



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE TECNOLOGIA E RECURSOS NATURAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
ENGENHARIA E GESTÃO DE RECURSOS NATURAIS



**RIVAILDO DA SILVA FILHO**

**CLASSIFICAÇÃO DA COBERTURA DO SOLO COM BASE EM IMAGENS RGB E  
MULTIESPECTRAIS OBTIDAS POR VANT: ESTUDO DE CASO PARA O BIOMA  
CAATINGA.**

Campina Grande - Paraíba (Brasil)

Fevereiro de 2024

RIVAILDO DA SILVA FILHO

**CLASSIFICAÇÃO DA COBERTURA DO SOLO COM BASE EM IMAGENS RGB E  
MULTIESPECTRAIS OBTIDAS POR VANT: ESTUDO DE CASO PARA O BIOMA  
CAATINGA.**

Tese de doutorado apresentada à Coordenação do Programa de Pós-Graduação em Engenharia e Gestão de Recursos Naturais da Universidade Federal de Campina Grande, para a obtenção do título de Doutor em Engenharia e Gestão de Recursos Naturais.

**Área de concentração:** Engenharia de Recursos Naturais

**Linha de pesquisa:** Recursos Hídricos e Saneamento Ambiental

**Orientador:** Prof. Dr. Vicente de Paulo Rodrigues da Silva

Campina Grande - Paraíba (Brasil)

Fevereiro de 2024

S586c

Silva Filho, Rivaildo da.

Classificação da cobertura do solo com base em imagens RGB e multiespectrais obtidas por VANT: estudo de caso para o bioma caatinga / Rivaildo da Silva Filho. – Campina Grande, 2024.

92 f. : il. color.

Tese (Doutorado em Engenharia e Gestão de Recursos Naturais.) – Universidade Federal de Campina Grande, Centro de Tecnologia e Recursos Naturais, 2024.

"Orientação: Prof. Dr. Vicente de Paulo Rodrigues da Silva".

Referências.

1. Sensoriamento Remoto. 2. Recursos Hídricos. 3. Classificador Random Forest (RF). 4. Drone. 5. Florestas Tropicais Sazonalmente Secas (FTSS). 6. Saneamento Ambiental. 7. Engenharia de Recursos Naturais. I. Silva, Vicente de Paulo Rodrigues da. II. Título.

CDU 528.8(043)

RIVAILDO DA SILVA FILHO

**CLASSIFICAÇÃO DA COBERTURA DO SOLO COM BASE EM IMAGENS RGB E MULTIESPECTRAIS OBTIDAS POR VANT: ESTUDO DE CASO PARA O BIOMA CAATINGA.**

tese aprovada em 27/02/2024

Banca Examinadora

---



Prof. Dr. Vicente de Paulo Rodrigues da Silva

Orientador

Documento assinado digitalmente



MADSON TAVARES SILVA

Data: 21/03/2024 21:39:06-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Madson Tavares Silva

Examinador Interno

Documento assinado digitalmente



VIVIANE FARIAS SILVA

Data: 22/03/2024 10:44:50-0300

Verifique em <https://validar.iti.gov.br>

---

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Viviane Farias Silva

Examinadora Interna

Documento assinado digitalmente



JOAO HUGO BARACUY DA CUNHA CAMPOS

Data: 16/03/2024 09:26:37-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. João Hugo Baracuy da Cunha Campos

Examinador Externo

Documento assinado digitalmente



EMERSON MARIANO DA SILVA

Data: 14/03/2024 14:05:52-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Emerson Mariano da Silva

Examinador Externo

Campina Grande - Paraíba (Brasil)

Fevereiro de 2024

*Dedico este trabalho aos meus pais, Rivaildo da Silva e Cleonice Sousa e Silva  
pelo carinho e apoio em todos os sonhos e desafios da minha vida.*

