

MAPEAMENTO DA COBERTURA DA TERRA NO PERÍODO SECO NOS SERTÕES CEARENSES UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

MAPPING OF LAND COVERAGE IN THE DRY PERIOD IN THE SERTÃOS OF CEARENS USING MACHINE LEARNING TECHNIQUES

MAPEO DE LA COBERTURA DEL SUELO EN EL PERIODO SECO EN LOS SERTÃOS DE CEARENS UTILIZANDO TÉCNICAS DE APRENDIZAJE MÁQUINA

ADINAN MARZULO MAIA MARTINS¹
DIEGO VICENTE SPERLE DA SILVA²
JOÃO PEDRO DAS NEVES CARDOSO PEDREIRA³
CARLA BERNADETE MADUREIRA CRUZ⁴

¹Universidade Federal do Rio de Janeiro, Centro de Ciências Matemáticas e da Natureza - Av. Athos da Silveira Ramos, 274, Cidade Universitária, Ilha do Fundão, Rio de Janeiro, Brasil, Mestrando em Geografia, UFRJ E-mail: adinanmaia@gmail.com ORCID: <https://orcid.org/0000-0002-9031-7153>

²Doutorando em Geografia, Universidade Federal do Rio de Janeiro, UFRJ, E-mail: diegovsperle@gmail.com ORCID: <https://orcid.org/0000-0003-0036-0463>

³Pós-graduando em Análise Ambiental e Gestão do Território, ENCE, E-mail: neves.pedreira@outlook.com ORCID: <https://orcid.org/0000-0002-5740-1396>

⁴Professora Doutora titular da Universidade Federal do Rio de Janeiro, UFRJ, E-mail: carlamad@gmail.com ORCID: <https://orcid.org/0000-0002-3903-3147>

RESUMO

As plataformas em nuvem associadas a métodos de aprendizado de máquina vêm contribuindo para bons resultados em mapeamentos digitais, seja pelo volume de dados ou pela alta performance. Este trabalho busca o mapeamento da cobertura da terra da mesorregião Sertões Cearenses presente no semiárido brasileiro utilizando o algoritmo Random Forest na plataforma em nuvem Google Earth Engine no período seco (agosto, setembro e outubro) com influência do fenômeno global *El Niño* de 2016. Resultados indicam as métricas kappa e exatidão global 82% e 85% respectivamente, significando alta acurácia para o mapeamento gerado. As classes que mais acertaram no processo de validação foram as classes “Corpos Hídricos” e “Vegetação Umidade Alta”, por outro lado, a classe “Vegetação de Umidade Moderada” foi a que mais errou pela dificuldade do modelo em separá-la da classe “Vegetação de Umidade Baixa”. A metodologia empregada neste trabalho contribuirá para outros mapeamentos na região possibilitando um melhor entendimento das dinâmicas da cobertura terrestre no semiárido brasileiro.

Palavras-chave: semiárido, seca, el niño, gee.

ABSTRACT

Cloud platforms associated with machine learning methods have contributed to good results in digital mappings, either by data volume or high performance. This work seeks to map the land cover of the mesoregion Sertões Cearenses present in the Brazilian semi-arid using the algorithm Random Forest on the cloud platform Google Earth Engine in the dry period (August, September and October) with influence of the global phenomenon El Niño 2016. Results indicate the kappa metrics and global accuracy 82% and 85% respectively, meaning high accuracy for the generated mapping. The classes that were most correct in the validation process were the classes "Water Bodies" and "High Humidity Vegetation", on the other hand, the class "Moderate Humidity Vegetation" was the one that was most wrong by the difficulty of the model to separate it from the class "Low Humidity Vegetation". The methodology used in this work will contribute to other mappings in the region enabling a better understanding of the dynamics of land cover in the Brazilian semi-arid region.

Keywords: semi-arid, drought, el niño, gee.

RESUMEN

Las plataformas en la nube asociadas a métodos de aprendizaje automático vienen contribuyendo a buenos resultados en la cartografía digital, ya sea por el volumen de datos o por el alto rendimiento. Este trabajo busca mapear la cobertura terrestre de la mesorregión Sertões Cearenses presente en la región semiárida brasileña utilizando el algoritmo Random Forest en la plataforma de nube Google Earth Engine en la estación seca (agosto, septiembre y octubre) influenciada por El Niño global. fenómeno de 2016. Los resultados indican kappa y métricas de precisión global del 82% y 85% respectivamente, lo que

significa una alta precisión para el mapeo generado. Las clases que obtuvieron más aciertos en el proceso de validación fueron las clases “Cuerpos de Agua” y “Vegetación de Alta Humedad”, por otro lado, la clase “Vegetación de Humedad Moderada” fue la que cometió más errores debido a la dificultad de el modelo al separarlo de la clase “Vegetación de baja humedad”. La metodología utilizada en este trabajo contribuirá a otros mapeos en la región, permitiendo una mejor comprensión de la dinámica de la cobertura del suelo en la región semiárida brasileña.

Palabras clave: semiárido, sequía, el niño, caramba.

INTRODUÇÃO

Desde a ocupação humana no semiárido brasileiro sempre houve registros de longos períodos de estiagem na região. Tal característica regional por muito tempo foi utilizada como fator de desvantagem regional (Nobre, 2011) e a não compreensão do homem frente a essa realidade climática propiciou em práticas históricas extensivas em agricultura e pecuária que vem contribuindo gradativamente para a perda da produtividade do solo (Giongo, 2011).

A ação do homem durante séculos sem a preocupação da utilização de práticas conservacionistas no uso da terra acaba por contribuir para a aceleração das mudanças climáticas. O último levantamento do IPCC (2021) indica que haverá aumento da ocorrência de eventos extremos no semiárido e, diante disso, a vegetação da *caatinga* poderá sofrer com o aumento da temperatura regional, onde estudos relacionados a essa temática indicam que a *caatinga* está entre os biomas mais ameaçados (Oyama; Nobre, 2003) e as práticas antrópicas extensivas poderão contribuir para a degradação nessas áreas. Por isso, diante dessa realidade, faz-se tão necessário o monitoramento do semiárido que contribuirá de maneira sistemática para o fomento de políticas públicas na região, seja no âmbito social ou ambiental.

O surgimento da plataforma em nuvem *Google Earth Engine* (GEE) tem agilizado de forma significativa a observação terrestre (Kumar e Mutanga, 2018) onde foi projetada com o intuito armazenar e processar um grande volume de dados com o objetivo de tomada de decisão (Kumar e Mutanga, 2018). A aplicação de metodologias com essa abordagem possibilita abordar temáticas em sensoriamento remoto com robustez, contribuindo para entendermos a dinâmica da superfície terrestre (Tamiminia, 2020).

