

ANÁLISE DE ATRIBUTOS ESPECTRAIS USANDO UMA CENA AVNIR-2 PARA O MAPEAMENTO DE USO E COBERTURA DO SOLO EM BACIA HIDROGRÁFICA

Analysis of spectral attributes using an AVNIR-2 scene for mapping land use and cover in a water basin

Admilson da Penha Pacheco

Doutor em Geofísica. Universidade de São Paulo (USP). Departamento de Engenharia Cartográfica e de Agrimensura (UFPE/Brasil).

admilson.pacheco@ufpe.br

Juarez Antônio da Silva Júnior

Engenheiro Cartógrafo e Agrimensor (UFPE). Departamento de Engenharia Cartográfica e de Agrimensura (UFPE/Brasil).

juarez.silvajunior@ufpe.br

Ubiratan Joaquim da Silva Junior

Mestre em Ciências Geodésicas e Tecnologias da Geoinformação (UFPE). Departamento de Engenharia Cartográfica e de Agrimensura (UFPE/Brasil).

ubiratan.joaquim@ufpe.br

Recebido: 17/04/2024

Aceito: 23/10/2024

Resumo

O mapeamento de uso e cobertura da terra é um elemento crucial na análise geoespacial para a gestão de bacias hidrográficas. A utilização de imagens de sensoriamento remoto, aliada a técnicas de classificação por aprendizado de máquinas, permite identificar e monitorar o ambiente e os componentes da paisagem. Neste estudo, foram desenvolvidos nove modelos de dados para a classificação do uso e cobertura da terra utilizando o algoritmo Random Forest em uma sub-bacia hidrográfica do Rio Terra Nova, no estado de Pernambuco. Para isso, foi utilizada uma imagem do sensor AVNIR-2 (Advanced Visible and Near-Infrared Radiometer-2) do satélite ALOS-2. Os modelos foram baseados nas bandas AVNIR-2, em quatro índices de vegetação, e no ranking de importância das variáveis utilizando o índice de Gini. A avaliação dos resultados foi realizada por meio da quantificação espacial e análise de acurácia. De forma geral, as variações nos recursos espectrais não resultaram em mudanças significativas na qualidade da classificação. As classes de Vegetação apresentaram acurácia do produtor superior a 85% em todos os modelos, enquanto as classes de Área Não Vegetada e Água mostraram uma queda nos valores de acurácia do produtor e do usuário à medida que os recursos espectrais diminuam. Este estudo é pioneiro na adaptação de dados espectrais do AVNIR-2 e em técnicas de aprendizado computacional para a aplicação de mapeamento de cobertura do solo de alta precisão, contribuindo ainda para o monitoramento ambiental em bacias hidrográficas que requerem informações espaciais precisas sobre o estado de conservação.

Palavras-chave: Random Forest, Sensoriamento Remoto, Índices de Vegetação, Bacia Hidrográfica.

Abstract

Land use and land cover mapping are crucial in geospatial analysis for watershed management. The use of remote sensing images, combined with machine learning classification techniques, allows the identification and monitoring of the environment and landscape components. This study developed nine data models for land use and land cover classification using the Random Forest algorithm in a sub-watershed of the Terra Nova River in Pernambuco. For this purpose, an image from the AVNIR-2 (Advanced Visible and Near-Infrared Radiometer-2) sensor of the ALOS-2 satellite was used. The models were based on the AVNIR-2 bands, four vegetation indices, and the ranking of importance of the variables using the Gini index. The results were evaluated through spatial quantification and accuracy analysis. In general, variations in spectral resources did not result in significant changes in the classification quality. Vegetation classes showed producer accuracy above 85% in all models, while Non-Vegetated Area and Water classes showed a decrease in producer and user accuracy values as spectral resources decreased. This study is a pioneer in adapting AVNIR-2 spectral data and machine learning techniques for applying high-precision land cover mapping, further contributing to environmental monitoring in watersheds that require accurate spatial information on conservation status.

Keywords: Random Forest, Remote Sensing, Vegetation Indices, Hydrographic Basin.

1. INTRODUÇÃO

O uso/cobertura da terra desempenha um papel importante nos processos ambientais, características climáticas e biodiversidade. Portanto, muitos estudos foram realizados sobre monitoramento e previsão da dinâmica da paisagem (Jalayer *et al.*, 2022). O uso/cobertura da terra é o resultado de uma combinação de forças naturais e artificiais, refletindo os atributos naturais da terra e o impacto das atividades humanas. O padrão de distribuição espacial e as mudanças dinâmicas no uso/cobertura da terra não afetam apenas o desenvolvimento econômico-social regional, mas também as mudanças ambientais e climáticas regionais (Wang *et al.*, 2022). Avaliar os impactos das mudanças no uso e cobertura do solo na hidrologia é a base para a gestão de bacias hidrográficas e restauração ecológica (Yulianto *et al.*, 2020). A avaliação geralmente inclui avaliação de padrões espaciais de consequências hidrológicas para diferentes mapas uso e cobertura do solo, comparação de valores basais de componentes hidrológicos simulados para mudanças uso e cobertura do solo na escala da bacia e exame de respostas temporais na descarga do canal com mudanças em uso e cobertura do solo, por exemplo, Nie *et al.*, (2011) e Shekar e Mathew (2023).

A dinâmica da mudança de uso e cobertura do solo pode ser monitorada usando dados de sensoriamento remoto. A disponibilidade de dados de sensoriamento remoto forneceu uma variedade de resolução temporal, espectral e espacial, que pode ser usada para detectar mudanças na superfície da Terra (Cunha *et al.*, 2020; Wang *et al.*, 2022).

Além disso, as utilizações de dados de sensoriamento remoto foram usadas como entrada na geração de um mapa de uso e cobertura da terra (Silva *et al.*, 2022; Yulianto *et al.*, 2020).

