Multispectral Instrument* (MSI) com cobertura de nuvens inferior a 3%, em 10 metros de resolução espacial, entre 2021 e 2023. As classificações foram realizadas em dois níveis para isolar *pixels* de F dos demais tipos de uso e cobertura (Nível 1) e, em um segundo nível, para refinar a identificação de FN e SILV (Nível 2). Para ajuste de parâmetros, diversos testes foram realizados com número de árvores (*nTree*) entre 50 e 300, e divisão do conjunto de amostras (*split*) entre treinamento e validação (0,4; 0,5; 0,6 e 0,7). Os melhores

resultados foram obtidos no Nível 1 com $nTree = 50$ e $split = 0,6$ (60% teste e 40% validação). No Nível 2, os melhores modelos foram produzidos com 100 árvores e 0,7 de $split$ (70% teste e 30% validação). Para avaliação do desempenho do classificador RF, foram considerados os valores de matriz de confusão, Exatidão Global [17] e índice Kappa [18].

Os resultados dos mapeamentos foram exportados do GEE para posterior análise em ambiente de Sistema de Informações Geográficas (SIG), onde foi realizado o cruzamento com as áreas de APP relacionadas aos cursos d'água (APP-hidrografia).

3. RESULTADOS E DISCUSSÕES

A área mapeada de F totalizou 8.586 hectares, dos quais 7.918 hectares correspondem à FN (92,2%) e 668 hectares à SILV (7,8%) (Figuras 2 e 3). Em comparação, os dados do MapBiomas [2], para o mesmo ano, indicaram uma área de cobertura florestal de 6.972 hectares, sendo 6.145 hectares de FN (88,1%) e 827 (11,9%) hectares de SILV. Cabe ressaltar que a metodologia do MapBiomas utiliza imagens Landsat, com resolução espacial de 30 metros, enquanto na metodologia aqui proposta foi baseada em imagens Sentinel-2, que possuem uma resolução espacial de 10 metros.

Segundo os dados obtidos no Sistema de Cadastro Ambiental Rural (SICAR), a área de APP-hidrografia no município de Lagoinha (SP) é de 2.487 hectares, aproximadamente 10% do total do município. Deste total, 1.416 hectares (56,9%) foram classificados como ARC, restando 1.072 hectares de APP-hidrografia sem nenhum tipo de ARC (APP-hidrografia-ARC). O cruzamento entre o mapa de FN e a área de APP-hidrografia-ARC, totalizou 777 hectares, ou seja, 72,5% das áreas de APP-hidrografia-ARC possuem área de FN.

A avaliação da acurácia da classificação no Nível 1 resultou em uma Exatidão Global de 91,0%. A avaliação no Nível 2 apresentou valores de Exatidão Global de 93,8%. Em ambos os níveis, o índice Kappa foi de 0,81. Com base nas métricas obtidas, o mapeamento obtido em ambos os níveis pode ser considerado de qualidade muito boa, com alto valor de exatidão. Esses resultados são comparáveis aos resultados obtidos por Peng et al. [19], Waśniewski et al. [20] e Hościło [21], que obtiveram Exatidão Global de 89,3%, 94% e, 98,3% respectivamente, no mapeamento de áreas de Floresta (F) e Não-Floresta (NF) utilizando o algoritmo *Random Forest* em imagens Sentinel-2.

A Exatidão Global da classificação de F foi de 92,1% e a da área NF foi de 89,6%, no Nível 1. No Nível 2, as áreas de FN e SILV obtiveram Exatidão Global de 94,1% e 92,3%, respectivamente. As matrizes de confusão dos Níveis 1 e 2 são mostradas nas Tabelas 1 e 2, respectivamente.