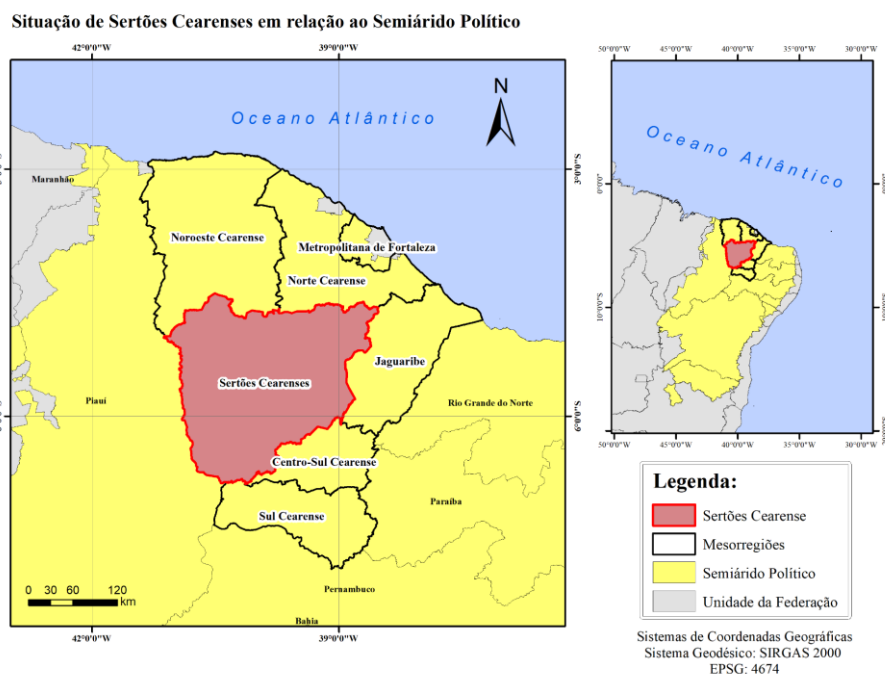
Muitos trabalhos utilizam-se dessa plataforma para mapeamentos de uso ou cobertura da terra na região semiárida (p.ex Sousa et al., 2023; Souza et al., 2020; Cardoso et al., 2021) mas poucos exploram a potencialidade de mapeamentos no período de baixo déficit hídrico. O monitoramento da vegetação seca em épocas de estiagem pode ser a solução para entender os fenômenos de degradação por perda florestal uma vez que a grande variabilidade da vegetação, devido às mudanças fenológicas, pode trazer incertezas diante do uso da terra na região. Contudo, no presente trabalho objetiva-se o mapeamento da cobertura da terra do ano de 2016 no período seco na mesorregião dos sertões cearenses no Estado do Ceará utilizando técnicas de aprendizado de máquina com dados de sensores orbitais na plataforma GEE. Esse trabalho em pesquisa servirá como ponto de partida para outros mapeamentos multitemporais para a compreensão mais ampla do histórico de degradação na região.

MATERIAL E MÉTODO

A Região Sertões Cearenses é localizada na porção central e oeste do Estado do Ceará, sendo composta por 30 municípios (IBGE, 2022), sendo eles: Acopiara, Aiuaba, Ararendá, Arneiroz, Banabuiú, Boa Viagem, Catarina, Choró, Crateús, Deputado Irapuan Pinheiro, Ibaretama, Independência, Iraporanga, Madalena, Milhã, Mombaça, Monsenhor Tabosa, Nova Russas, Novo Oriente, Parambu, Pedra Branca, Piquet Carneiro, Quiterianópolis, Quixadá, Quixeramobim, Saboeiro, Senador Pompeu, Solonópole, Tamboril e Tauá. A região se encontra inserida no semiárido político (Figura 1) que é caracterizado por longos períodos de estiagem e chuvas irregulares, o que permitiu que a vegetação local se adaptasse às condições de estresse

hídrico. A caracterização florística é predominantemente composta por vastas áreas de *caatinga* arbustiva aberta e arbórea, sendo também observada a presença de cerradão e carrasco (FUNCEME, 1994).

Figura 1. Mapa de localização da mesorregião sertões cearenses



Fonte: Elaborado por autores com dados do IBGE (2022).

A intensidade das secas na região semiárida brasileira é muitas vezes determinada pelos eventos de ENOS (El Niño - Oscilação do Sul) (Quadro 1). Embora seja um evento de escala global pelo aquecimento diferenciado no pacífico equatorial (Melo, 1999), provoca alterações nos regimes de precipitações atmosféricas em grande parte do mundo. Especificamente no Brasil, os efeitos mais sentidos são as reduções de chuvas nas regiões Norte e Nordeste e o aumento na região Sul.

Quadro 1: Relação entre as secas no semiárido e o ENOS.

ENOS (Século XXI) - Até 2016	Classificação ENOS	Secas (Século XXI) - Até 2016
2002 - 2003	Moderada	2001 - 2002
2006 - 2007	Moderada	2005 e 2007 - 2008
2009 - 2010	Moderada	2010
2015 - 2016	Forte	2012 - 2016

Fonte: Adaptado de INPE, Melo (2016) e Marengo et al. (2017).

Para a elaboração do mapeamento da cobertura da terra, primeiramente foi necessária a identificação dos limites territoriais da mesorregião Sertões Cearenses em formato shapefile (formato vetorial) através do Instituto Brasileiro de Geografia e Estatística (IBGE). Este recorte espacial define a nossa área de estudo no ambiente GEE. Ao executar o *upload* das imagens na plataforma em nuvem foi criada uma função para recortar automaticamente todas as cenas para os nossos limites de interesse para os sensores orbitais utilizados.

Foram incorporados à plataforma GEE dados do Landsat-8, do sensor *Operational Land Imager* (OLI), que serviram de base para o modelo de aprendizado de máquina *Random Forest* (RF). A fim de adquirir dados ópticos com a menor interferência atmosférica possível foram elaborados diversos filtros. O primeiro filtro é referente a escolha de cenas para a faixa temporal, com isso, foram determinados os mosaicos de dados para os meses praticamente sem chuva (Figura 2) nos sertões cearenses sendo escolhidos os meses agosto, setembro e outubro no ano de 2016. Segundo a FUNCEME, neste ano em específico, foi um dos menos chuvosos da história desde 1910 impactando principalmente a disponibilidade pela água pelo baixo volume dos reservatórios, açudes e entre outros. Portanto, entender a complexidade através de mapeamentos e monitoramentos em épocas de longos períodos de estiagem com auxílio das geotecnologias na região podem elucidar estratégias de tomada de decisão na região estudada.

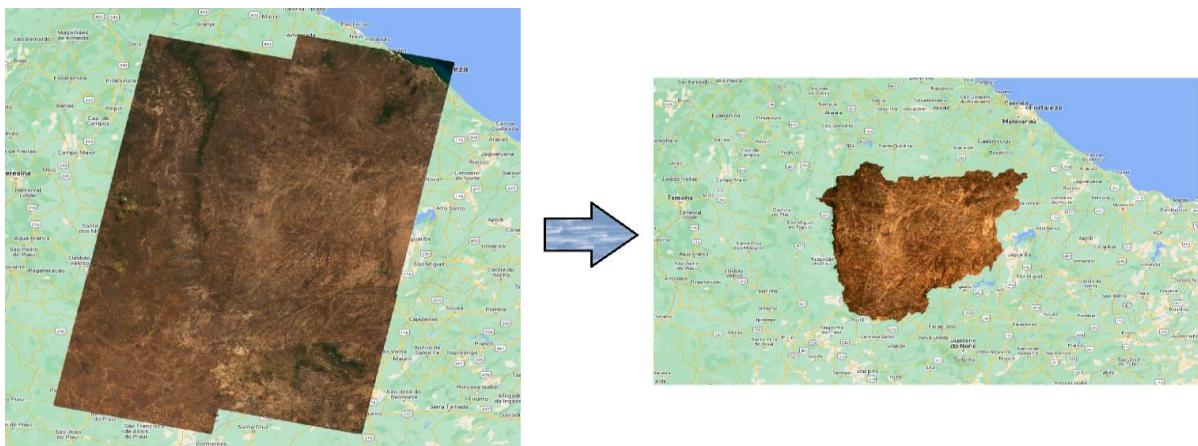
Figura 2 - Médias mensais de chuvas nos sertões cearenses. 1 - agosto; 2 - setembro; 3 - outubro.