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, pelo dom da vida. A minha família, em especial aos meus pais, Rivaildo e Cleonice, minha irmã Nelmara, minha companheira de vida Ednara e aos demais familiares próximos por todo apoio, incentivo e força que me mantiveram neste sonho. Ao meu primo João Costa por toda ajuda prestada, ao meu orientador, Prof.º Dr. Vicente de Paulo Rodrigues da Silva, pela sua orientação, confiança e paciência para realização deste trabalho. Aos membros da banca examinadora, pela dedicação de seu tempo na contribuição para a versão final desta tese. A todos os professores, funcionários e amigos do Curso de Engenharia Civil da UFCG, do PPGECA e do PPGEGRN, aos quais tenho enorme consideração. Aos meus amigos, que contribuíram direta ou indiretamente para a concretização desse sonho, obrigado.

## RESUMO

A Caatinga Brasileira é um bioma que corresponde à uma das maiores Florestas Tropicais Sazonalmente Secas (FTSS) do mundo e o entendimento da cobertura do solo nesta área contribui para estudos voltados a preservação ambiental e a mitigação de impactos ambientais. Um dos desafios na classificação da cobertura do solo nessa área é a distinção entre os diferentes padrões vegetativos. Nesse contexto, surgem os Veículos Aéreos Não Tripulados (VANT's) capazes de obter dados com alta resolução espacial e superar esta limitação. Esta pesquisa avalia a precisão da classificação da cobertura do solo através de dados obtidos a partir de VANT com câmera multiespectral embarcada. Duas abordagens utilizando o classificador Random Forest (RF) foram aplicadas, a primeira baseada nas bandas Red, Green, Nir e no NDVI obtidos da câmera multiespectral e a outra baseada nas bandas Red, Green, Blue e no NGRDI obtidos da câmera RGB nativa do VANT em voos realizados à 120 metros de altura com condição de céu claro e sol centrado, na Bacia do Riacho do Frango em Patos-PB. O período de coleta foi de 22/09/2023 à 22/10/2023, totalmente compreendido na estação seca. Os resultados das classificações foram avaliados com base em amostras de validação do modelo e indicam que a primeira abordagem apresentou o melhor desempenho com Acurácia Global média de 95,33% e Coeficiente kappa médio de 0,93. A variável mais importante, segundo índice MDG, foi o NDVI, o que mostra a importância do uso de índices de vegetação para realçar as características vegetativas. A partir dos resultados, conclui-se que o uso de VANT com a câmera Mapir Survey 3w embarcada é uma ferramenta eficiente para classificação da cobertura do solo na Caatinga, com alta capacidade de distinção dos padrões vegetativos, enquanto que a classificação para área em estudo feita com base apenas em dados RGB não fornece uma boa precisão.

**Palavras-chave:** Sensoriamento remoto; Random Forest; Drone; Semiárido; FTSS.

## ABSTRACT

The Brazilian Caatinga is a biome that corresponds to one of the largest Seasonally Dry Tropical Forests (SDTF) in the world and understanding the land cover in this area contributes to studies aimed at environmental preservation and mitigation of environmental impacts. One of the challenges in classifying land cover in this area is the distinction between different vegetative patterns. In this context, Unmanned Aerial Vehicles (UAVs) emerge, capable of obtaining data with high spatial resolution and overcoming this limitation. This research evaluates the accuracy of land cover classification using data obtained from a UAV with an onboard multispectral camera. Two approaches using the Random Forest (RF) classifier were applied, the first based on the Red, Green, Nir and NDVI bands obtained from the multispectral camera and the other based on the Red, Green, Blue bands and NGRDI obtained from the RGB camera's native UAV in flights carried out at 120 meters high with clear sky conditions and centered sun, in the Riacho do Frango Basin in Patos-PB. The collection period was from 09/22/2023 to 10/22/2023, fully included in the dry season. The classification results were evaluated based on model validation samples and indicate that the first approach presented the best performance with Global Accuracy of 95.33% and Kappa Coefficient of 93.33%. The most important variable, according to the MDG index, was the NDVI, which shows the importance of using vegetation indices to highlight vegetative characteristics. From the results, it is concluded that the use of UAV with the onboard Mapir Survey 3w camera is an efficient tool for classifying land cover in the Caatinga, with a high capacity for distinguishing vegetative patterns, while the classification for the area under study made based on RGB data alone does not provide good accuracy.