A classificação de imagens é a técnica mais significativa usada em sensoriamento remoto para criar mapas temáticos a partir de imagens de satélite (Duong *et al.*, 2018; Solanki *et al.*, 2022; Jalayer *et al.*, 2022). Vários métodos foram desenvolvidos para classificar os tipos uso e cobertura do solo. Por exemplo, Talukdar *et al.*, (2022) utilizaram seis algoritmos de aprendizado de máquina para mapeamento uso e cobertura do solo, ou seja, fuzzy ARTMAP, RF, ANN, distância de Mahalanobis, SAM e SVM, para avaliar a precisão de diferentes algoritmos e identificar o melhor classificador de uso e cobertura do solo de aprendizado de máquina. Din e Mak (2021) analisaram as mudanças no uso e cobertura do solo em Hyderabad durante os anos 1979-2020, e o algoritmo SVM foi aplicado para a classificação supervisionada de diferentes dados Landsat adquiridos. Em seu estudo, Liu *et al.*, (2020) usaram imagens Landsat e um classificador de RF baseado na plataforma GEE para mapeamento uso e cobertura do solo no Tibete. Os autores utilizaram, os índices espectrais (NDVI, NDBI e MNDWI) foram calculados para cada imagem e 11 fatores potenciais foram especificados para avaliar suas influências na dinâmica espacial na região. No Brasil, o IBGE (Instituto Brasileiro de estatística e geografia) realizou um mapeamento de uso e cobertura do solo para todo o país, utilizando imagens as do sensor AVNIR-2 do satélite ALOS-2 para a geração de uma base de referência formada por imagens de referência geométrica compatível com a exatidão esperada para os produtos do sistema cartográfico nacional detalhado em (IBGE, 2010).

O satélite japonês ALOS foi lançado em 24 de janeiro de 2006 pela JAXA (Agência Espacial Japonesa), visando a aquisição de imagens de todo o planeta para o monitoramento de desastres ambientais, levantamento de recursos naturais e suporte à cartografia (IBGE, 2008;). Depois de 9 meses de calibração, o satélite passou a ser considerado operacional em 20 de outubro de 2006, adquirindo imagens com período de revisita (ciclo) de 46 dias (Liu *et al.*, 2016; Wiederkehr *et al.*, 2020). O ALOS possui 3 sensores a bordo: PRISM (Panchromatic Remote-sensing Instrument for Stereo Mapping), AVNIR-2 (Advanced Visible and Near-Infrared Radiometer – Type 2) e PALSAR (Phased Array L-band Synthetic Aperture Radar). PRISM e AVNIR-2 são sensores ópticos e o PALSAR é um radar imageador (Parida *et al.*, 2020; Liu *et al.*, 2016; Wiederkehr *et al.*, 2020).

Este estudo tem como objetivo analisar a influência dos recursos de dados de sensoriamento remoto para o mapeamento e avaliação da qualidade da classificação de

uso e cobertura do solo numa subbacia no sertão pernambucano. Para isso, foi utilizado uma cena do sensor AVNIR-2 do satélite ALOS e os índices espectrais, sendo o SR (Simple Ratio), NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index) e EVI (Enhanced Vegetation Index) e o classificador Random Forest (RF). Foram gerados nove produtos de uso e cobertura do solo obtidos pela seleção de diferentes recursos espectrais de importância da variável do modelo RF, bem como a análise de acurácia da classificação para cada produto.

2. MATERIAIS E MÉTODOS

2.1. Dados e Área de estudo

Neste estudo foi utilizado uma cena ALOS AVNIR-2 com nível de pré-processamento nível 1B2G na data de 13 de novembro de 2010 cobrindo integralmente toda a área de estudo, disponibilizado pela plataforma ASF Data Search - Alaska Satellite Facility (<https://search.asf.alaska.edu>).

A cena AVNIR-2 cobriu integralmente a subbacia hidrográfica selecionada a partir da delimitação de Bacias Hidrográficas Ottocodificadas Nivel 6 obtida através do módulo de Inteligência Geográfica do Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) da Agência Nacional das Águas e Saneamento (ANA), localizada no semi-árido de Pernambuco com área de 269,36 km², inserida na bacia hidrográfica Rio Terra Nova (Figura 1).

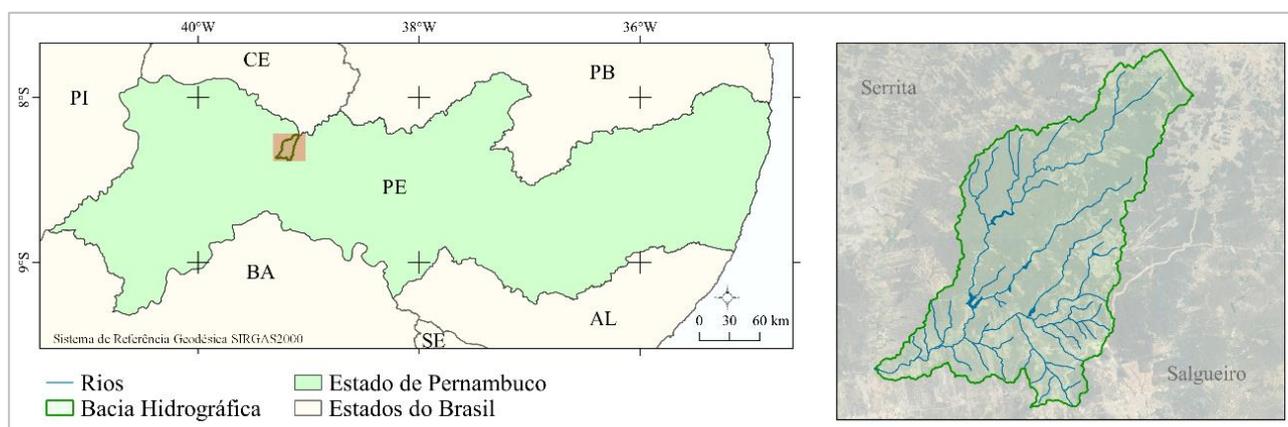


Figura 1 - Área de estudo.

Fonte: Autores, 2024.

A cena AVNIR-2 apresenta resolução espacial de 10 metros, em 4 bandas espectrais equivalentes ao Azul, Verde, Vermelho e Infravermelho Próximo. O Quadro 1 detalha as características do sensor. O sensor AVNIR-2 é capaz de operar em visado nadir ou visada

inclinada transversal à direção de órbita, com ângulo de inclinação que pode variar até +/- 44 graus, o que permite reduzir o tempo de revisita em áreas específicas. Tal flexibilidade é especialmente interessante para o imageamento emergencial de desastres naturais e monitoramento ambiental.

