

USO DE SÉRIES TEMPORAIS DE DADOS DE SENSORIAMENTO REMOTO E ALGORITMO DE MACHINE LEARNING PARA MAPEAMENTO DE WETLANDS NO NORTE FLUMINENSE - RJ

Costa, E.C.P. (UFF) ; Santos, M.P. (UFF) ; Almeida, P.M.M. (UFF) ; Vicens, R.S. (UFF)

RESUMO

O artigo tem o objetivo de contribuir metodologicamente para o mapeamento de wetlands, ambientes que demandam critérios específicos para seu mapeamento, devido a sua sazonalidade. O estudo se concentra na identificação desses objetos em um recorte no estado do Rio de Janeiro, utilizando técnicas de sensoriamento remoto multitemporal em imagens de sensores óticos (Landsat 8) e imagens de radar (Sentinel 1). O algoritmo Random Forest de aprendizado de máquina foi utilizado, na plataforma GEE, para a classificação desses ambientes. Para isso, foram utilizadas séries temporais mensais de imagens de radar e do índice NDWI, do ano de 2020, sendo possível testar suas potencialidades na identificação desses objetos. Como resultados obteve-se a classificação dos wetlands, considerando o período de inundação no ciclo anual. A validação, através do índice kappa da classificação de wetlands permanentes foi de 0,82 e de wetlands temporários foi de 0,76.

PALAVRAS CHAVES

Série temporal; Aprendizado de Máquina; NDWI; Wetlands ; Radar

ABSTRACT

The article aims to contribute methodologically to the mapping of wetlands, which require specific criteria for their mapping due to their seasonality. The study focuses on identifying these objects in a specific area in the state of Rio de Janeiro, using multitemporal remote sensing techniques on optical sensors (Landsat 8) and radar images (Sentinel 1). The Random Forest machine learning algorithm was used on the GEE platform for the classification of these environments. Monthly time series of radar images and the NDWI index from 2020 were used to test their potential in identifying these objects. The results obtained the classification of wetlands, considering the flooding period in the annual cycle. The validation, through the kappa index of the classification of permanent wetlands, was 0.82 and of temporary wetlands was 0.76.

INTRODUÇÃO

Estudos que se propõem a identificar, compreender e analisar as questões ambientais são relevantes no cenário de crise ambiental global. Tais abordagens dão suporte no entendimento de dinâmicas, no diagnóstico de problemas, assim como são ferramentas para gerenciá-los. Nesse contexto, o sensoriamento remoto e seus aportes metodológicos são ferramentas muito importantes, a prospecção da metodologia mais apropriada a ser usada, considerando as especificidades de cada objeto de pesquisa ou problema científico, pode otimizar o tempo de pesquisa, assim como a qualidade dos resultados a serem gerados. A água cobre cerca de 74% da superfície terrestre. Esse recurso natural existe em vários estados na Terra, que incluem água doce, água salgada, água em estado gasoso, chuva, neve e gelo (JENSEN, 2009). Apesar dos wetlands estarem presentes em diversas áreas da superfície terrestre, esse recurso natural, muitas das vezes é invisibilizado e ignorado por parte das entidades gestoras e até mesmo pela sociedade em geral, sendo suprimido e degradado. A Convenção de Ramsar (2013) define os wetlands como áreas cobertas com água, sejam naturais ou artificiais, permanentes ou temporárias, estagnadas ou em funcionamento, frescas, doce ou salgadas. No Brasil, os wetlands são definidos como ecossistemas na interface entre ambientes terrestres e aquáticos, continentais ou costeiros, naturais ou artificiais, permanente ou periodicamente inundados ou com solos encharcados, as águas podem ser doces, salobras ou salgadas, com comunidades de plantas e animais adaptados à sua dinâmica hídrica (BRASIL, 2015). É importante mencionar que as áreas de wetlands não podem ser consideradas

