ANÁLISE TEMPORAL DO COMPORTAMENTO ESPECTRAL DA VEGETAÇÃO PRIMÁRIA E SECUNDÁRIA NO CERRADO

Juliana Santos da Mota[1], *Marcos Adami*[2], *Mayrine Silva*[3], *Lívia Caroline César Dias*[3], *Neil Damas de Oliveira Junior*[3], *Mayara Teodoro*[3]

[1],[2],[3] Instituto Nacional de Pesquisas Espaciais, Avenida dos Astronautas, 1758 – Jardim da Granja, São José dos Campos,SP – CEP: 12227-010; juliana.mota@inpe.br[1], marcos.adami@inpe.br[2], mayrine.silva@inpe.br[3], diascclivia@gmail.com[3], neil.oliveira.jr@gmail.com[3], mayara.teodoro@inpe.br[3]

RESUMO

Este estudo analisou o comportamento espectral da vegetação primária e secundária em uma área de transição entre os biomas Amazônia e Cerrado, utilizando séries temporais de imagens Sentinel-2. Foram coletadas amostras de vegetação secundária, natural, pastagem e corpos d'água, processadas com o algoritmo *Self Organizing Maps* (SOM) e *Random Forest*. Os resultados mostraram alta precisão na classificação de pastagem e corpos d'água, mas a distinção entre vegetação primária e secundária foi dificultada pela similaridade espectral, especialmente em áreas de transição. A vegetação secundária, que pode estar em estágios avançados de sucessão ecológica, apresentou comportamento espectral semelhante ao da vegetação primária. Apesar desses desafios, o estudo demonstrou a viabilidade do uso de sensoriamento remoto para monitorar a regeneração da vegetação no Cerrado, com potencial para aprimoramento com mais amostras e técnicas avançadas.

Palavras-chave **–** vegetação secundária, vegetação primária, análises espectrais, sensoriamento remoto, cerrado.

ABSTRACT

This study analyzed the spectral behavior of primary and secondary vegetation in a transitional area between the Amazon and Cerrado biomes, using time series from Sentinel-2 images. Samples of secondary vegetation, natural vegetation, pasture, and water bodies were collected and processed using the Self Organizing Maps (SOM) and Random Forest algorithms. The results showed high accuracy in classifying pasture and water bodies, but distinguishing between primary and secondary vegetation was challenging due to spectral similarity, especially in transition zones. Secondary vegetation in advanced stages of ecological succession exhibited spectral behavior similar to primary vegetation. Despite these challenges, the study demonstrated the viability of remote sensing to monitor vegetation regeneration in the Cerrado, with potential for improvement through increased sample size and advanced techniques.

Key words **–** *Secondary vegetation, Primary vegetation, Spectral analysis, Cerrado, Remote sensing.*

1. INTRODUÇÃO

O Cerrado brasileiro ocupa cerca de 2 milhões de km² na região central do Brasil e conecta quatro dos seis biomas brasileiros. Representa 24% do território brasileiro, o segundo maior bioma em área no território nacional [1]. Considerado um ecossistema de Savana, apresenta grande heterogeneidade nas formações vegetacionais, resultado da variação das condições edáfico-climáticas, bem como da ocorrência e frequência de incêndios [2].

Essa variação influencia a estrutura da vegetação, a distribuição espacial de indivíduos lenhosos na paisagem e na composição florística do Cerrado. Assim, a vegetação primária do bioma é caracterizada por três formações distintas: Florestais (Mata Ciliar, Mata de Galeria, Mata Seca e Cerradão), savânicas (Cerrado sentido restrito, Parque de Cerrado, Palmeiral e Vereda) e campestres (Campo Sujo, Campo Limpo e Campo Rupestre) [3].

