

DETECÇÃO DE FEIÇÕES EROSIVAS EM BACIAS HIDROGRÁFICAS POR TÉCNICAS DE SENSORIAMENTO REMOTO E CLASSIFICAÇÃO SUPERVISIONADA COM BASE EM ALGORITMO DE APRENDIZADO DE MÁQUINA

Mateus Benchimol Ferreira de Almeida ¹, Margareth Simões ^{1,2}, Rodrigo Peçanha Demonte Ferraz ², Patrick Calvano Kuchler ³, Marcela Antunes Meira ⁴, Diego Escobar Alves ⁵

¹Programa de Pós-graduação em Meio Ambiente/Doutorado Multidisciplinar (PPGMA)/FEN/DESC/UERJ, Rua São Francisco Xavier, 524, Pavilhão João Lyra Filho, 12 Andar, Bloco F, Sala 12005, Bairro Maracanã, RJ, e-mail: mateusbenchimol@hotmail.com; ²Embrapa Solos, Rua Jardim Botânico, 1024, Jardim Botânico, RJ, e-mail: margareth.simoess@embrapa.br, e-mail: rodrigo.demonte@embrapa.br; ³Departamento de Geografia Física (DGF) e do Programa de Pós-graduação em Geografia (PPGEO) da Universidade do Estado do Rio de Janeiro (UERJ), e-mail: geocalvano@gmail.com; ⁴Department of Civil Engineering, Swansea University Bay Campus, Fabian Way, Swansea SA1 8EN, UK, e-mail: marcelameira.ufcg@gmail.com; ⁵Departamento de Geografia, Faculdade de Educação da Baixada Fluminense, Universidade do Estado do Rio de Janeiro (UERJ), e-mail: escobartdiego@gmail.com

RESUMO

O presente estudo teve como finalidade o monitoramento e mapeamento de feições erosivas, tendo como área piloto a microbacia do Caiador localizada no município de Silvânia, Estado de Goiás (GO). Deste modo, recorreu-se ao uso de séries temporais de imagens de satélite Sentinel-2 (3 anos), amostras coletadas em tela do computador e técnicas de Sensoriamento Remoto, como processamento digital de imagens e classificação supervisionada *Random Forest*. O VIgreen e o EVI foram os índices de vegetação mais importantes na classificação supervisionada *Random Forest*. As séries temporais de imagens Sentinel-2 (3 anos) foram fundamentais para identificar áreas em pousio, sem confundir com feições erosivas, tornando o monitoramento e o mapeamento mais precisos. Com base nos resultados alcançados, a metodologia desenvolvida neste estudo atendeu com satisfação os objetivos almejados. A classificação supervisionada *Random Forest* apresentou as seguintes acurácias: $K = 0.99$, $AG = 0.99$ e $F1\text{-score} = 0.99$.

Palavras-chave — *Sentinel-2, Random Forest, Séries Temporais, Índices de Vegetação, Google Earth Engine*

ABSTRACT

The present study had the goal of monitoring and mapping erosive features, with the pilot area the Caiador watershed located in the Silvânia municipality, Goiás (GO) State. Thus, using Sentinel-2 images time series (3 years), samples collected on computer screen, and remote sensing techniques such as digital image processing and supervised Random Forest classification. VIgreen and EVI were the most important vegetation indices in the Random Forest

supervised classification. Sentinel-2 images (3 years) time series were fundamental to identify areas in fallow, without confusing them with erosive features, making monitoring and mapping more accurate. Based on the results achieved, the methodology developed in this study met the desired objectives with satisfaction. Random Forest supervised classification reached the following accuracies: $K = 0.99$, $OA = 0.99$ and $F1\text{-score} = 0.99$.

Key words — *Sentinel-2, Random Forest, Time Series, Vegetation Indices, Google Earth Engine.*

1. INTRODUÇÃO

Globalmente, a erosão do solo tem sido considerada um dos riscos mais destrutivos de degradação do solo, apontada como uma ameaça dinâmica com atributos físicos e socioeconômicos, portanto, suas avaliações precisam ser atualizadas com frequência [1]. Desta forma, o conhecimento sobre a dinâmica erosiva se torna fundamental para que se estabeleçam metodologias que possam reverter a situação de degradação dos solos e do ambiente [2].