A imagem AVNIR-2 pode ser adquirida em 4 diferentes níveis de processamento: 1A, 1B1, 1B2R e 1B2G. As imagens nível 1B2 são calibradas radiométrica e geometricamente. No nível 1B2G, esta rotação já vem aplicada, ou seja, a imagem está georreferenciada e, conseqüentemente, seus pixels alinhados com a grade da projeção UTM (IBGE, 2009). A correção geométrica aplicada no nível 1B2G é efetuada utilizando os dados obtidos pelos instrumentos do Sistema de Controle da Órbita e Atitude da plataforma ALOS (AOC), o que significa que o usuário não necessita fornecer pontos de controle medidos no terreno para obter uma imagem georreferenciada.

2.2. Índices Espectrais

Após a obtenção bandas em níveis de refletância, quatro índices de vegetação diferentes foram calculados para esta pesquisa e comumente preferidos na literatura, tais o SR, NDVI, NDWI e EVI. As especificações dos índices são apresentadas na Tabela 1.

Tabela 1 – Especificações dos índices espectrais.

Nome	Formula	Referência
Índice de Vegetação Aprimorado (EVI)	$G * \frac{NIR - Red}{NIR + C1 * Red - C2 * Blue + L}$	Huete <i>et al.</i> , (2002)
Índice de Razão Simples (SR)	$\frac{NIR}{Red}$	Birth e Mcvey (1968)
Índice de Vegetação por Diferença Normalizada (NDVI)	$\frac{NIR - Red}{NIR + Red}$	Rouse <i>et al.</i> , (1973)
Índice de Diferença Normalizada da Água (NDWI)	$\frac{Green - NIR}{Green + NIR}$	Mcfeeters (1996).

Fonte: Autores, 2024

2.3. Seleção de amostras

A seleção de amostras para a classificação de imagem é uma etapa importante pois definem as classes consideradas na classificação e são os principais dados de entrada para o treinamento do modelo. Neste artigo foi adotado o método de seleção amostral sugerido por Körting *et al.* (2014), que descreve de coleta de amostras deve procurar dentro das classes características com um pequeno desvio padrão, calculando a homogeneidade

intraclasse. Dessa forma, A seleção de amostras foi realizada de forma aleatória e estratificada por meio de interpretação visual baseados em polígonos vetoriais categorizados para cada classe. Com a finalidade de reduzir inconstâncias espaciais e temáticas, foi utilizada um recorte do produto Mapbiomas com a identificação das classes de cobertura do solo em relação ao reconhecimento de padrões na imagem fusionada AVNIR-2 em cores naturais.

2.4. Análise do Coeficiente de Correlação de Pearson

A associação espacial entre os parâmetros espectrais para cada tipo de cobertura da terra foi avaliada através da correlação de Pearson (r) baseados nos dados de amostras de treinamento por pixel (Eq.1) (Al-Taei; Alesheikh; Boloorani, 2023).

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2(y_i - \bar{y})^2}} \quad (1)$$

Onde r é o coeficiente de correlação, x_i e y_i são as variáveis entre as quais a correlação deve ser encontrada \bar{x} e \bar{y} são os valores médios das variáveis da amostra. Este coeficiente varia de -1 a 1, onde $r = 1$ significa uma correlação perfeita positiva entre os dois recursos, $r = -1$ significa uma correlação negativa perfeita entre as duas variáveis - Isto é, se uma aumenta, a outra sempre diminui e $r = 0$ nenhuma correlação foi detectada entre os recursos de dados. Dessa forma, o par de recursos espectrais com correlação próxima ou igual 0 (zero), por exemplo, possui maior capacidade de identificação do alvo a ser classificado. Por sua vez, um par de recursos espectrais com correlação próxima ou igual a 1 (um), podem estar autocorrelacionados espacialmente e, portanto, pode não ser adequado no processo de classificação de uma determinada classe. Todos os resultados foram inferidos com nível de significância de 5%.

2.5. Classificador Random Forest e Score de Gini

O classificador de floresta aleatória consiste em uma combinação de classificadores de árvore onde cada classificador é gerado usando um vetor aleatório amostrado independentemente do vetor de entrada, e cada árvore lança um voto unitário para a classe mais popular para classificar um vetor de entrada (Breiman, 2001). O treinamento e classificação por RF foi aplicado através da combinação de produtos espectrais e as

amostras de treinamento, trazendo como resultado, um modelo raster categorizados com as classes de uso e ocupação do solo utilizando o software SNAP.

Para implementar o RF, dois parâmetros precisam ser configurados: o número de árvores (*ntree*) e o número de recursos em cada divisão (*mtry*). Este estudo seguiu a recomendação de Noi e Kappas (2017) que afirmaram que o modelo RF implementado para classificação da cobertura do solo alcança valores de acurácia estáveis quando *ntree* = 500. Com relação ao parâmetro *mtry*, foi utilizado o valor padrão $mtry = \sqrt{p}$, onde *p* é o número de variáveis preditoras (Noi; Kappas, 2017; Breiman, 2001).

Após o processamento do classificador RF é gerado a proporcionalidade importância de predição de cada variável no modelo gerado. O Score Gini calcula a importância de cada recurso como a soma do número de divisões (em todas as árvores) que incluem o recurso, proporcionalmente ao número de amostras que ele divide (Archer; Kimes, 2008). Matematicamente, o Índice de Gini pode ser expresso como:

$$Gini\ index = 1 - \sum_{i=1}^n (p_i)^2 \quad (2)$$

Onde p_i denota a probabilidade de um elemento ser classificado para uma classe distinta. A diminuição média no coeficiente de Gini é uma medida de como cada variável contribui para a homogeneidade dos nós e folhas na floresta aleatória resultante (Behnamian *et al.*, 2017). Quanto maior o valor do *score de Gini*, maior a importância da variável no modelo (Breiman, 2001).

Baseados nos resultados de importância, as análises foram definidas através 6 diferentes produtos de dados denominados (R1-R9), selecionadas com base no nível de importância de cada variável. Dessa forma, inicialmente, o produto P1 é composta por todos os recursos espectrais (Azul, Verde, Vermelho, Infravermelho Próximo, SR, NDVI, NDWI e EVI), os demais produtos serão selecionados através na análise de ranking no Score de Gini.