Todavia, esse ecossistema brasileiro sofreu nos últimos anos com reiteradas pertubações antrópicas provenientes dos avanços tecnológicos que permitiram a expansão agrícola, em detrimento das formações naturais do Bioma, gerando assim intensas transformações, interferindo na dinâmica da regeneração e no estabelecimento de uma vegetação secundária [4].

A vegetação secundária surge após a remoção da vegetação primária devido a intervenções antrópicas como desmatamento e incêndios [5]. Sem manejo, essas áreas desmatadas podem iniciar um processo de sucessão ecológica, onde espécies pioneiras dão lugar a uma comunidade vegetal mais complexa e diversificada, sendo importantes para a mitigação da perda de biodiversidade, melhorando serviços ecossistêmicos e formando novas dinâmicas ecológicas.

O uso de séries temporais de dados de sensoriamento remoto tem se mostrado eficaz no monitoramento e análise da dinâmica da vegetação, permitindo a observação contínua de grandes áreas e fornecendo informações valiosas sobre mudanças na cobertura vegetal e sua fenologia [6]. Neste sentido, este estudo propõe analisar e comparar o comportamento espectral de amostras de vegetação primária e secundária do Cerrado, utilizando séries temporais de imagens obtidas pelo Sentinel-2/MSI. A hipótese central é que, devido aos processos de sucessão ecológica, o comportamento espectral da vegetação secundária e da vegetação primária são semelhantes. Para testar essa hipótese, foram aplicadas técnicas avançadas de processamento de imagens e aprendizado de máquina para classificar e avaliar a similaridade espectral entre essas classes.

2. MATERIAL E MÉTODOS

2.1. Área de estudo

A área de estudo está localizada no Estado do Maranhão, no bioma Cerrado (Figura 1). A escolha dessa área se justifica pela relevância ecológica das ecorregiões presentes, como a Floresta de Cocais e Alto Parnaíba, ambas representativas da vegetação nativa do Cerrado.

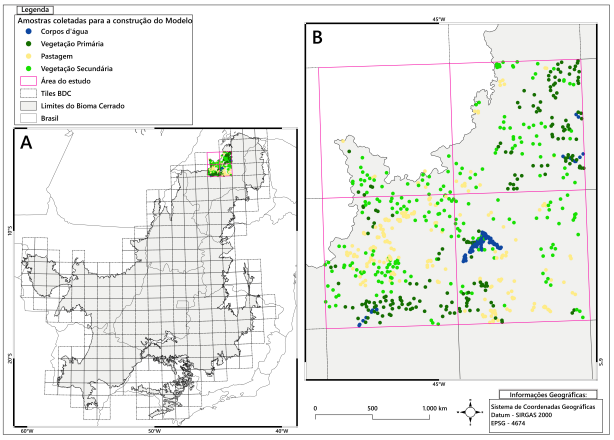

Figura 1: Distribuição das amostras coletadas manualmente na grade BDC_SM_V2 (A) e a área de estudo (tiles BDC 031010, 031011, 031010, 032011), localizada em rosa (B).

Essa região enfrenta desafios como o desmatamento e a degradação ambiental, temas centrais abordados em nossa pesquisa. O clima predominante é tropical sazonal com estação seca bem definida, a precipitação média anual para Alto Parnaíba e Floresta cocais fica em torno de 1.160(mm) e 1.556(mm), respectivamente. Os solos na Floresta de cocais são majoritariamente Plintossolos, enquanto em Alto Parnaíba são da classe Latossolos [7]. As fitofisionomias presentes na área de estudo são: Floresta Estacional Decidual e Semidecidual, Floresta Ombrófila Aberta e Densa e regiões de contato, onde duas ou mais regiões fitoecológicas se tocam, justapondo-se ou interpenetrando-se [8].