A análise espacial da distribuição da suscetibilidade à erosão e os mapas que apontam o potencial deste processo constitui um produto cartográfico essencial para o planejamento ambiental, servindo assim como orientação para os parâmetros de ocupação de um determinado espaço geográfico, uma vez que destaca as potencialidades e fragilidades de um determinado território, seja ele uma bacia hidrográfica ou um município [3].

Deste modo, este estudo objetiva o desenvolvimento de uma metodologia para o monitoramento e mapeamento de feições erosivas em áreas agrícolas, com o uso de técnicas de Sensoriamento Remoto e algoritmos de inteligência artificial e aprendizado de máquina, com especial destaque ao *Random*

Forest. A metodologia foi validada na microbacia do Caiador, localizada no município de Silvânia, Estado de Goiás (GO), onde há grande ocorrência de erosão superficial e erosão severa (voçorocas).

2. MATERIAL E MÉTODOS

As bandas espectrais Sentinel-2 utilizadas inicialmente foram os seguintes: B2 (azul), B3 (verde), B4 (vermelho), B5 (borda vermelha), B6 (borda vermelha), B7 (borda vermelha), B8 (infravermelho próximo), B8A (infravermelho próximo), B11 (infravermelho médio 1) e B12 (infravermelho médio 2). Já os índices de vegetação utilizados inicialmente são apresentados na Tabela 1.

Para a execução da classificação supervisionada *Random Forest* foram utilizadas 500 árvores de decisão sobre o conjunto de dados ou cubo de dados (*dataset* ou *datacube*), isto é, a série temporal das imagens de satélite Sentinel-2 (3 anos), índices de vegetação (3 anos) e amostras das classes de interesse. O grande número de árvores de decisão aplicadas deve-se a necessidade de se realizar o processamento com a melhor precisão possível, pois quanto maior o número de árvores de decisão maior a acurácia do mapa final.

O treinamento consistiu na separação de 70% das amostras coletadas, que, em conjunto com a série temporal de imagens Sentinel-2 (3 anos) e os índices de vegetação correspondentes (3 anos) foram processados por meio da classificação supervisionada *Random Forest*. Nesta fase, com

base nas amostras, extraiu-se os descritores, ou seja, os valores dos pixels nas áreas das amostras. Os descritores correspondem tanto aos valores dos pixels da série temporal de imagens Sentinel-2 (3 anos) quanto aos valores dos índices de vegetação correspondentes (3 anos). Desta forma, cada amostra é associada a uma resposta espectral (padrão) que a caracteriza. O *Random Forest* procura encontrar padrões espectrais neste conjunto de dados, de modo a acertar a separabilidade das classes de interesse.

Neste estudo adotou-se o método de classificação supervisionada *Random Forest*. Esta escolha teve como base a robustez deste algoritmo, bem como os inúmeros casos de sucesso que se tem observado em diversos estudos científicos ao redor do mundo.

Foram essenciais os seguintes materiais:

- Plataforma *Google Earth Engine* com base em programação de alto desempenho com o uso de linguagem de programação *Javascript*, computação em nuvem e computação paralela;
- 3 computadores notebook;
- 200 amostras coletadas em tela do computador;
- Shapefile da delimitação geográfica da Bacia do Caiador (IBGE).