	Normal (mm)	Observado (mm)	Desvio (%)	Categoria
1)	4.9	0	-100	
2)	1.9	0.2	-91.7	
3)	3.1	0.1	-97.1	

Fonte: Adaptado por autores com dados de chuvas da FUNCEME.

O segundo filtro tem por objetivo a filtragem de cenas por cobertura de nuvens. Foi necessário recorrer aos metadados das imagens e aplicar funções de filtragens para captar cenas com menos de 1% de cobertura de nuvens (Figura 3).

Figura 3 - Composição colorida RGB para as cenas utilizadas e o recorte espacial para os limites territoriais da Mesoregião Sertões Cearenses.



Fonte: GEE.

Após a seleção final das imagens a serem utilizadas, aplicou-se o fator de escala de acordo com os parâmetros disponibilizados pelo fornecedor do produto, com o objetivo da transformação de níveis de cinza em valores de reflectância de superfície. Após o processo de

tratamento dos dados ópticos é possível observar ao todo 8 cenas utilizadas para o período seco (Quadro 2).

Quadro 02: Cenas filtradas pela cobertura de nuvens e pela área de estudo em ambiente GEE.

Identificador da cena Landsat - 8	Data
LANDSAT/LC08/C02/T1_L2/LC08_218063_20160801	01/08/2016
LANDSAT/LC08/C02/T1_L2/LC08_218064_20160801	01/08/2016
LANDSAT/LC08/C02/T1_L2/LC08_217065_20160810	10/08/2016
LANDSAT/LC08/C02/T1_L2/LC08_217063_20160826	26/08/2016
LANDSAT/LC08/C02/T1_L2/LC08_217064_20160826	26/08/2016
LANDSAT/LC08/C02/T1_L2/LC08_218065_20160902	02/09/2016
LANDSAT/LC08/C02/T1_L2/LC08_217065_20161029	29/10/2016
LANDSAT/LC08/C02/T1_L2/LC08_217064_20161029	29/10/2016

Fonte: GEE.

A partir do mosaico de imagens criado foram utilizados para o modelo de aprendizado de máquina RF as bandas do azul (BLUE), verde (GREEN), vermelho (RED), infravermelho próximo (NIR), infravermelho de ondas curtas 1 e 2 (SWIR 1 e 2). Além disso, foram executadas operações matemáticas para a geração de índices espectrais (Quadro 3).

Quadro 3. Índices espectrais e as suas funcionalidades.

SAVI (<i>Soil Adjusted Vegetation Index</i>)	Ressaltar as áreas vegetadas.	$\frac{(1 + 0,5) \times (NIR - RED)}{(NIR + RED + 0,5)}$
NDWI (<i>Normalized Difference Water Index</i>)	Ressaltar as áreas de corpos hídricos.	$\frac{(GREEN - SWIR 1)}{(GREEN + SWIR 1)}$
NDBI (<i>Normalized Difference Built-Up Index</i>)	Ressaltar as áreas antrópicas.	$\frac{(SWIR 1 - NIR)}{(SWIR 1 + NIR)}$
NDDI (<i>Normalized Difference Drought Index</i>) adaptado	Ressaltar as áreas mais secas.	$\frac{(SAVI - NDWI)}{(SAVI + NDWI)}$
MBI (<i>Modified Bare Soil Index</i>)	Ressaltar áreas de solos expostos.	$\frac{(SWIR 1 - SWIR 2 - NIR)}{(SWIR 1 + SWIR 2 + NIR)} + 0,5$

Fonte: Adaptado de MARTINS, A. M. M et al., (2023); GU et al., (2007); SILVA et al., (2021); plataforma INDEX DataBase.

Empregou-se a operação matemática de média entre as cenas selecionadas, sendo necessário devido a ocorrência de sobreposição, principalmente nas áreas limítrofes. Além disso, a aplicação da aplicação de média teve como objetivo mitigar, embora de forma mínima, os efeitos atmosféricos oriundos das nuvens ainda remanescentes nas cenas. A formação da

imagem sintética para a classificação supervisionada é resultado da junção das bandas espectrais e índices espectrais calculados no recorte espacial da área de interesse.

Para dar suporte a classificação RF da imagem sintética do período seco, as variáveis estatísticas média, máximo, mínimo, desvio padrão e variância para as bandas e índices espectrais foram aplicados aos dados ópticos do ano de 2016 utilizando também os mesmos filtros supracitados. Dados de elevação e derivadas geomorfológicas oriundos do produto de sensoriamento remoto *NASADEM* da mesma resolução espacial do Landsat 8, e também, para a mesma área de interesse foram também incluídos no processo de classificação de imagem. O algoritmo *ee.Algorithms.Terrain*, presente na plataforma GEE permitiu os cálculos das derivadas de terreno declividade (*slope*), aspecto (*aspect*) e sombreamento (*hillshade*) que podem ser importantes na distinção das classes mapeadas.

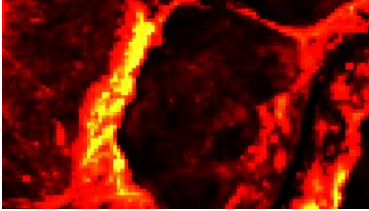
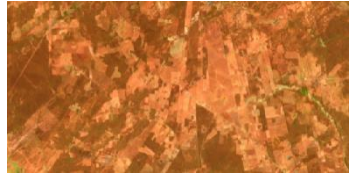
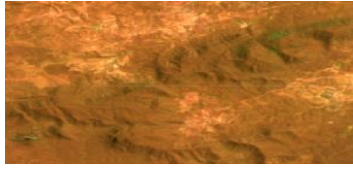



Ao todo definiu-se as classes Areia ou Degradação, Agropecuária, Vegetação de Umidade Baixa, Vegetação de Umidade Moderada, Vegetação de Umidade Alta e Corpos Hídricos, tendo sido coletadas, respectivamente, 47, 41, 46, 50, 48 e 40 amostras a partir da chave de interpretação (Quadro 4) das classes.

A partir da seleção das amostras e das variáveis a serem utilizadas foi possível a construção do modelo de classificação supervisionada RF. Foi utilizado o algoritmo *ee.Classifier.smileRandomForest* presente na plataforma GEE. O classificador RF é formado pela junção de preditores em forma de árvores, onde cada árvore é influenciada pelos valores de um vetor aleatório que é amostrado de maneira independente, seguindo a mesma distribuição para todas as árvores presentes na floresta e com a mesma distribuição para todas as árvores (Breiman, 2001). É importante citar que o parâmetro correspondente ao número de árvores determinado pelo valor 250.