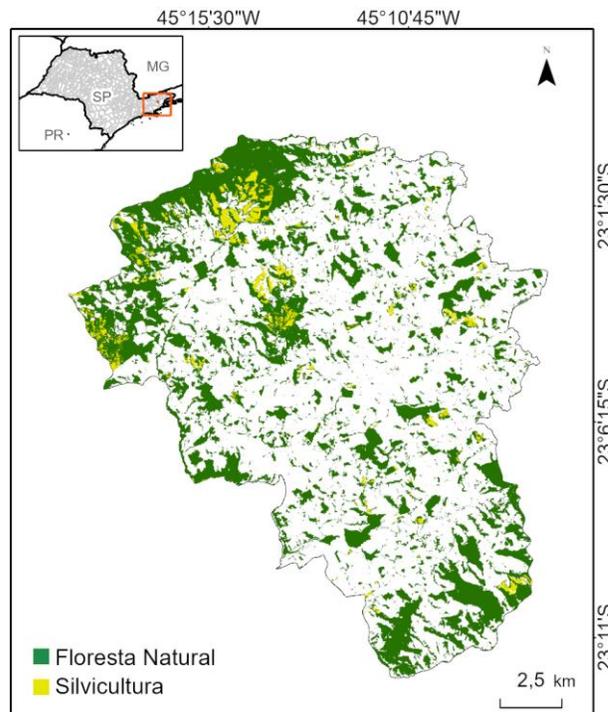


Figura 2. Distribuição espacial das áreas de Floresta Natural e Silvicultura no município de Lagoinha (SP), no ano de 2023



Figura 3. Exemplo do resultado da classificação das áreas de Floresta Natural (verde escuro) e Silvicultura (verde claro) sobrepostas à imagem Sentinel-2, R(3)G(2)B(1), de 18/09/2023.

Nível 1	Floresta (F)	Não-Floresta (NF)
Floresta (F)	93	8
Não-Floresta (NF)	8	69
Total	101	77

Tabela 1. Matriz de confusão do mapeamento no Nível 1 da classificação hierárquica, em valores absolutos.

Nível 2	Floresta Natural (FN)	Silvicultura (SILV)
Floresta Natural (FN)	49	1
Silvicultura (SILV)	3	12
Total	52	13

Tabela 2. Matriz de confusão do mapeamento no Nível 2 da classificação hierárquica, em valores absolutos.

4. CONCLUSÕES

Os resultados deste estudo demonstraram que a aplicação do algoritmo *Random Forest* em séries temporais de bandas multiespectrais e de índices de vegetação NDVI e NDWI, derivados do Sentinel-2, permite mapear, de forma eficiente e precisa, as áreas de cobertura de Floresta Natural e Silvicultura. O método proposto pode ser sistematizado e implementado de forma operacional, e assim contribuir com o mapeamento, monitoramento e diagnóstico da cobertura florestal em nível municipal.

5. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - processos 2022/09319-9, 2024/05205-4 e 2024/13150-5).

6. REFERÊNCIAS

[1] M. C. Ribeiro; J. P. Metzger, A. C. Martensen, F. J. Ponzoni, M. M. Hirota. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, v. 142, p. 1141–1153, 2009.

[2] C. M. Souza et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and earth engine. *Remote Sensing*, v. 12, 2735, 2020.

[3] G. Buffa, S. Del Vecchio, E. Fantinato, V. Milano. Local versus landscape-scale effects of anthropogenic land-use on forest species richness. *Acta Oecologica*, v. 86, p. 49-56, 2018.

[4] E. E. Sano, P. Rizzoli, C.N. Koyama, M. Adami, Y. E. Shimabukuro, G. Bayma, D. M. Freitas. Comparative analysis of the global forest/non-forest maps derived from SAR and optical sensors. Case studies from Brazilian Amazon and Cerrado biomes. *Remote Sensing*, v. 13, 367, 2021.

[5] S. N. Oliveira, O. A. de Carvalho Júnior, R. A. T. Gomes, R. F. Guimarães, C. M. McManus. Deforestation analysis in protected areas and scenario simulation for structural corridors in the agricultural frontier of Western Bahia, Brazil. *Land Use Policy*, v. 61, p. 40-52, 2017.