**Keywords:** Remote Sensing; Random Forest; Drone; Semiarid; SDTFs

## LISTA DE FIGURAS

Figura 1: Localização da área de estudo: 1, 2 e 3 referem-se as áreas selecionadas.....	27
Figura 2: Fluxograma representativo das etapas metodológicas para classificação da cobertura do solo na Caatinga utilizando VANT e classificador RF.....	28
Figura 3: Equipamentos utilizados na pesquisa: Controle RC Mavic Pro (A), Drone DJI Mavic Pro com câmera e GPS acoplados (B) e Câmera multiespectral Mapir Survey 3W (C).....	29
Figura 4: Grupos e classes para a vegetação Caatinga.....	32
Figura 5: Classes de cobertura identificadas na área de estudo: Aurbustiva arbórea densa (A), Subarbustiva muito rala (B) e Solo exposto (C).....	32
Figura 6: Classes de cobertura de solo para a área 1: método 1 (A) e método 2 (B).....	37
Figura 7: Distribuição percentual das classes de cobertura para a área 1: método 1 (A) e método 2 (B).....	39
Figura 8: Importância das variáveis preditoras por <i>MDG</i> na área 1: método 1 (A) e método 2 (B).....	40
Figura 9: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 1: método 1 (A) e método 2 (B).....	41
Figura 10: Comportamento dos IV's para os dados de treinamento em cada classe na área 1.....	42
Figura 11: Composições para a área 1: RNG (A) e RGB (B).....	43
Figura 12: Classes de cobertura de solo para a área 2: método 1 (A) e método 2 (B).....	45
Figura 13: Distribuição percentual das classes de cobertura para a área 2: método 1 (A) e método 2 (B).....	47
Figura 14: Importância das variáveis preditoras por <i>MDG</i> na área 2: método 1 (A) e método 2 (B).....	48
Figura 15: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 2: método 1 (A) e método 2 (B).....	49
Figura 16: Comportamento dos IV's para os dados de treinamento em cada classe na área 2.....	50
Figura 17: Composições para a área 2: RNG (A) e RGB (B).....	51
Figura 18: Classes de cobertura de solo para a área 3: método 1 (A) e método 2 (B).....	53
Figura 19: Distribuição percentual das classes de cobertura para a área 3: método 1 (A) e método 2 (B).....	55

Figura 20: Importância das variáveis preditoras por <i>MDG</i> na área 3: método 1 (A) e método 2 (B).....	56
Figura 21: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 1: método 1 (A) e método 2 (B).....	57
Figura 22:Comportamento dos IV's para os dados de treinamento em cada classe na área 3.....	58
Figura 23: Composições para a área 3: RNG (A) e RGB (B).....	59
Figura 24: Índices de precisão da classificação: Acurácia Geral – AG (A) e Coeficiente kappa – k (B).....	61