Índice de Vegetação	Fórmula equivalente (Sentinel-2)	Referências
Cellulose Absorption Index (CAI)	B12/B11	(Guerschman et al., 2009; Hill, 2013)
Carotenoid Reflectance Index 1 (CRI1)	$(1/B2) - (1/B3)$	(Gitelson et al., 2002; Hill, 2013)
Enhanced Vegetation Index (EVI)	$(2.5(B8 - B4))/(B8 + 6B4 - 7.5B2 + 1)$	(Huete, 1999)
Green Vegetation Index (VIgreen)	$(B3 - B4)/(B3 + B4)$	(Daughtry et al., 2005; Peña-Barragán et al., 2011).
Normalized Canopy Index (NCI)	$(B11 - B3)/(B11 + B3)$	(Li and Guo, 2016; Ren and Zhou, 2019; Vescovo and Gianelle, 2008)
Normalized Difference Index (NDI7)	$(B8 - B12)/(B8 + B12)$	(Dinga and Tsubo, 2019; Li and Guo, 2016)
Normalized Difference Tillage Index (NDTI)	$(B11 - B12)/(B11 + B12)$	(Dinga and Tsubo, 2019; Li and Guo, 2016; Van Deventer et al., 1997)
Normalized Difference Vegetation Index (NDVI)	$(B8 - B4)/(B8 + B4)$	(Huete et al., 2002; Rouse et al., 1973)
Soil Adjusted Total Vegetation Index (SATVI)	$2(B11 - B4)/(B11 + B4 + 1) - B12/2$	(Li and Guo, 2016; Marsett et al., 2006)
Soil Tillage Index (STI)	B11/B12	(Dinga and Tsubo, 2019)

Tabela 1. índices de vegetação (Sentinel-2).

3. RESULTADOS E DISCUSSÕES

A fim de se poder mapear com maior rigor as feições erosivas houve a necessidade de se utilizar uma série temporal de imagens de satélite Sentinel-2 de 3 anos (2021-2023), devido ao fato de que algumas regiões se encontravam em pousio; e a fim de se evitar confusão com as regiões que apresentavam indícios de feições erosivas, foi necessário recorrer a uma

série temporal mais longa. Esta técnica de adotar uma série temporal de 3 anos garantiu maior sucesso no monitoramento e mapeamento das feições erosivas, bem como na sua acurácia.

O período de 3 anos foi definido com base em experimentos de tentativa e erro. A combinação destes experimentos com a análise dos perfis temporais do NDVI (Figura 1) e a coleta sucessiva de amostras nas áreas mais

críticas foram cruciais para o sucesso desta metodologia, pois percebe-se que a partir de março de 2023 o valor do NDVI é inferior a 0.5, dificultando a caracterização espectral do alvo. O algoritmo de classificação supervisionada *Random Forest* possui uma função estatística que detecta as contribuições mais significativas durante o processo de classificação. Deste modo, com base nos produtos de sensoriamento remoto de entrada, foi possível determinar as contribuições mais significativas no processo de classificação supervisionada *Random Forest*.

Definitivamente, as bandas espectrais e os índices de vegetação mais significativos e que serviram para o prosseguimento do processo de classificação supervisionada *Random Forest* foram os seguintes: B3, B4, B5, B12, NCI, VIgreen, EVI e NDVI.

Os índices de vegetação mais importantes foram o VIgreen e o EVI, conforme a Figura 2.

O índice VIgreen utiliza as bandas espectrais B3 (verde) e B4 (vermelho), considerando que a vegetação saudável absorve a energia eletromagnética na região do vermelho e reflete a quase totalidade da energia eletromagnética na região do verde, é interessante associar este índice com o predomínio de usos e coberturas vegetais existentes na Bacia do Caiador, contrastando com coberturas não vegetadas com o solo exposto. Ademais, O VIgreen tem a finalidade de favorecer o monitoramento das condições da vegetação, a estrutura da planta e diferenciar cultivos com dossel aberto ou fechado [4][5].

Já o EVI é um índice que utiliza as bandas espectrais B2 (azul), B4 (vermelho) e B8 (infravermelho). Quanto ao uso da banda espectral B4, o EVI é similar ao VIgreen. Quanto ao uso da banda espectral B2 no cálculo do EVI, é importante enfatizar que a vegetação saudável absorve a energia eletromagnética na região do azul, podendo contrastar com outros usos e coberturas não vegetadas. Contudo, o EVI também utiliza a banda espectral B8, que, tem a propriedade de distinguir o conteúdo de água existente na vegetação, o que pode ser bastante útil neste caso em que existe uma grande variedade de usos e coberturas vegetais na Bacia do Caiador. Ademais, o EVI é um índice que favorece o monitoramento da vegetação, tendo maior sensibilidade em regiões de alta biomassa [6]. Considerando a existência de regiões de alta biomassa na Bacia do Caiador, é perceptível que o EVI tenha sido o segundo índice de vegetação mais importante na classificação supervisionada *Random Forest*.