como ecossistemas estáticos, mas que estão sempre em constante alteração, sendo a dinâmica algo inerente a essas áreas. Dessa forma, há nos wetlands um mosaico complexo de distintos estágios de desenvolvimento, aumentando a diversidade da paisagem e a variedade de espécies encontradas (MOORE, 2008). Sendo assim, é extremamente relevante esforços que deem conta de mapear esses objetos complexos, mas de extrema relevância ecossistêmica. Como aporte metodológico para mapeamentos e diagnósticos de variáveis na superfície terrestre, temos as imagens de radar. O radar é considerado como uma grande ferramenta metodológica para este trabalho devido a disponibilidade de suas séries mensais de imagens. Outra relevância metodológica para este trabalho é a utilização do algoritmo de classificação Random Forest em séries temporais. Esse algoritmo de classificação é um dos mais utilizados para classificação de uso e cobertura da Terra usando dados de Sensoriamento Remoto, possuindo maior precisão do que outros classificadores e uma otimização da velocidade de processamento selecionando variáveis importantes (PHAN et al. 2020). Também chamamos a atenção para a utilização de uma plataforma que tem auxiliado de maneira muito efetiva os estudos e pesquisas na área das geotecnologias, o uso de programação para mapeamentos de sistemas terrestres, a exemplo, a plataforma Google Earth Engine (GEE), que é baseada em nuvem para processamento computacional de dados geoespaciais em grande escala (GOOGLE, 2022). Outra importância se deve ao uso de imagens óticas, que propiciam o uso de índices radiométricos na identificação de wetlands, tendo em vista sua dinâmica de alagamento. Neste trabalho foi usada uma serie temporal mensal de imagens-índice do grau de umedecimento/alagamento da superfície: o índice de umedecimento por diferença normalizada (NDWI pelas suas siglas em inglês), utilizado como descritor na classificação de wetlands, buscando então avaliar o desempenho desse índice na identificação de objetos complexos, como os wetlands. Por fim, o objetivo deste trabalho é contribuir para o mapeamento de wetlands, ambientes que demandam critérios específicos para seu mapeamento, devido a sua sazonalidade.

MATERIAL E MÉTODOS

A área de estudo definida para esta pesquisa possui uma extensão de 9.917 km², e se configura como uma área no Norte Fluminense, no Estado do Rio de Janeiro. Ela comporta uma abundância de wetlands, com grande diversidade ecossistêmica. A pesquisa foi desenvolvida na plataforma do Google Earth Engine (figura 1), que permite o acesso e processamento das imagens de satélite por meio de códigos em linguagem de programação, tais como Javascript e Python. Cada vez mais o GEE tem se tornado uma plataforma usual para pesquisas acadêmicas, tendo em vista o crescimento e a influência da programação na área de geotecnologias. Na plataforma foram então desenvolvidos dois scripts separadamente. O primeiro, para classificar corpos d'água, sem cobertura vegetal, e o segundo, para classificar wetlands vegetados. Em ambos os casos, as classificações foram binárias e posteriormente combinadas no mapeamento final. Essa divisão foi necessária devido à natureza das imagens a serem processadas e os tipos de amostras veiculadas aos códigos. No algoritmo utilizado para a classificação da variação sazonal da água, foram utilizadas imagens Sentinel 1B (SAR) do ano de 2020, com resolução espacial de 10 metros, imagens de órbita descendente, IW, banda VH. Foram adotadas 4.666 amostras de todas as classes existentes na área de estudos. Para ser possível realizar o processamento anual, foram elencados todos os dias de cada mês do ano de 2020, sendo processadas 12 classificações diferentes, correspondendo a uma para cada mês do ano. Como métricas da classificação Random Forest no GEE foram utilizados a média, mediana, percentis, desvio padrão e uma máscara de relevo, advinda do SRTM. Na classificação mensal foi adotado um filtro espacial denominado "majority", que utiliza de padrões de vizinhança para evitar a presença de ruídos na classificação. No algoritmo utilizado na classificação de wetlands vegetados foram adotadas imagens Landsat 8, sensor OLI, do ano de 2020, com resolução espacial de 30 metros. As imagens foram pré processadas, adotando filtro de nuvens e inserindo o índice radiométrico NDWI proposto por McFeeters (1996), que corresponde a $(NIR - G)/(NIR + G)$. Esse índice foi adotado como banda espectral e inserido no script. Como métricas das imagens foram adotadas média, mediana, desvio padrão e uma máscara de relevo, originada pelo SRTM. Como resultados, obteve-se a classificação mensal de 2020, totalizando 12 classificações binárias de wetlands ao longo do ano. Ambos os resultados foram inseridos no ArcGis, utilizando as ferramentas do repositório "Data Management Tools": Copy Raster, Tabulate Area e Combine para processar os dados e obter valores para a representação final. A validação do mapeamento ocorreu através de metodologia de