[6] J. M. F. Souza, E. F. Reis, A. S. Martins, A. L. F. Santos. Avaliação dos conflitos no uso da terra na bacia hidrográfica do ribeirão Lamarão, Distrito Federal. *Ciência Florestal*, v. 29, p. 950–964, 2019.

[7] D. E. G. Furuya; J. A. F. Aguiar; N. V. Estrabis; M. M. F. Pinheiro; M. T. G. Furuya, D. R. Pereira, W. N. Gonçalves; V. Liesenberg; J. Li; J. Marcato Júnior, L. P. Osco; A. P. M. Ramos. A Machine learning approach for mapping forest vegetation in riparian zones in an Atlantic biome environment using Sentinel-2 imagery. *Remote Sensing*, v. 12, 4086, 2020.

[8] G. A. V. Mataveli, G. Pereira, M. E. D. Chaves, F. S. Cardozo, S. C. Stark, Y. E. Shimabukuro, L. E. O. C. Aragão, G. Oliveira, J.

M. Chen, J. M. Deforestation and land use and land cover changes in protected areas of the Brazilian Cerrado: Impacts on the fire-driven emissions of fine particulate aerosols pollutants. *Remote Sensing Letters*, v. 12, p. 79–92, 2021.

[9] Brasil. Lei Federal nº 12.651, de 25 de maio de 2012. Dispõe sobre a proteção da vegetação nativa.

[10] São Paulo. Lei Estadual nº 15.684, de 14 de janeiro de 2015. Dispõe sobre o Programa de Regularização Ambiental (PRA) de imóveis rurais no Estado de São Paulo.

[11] Embrapa. Empresa Brasileira de Pesquisa Agropecuária. Centro de Ciência para o Desenvolvimento da Agricultura Digital, 2023.

[12] C. A. Alvares, J. L. Stape, P. C. Sentelhas, J. D. Gonçalves, G. Sparovek. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, n. 6, p. 711-728, 2013.

[13] MMA. Ministério do Meio Ambiente. Regularização Ambiental - Cadastro Ambiental Rural. Disponível em: <<https://www.car.gov.br/#/>>. Acesso em: 05 out. 2024.

[14] D. G. S. W. LOEBMANN, R.P. MAÇORANO, G. BAYMA, D.C. VICTORIA. Interpretação de alvos a partir de imagens de satélite de média resolução espacial. Campinas, SP: Embrapa Monitoramento por Satélite, 2012. 24 p. (Embrapa Monitoramento por Satélite. Circular Técnica, 21).

[15] J. W. Rouse, R. H. Haas, J. A. Schell, D.W. Deering. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.*, v. 351, n. 1, 309 p., 1973.

[16] B. C. Gao. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, v. 58, p.257-266, 1996.

[17] R. G. Congalton, K. Green. Assessing the accuracy of remotely sensed data: Principles and practices. Boca Raton: CRC Press, 160 p., 1999.

[18] R. G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, v. 37, p. 35-46, 1991.

[19] X. Peng, G. He, W. She, X. Zhang, G. Wang, R. Yin, T. Long. A comparison of random forest algorithm-based forest extraction with GF-1 WFV, Landsat 8 and Sentinel-2 images. *Remote Sensing*, v. 14, n. 21, 5296, 2022.

[20] A. Waśniewski, A. Hościło, B. Zagajewski, D. Moukétou-Tarazewicz. Assessment of Sentinel-2 satellite images and random forest classifier for rainforest mapping in Gabon. *Forests*, v. 11, n. 9, 941, 2020.

[21] A. Hościło, A. Lewandowska. Mapping forest type and tree species on a regional scale using multi-temporal Sentinel-2 data. *Remote Sensing*, v. 11, n. 8, 929, 2019.