2.6. Validação do uso e cobertura da terra

Para avaliar a precisão dos conjuntos de dados de uso e cobertura do solo, contamos com a construção de uma matriz de erro. De acordo com a matriz de erro, foi calculado os parâmetros de qualidade sendo a Acurácia do produtor (AP) (eq.3), Acurácia do usuário (AU) (eq.4) (Islami *et al.*, 2022). As fórmulas são as seguintes:

$$AU = \frac{\text{Número de pixels corretamente classificados em cada categoria}}{\text{Número total de pixels de referência nessa categoria (o total da linha)}} \quad (3)$$

$$AP = \frac{\text{Número de pixels corretamente classificados em cada categoria}}{\text{Número total de pixels de referência nessa categoria (total da coluna)}} \quad (4)$$

A precisão do produtor refere-se a amostras que não foram classificadas corretamente como pertencentes a essa categoria e omitidas de sua categoria correta, e a precisão do usuário indica a probabilidade de que um pixel classificado na imagem realmente represente essa categoria (Congalton *et al.*, 2009; Wang *et al.*, 2022).

Dessa forma, os parâmetros de acurácia são obtidos pelos elementos da tabela de contingência gerados pela comparação do produto classificado e dado de referência. Uma cena UAV do Programa Pernambuco Digital (PE3D) foi adotada como referência recobrando integralmente a área de estudo. O PE3D é o mapeamento aerofotogramétrico de todo território do Estado de Pernambuco realizado entre 2010 e 2011, que disponibiliza ortofotos georreferenciados na escala 1:5.000, resolução espacial de 1m, atendendo ao mais alto nível de qualidade cartográfica. As imagens estão disponíveis gratuitamente em <http://www.pe3d.pe.gov.br/>.

3. RESULTADO E DISCUSSÃO

3.1. Correlação de Pearson por classe de uso e cobertura

A Figura 2 apresenta um correlograma exibindo os coeficientes de correlação de Pearson entre os recursos espectrais para cada classe de uso e cobertura do solo. Essa análise permite identificar, de forma preliminar, como as características de cada conjunto de dados influenciam sua capacidade de detecção em relação a uma classe específica.

Com uma média geral de correlação de -0,06, observada na Figura 2a, que representa a classe de vegetação arbórea, foram identificadas correlações significativas entre os índices EVI, NDVI, SAVI e a banda do NIR. Em alguns casos, essas correlações foram tão fortes que os valores chegaram a ser iguais a um. No entanto, em contraste, esses índices apresentaram correlações inversas com o NDWI e as bandas azul e vermelha. A banda verde, por sua vez, exibiu as correlações mais baixas com todos os índices espectrais. Vale destacar a relação entre a banda verde e o NIR, que registrou uma correlação de -0,02, sugerindo que essa banda é adequada para detectar vegetação densa. A maior correlação

negativa foi observada com o NDWI, com uma média de -0,267, enquanto o EVI apresentou o melhor desempenho, com uma média de correlação de -0,001.

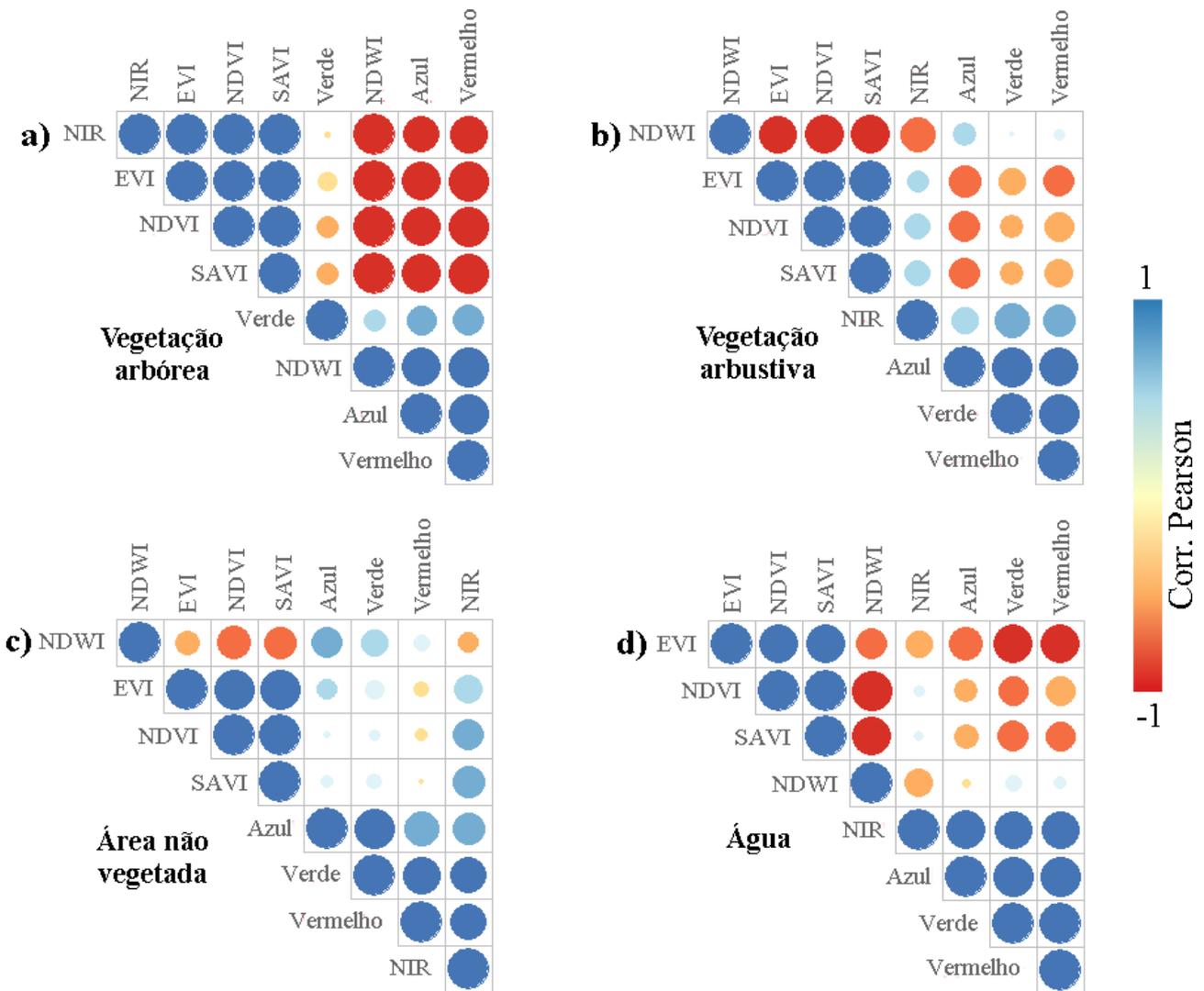


Figura 2 - Coeficientes de correlação de Pearson entre classes de uso e cobertura do solo em relação aos recursos espectrais. **Fonte:** Autores, 2024.