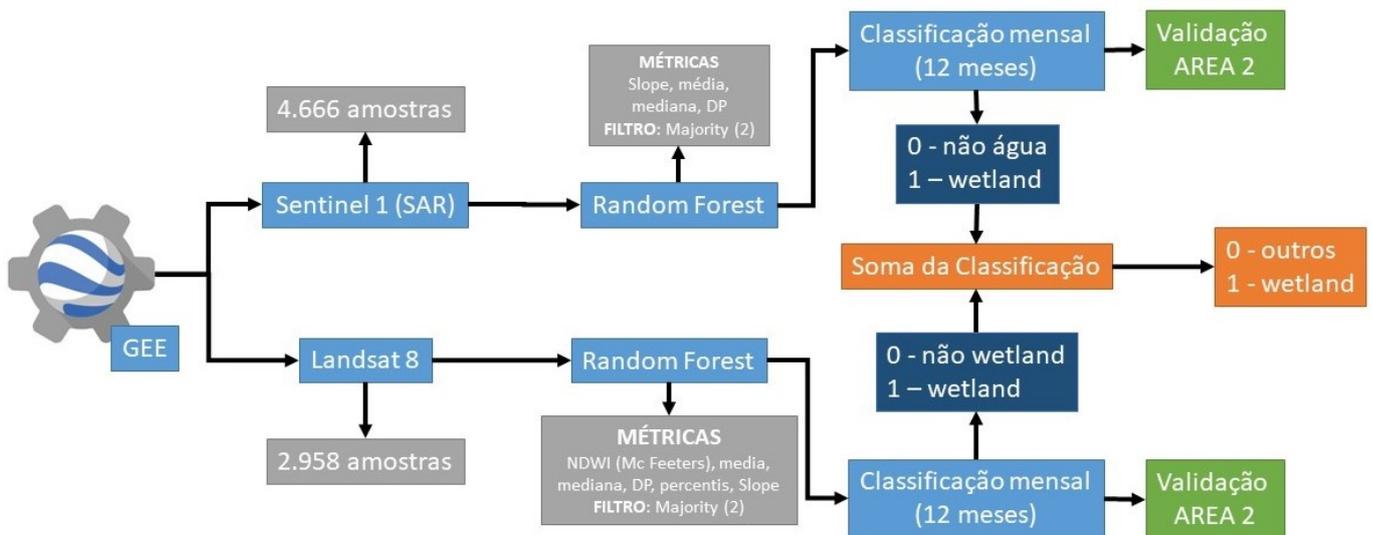
validação estratificada e aleatória, no software GEE. Devido a proporcionalidade de áreas de wetlands serem menores na área de estudos, foi utilizado o desenho amostral do AREA 2. O objetivo principal dessa metodologia de validação é estimar precisão específica da classe. A amostragem aleatória estratificada oferece a opção de aumentar o tamanho da amostra em classes que ocupam uma pequena proporção da área para reduzir os erros padrão das estimativas de precisão específicas da classe (OLOFSSON et al, 2014). Foram então selecionados 100 pontos de não wetland e 20 de wetland permanente, bem como 100 pontos de não wetlands e 20 de wetland temporário, para realizar a validação visual no Google Earth Engine. As validações ocorreram separadamente, para avaliar o desempenho das duas classificações de sensores diferentes. Os resultados, que possuem resoluções espaciais diferentes (10 metros e 30 metros), tiveram que ser exportados na resolução de 30 metros, a fim de adequar os mapeamentos e evitar os ruídos de escala.