O mapa final possui duas classes, as quais são representadas pela cor vermelha (Áreas com Evidências de Feições Erosivas Acentuadas) e incolor (outros usos e coberturas da área de estudo), conforme Figura 3 e Figura 4.

A classificação supervisionada *Random Forest* apresentou um alto desempenho tanto em termos visuais (qualidade visual do mapa final) quanto em termos de acurácia, isto é, o controle estatístico e/ou validação. As acurácias alcançadas na classificação supervisionada *Random Forest* foram o índice *Kappa* (K), a Acurácia Global (AG) e o *F1-score*, isto é: $K = 0.99$; $AG = 0.99$; $F1\text{-score} =$

$[0.99, 0.99]$.

Na Bacia do Caiador obteve-se as seguintes estatísticas em termos de área de superfície (Figura 5):

- Classe 1 – Áreas com Evidências de Feições Erosivas Acentuadas correspondem a 0.781 km²;
- Classe 2 – Áreas sem Evidências de Feições Erosivas Aparentes correspondem a 40.824 km².

As feições erosivas foram mapeadas satisfatoriamente e com grande exatidão e/ou precisão, conforme já foi abordado anteriormente. Porém, o *Random Forest* agregou na mesma classe as feições erosivas e as estradas de terra, devido à grande semelhança entre as respostas espectrais destes alvos, uma vez que a diferença entre ambos consiste apenas na forma.

Com o alcance de resultados bastante satisfatórios, estabelece-se assim uma metodologia robusta, eficaz e eficiente capaz de permitir o monitoramento e o mapeamento das feições erosivas da Bacia do Caiador, podendo esta metodologia ser extrapolada para outras regiões do Brasil e/ou do mundo.

Estudos mais avançados serão necessários para a separabilidade dos diferentes tipos de feições erosivas, como por exemplo, separar as feições erosivas superficiais das voçorocas, bem como das estradas de terra. Outros algoritmos de inteligência artificial, com outros recursos capazes de distinguir a forma dos alvos (informação contextual), poderão ser implementados para o aprimoramento da metodologia desenvolvida neste estudo.



Figura 1. Perfil temporal do NDVI no período de 2021 a 2023 de um pivô de irrigação.

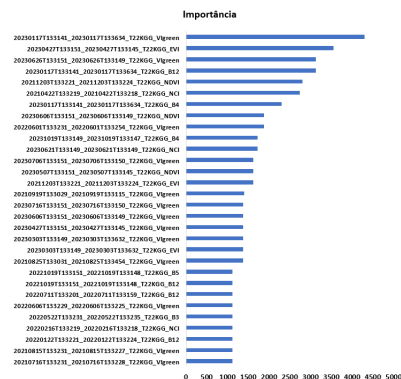


Figura 2. Ordem e proporção relativa de importância de contribuição das bandas espectrais e índices de vegetação na classificação supervisionada *Random Forest*.

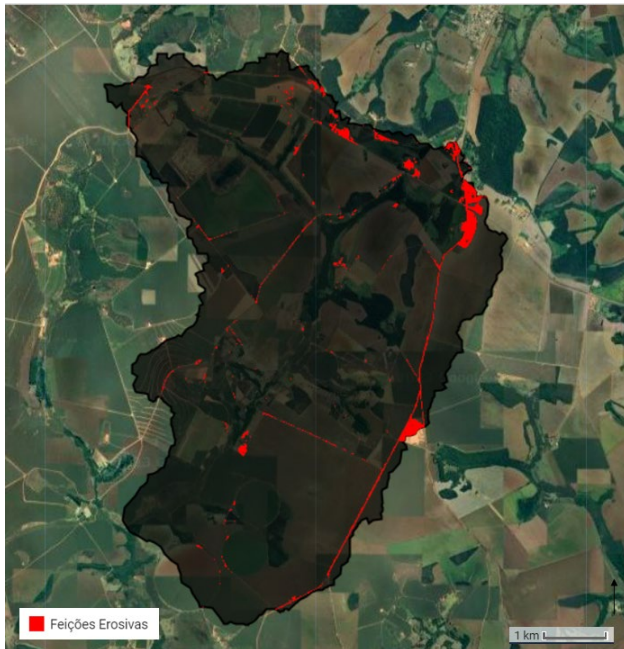


Figura 3. Mapa da classificação supervisionada *Random Forest* da série temporal de índices de vegetação e bandas espectrais da Bacia do Caiador. A cor vermelha representa as feições erosivas.