Já na classe de vegetação arbustiva Figura 1b, mostrou-se um comportamento diversificado em relação a classe de vegetação arbórea. O NDWI mostrou baixa correlação com as bandas do visível, principalmente as bandas do verde e do vermelho, enquanto para os índices espectrais foi visto uma correlação inversa significativa, com valores concentrados abaixo 0,8. No caso dos índices EVI, NDVI e SAVI, a correlação entre si foram altos, já na relação entre as bandas do visível, foram negativas e com o NIR foram positivas, porém moderadas. A bandas do visível entre si apresentaram valores próximos a 1. No geral a correlação média geral entre os atributos para a classe de vegetação arbustiva foi

de 0,05, com a menor correlação média para a banda do verde (-0,31) e o NDWI e o NDVI ambas com 0,98.

Na área não vegetada (Figura 2c), a banda do NIR apresentou correlações positivas com todos os recursos espectrais, exceto com o NDWI. Esse mesmo comportamento foi observado nas correlações entre os índices de vegetação. De maneira geral, predominam correlações médias positivas entre os recursos espectrais, exceto no caso do NDWI, que apresentou um valor de -0,1. A maior correlação média foi registrada para o NIR (0,5), enquanto a correlação média geral foi de 0,31. A Classe de Água, por sua vez, exibiu a menor correlação geral média, -0,0006, indicando que a detecção dessa classe pode ser eficaz com os recursos espectrais utilizados (Figura 2d). O NDWI apresentou a maior correlação média inversa, com -0,37, enquanto o NDVI mostrou a menor correlação média geral (-0,04). A menor correlação foi encontrada entre o NDWI e a banda azul (-0,04). Por outro lado, assim como nas outras classes, as bandas do visível demonstraram alta correlação entre si, além de uma correlação significativa com o NIR. O NDWI mostrou correlações negativas elevadas em relação aos índices de vegetação, o que pode estar associado à sua alta capacidade de detecção de corpos hídricos.

3.2. Análise de importância

A figura 3 mostra as variáveis mais importantes para a classificação RF de uso e cobertura do solo na bacia hidrográfica. Através desta lista, as novas classificações de uso e cobertura do solo por RF foram separadas baseado da supressão de cada variável em relação a ordem decrescente de importância, resultando em nove conjuntos de variáveis “rasters” (R1-R9) organizadas na tabela 2.

A Figura 3 ilustra a importância das variáveis utilizadas no modelo de classificação Random Forest (RF) para prever as classes de uso e cobertura do solo. De forma surpreendente, a banda do vermelho foi identificada como a variável mais importante na classificação, o que é um comportamento atípico, já que, normalmente, as bandas multiespectrais individuais apresentam baixa variabilidade entre as classes. A segunda variável mais relevante, embora com uma diferença significativa em relação à banda do vermelho, foi o índice NDWI (Índice de Diferença Normalizada da Água). Este índice, que é especializado na detecção de corpos d'água, destacou-se como a segunda variável mais importante no processo de classificação de uso e cobertura do solo.

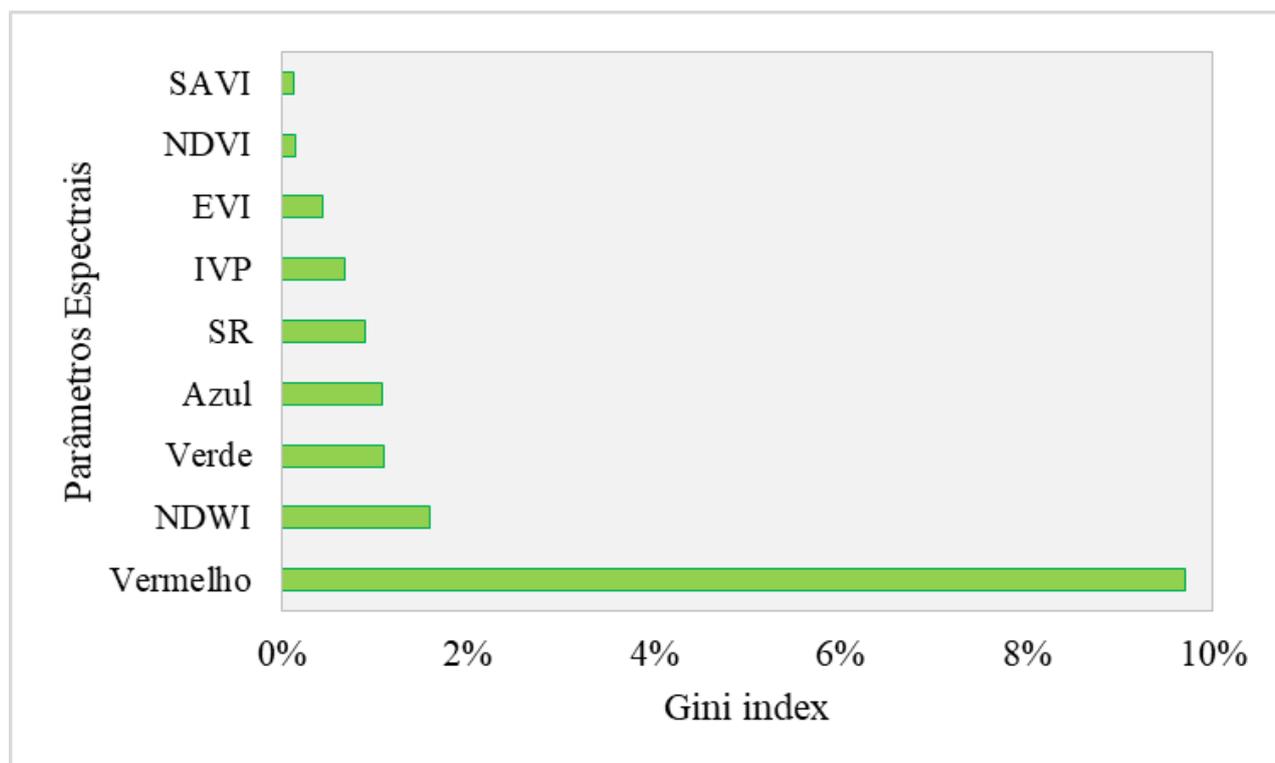


Figura 3 - Ranking de importância da variável com os respectivos valores de gini index implementado pela classificação RF de uso e cobertura do solo. O maior valor de gini index indica a variável de maior importância no modelo, enquanto baixos valores de gini index, conseqüentemente indica a variável de menor importância na classificação. **Fonte:** Autores, 2024.