RESULTADOS E DISCUSSÃO

Se tratando dos resultados, é relevante dizer que pesquisas desenvolvidas nas geotecnologias cada vez mais tem ganhado a contribuição das chamadas IA (inteligência artificial), que tem otimizado o processamento de dados e a geração de resultados. Isso também é tocante a utilização de machine learning, como no caso deste trabalho, que auxiliou em um mapeamento e processamento muito mais automatizado, permitindo com que fossem feitos mais testes de aperfeiçoamentos, com curtos períodos, até que se chegasse a resultados mais satisfatórios. Há uma grande diversidade de wetlands existentes no Norte Fluminense, tais como: costeiros, lagunares, fluviomarinhos, mangues, antrópicos, antropizados ou com cobertura vegetal inundada. No processo de classificação foram utilizadas toda essa diversidade amostral, que permite observar uma maior amplitude na classificação. A metodologia da pesquisa permitiu a identificação de wetlands no Norte Fluminense, sendo possível caracterizá-los segundo o seu regime de inundação. Para tanto, o resultado obtido foi processado e representado em duas classes principais de wetlands: wetlands permanentes, que corresponde aos wetlands que permanecem durante os 12 meses do ano inundados; e wetlands temporários, equivalente aos wetlands que possuem cobertura de água sazonalmente, durante o ano de 2020. Como principal resultado deste trabalho, foi obtido um mapa da classificação de wetlands de inundação permanentes e temporários (figura 2). O mapeamento elaborado foi representado a partir da soma dos 12 meses do ano de 2020. Logo, a figura representa uma síntese dos dados gerados, com o objetivo de representar o limite espacial de maior incidência desse fenômeno para o ano de 2020, somando assim, a área total de wetlands do recorte de estudos. A partir do mapa é possível identificar os wetlands permanentes presentes na área de estudos (figura 2). A delimitação desta classe está diretamente associada ao desempenho das imagens SAR. Os wetlands temporários foram classificados pelo desempenho das imagens SAR e das imagens ópticas, assim como do uso de índices radiométricos, como o NDWI, que apoiou o classificador na detecção da umidade. Os wetlands temporários são encontrados ao redor de grandes corpos hídricos como a Lagoa Feia e a Lagoa do Campelo, assim como em áreas próximas aos rios da região, apontando também essa estreita relação entre as wetlands permanentes e temporários. Já no que tange a realização de uma análise mais quantitativa do estudo, apresenta-se como resultado uma extensão estimada de 322 km² de wetlands permanentes e 2.510 km² de wetlands temporários, em uma área de estudos que corresponde a 9.917 km². Estima-se que os wetlands permanentes correspondam a cerca de 3% da área em estudos, enquanto os wetlands temporários alcançam cerca de 25% da área. Em relação ao resultado da validação (figura 3) da classificação de wetlands permanente inundados, obteve-se um valor de kappa de 0,82; Exatidão Global de 0,950; Acurácia do Produtor (Precision) 0,900; Acurácia do Usuário (Recall) 0,818; e F-score de 0,857. Já se tratando do resultado da validação da classificação de wetlands temporariamente inundado, o valor do kappa foi de 0,76; Exatidão Global de 0,933; Acurácia do Produtor 0,800; Acurácia do Usuário 0,800; e F-score de 0,80. É importante destacar que os erros de omissão encontrados durante a validação estão relacionados a áreas de wetlands que foram classificadas, erroneamente, como solo exposto, afloramento rochoso ou áreas de agricultura. Uma análise associada ao regime de inundação dos wetlands temporários foi realizada através dos dados espaciais gerados, conforme dados do gráfico do mapa. Ressalta-se que os meses de maior inundação são os meses de novembro a março, sendo estes caracterizados por obterem maiores irradiações solares e conseqüentemente maiores índices pluviométricos, conforme as características climáticas do Rio de Janeiro. Esses meses correspondem