2.2. Material

Ao todo foram coletadas 730 amostras por meio de interpretação visual, sendo 210 amostras de vegetação primária e 100 amostras representativas de corpos d'água. Essas classes foram selecionadas para a melhoria do potencial discriminatório do modelo para as classes de interesse. A seleção ocorreu a partir dos mapas do programa de monitoramento anual da supressão da vegetação nativa no Cerrado (Prodes Cerrado) obtidos na plataforma Terrabrasilis [9] que divulga conjuntos de dados e taxas anuais de desmatamento para os biomas brasileiros, auxiliando o governo na tomada de decisões em políticas [10].

Também foram selecionadas 210 amostras de vegetação secundária e 210 de pastagem, oriundas do mapeamento feito pelo projeto TerraClass Cerrado. Os dados desse projeto estão disponíveis para os anos de 2018, 2020 e 2022 por meio do sistema de geoinformações web GeoPortal TerraClass [11]. Este projeto identifica as classes de uso e cobertura da terra nas áreas desmatadas detectadas pelo Prodes [5].

Na fase de preparação de dados, a abordagem usada se concentrou no processamento das imagens de satélite do Sentinel-2/MSI para o período de 01-01-2018 a 30-09-2022, com resolução espacial de 10 metros. Essas imagens foram adquiridas e processadas no ambiente BDC-Lab por meio do projeto *Brazil Data Cube* (BDC), sendo transformadas em cubos de dados com composição temporal de 16 dias [12]. Os cubos de dados do projeto BDC são organizados em uma grade própria de tiles. Cada estrutura dessa possui o tamanho de 10560×10560 que corresponde um total de $111.513,600$ pixels, a amostragem foi feita separadamente para cada tile, assim, foram selecionados 3 tiles na Ecorregião de Floresta de Cocais e 1 tile na Ecorregião de Alto Parnaíba, ambas no Cerrado. Os cubos foram constituídos com composições de 16 dias obtidas para o período janeiro de 2018 à setembro de 2022, totalizando 112 imagens do Sentinel-2, contendo as bandas espectrais B03, B04, B08, B8A, os índices de vegetação *Enhanced Vegetation Index* (EVI), *Normalized Difference Vegetation Index*(NDVI), e o produto de mascaramento de nuvens CLOUD.

2.3. Métodos

A avaliação da qualidade das amostras foi feita pelo método *Self-Organizing Map* (SOM) que utiliza redes neurais não supervisionadas para reconhecimento de padrões em séries temporais de dados espaciais extraídos de cubos de dados e inferência bayesiana para filtragem das amostras [13]. No pacote *Satellite Image Time Series* (SITS-R) as séries temporais de cada amostra foram extraídas de cubos de dados Sentinel-2 disponibilizados pelo BDC.

Para a classificação foi aplicado o algoritmo *Random Forest* (RF) usando como parâmetro um modelo de *Machine Learning* com 300 árvores de decisão através do (*sits_rfor*) função disponibilizada pelo SITS, um filtro Bayesiano pós-classificação pela ferramenta (*sits_smooth*) com uma janela 5x5 afim de refinar os mapeamentos. A escolha do algoritmo baseou-se no desempenho do classificador e no tempo de processamento para mapeamentos de áreas muito extensas. Ademais, foi utilizada a função (*sits_kfold_validate*) para gerar a validação cruzada e a matriz de confusão baseada nas classes definidas na pesquisa e fornece a matriz de confusão e as estatísticas de precisão geral e por classe, permitindo estimar o erro de predição e avaliar o desempenho do nosso modelo.

3. RESULTADOS E DISCUSSÃO

As 730 amostras foram processadas pelo SOM. Destas, 4 de vegetação secundária, 31 de vegetação primária, 7 de pastagem e 13 de água foram consideradas ruídos *noisy samples*. Após a análise elas foram removidas por serem coletadas em áreas transicionais podem apresentar muita variação. Portanto, o modelo final foi gerado com base nas amostras restantes (n=675) resultando no mapa SOM (Figura 2) com acurácia global igual a 0,90% e intervalo de confiança (IC) a 95% de 1,17%.