Tabela 2 – Conjunto de dados (R-R9) baseado na ordem de importância da classificação de uso e cobertura por RF.

Raster	Conjunto de dados
R1	Vermelho~NDWI~Verde~Azul~SR~IVP~EVI~NDVI~SAVI
R2	NDWI~Verde~Azul~SR~IVP~EVI~NDVI~SAVI
R3	Verde~Azul~SR~IVP~EVI~NDVI~SAVI
R4	Azul~SR~IVP~EVI~NDVI~SAVI
R5	SR~IVP~EVI~NDVI~SAVI
R6	IVP~EVI~NDVI~SAVI
R7	EVI~NDVI~SAVI
R8	NDVI~SAVI
R9	SAVI

Fonte: Autores, 2024.

As bandas do verde, azul e o índice SR demonstraram ser as variáveis moderadamente importantes na detecção, com valores de *Tp-score* próximos (Figura 2). Surpreendentemente, os índices NDVI e SAVI, que são amplamente fundamentados na literatura como indicadores de vegetação, obtiveram as menores colocações no ranking. Isso indica que, neste contexto específico do modelo implementado, esses índices de

vegetação não foram tão eficazes na predição das classes de uso e cobertura do solo quando comparados com outras variáveis, como as bandas do espectro visível, infravermelho próximo e o índice EVI.

Essa análise sugere que o modelo desenvolvido pode ter particularidades e requerer uma abordagem diferenciada em relação ao uso de índices de vegetação tradicionais. Além disso, ressalta a importância de avaliar cuidadosamente as variáveis utilizadas em cada contexto específico de classificação, a fim de obter resultados mais precisos e relevantes para o objetivo do estudo de uso e cobertura do solo.

3.3. Análise espacial

Visualmente não há variações importantes sobre as classes de cobertura do solo na bacia, onde majoritariamente a vegetação arbórea concentra-se no setor norte e vegetação arbustiva nos setores centro-sul com pequenas feições de área não vegetada (Figura 4). Esse comportamento preliminar mostra que a variação de dados espectrais AVNIR-2 na classificação RF de forma visual, não interfere significativamente a distribuição espacial de cobertura do solo na bacia hidrográfica.

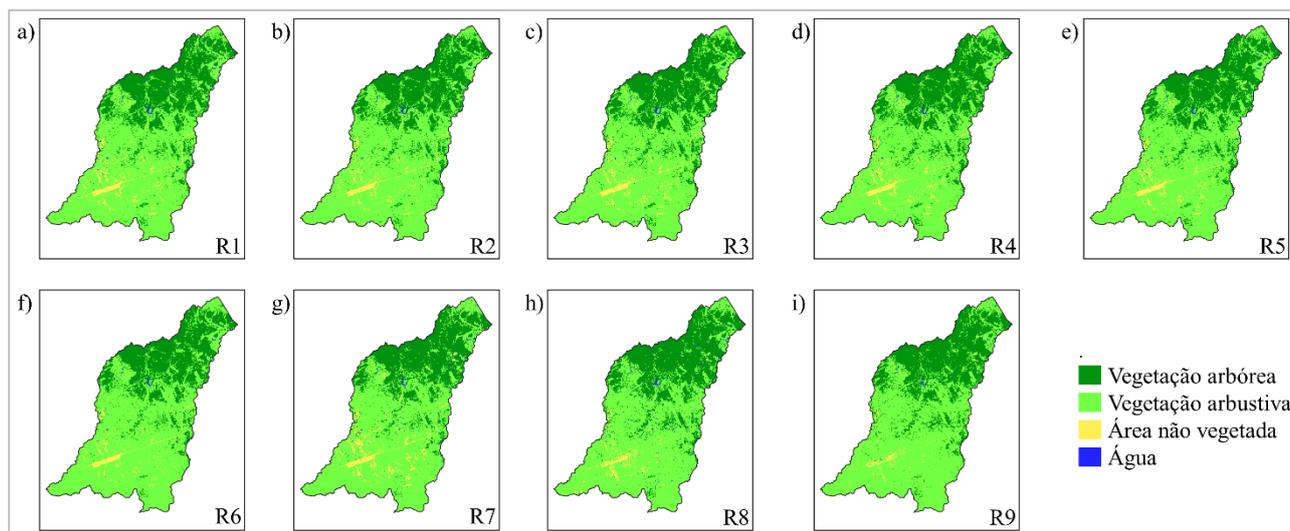


Figura 4 - Distribuição espacial das classes de uso e cobertura do solo baseado no conjunto de dados R1-R9. **Fonte:** Autores, 2024.

A classe de vegetação arbórea mostra-se estáveis em relação a mudanças das variáveis de entrada na classificação RF, com espacialização variando entre 79,18~88,75 Km², embora o mapa de referência registrou uma área de 66,84 Km² nesta classe. O mesmo comportamento foi visto para a classe de vegetação arbustiva, sendo a classe de maior predominância da bacia, correspondente aproximadamente 73% de toda a área da bacia

segundo o mapa de referência. Para as classificações, as espacializações variaram entre 182,21~193,69 Km², sendo a menor área registrada para o conjunto R8 e a menor R6. O conjunto R9, formado apenas pelo índice SAVI mapeou aproximadamente uma área de 5,04 Km² para a classe de Área Não Vegetada enquanto as outras variáveis encontraram uma área variando entre 10~13Km², exceto para o conjunto R6 e R8, ambos com área de aproximadamente 7,4Km².