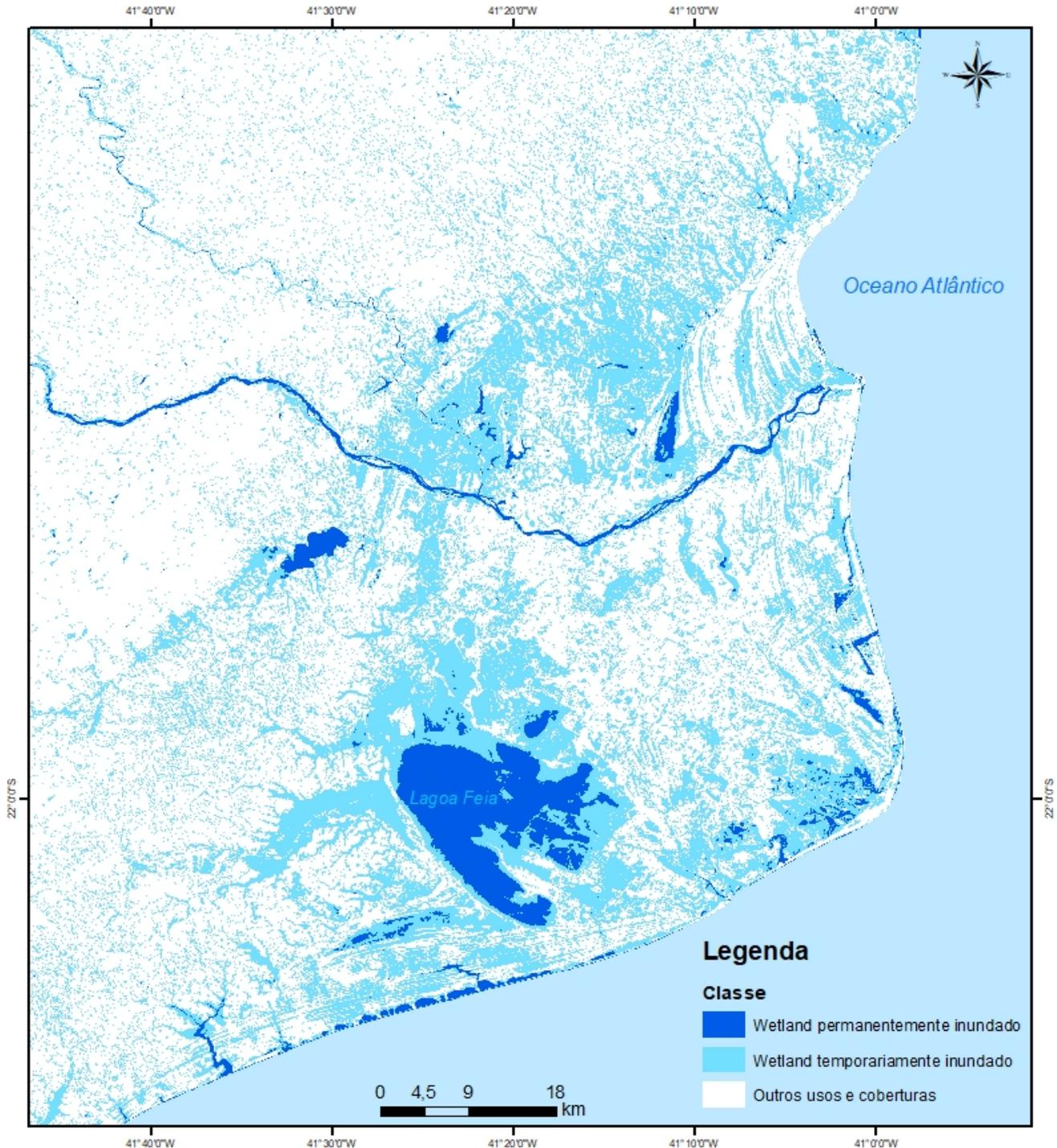
a áreas de wetlands temporários de cerca de 20% a 25% da área de estudos. Já nos meses de menores índices pluviométricos, encontram-se as menores dimensões espaciais de wetlands, correspondendo a cerca de menos de 20% da área de estudos nos períodos entre abril e outubro. Dentre as análises espaciais que podem ser evidenciadas por esta classificação, destaca-se a disposição espacial do fenômeno estudado, que se localiza em áreas ao entorno de wetlands permanentes, como grandes lagoas. As suas formas geométricas e dimensão espacial variam de acordo com o corpo hídrico associado ao wetland temporário. A disposição dos wetlands também está associada a geomorfologia da região estudada, onde predominam planícies costeiras e fluviomarinhas. Conforme a classificação aponta, os wetlands ao sul da área de estudo são caracterizados por zonas úmidas entre feixes de cordões arenosos do Paraíba do Sul. Há outras áreas de planície que apresentam características úmidas, tal como áreas no entorno da Lagoa Feia, Lagoa de Cima e a Lagoa do Campelo, corpos hídricos de maiores dimensões da área de estudos, que são caracterizadas por este mapeamento como wetlands permanentes. Outros wetlands permanentes estão localizados em áreas próximos a linha de costa, localizados nas áreas próximas à UCs como Restinga de Jurubatiba e Parque Estadual da Lagoa do Açu (PELAG). Ressalta-se que uso da máscara de relevo, originada pelo SRTM e adotada como métrica, auxiliou na classificação dos wetlands, ajudando a minimizar os erros de áreas úmidas que se confundiam com áreas de afloramentos rochoso, sombras e pastagens. Ou seja, com o uso do da máscara de relevo, o classificador pode entender a restrição dos wetlands dentro do limiar de superfícies planas.