Figura 4. Extrato em grande escala das feições erosivas (em vermelho) do mapa da classificação supervisionada *Random Forest* da série temporal de índices de vegetação e bandas espectrais da Bacia do Caiador.

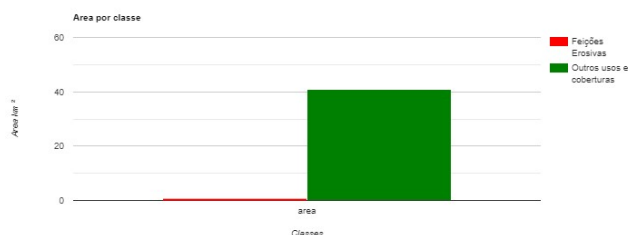


Figura 5. Cálculo das áreas das classes de interesse na Bacia do Caiador.

Finalmente, este estudo atingiu com sucesso o objetivo de desenvolver uma metodologia para o monitoramento e mapeamento de feições erosivas em áreas agrícolas, com o uso de técnicas de Sensoriamento Remoto e algoritmo de inteligência artificial e aprendizado de máquina (*Random Forest*), sendo validada na microbacia do Caiador.

4. CONCLUSÕES

A classificação supervisionada *Random Forest* apresentou um alto desempenho no mapeamento das feições erosivas da Bacia do Caiador localizada no município de Silvânia, Estado de Goiás (GO). Este alto desempenho foi mensurado com base em controle estatístico que indica os valores das acurácias alcançadas, isto é, $K = 0.99$, $F1\text{-score} = 0.99$ e $AG = 0.99$.

O VIgreen e o EVI foram os índices de vegetação mais importantes no processo de classificação supervisionada *Random Forest*.

Apesar do alto desempenho do *Random Forest*, não foi possível separar as erosões superficiais das voçorocas. Para estudos futuros, recomenda-se o uso de algoritmos de inteligência artificial que considerem a informação contextual para solucionar este problema.

A metodologia desenvolvida neste estudo é robusta, eficaz e eficiente para o monitoramento e mapeamento de feições erosivas por meio de imagens satélite Sentinel-2, séries temporais, algoritmo de inteligência artificial e aprendizado de máquina *Random Forest*, bem como técnicas de Sensoriamento Remoto, podendo esta metodologia ser extrapolada para outras regiões do Brasil e/ou do mundo.

8. REFERÊNCIAS

- [1] P. Panagos, C. G. Karydas, I. Z. Gitas and L. Montanarella. Monthly soil erosion monitoring based on remotely sensed biophysical parameters: a case study in strymonas river basin towards a functional pan-european service, *International Journal of Digital Earth*, 5:6, 461-487, 2012.
- [2] A. J. T. Guerra. O início do processo erosivo. In: guerra, a. J. T.; Silva, A. S. Da; Botelho, R. G. M. (Org.). *Erosão e Conservação dos Solos: Conceitos, temas e aplicações*. 7. ed. Rio de Janeiro: Bertrand Brasil. Cap. 1. p. 17-50, 2012.
- [3] A. B. Nunes and E. F. Leite. Geoprocessamento aplicado ao estudo de caracterização física e ambiental da bacia hidrográfica do rio areias, no estado do Tocantins. *Espaço em revista* 19, 104-127, 2017.
- [4] C. S. T. Daughtry, E. R. Hunt, P. C. Doraiswamy and J. E. McMurtrey. Remote sensing the spatial distribution of crop residues. *Agronomy Journal*, vol. 97, nº 3, p. 864-871, 2005.
- [5] J. M. Peña-Barragán, M. K. Ngugi; R. E. Plant and J. Six. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, vol. 115, nº 6, p. 1301-1316, 2011.
- [6] Huete, A. R. MODIS vegetation index algorithm theoretical basis v3. *Environmental Sciences*, nº 13, 1999.