Entre R1-R6 a classe de Água espacialmente ficou concentrada de 0,7~0,8 Km², correspondente a riachos efêmeros, e em conformidade com o mapa de referência que registrou uma área de 0,77 Km². O valor mais próximo com a referência foi visto para o R2 com área de 0,79 Km², conjunto formado pela variável mais importante (NDWI) após a remoção da banda do vermelho, destacando assim, o NDWI na detecção de corpos hídricos, apesar da baixa variação em relação aos demais conjuntos. Ainda para a classe de Água, com a remoção da banda do infravermelho próximo e o EVI do modelo, respectivamente atribuídos aos rasters R7 e R8, mostraram baixo desempenho com valores de área 2,07 Km² e 2,84 Km². Enquanto o R9 mostrou a menor estimativa de 0,5 Km².

Outros estudos abordaram propostas eficazes e um bom desempenho em análises de uso e cobertura do solo por meio de cenas AVNIR-2, por exemplo, Jiang *et al.* (2015) que obtiveram um mapa refinado de áreas úmidas costeiras com 15 categorias de cobertura do solo no estado de Washington, EUA e Solanki *et al.* (2022) revelaram uma devastação de aproximadamente 65% de área de mangue no período de 2009 a 2019 no oeste da Índia utilizando o NDVI e classificação de imagens. A Figura 4(a) e (b) resume as estimativas de (AP) e (AU) correspondente ao modelo de dados R1-R9.

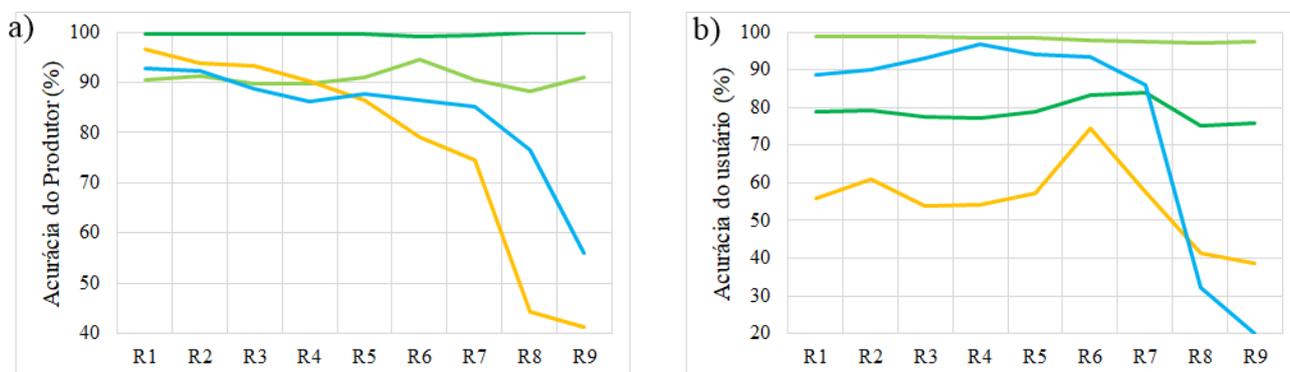


Figura 5 - Estimativa de Acurácia do Produtor e do usuário para o conjunto de dados R1-R9 diante das estimativas de classes de uso e cobertura da terra.

Fonte: Autores, 2024.

Na análise da acurácia do produtor, observa-se que as classes de vegetação apresentaram estimativas consistentemente acima de 87% em todos os conjuntos de dados, indicando subestimações mínimas e estáveis, mesmo com a variação na estrutura dos dados de classificação (Figura 5a). Em contraste, as classes de Água e, especialmente, de Área Não Vegetada, mostraram-se mais sensíveis ao modelo de classificação, com uma redução significativa da acurácia a partir dos conjuntos R6-R7, acompanhando a diminuição dos atributos espectrais. De forma resumida, a remoção do EVI e da banda do infravermelho próximo foi crucial para a queda na qualidade de detecção dessas classes, principalmente para a Área Não Vegetada e Água.

A classe de vegetação arbustiva na acurácia do usuário se mostrou estável com valores próximos a 100% para os conjuntos de dados. Já a classe de vegetação arbórea, apesar do decréscimo em relação a vegetação arbustiva, ainda mostrou bom desempenho com valores de acurácia concentrados acima de 70% (Figura 5b). Entre de R7-R9 houve uma abrupta diminuição na estiva de acurácia para a classe de água, partindo de valores próximo a 90% para aproximadamente 20%. Esses valores corroboram com a disparidade espacial encontrada nas estimativas de área na análise anterior, provavelmente relacionada a superestimações na classificação, ou seja, pixels pertencentes a outras classes pela referência, foram erroneamente classificadas como Água. Ainda na acurácia do usuário, a maior instabilidade foi vista na classe de Área Não Vegetada com um pico no conjunto R6 e um decréscimo de R7 a R9, estando na faixa de estimativas entre 80% e 37%, sendo a menor no conjunto R9 (Figura 4b).

Valores semelhantes para as classes de vegetação foram encontrados em Hashimoto *et al.* (2013), no mapeamento de uso e cobertura do solo para todo o Japão usando dados AVNIR-2, que obtiveram uma acurácia geral 65,4% enquanto a classes de Área não vegetada Água obtiveram estimativas 7,8 e 7,5% respectivamente, embora obtiveram uma precisão geral de 89,3%. No estudo de Liu *et al.* (2016) a combinação dos índices NDVI e NDWI para o mapeamento de uso e cobertura numa área de mineração no distrito de Pangzhuang, na China melhoraram a acurácia global passando de 86,45% para 89,55% quando comparado com a classificação usando apenas as bandas espectrais AVNIR-2. Enquanto Duong *et al.* (2018) obtiveram um precisão geral de 90% usando imagens AVNIR-2 e oito índices de vegetação para o mapeamento de uso e cobertura do solo no Vietnã Central de 2007 a 2017. Os autores ainda constataram que os mapas gerados pelos sensor AVNIR-2 podem fornecer aos formuladores de políticas e associações científicas

dados de entrada para uma discussão mais aprofundada sobre gestão ambiental, em particular balanço hídrico, estimativa de sedimentos e segurança alimentar.