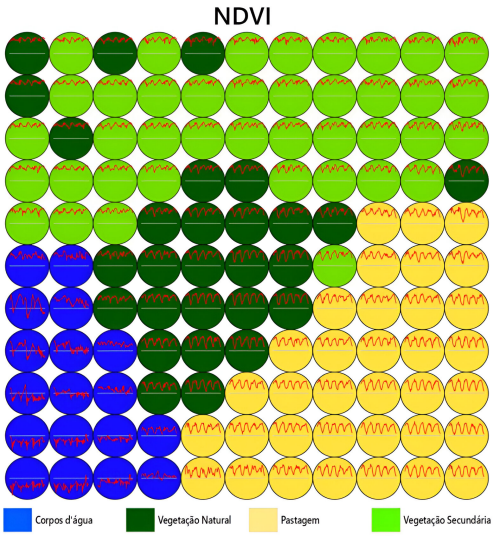

Figura 2: Distribuição de neurônios gerada pelo método SOM (Self-Organizing Maps) utilizado para a construção do modelo.

O contraste entre os perfis de séries temporais das classes de cobertura do solo avaliadas neste estudo para os índices espectrais propostos é apresentado na (Figura 3). O classificador teve ótima desempenho em diferenciar áreas com pastagem e água das demais classes. Apesar dos padrões das curvas espectrais serem bastante similares, outros fatores influenciam as medidas de reflectância da vegetação, dificultando uma correspondência direta de análise entre as curvas espectrais e a imagem formada nos sensores imageadores [14].

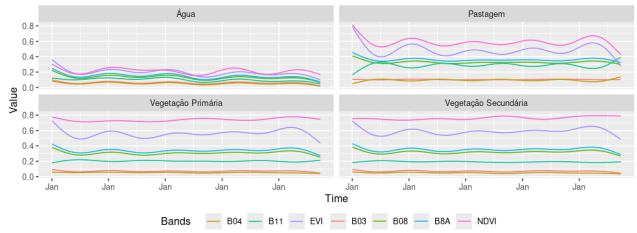

Figura 3: Análise espectral do comportamento das classes na região de estudo

Os percentuais de acerto da classificação em relação à referência (i.e., acurácia do usuário) para vegetação secundária, vegetação primária, pastagem e água foram de 81.30%, 86.36%, 98.54% e 100%, respectivamente. Enquanto os erros de omissão apresentaram valores de 9.22% 25%, 0.49%, e 0%. A confusão entre as classes está associada geralmente ao gradiente de similaridade entre elas, sendo assim, o erro de comissão das classes de vegetação secundária e vegetação natural foi de 18.70% e 13.64% respectivamente (Figura 4).

Porém, a acurácia do mapa foi superior àquelas obtidas nos estudos de [15] que obteve 74% ao usar aprendizado de máquina e os algoritmos classificação supervisionada RF e Support Vector Machine (SVM) na discriminação de classes de cobertura vegetal em uma região de transição Amazônia/Cerrado no Estado de Mato Grosso por meio de Imagens do Satélite ALOS-2/PALSAR-2. A regeneração no Cerrado é afetada por variáveis de manejo e biofísicas, em vez da cobertura vegetal nativa em escala de paisagem [16]. A precipitação foi o fator mais representativo para determinar as variações no crescimento da vegetação do bioma [17].

		Classes						
		Vegetação Secundária	Pastagem	Vegetação Natural	Corpos d'água	Total	Acurácia do Usuário(%)	Erro de Omissão (%)
Predição	Vegetação Secundária	187	0	43	0	230	81.30	9.22
	Pastagem	0	202	3	0	205	98.54	0.49
	Vegetação Natural	19	1	133	1	154	86.36	25.70
	Corpos d'água	0	0	0	86	86	100.00	0
	Total Geral	206	203	179	87			
	Acuracia do Produtor (%)	90.78	99.51	74.30	98.85			
	Erro de Comissão (FP)	18.70	1.46	13.64	0			
	Acurácia global (%)	90.07						

Figura 4: Matriz de confusão e métricas de acurácia.