A classificação da cobertura da terra gerada foi exportada para o software ArcGIS Pro, onde foi transformado em vetor (*shapefile*) pela função *Raster to Polygon*, e após isso, foi criado o mapa de cobertura da terra para os Sertões Cearenses. Foram utilizados também os recortes espaciais das áreas edificadas como máscara (buffer de 250 metros) disponibilizados pelo IBGE na escala de 1:250.000. Para a validação foram criados 30 pontos aleatórios pela função *Random Points* para cada classe mapeada onde foram comparadas com imagens de referência do *Google Earth Pro* de alta resolução espacial. Tal comparação entre as classes mapeadas e as imagens de referência, permitiu a geração da matriz de confusão para o cálculo das métricas de validação kappa e exatidão global (Fórmula 1 e 2) do mapeamento gerado.

As fórmulas 1 e 2 representam as métricas de avaliação: Exatidão Global e Índice Kappa. Define-se um nível mínimo de acurácia como Exatidão Global $\geq 0,85$, considerando-se que a acurácia é baixa se o Índice Kappa for inferior a 0,67, média se estiver no intervalo de 0,67 a 0,80, e alta se Kappa $\geq 0,80$.

Quadro 04: Chave de interpretação.

CLASSES	LANDSAT 8 - OLI	CARACTERÍSTICAS
Areia ou Degradação		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4). Parâmetros utilizados: mínimo: 0,0924; máximo: 0,4; gamma:0,14. Tons de amarelo brilhantes. É possível observar essa classe em áreas próximas a rios, áreas de agricultura e próximas de áreas construídas.
Agropecuária		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4). Tons de marrom muito claro. Normalmente formatos retangulares caracterizados pela ação antrópica. Solos expostos ou com uma leve cobertura vegetal.
Vegetação de Baixa Umidade		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4). Apresenta tons de marrom terroso e diversos formatos, às vezes influenciados pela ação antrópica.
Vegetação de Umidade Moderada		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4) e SAVI para valores entre 0,3 e 0,4. Apresenta tons de verdes claros, porém, não tanto quanto a classe “Vegetação Úmida”, apresenta também rugosidade em sua superfície.
Vegetação de Alta Umidade		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4) e SAVI para valores maiores que 0,4. Apresenta tons de verdes mais claros apresentando rugosidade em sua superfície. Geralmente a sua presença é próxima de corpos hídricos.
Corpos Hídricos		Composição falsa cor: Infravermelho de ondas 1 (B6), Infravermelho próximo (B5) e Vermelho (B4) e NDWI acima de 0,2. Caracterizado por tons mais escuros da cor azul. Não apresenta rugosidade em sua superfície.

Fonte: Elaborado por autores com imagem Landsat 8 na plataforma GEE.

$$Exatidão\ Global = \sum_{i=1}^m \frac{n_{ii}}{n} (1)$$

$$Kappa = \frac{n \sum_{i=1}^m n_{ii} - \sum_{i=1}^m n_i \cdot n_i}{n^2 - \sum_{i=1}^m n_i \cdot n_i} (2)$$

n: área total; n_{ii}: área pertencente à classe i do mapa modelo e do mapa de referência; n_{i.}: área pertencente à classe i do mapa modelo; n_{.i}: área pertencente à classe i do mapa de referência; m: número de classes.

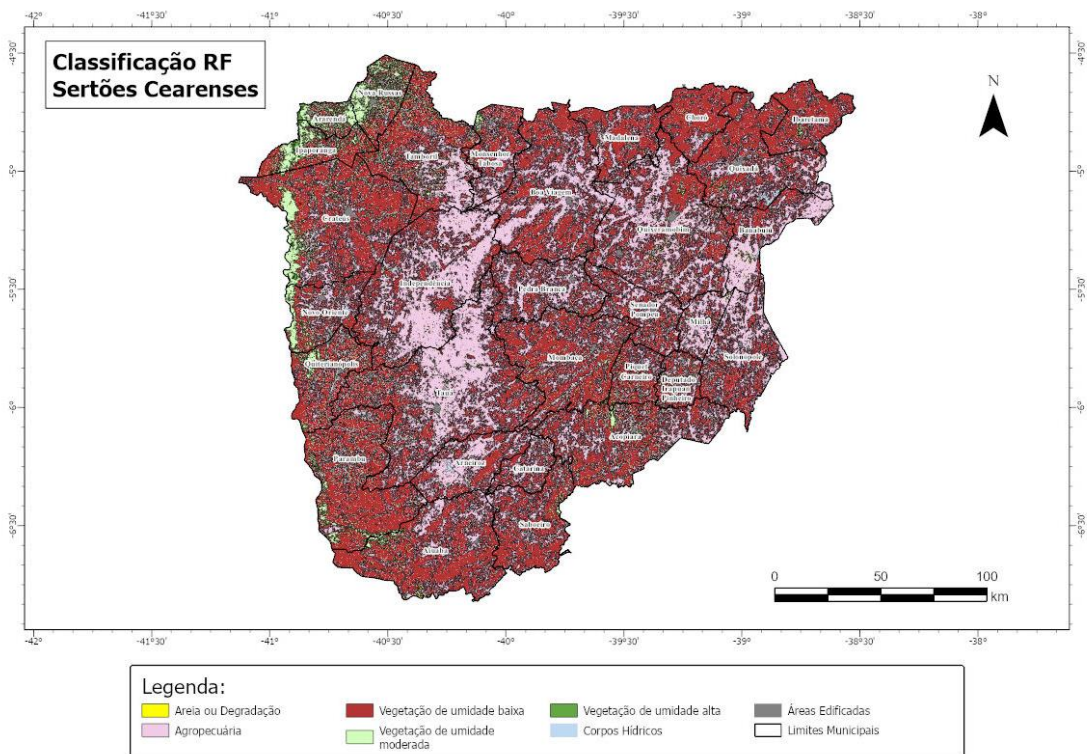
Fonte: Adaptado de Kestring et al. (2015), Anderson et al. (2001) e Krippendorff (2004).

RESULTADOS E DISCUSSÃO

O mapa de cobertura da terra gerado a partir da classificação RF para o período seco do ano de 2016 (Figura 4) indica a predominância da classe “Vegetação de Umidade Baixa” em todo o recorte espacial. Essa classe, em grande maioria, é composta por florestas secas e de vegetação rala com uma variedade de vegetações associadas ao bioma *caatinga*. Artíficos como o acúmulo de água em seus tecidos e raízes, redução do tamanho das folhas e folíolos, esclerofilia e emurchecimento das folhas (Alves, 2007) são estratégias que a vegetação apresenta para a sobrevivência na ausência de chuvas.

Outra classe bastante presente no mapa gerado é a “Agropecuária”, composta por grandes extensões de campos abertos e áreas características com a agricultura e pecuária. É bastante presente na porção leste do mapa abrangendo parte significativa da cobertura dos municípios de Independência e Tauá (Quadro 5). Os municípios Quixeramobim, Milhã e Banabuiú, presentes na porção norte e oeste, apresentaram predominância também para essa classe com padrões característicos de agropecuária em suas coberturas. No município de Banabuiú são observadas grandes extensões de campos abertos de agropecuária, embora chame atenção a presença de áreas classificadas como “Areia ou Degradação”, concentradas nas proximidades do açude Banabuiú, que aparece praticamente seco em função das ausências de chuvas na região. Essas áreas requerem atenção pois podem ser indícios de uma possível degradação local, podendo contribuir com a erosão e o assoreamento de rios. Na porção extremo sul do município de Aiuaba é possível observar a presença dessa classe, muito por conta do predomínio de solos expostos mais arenosos.