Figura 1



Fluxograma metodológico das etapas das classificações

Figura 2



Mapa da cobertura de wetlands na área de estudo

Figura 3

Parâmetros	SAR	Landsat
	Wetland permanentemente inundado	Wetland temporariamente inundado
Exatidão Total	0,95	0,93
Índice Kappa	0,83	0,76
Acurácia Total (Overall accuracy)	0,95	0,93
Acurácia do Produtor (Precision)	0,90	0,80
Acurácia do Usuário (Recall)	0,82	0,80
F-score	0,86	0,80

Resultados da validação das classificações

CONSIDERAÇÕES FINAIS

É importante mencionar que as classificações desenvolvidas em aprendizado de máquina foram essenciais para a realização deste estudo, pois o processamento em nuvem otimizou as 12 classificações correspondentes aos meses do ano de 2020 e a representação da dimensão temporal do fenômeno. O uso de radar em classificações de wetlands permanentes, alvos complexos no sentido de sua correta identificação, pode ser considerado como um diferencial metodológico e, a abertura de caminhos para diversos outros testes e análises, objetivando um mapeamento cada vez mais detalhado, dando conta de representar corretamente a cobertura desses objetos e suas dinâmicas sob a superfície terrestre. O uso de imagens multiespectrais, como as imagens Landsat-8, foi satisfatório para essa pesquisa, tendo em vista que possibilitou a utilização do índice radiométrico NDWI, auxiliando na identificação dos wetlands, que são alvos bastante complexos, tendo em vista que sua resposta espectral muitas vezes se confunde com demais classes, como áreas agropastoris e afloramentos rochosos. Ressalta-se que o uso de algoritmos computacionais que adotam a inteligência artificial como método são de grande relevância para estudos de ambientes complexos, tais como os wetlands. O uso destes recursos pelas geotecnologias otimiza o tempo de processamento de dados para investigações no âmbito das geociências.

AGRADECIMENTOS

Agradecemos a CAPES pelo apoio com a concessão de bolsas de estudos, e ao PROEX e POSGEO-UFF pelo financiamento aos trabalhos de campo da pesquisa.

REFERÊNCIAS BIBLIOGRÁFICA

BRASIL. Recomendação CNZU n.º 07 de 11 de junho de 2015. Brasília, DF. 2015. GOOGLE, Google Earth Engine. 2022. JENSEN, J. R. Sensoriamento Remoto do Ambiente: uma perspectiva em recursos terrestres. Tradução José Carlos Neves Epiphanyo (Cor.) et al. São José dos Campos, SP. 2ª Edição. 2009. 600p. MOORE, Peter D. Ecosystem: Wetlands. Revised Edition. New York, NY, 2008. 270p. OLOFSSON, P. et al. Good practices for estimating area and assessing accuracy of land change. Remote sensing of Environment, v. 148, p. 42-57, 2014. DOI: <https://doi.org/10.1016/j.rse.2014.02.015>. PHAN, T. N.; KUCH, V.; LEHNERT, L. W. Land Cover Classification using Google Earth Engine and Random Forest Classifier – The Role of Image Composition. Remote Sensing. 2020, 12, 2411; DOI:10.3390/rs12152411 SECRETARÍA DE LA CONVENCION DE RAMSAR. Manual de la Convención de Ramsar: Guía a la Convención sobre los Humedales (Ramsar, Irán, 1971), 6ª Ed. edición. Secretaría de la Convención de Ramsar, Gland (Suiza), 2007. 116p.