Esses valores sugerem ótima precisão para identificação das classes de Pastagem e Água, porém, existe bastante confusão entre as classes de vegetação natural e vegetação secundária. De tal modo que o modelo tem dificuldades significativas para diferenciar estas duas classes. Estas confusões ocorreram, geralmente, em áreas de transição de vegetação arbórea (Figura 5).

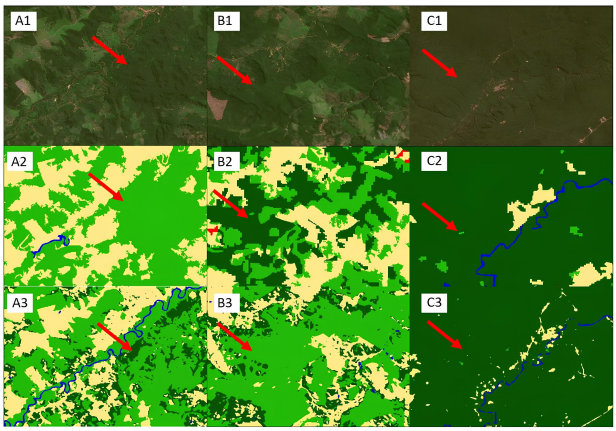

Figura 5: Imagem Sentinel-2 (A1, B1 e C1), mapeamento TerraClass (A2, B2 e C2), e a classificação de área de estudo (A3, B3 e C3). Setas na cor vermelha indica a área onde houve inclusão de vegetação natural em áreas de vegetação secundária ou vice e versa.

A confusão entre os estratos de vegetação primária e secundária pode estar associada a características espectrais e espaciais das fitofisionomias, especialmente em regiões de contato, onde coexistem diferentes tipos de vegetação, formando um mosaico de savanas e florestas. Além disso, a vegetação secundária em estágio avançado de sucessão ecológica pode apresentar características espectrais semelhantes às da vegetação primária, dificultando a distinção entre essas classes [18]. Entretanto, o TerraClass não enfrenta esse problema, pois trabalha mapeando apenas a vegetação secundária e mantendo a vegetação primária originária do Prodes [5] e o modelo sugere que ela pode ser facilmente distinguida das demais classes. Por outro lado, o Prodes

mapeia o desflorestamento que ocorreu em áreas de vegetação original [10], não tendo dificuldade em diferenciar das demais classes à exceção da vegetação secundária, que não é objeto do mapeamento do projeto.

4. CONCLUSÕES

O uso de séries temporais de imagens Sentinel-2, aliado a algoritmos de aprendizado de máquina, demonstrou ser uma ferramenta promissora para o monitoramento da vegetação do Cerrado, especialmente na discriminação de pastagens e corpos d'água. Contudo, a similaridade espectral entre a vegetação primária e secundária, possivelmente em estágios avançados de sucessão ecológica, representou um desafio para a classificação precisa dessas classes. Os resultados obtidos, apesar das limitações impostas pelo tamanho amostral e pela complexidade da área de estudo, reforçam o potencial do sensoriamento remoto para monitorar a dinâmica da vegetação em regiões ecotonais, ao mesmo tempo em que indicam a necessidade de ajustes metodológicos para lidar com a heterogeneidade da paisagem e a complexidade das relações entre as diferentes classes de cobertura do solo.

5. AGRADECIMENTOS

Os autores agradecem o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por meio do projeto número 422354/2023-6 (Monitoramento e Alertas de Mudanças de Cobertura do Solo nos Biomas Brasileiros – Capacitação e Semi-Automação do Programa BiomasBR), bolsa [384897/2023-1], ao BDC-Lab do Brazil Data Cube e ao Instituto Nacional de Pesquisas Espaciais (INPE).