Figura 4. Classificação Random Forest para os Sertões Cearenses no período seco de 2016.



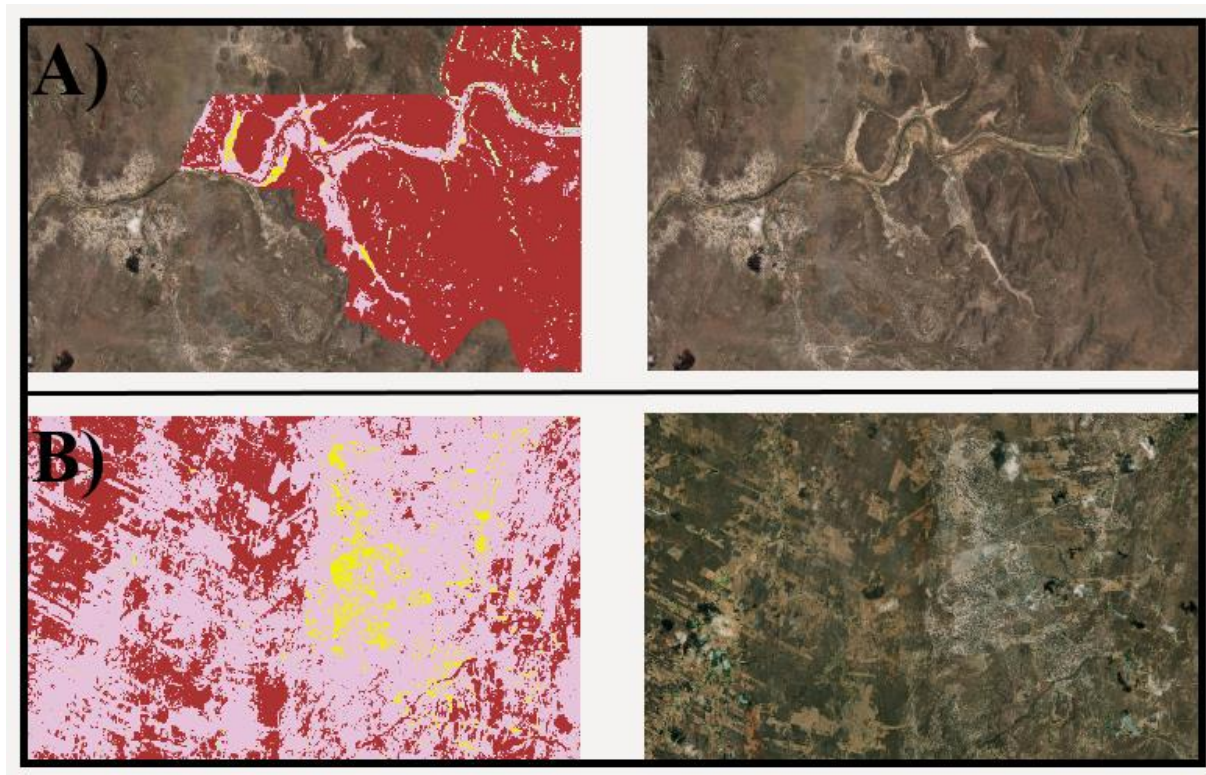
Fonte: Elaborado por autores a partir de dados do Landsat OLI e NASADEM.

Em Cratêus, no extremo leste próximo ao Estado do Piauí, é possível observar grande faixas de solos mais arenosos no entorno de rios (Figura 5 - A) podendo acentuar o assoreamento devido aos processos erosivos presentes na região. Outro exemplo também é visto em Milhã (Figura 5 - B) onde é possível ver uma grande extensão arenosa na porção nordeste

do município, localizada nas proximidades de áreas de agricultura. Tais presenças podem ser indicativos de formação de núcleos de desertificação, demandando monitoramento e ações mitigadoras.

Importante mencionar que as áreas de degradação e de areia foram agrupadas em virtude da possibilidade de confusão entre classes, pois em muitos casos, devido à alta reflectância dessas coberturas em função do solo descamado por erosão, se assemelha. Como alternativa de resolução metodológica, o agrupamento de classes foi a melhor solução. Para uma distinção mais precisa entre essas classes, seria necessário recorrer a uma classificação de imagem de referência mais antiga, a fim de destacar as áreas de mudanças.

Figura 5. Possíveis áreas de núcleos de desertificação. (A) - Porção leste do município de Crateús e (B) - Porção nordeste do município de Milhã.

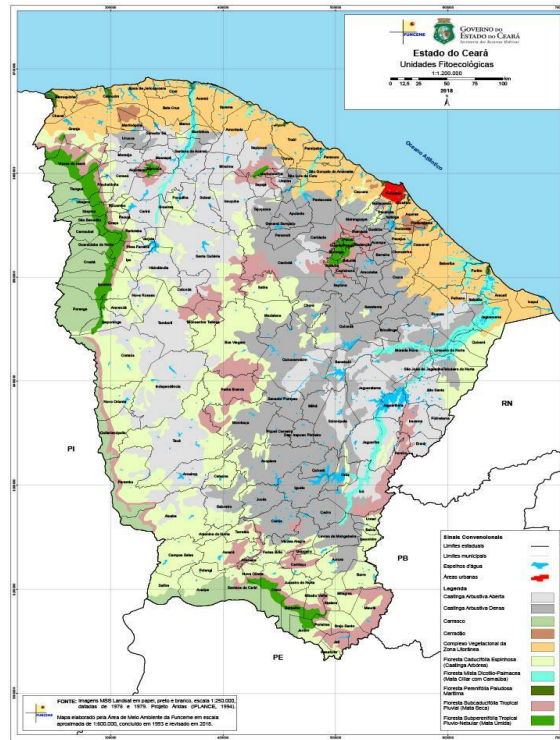


Fonte: Elaborado por autores usando o basemap da ESRI de referência presente no software ArcGIS Pro.

É possível observar também grandes áreas de coberturas de areia de origem natural sendo classificadas como “Areia ou Degradação”, presentes em Quixeramobim sendo um dos municípios que mais apresentaram essa classe em sua cobertura.

As vegetações mais úmidas correspondentes pelas classes “Vegetação Umidade Alta” e “Vegetação de Umidade Moderada”, muitas das vezes, são associadas pelas vertentes mais úmidas do terreno. Segundo o mapa fitoecológico (Figura 6) da FUNCEME (1994) as áreas da porção extremo oeste e noroeste dos sertões cearenses são caracterizadas por vegetação de carrasco e do cerrado sendo caracterizadas por tipos de vegetação mais úmidos que a *caatinga*. Além disso, a vegetação de carrasco são áreas de transição entre as vegetações da *caatinga* e do cerrado. Contudo, em termos de predominância de área são encontradas em Ararendá, Crateús, Ipaoranga, Novas Russas, Novo Oriente, Parambu, Quiterianópolis e Tamboril.