4. CONSIDERAÇÕES FINAIS

A aplicação de modelos de Random Forest (RF) com base nas bandas AVNIR-2 e índices espectrais demonstrou ser uma metodologia eficaz na modelagem de dados de sensoriamento remoto, preenchendo uma lacuna em relação às atribuições espectrais no mapeamento de uso e cobertura do solo. Este estudo também destacou o potencial do sensor AVNIR-2, que superou os índices espectrais na detecção de quatro classes distintas de uso e cobertura do solo. Em conjunto com o índice NDWI, esses elementos proporcionaram valores estáveis de acurácia temática. No entanto, visualmente, a ausência desses fatores não apresentou variações significativas. As classes de Água e Área Não Vegetada foram as mais sensíveis à remoção de dados espectrais, com uma diminuição considerável nos valores de acurácia, enquanto as classes de vegetação mantiveram-se estáveis. Assim, este estudo demonstrou que o sensoriamento remoto integrado a métodos de Aprendizado de Máquina constitui uma ferramenta avançada para a análise e quantificação de atributos espaciais, além de evidenciar que as tecnologias tradicionais de mapeamento podem ser utilizadas de maneira inovadora.

REFERÊNCIAS

AL-TAEI, A. I.; ALESHEIKH, A. A.; BOLOORANI, A. D. Land Use/Land Cover Change Analysis Using Multi-Temporal Remote Sensing Data: a case study of tigris and euphrates rivers basin. **Land**, v. 12, n. 5, p. 1101, 2023.

BIRTH, G. S.; MCVEY, G. R. Measuring the Color of Growing Turf with a Reflectance Spectrophotometer 1. **Agronomy Journal**, v. 60, n. 6, p. 640-643, 1968.

CUNHA, E. R. Mapping LULC types in the Cerrado-Atlantic Forest ecotone region using a Landsat time series and object-based image approach: a case study of the prata river basin, mato grosso do sul, brazil. **Environmental Monitoring and Assessment**, v. 192, n. 2, p. 01-20, 2020.

DIN, S. U.; MAK, H. W. L. Retrieval of Land-Use/Land Cover Change (LUCC) Maps and Urban Expansion Dynamics of Hyderabad, Pakistan via Landsat Datasets and Support Vector Machine Framework. **Remote Sensing**, v. 13, n. 16, p. 3337, 2021.

DUONG, P. *et al.* JAXA High-Resolution Land Use/Land Cover Map for Central Vietnam in 2007 and 2017. **Remote Sensing**, v. 10, n. 9, p. 1406, 2018.

HASHIMOTO, S. *et al.* Land use and land cover inference in large areas using multi-temporal optical satellite images. 2013 **IEEE International Geoscience and Remote Sensing Symposium – Igarss**, p. 27-37, 2013.

HUETE, D. *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 195-213, 2002.

IBGE. **Avaliação Planimétrica de Imagens ALOS/AVNIR-2**. Estudo de caso: Uberlândia-MG. Relatório 2009/2010. Coordenação de Cartografia. Diretoria de Geociências. 2011.

ISLAMI, F. *et al.* Accuracy Assessment of Land Use Change Analysis Using Google Earth in Sadar Watershed Mojokerto Regency. Iop Conference Series: **Earth and Environmental Science**, v. 950, n. 1, p. 012091. 2022.

JALAYER, S. *et al.* Modeling and Predicting Land Use Land Cover Spatiotemporal Changes: a case study in chalus watershed, iran. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 15, p. 5496-5513, 2022.

JIANG, Y *et al.* Study of coastal wetland classification based on decision rules using ALOS AVNIR-2 images and ancillary geospatial data. **Geocarto International**, v. 30, n. 10, p. 1172-1188, 2015.

LIU, C. *et al.* Land Use/Land Cover Changes and Their Driving Factors in the Northeastern Tibetan Plateau Based on Geographical Detectors and Google Earth Engine: a case study in gannan prefecture. **Remote Sensing**, v. 12, n. 19, p. 3139, 2020.

LIU, P.; Using Optical and SAR Satellite Data to Classify Land Use and Land Cover (LULC) of Mining Area. **16TH International Congress for mine Surveying**, Brisbane, Australia, 12-16, 2016.

NIE, W. Assessing impacts of Landuse and Landcover changes on hydrology for the upper San Pedro watershed. **Journal of Hydrology**, v. 407, n. 1-4, p. 105-114, 2011.

PARIDA, B. R; MANDAL, S. P. Polarimetric decomposition methods for LULC mapping using ALOS L-band PolSAR data in Western parts of Mizoram, Northeast India. **Applied Sciences**, v. 2, n. 6, p. 100-115, 2020.

ROUSE, J. W. *et al.* Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation; **Final Report**; NASA/GSFC: Greenbelt, MD, USA, 1973; p. 1–137.

SOLANKI, J. B. *et al.* Assessment of mangrove cover dynamics and its health status in the Gulf of Khambhat, Western India, using high-resolution multi-temporal satellite data and Google Earth Engine. **Environmental Monitoring and Assessment**, v. 194, n. 12, p. 100-120, 2022.

SILVA, M. V. *et al.* Remote sensing techniques via Google Earth Engine for land degradation assessment in the Brazilian semiarid region, Brazil. **Journal Of South American Earth Sciences**, [S.L.], v. 120, p. 104061, 2022.

SHEKAR, P. R.; MATHEW, A. Detection of land use/land cover changes in a watershed: a case study of the murredu watershed in telangana state, india. **Watershed Ecology and The Environment**, v. 5, p. 46-55, 2023.

TALUKDAR, S. *et al.* Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. **Remote Sensing**, v. 12, n. 7, p. 1135, 2 abr. 2020.

YULIANTO, F. Spatial-Temporal Dynamics Land Use/Land Cover Change and Flood Hazard Mapping in the Upstream Citarum Watershed, West Java, Indonesia. **Quaestiones Geographicae**, v. 39, n. 1, p. 125-146, 2020.

WANG, H. *et al.* Consistency and Accuracy of Four High-Resolution LULC Datasets—Indochina Peninsula Case Study. **Land**, v. 11, n. 5, p. 758, 2022.

WIEDERKEHR, N. C. Discriminating Forest Successional Stages, Forest Degradation, and Land Use in Central Amazon Using ALOS/PALSAR-2 Full-Polarimetric Data. **Remote Sensing**, v. 12, n. 21, p. 3512, 2020.

Recebido: 17/04/2024

Aceito: 23/10/2024