6. REFERÊNCIAS

[1] E. E. Sano and et al. Cerrado ecoregions: a spatial framework to assess and prioritize brazilian savanna environmental diversity for conservation. *Journal of Environmental Management*, 232:818–828, 2019.

[2] R. Marchant. Understanding complexity in savannas: climate, biodiversity and people. *Current Opinion in Environmental Sustainability*, 2(1-2):101–108, 2010.

[3] J. F. Ribeiro and B. M. T. Walter. Fitofisionomias do bioma cerrado. In S. M. Sano and S. P. Almeida, editors, *Cerrado: ambiente e flora*, pages 87–166. Embrapa Cerrados, Brasília, 1998.

[4] C. A. Klink and et al. The role of vegetation on the dynamics of water and fire in the cerrado ecosystems: implications for management and conservation. *Plants*, 9(12):1803, 2020.

[5] C. A. Almeida and et al. High spatial resolution land use and land cover mapping of the brazilian legal amazon in 2008 using landsat-5/tm and modis data. *Acta Amazonica*, 46:291–302, 2016.

[6] A. Garioud and et al. Recurrent-based regression of sentinel time series for continuous vegetation monitoring. *Remote Sensing of Environment*, 263:112419, 2021.

[7] E. Sano and et al. *Características gerais da paisagem do bioma Cerrado*, pages 21–37. Embrapa, 2020.

[8] Instituto Brasileiro de Geografia e Estatística IBGE. *Banco de Dados e Informações Ambientais (BDiA): Mapeamento de Recursos Naturais (MRN): escala 1:250 000 : versão 2023 : nota metodológica / IBGE, Coordenação de Meio Ambiente*, page 64. IBGE, 2023.

[9] L. F. F. G. Assis and et al. Terrabrasilis: A spatial data analytics infrastructure for large-scale thematic mapping. *ISPRS International Journal of Geo-Information*, 8:519, 2019.

[10] C. A. Almeida and et al. *Methodology for forest monitoring used in PRODES and DETER projects*. http://mtc-m21c.sid.inpe.br/col/sid.inpe.br/mtc-m21c/2021/01.25.19.14/doc/publicacao.pdf, 2021.

[11] Embrapa (Empresa Brasileira de Pesquisa Agropecuária) and INPE (Instituto Nacional de Pesquisas Espaciais). Geoportal terraclass. In *Projeto TerraClass*. Embrapa Cerrados, Brasília, 2022.

[12] K. R. Ferreira and et al. Earth observation data cubes for brazil: Requirements, methodology and products. *Remote Sensing*, 12:4033, 2020.

[13] L. A. Santos and et al. Identifying spatiotemporal patterns in land use and cover samples from satellite image time series. *Remote Sensing*, 13:974, 2021.

[14] T. P. Lima and et al. Correlação entre as transformações da cobertura e uso da terra com variáveis climáticas e ambientais na região do matopiba, brasil. *Geography Department University Of Sao Paulo*, 43:1–13, 2023.

[15] V. S. Silva, E. E. Sano, T. Almeida, and H. N. Mesquita Júnior. Discriminação de classes de cobertura vegetal em uma região de transição amazônia/cerrado no estado de mato grosso por meio de imagens do satélite alos-2/palsar-2. *Revista Brasileira De Cartografia*, 73:1–16, 2021.

[16] T. R. Silva and et al. Factors affecting savanna and forest regeneration in pastures across the cerrado. *Journal of Environmental Management*, 330:117185, 2023.

[17] H. A. Barbosa, T. V. Kumar, and L. R. M. Silva. Recent trends in vegetation dynamics in the south america and their relationship to rainfall. *Natural Hazards*, 77:883–899, 2015.

[18] J. H. Everitt and et al. Reflectance characteristics and remote sensing of a riparian zone in south texas. *The Southwestern Naturalist*, 47:433, 2002.