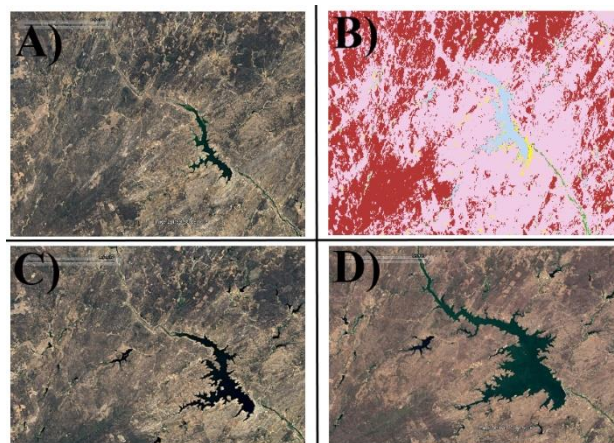
Figura 6. Unidades Fitoecológicas do Estado do Ceará (CE)



Fonte: FUNCEME (1994).

A classe “Corpos Hídricos” é influenciada pelo fomento da construção de açudes, política pública bastante importante em todo o século XX (Melo, 1999) e disseminada até os dias de hoje. É possível visualizar no mapa de cobertura da terra diversos corpos hídricos com formatos semelhantes às de açudes espalhados em todo o sertão cearense, porém mais secos devido ao déficit hídrico, apesar disso, apresentam grande impacto na cobertura dos municípios como é o caso do açude Arneiroz II estudado por Brasil (2017) que trouxe a relevância das áreas vegetadas como inibidor da erosão do solo nas proximidades da massa d’água.

Figura 7. Açude Arneiroz II presente no Município de Arneiroz. (A) - Dezembro de 2015/Ano seco, (B) - Classificação RF Agosto, Setembro e Outubro de 2016, (C) - Dezembro de 2016/Ano seco e (D) - Dezembro de 2020.



Fonte: Elaborado por autores através de imagens disponibilizadas pelo *Google Earth Pro*.

Durante anos de seca (Figura 7 - A, B e C), é possível observar as alterações na extensão da superfície do açude Arneiroz II em comparação com anos normais, nos quais a área está mais preenchida de água (Figura 7 - D).

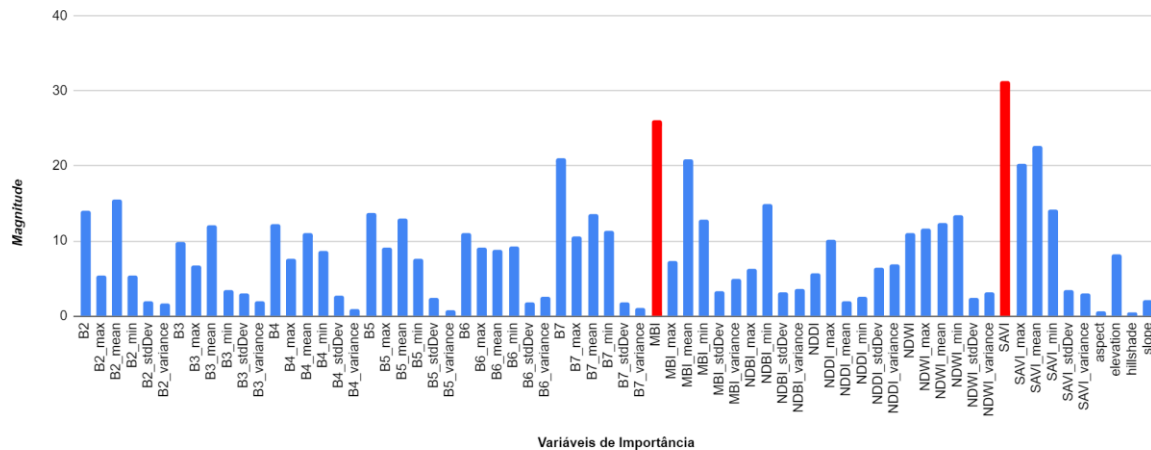
Quadro 5. Áreas (%) das classes mapeadas pelo modelo RF.

Sertões Cearenses - Municípios	Areia e Degradação	Vegetação de baixa umidade	Agropecuária	Corpos Hídricos	Vegetação de média umidade	Vegetação de alta umidade
Acopiara	0,45	60,33	34,84	0,26	3,43	0,36
Aiuaba	1,35	68,41	27,21	0,15	2,48	0,11
Ararendá	0,01	40,9	10,03	0,37	42,54	4,37
Arneiroz	0,35	48,3	49,42	0,96	0,51	0,17
Banabuiú	1,96	37,72	58,61	0,37	0,42	0,21
Boa Viagem	0,06	56,74	42,09	0,01	0,47	0,07
Catarina	0,25	54,72	43,85	0,11	0,55	0,15
Choró	0,07	82,57	15,58	0,19	0,39	0,23
Crateús	0,09	66,74	21,18	0,23	9,6	0,88
Deputado Irapuan Pinheiro	0,62	58,62	36,13	0,27	2,14	0,63
Ibaretama	0,2	79,52	17,94	0,07	1,4	0,22
Independência	0,43	37,78	60,36	0,11	0,77	0,09
Ipaporanga	0,02	56,13	8,19	0,33	30,14	4,47
Madalena	0,08	68,53	29,24	0,03	0,74	0,37
Milhã	1,01	33,65	62,98	0,03	0,65	0,17
Mombaça	0,08	70,03	26,85	0,03	2,02	0,2
Monsenhor Tabosa	0,09	57,48	38,7	0,03	2,91	0,18
Nova Russas	0,12	49,74	11,16	0,46	34,74	1,71
Novo Oriente	0,08	55,29	30,03	0,41	12,53	0,38
Parambu	0,17	74,39	17,42	0,25	6,87	0,37
Pedra Branca	0,09	50,17	48,11	0,02	0,59	0,05
Piquet Carneiro	0,16	65,93	28,65	0,17	3,7	0,43
Quiterianópolis	0,07	67,12	21,37	0,4	9,62	0,58
Quixadá	0,65	62,93	32,7	0,63	1,01	0,25
Quixeramobim	1,06	48,09	48,71	0,03	0,93	0,13
Saboeiro	0,1	69,29	27,85	0,15	2,07	0,2
Senador Pompeu	0,44	55,58	41,68	0,1	0,92	0,18
Solonópole	0,71	49,5	48,3	0,09	0,49	0,27
Tamboril	0,2	59,41	33,17	0,2	5,86	0,57
Tauá	0,24	44,75	52,71	0,29	1,24	0,33

Fonte: Elaborado por autores com base no mapeamento e uso e cobertura da terra.

A importância relativa de cada variável preditora (Figura 8) indicou que o índice de vegetação SAVI foi a variável que mais contribuiu para a classificação do modelo RF. Outro índice espectral que teve resultados significativos foi o MBI. Em relação ao SAVI e o MBI (em vermelho na figura 8), destacam a variabilidade na cobertura da terra e que as condições sazonais, como o período seco, podem afetar essas variáveis, influenciando assim o resultado da classificação pelo modelo RF.