## LISTA DE TABELAS

Tabela 1: Matriz de confusão genérica.....	29
Tabela 2: Parâmetros do planejamento de voo no <i>Drone Deploy</i> .....	30
Tabela 3: Configurações para câmera Mapir Survey 3W (RGN).....	35
Tabela 4: Matrizes de confusão para a área 1.....	39
Tabela 5: Matrizes de confusão para a área 2.....	47
Tabela 6: Matrizes de confusão para a área 3.....	55
Tabela B.1 – Amostras de treinamento na área 1 pelo método 1.....	80
Tabela B.2 – Amostras de validação na área 1 pelo método 1.....	81
Tabela B.3 – Amostras de treinamento na área 1 pelo método 2.....	82
Tabela B.4 – Amostras de validação na área 1 pelo método 2.....	83
Tabela B.5 – Amostras de treinamento na área 2 pelo método 1.....	84
Tabela B.6 – Amostras de validação na área 2 pelo método 1.....	85
Tabela B.7 – Amostras de treinamento na área 2 pelo método 2.....	86
Tabela B.8 – Amostras de validação na área 2 pelo método 2.....	87
Tabela B.9 – Amostras de treinamento na área 3 pelo método 1.....	88
Tabela B.10 – Amostras de validação na área 3 pelo método 1.....	89
Tabela B.11 – Amostras de treinamento na área 3 pelo método 2.....	90
Tabela B.12 – Amostras de validação na área 3 pelo método 2.....	92

## LISTA DE SIGLAS

AESA	-	Agência Executiva de Gestão das Águas do Estado da Paraíba
AG	-	Acurácia geral
B	-	<i>Blue</i> (azul)
EMBRAPA	-	Empresa Brasileira de Pesquisa Agropecuária
FTSS	-	Floresta Tropical Sazonal Seca
G	-	<i>Green</i> (verde)
IBGE	-	Instituto Brasileiro de Geografia e Estatística
ISO	-	<i>International Organization for Standardization</i>
IV's	-	Índices de vegetação
k	-	Coefficiente Kappa
Km	-	Kilômetro
LANDSAT	-	<i>Land Remote Sensing Satellite</i>
m	-	Metro
MDG	-	<i>Mean Decrease Gini</i>
MODIS	-	<i>Moderate-Resolution Imaging Spectroradiometer</i>
NDVI	-	<i>Normalized Difference Vegetation Index</i>
NGRDI	-	<i>Normalized Green-Red Difference Index</i>
Nir	-	<i>Near Infrared</i> (Infravermelho próximo)
PP	-	Precisão do produtor
PU	-	Precisão do usuário
R	-	<i>Red</i> (vermelho)
REM	-	Radiação Eletromagnética
RF	-	<i>Random Forest</i>
RGB	-	<i>Red-Green-Blue</i>
RNG	-	<i>Red-Nir-Green</i>
SR	-	Sensoriamento Remoto
VANT	-	Veículo Aéreo Não Tripulado

## SUMÁRIO

1.	INTRODUÇÃO .....	13
2.	OBJETIVOS .....	15
2.1.	Objetivo geral .....	15
2.2.	Objetivos específicos .....	15
3.	FUNDAMENTAÇÃO TEÓRICA.....	15
3.1.	Bioma Caatinga.....	15
3.2.	Sensoriamento remoto (SR).....	17
3.2.1.	<i>Sensoriamento remoto da vegetação</i> .....	18
3.2.2.	<i>Índices de vegetação (IV's)</i> .....	20
3.2.3.	<i>Veículo Aéreo Não Tripulado (VANT)</i> .....	23
3.3.	Classificação da cobertura do solo com VANT's e Random Forest .....	24
4.	MATERIAL E MÉTODOS .....	26
4.1.	Área de estudo.....	26
4.2.	Etapas metodológicas.....	28
4.2.1.	<i>Aquisição de dados</i> .....	28
4.2.2.	<i>Processamento de dados</i> .....	30
4.2.3.	<i>Definição das classes de cobertura e coleta de amostras</i> .....	31
4.2.4.	<i>Classificação da cobertura do solo</i> .....	33
4.2.5.	<i>Importância das variáveis e precisão da classificação</i> .....	33
5.	RESULTADOS E DISCUSSÕES .....	36
5.1.	Classificação para a área 1 .....	36
5.2.	Classificação para a área 2.....	45
5.3.	Classificação para a área 3.....	53
5.4.	Análise comparativa do desempenho dos métodos avaliados .....	61
6.	CONCLUSÕES .....	64
7.	REFERÊNCIAS BIBLIOGRÁFICAS .....	65
	ANEXO A – <i>SCRIPT PARA MODELO RANDOM FOREST</i> .....	74
	ANEXO B – DADOS DE TREINAMENTO E VALIDAÇÃO .....	80