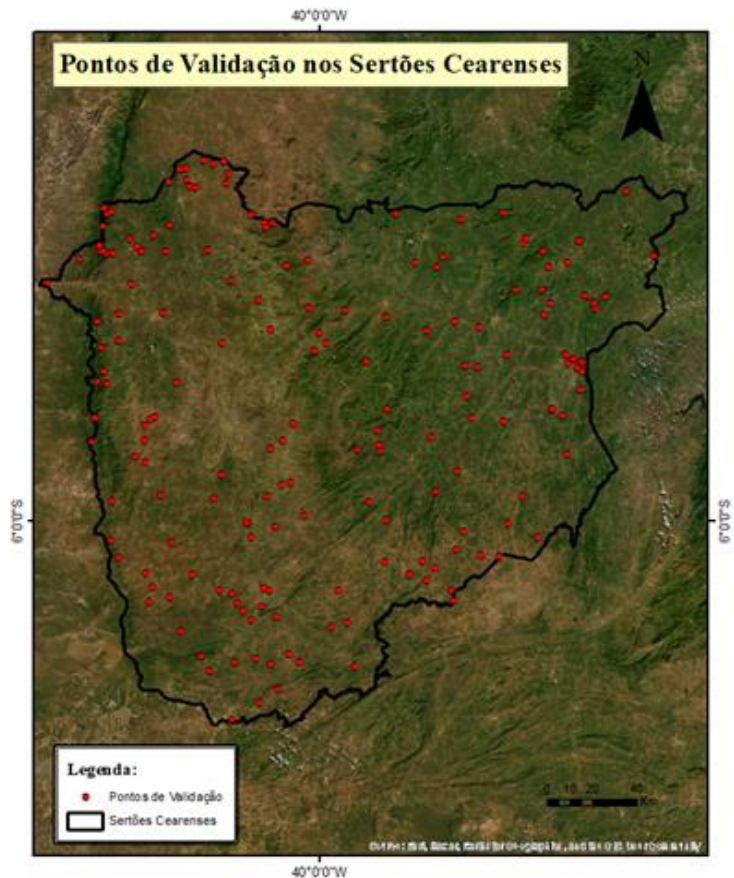
Figura 8. Gráfico de barras das variáveis de importância do modelo RF.



Fonte: Elaborado por autores na plataforma GEE.

Foram empregados um total de 30 pontos de validação por classe, totalizando 180 pontos, distribuídos aleatoriamente no recorte espacial dos Sertões Cearenses (Figura 9). O objetivo foi comparar o mapa gerado pela classificação RF com imagens de satélite de resolução espacial igual ou superior, capturadas durante o período seco na plataforma Google Earth Pro. Essa comparação foi fundamental para criar a matriz de confusão, que, por sua vez, possibilitou o cálculo das métricas para avaliação do mapeamento gerado.

Figura 9. Pontos de validação para a área de estudo.



Fonte: Elaborado por autores usando o basemap da ESRI.

A matriz de confusão (Quadro 6) e os respectivos cálculos das métricas de validação indicaram que o kappa e a exatidão global obtiveram valores de 82% e 85%, respectivamente. Segundo Krippendorff (2004), pelos valores das métricas o mapeamento pode ser avaliado como de alta acurácia. A classe “Vegetação de Umidade Moderada” foi a que apresentou maior incerteza, pois o modelo teve dificuldades em separar da classe “Vegetação de Umidade Baixa”. Isso ocorre frequentemente porque a classe “Vegetação de Umidade Moderada” é frequentemente encontrada em minoria em extensas áreas dominadas pela classe “Vegetação de Baixa Umidade”. Além disso, essa situação pode indicar o avanço da classe de vegetação moderada para áreas com vegetação mais seca. Em contrapartida, as classes de “Corpos Hídricos” e as de “Vegetação de Alta Umidade” foram as que mais acertaram em comparação com as demais.

Quadro 6. Matriz de Confusão.

Classes	Areia ou Degradação	Vegetação de Umidade Baixa	Agropecuária	Corpos Hídricos	Vegetação de Umidade Moderada	Vegetação de Umidade Alta	Total
Areia ou Degradação	26	3	1	0	0	0	30
Vegetação de Umidade Baixa	0	27	1	0	2	0	30
Agropecuária	0	6	24	0	0	0	30
Corpos Hídricos	0	0	1	28	1	0	30
Vegetação de Umidade Moderada	0	8	0	2	20	0	30
Vegetação de Umidade Alta	0	0	0	0	2	28	30
Total	26	44	27	30	25	28	180

Fonte: Elaborado por autores.

Além disso, ressalta-se a dificuldade de se encontrar imagens históricas no Google Earth Pro, pois muitas das vezes, foi necessário recorrer a interpretação de anos próximos ou avaliar o histórico da cobertura terrestre da área de estudo através das imagens, pois as condições de visualização e iluminação da imagem de satélite na plataforma não eram favoráveis.

CONCLUSÕES

O sensoriamento remoto em nível orbital, associado às técnicas de aprendizado de máquina, pode contribuir significativamente para o monitoramento e o mapeamento da cobertura da terra no semiárido brasileiro ainda mais em um contexto de eventos extremos de secas como ocorridos no ano de 2016.

O modelo RF indicou que a variável SAVI foi a mais importante no processo de classificação de imagem, devido à sua sensibilidade em captar a vegetação na área de estudo. As características físicas da região, especialmente durante a ausência de chuvas, resultam em extensões de solos descamados, o que favorece o destaque das áreas vegetadas na área de estudo

devido às diferenças em seus comportamentos espectrais. O processo inverso ocorre com a variável MBI, cujo destaque mais acentuado de comportamentos espectrais semelhantes ao solo permitiu que o índice espectral de solos expostos também tivesse importância no modelo RF na área de estudo.

A metodologia de baixo custo e do uso de plataforma em nuvem de alta performance de processamento de grande volume de dados desempenhou um mapeamento de cobertura terrestre com índice kappa e exatidão global satisfatórios possibilitando assim a replicação em outras áreas do semiárido.

Espera-se que este trabalho contribua para uma melhor caracterização da natureza física do semiárido brasileiro. A defasagem de estudos dessa natureza no sensoriamento remoto (GANEM, 2020) pode ser considerada um impedimento para compreender a dinâmica da vegetação, dos núcleos de desertificação, e o monitoramento da água, entre outros temas relevantes no semiárido brasileiro. Contudo, a metodologia aplicada neste estudo pode ser considerada um ponto de partida para a geração de mapas e a construção de séries temporais, sendo fundamental para proporcionar maior clareza quanto à degradação de terras na região.

AGRADECIMENTOS

Agradeço novamente à oportunidade que o Laboratório de Sensoriamento Remoto e Estudos Ambientais - ESPAÇO/UFRJ me concedeu para dar continuidade a uma pesquisa tão rica, que é trabalhar com o semiárido brasileiro. Agradeço também à minha orientadora, Carla Madureira, e aos outros coautores, Diego Sperle e João Neves, que se colocaram à disposição para contribuir para este trabalho fantástico de sensoriamento remoto em áreas semiáridas.

REFERÊNCIAS

ALVES, Jose Jakson Amancio. Geocologia da Caatinga no Semiárido do Nordeste Brasileiro. CLIMEP. Climatologia e Estudos da Paisagem, Rio Claro, SP, v. 2, n. 1, p. 58-71, jan./jun. 2007.

AMANI, Meisam; GHORBANIAN, Arsalan; AHMADI, Seyed Ali; KAKOOEI, Mohammad; MOGHIMI, Armin; MIRMAZLOUMI, S. Mohammad; MOGHADDAM, Sayyed Hamed Alizadeh; MAHDAVI, Sahel; GHAHREMANLOO, Masoud; PARSIAN, Saeid; WU, Qiusheng; BRISCO, Brian. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, [s. l.], v. 13, p. 5326-5350, 2020. DOI 10.1109/JSTARS.2020.3021052. Disponível em: <https://ieeexplore.ieee.org/document/9184118/authors>.

BRASIL, J. B.; ROSA, G. Q.; RIBEIRO FILHO, J. C.; SILVA, C. V. T. Estimativa de perda de solo na bacia do açude Arneiroz II por ferramentas de geoprocessamento. Revista Ambiente, Guarapuava, v. 13, n. 2, p. 503-517, 2017. Disponível em: <https://revistas.unicentro.br/index.php/ambiencia/article/view/4703/pdf>. Acesso em: 15 abr. 2023.

ANDERSON, J.F.; HARDY, E.E.; ROACH J.T.; WITMER R.E. A land use and land cover classification system for use with remote sensor data. Washington, DC: U.S. Geological Survey, 2001. 41p. (Professional Paper 964).

BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.

CARDOSO, P. V.; SEABRA, V. da S.; XAVIER, R. A.; RODRIGUES, E. de M.; GOMES, A. S. Mapeamento de Áreas de Caatinga Através do Random Forest: Estudo de caso na Bacia do Rio Taperoá. *Revista Geoaraguaia*, [S. l.], v. 11, n. Especial, p. 55–68, 2021. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/geo/article/view/12743>. Acesso em: 18 jan. 2024.

FUNCEME. Mapa Fitoecológico do Estado do Ceará. [S. l.: s. n.], 1994. Fitofisionomias. Disponível em: http://www.funceme.br/wp-content/uploads/2019/02/15-Mapa_CE_Fitoecologico_A2.pdf. Acesso em: 18 de Janeiro de 2024.

GANEM, Khalil Ali; DUTRA, Andeise Cerqueira; OLIVEIRA, Marcell Terra de; FREITAS, Ramon Morais de; GRECCHI, Rosana Cristina; VIEIRA, Rita Marcia da Silva Pinto; ARAI, Egidio; SILVA, Fabrício Brito; SAMPAIO, Cláudia Bloisi Vaz; DUARTE, Valdete; SHIMABUKURO, Yosio Edemir. Mapeamento da Vegetação da Caatinga a partir de Dados Ópticos de Observação da Terra – Oportunidades e Desafios. *Revista Brasileira de Cartografia*, [s. l.], v. 72, p. 829–854, 30 dez. 2020. DOI <https://doi.org/10.14393/rbcv72nespecial50anos-56543>. Disponível em: <https://seer.ufu.br/index.php/revistabrasileiracartografia/article/view/56543>. Acesso em: 18 jan. 2024.

GIONGO, Vanderlise. Desertificação e Mudanças Climáticas no Semiárido Brasileiro: Mudanças Climáticas e Sustentabilidade. In: *DESERTIFICAÇÃO e Mudanças Climáticas no Semiárido Brasileiro*. Campina Grande: Instituto Nacional do Semiárido, 2011. Capítulo Balanço de carbono no semiárido brasileiro: Perspectivas e desafios, p. 115-130. ISBN 978-85-64265-02-8.

KESTRING, F. B. F.; GUEDES, L. P. C.; BASTIANI, F. de.; URIBE-OPAZO, M. A.. Comparação de mapas temáticos de diferentes grades amostrais para a produtividade da soja. *Engenharia Agrícola*, v. 35, n. 4, p. 733–743, 2015. Disponível em: <https://doi.org/10.1590/1809-4430-Eng.Agric.v35n4p733-743/2015>. Acesso em: 18 jan. 2024.

KRIPPENDORFF, K. *Content analysis: an introduction to its methodology*. Beverly Hills: Sage Publications, 2004. 412 p.

KUMAR, L.; MUTANGA, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sensing*, v. 10, p. 1509, 2018.

INDEX DataBase: A database for remote sensing indices. [S. l.], 2011. Disponível em: <https://www.indexdatabase.de/>. Acesso em: 18 jan. 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Malhas Territoriais. Disponível em: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html>. Acesso em: 18 jan. 2024.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE. ENOS - Estação Nacional de Órbita e Sonda. Disponível em: <http://enos.cptec.inpe.br/>. Acesso em: 18 jan. 2024.

IPCC. Summary for Policymakers. Sixth Assessment Report, 2021. Disponível em: <https://www.ipcc.ch/report/ar6/wg1/>.

MARTINS, Adinan Marzulo Maia et al. Mapeamento do uso e cobertura do solo do município de Saboeiro (CE) através de processamento na nuvem no GEE. Revista Continentes, [S.l.], n. 22, p. 66-85, out. 2023. ISSN 2317-8825. Disponível em: <<https://www.revistacontinentes.com.br/index.php/continentes/article/view/410>>. Acesso em: 18 jan. 2024. doi: <https://doi.org/10.51308/continentes.v1i22.410>.

MARENGO, J.A.; ALVES, L.; AVALA, R.; BRITO, S.; MORAES, O. Climatic characteristics of the 2010-2016 drought in the semiarid Northeast Brazil region. Anais da Academia Brasileira de Ciências. São Paulo 2017.

MELO, Josemir Camilo de. O fenômeno El Niño e as secas no Nordeste do Brasil. Raízes: Revista de Ciências Sociais e Econômicas, [s. l.], v. 162, ed. 20, p. 13-21, Novembro 1999. DOI <https://doi.org/10.37370/raizes.1999.v.162>. Disponível em: <http://raizes.revistas.ufcg.edu.br/index.php/raizes/article/view/162>.

NASA JPL. (2020). NASADEM Merged DEM Global 1 arc second V001 [Conjunto de dados]. NASA EOSDIS Land Processes DAAC. Acessado em 30 de dezembro de 2020, de doi:10.5067/MEaSURES/NASADEM/NASADEM_HGT.

NOBRE, Paulo. Mudanças climáticas e desertificação: os desafios para o Estado Brasileiro. Desertificação e Mudanças Climáticas no Semiárido Brasileiro, Instituto Nacional do Semiárido, Campina Grande, p. 25-35, 2011. ISBN: 978-85-64265-02-8.

OYAMA, Marcos Daisuke; NOBRE, Carlos Afonso. A new climate-vegetation equilibrium state for Tropical South America. Geophysical Research Letters, [s. l.], v. 30, 5 dez. 2003. DOI <https://doi.org/10.1029/2003GL018600>. Disponível em: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2003GL018600>.

SOUZA, A. 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, n. 17, 2020. DOI: 10.3390/rs12172735.

TAMIMINIA, H. et al. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS Journal of Photogrammetry and Remote Sensing, v. 164, p. 152-170, 2020. ISSN 0924-2716. Disponível em: <<https://doi.org/10.1016/j.isprsjprs.2020.04.001>>.