## 1. INTRODUÇÃO

Segundo Jensen (2009), a cobertura do solo está relacionada aos materiais biofísicos que estão distribuídos nas áreas naturais. Muitos estudos em todo o mundo têm realizado o monitoramento da cobertura do solo através do sensoriamento remoto (Schulz *et al.*, 2017; Lu *et al.*, 2016; Beuchle *et al.*, 2015; Mendoza *et al.*, 2011). Estudos mais recentes sobre o comportamento da vegetação, com foco na análise de grandes áreas, têm utilizado índices de vegetação (IV), adquiridos por sensores de resolução espacial moderada acoplados à satélites, como MODIS e o LANDSAT, é o caso dos trabalhos na China (Li *et al.*, 2017; Lu *et al.*, 2016) e na África (Di Vittorio & Georgakakos, 2018). Contudo, em áreas pequenas e com heterogeneidade marcante, como na Caatinga, utilizar dados de sensores com resolução moderada podem resultar em interpretações incorretas sobre a dinâmica da vegetação (Liu *et al.*, 2017). Nesse sentido, segundo Boschetti (2015), imagens obtidas a partir de sensores de resolução espacial mais alta, são ideais por fornecerem informações espaciais muito mais detalhadas.

Além disso, a classificação da cobertura do solo comumente usa dados multiespectrais mono-temporais, ou seja, com apenas uma data de captura (Jia *et al.*, 2014; Mahdianpari *et al.*, 2018; Alhassan *et al.*, 2019). Esta abordagem pode induzir confusão na classificação dos diferentes padrões de cobertura do solo existentes nas Florestas Tropicais Sazonais Secas (FTSS) se realizada a partir de imagens de satélites com resolução espacial baixa, devido à similaridade da resposta espectral da vegetação em estágios fenológicos específicos, principalmente logo após o período de chuvas (Santos *et al.*, 2021; Karnieli, 2002).

A utilização de Veículos Aéreos Não Tripulados (VANT's) é uma alternativa para superar essa limitação, uma vez que é possível a obtenção de dados multiespectrais em alta resolução espacial e temporal, permitindo captar aspectos importantes da vegetação. O uso dos VANTs é cada vez mais presente em monitoramentos ambientais e em estudos agrícolas (Sampaio *et al.*, 2020; Gomes *et al.*, 2021). Aplicação de VANTs nesses estudos proporciona melhoria e flexibilidade dos dados obtidos, uma vez que apenas alterando a altura de voo é possível obter dados com resoluções espaciais diferentes, que comumente atingem a ordem dos centímetros. Quando aplicadas à vegetação, as imagens multiespectrais possibilitam estudos detalhados dos padrões espectrais dos alvos (assinaturas espectrais), os quais estão altamente associados a diversas propriedades fisiológicas das diferentes espécies de plantas, assim como fatores ambientais, fenológicos e antropogênicos que possuem influência sobre elas (Gomes *et*

*al.*, 2021).

A partir do processamento das imagens obtidas com sensores multiespectrais acoplados em VANT, é possível a obtenção dos IV's. Segundo Shiroma *et al.* (2020), é comum o uso desses índices espectrais em estudos voltados para a análise da cobertura vegetal. Os índices de vegetação são sensíveis as diferentes características da vegetação - clorofila, nutrientes, pigmentos, biomassa e teor de água nas folhas, dentre outros (Giovos *et al.*, 2021). Um dos índices mais utilizados para esta finalidade é o NDVI, ele tem sido aplicado em imagens obtidas em VANTs para diversos estudos ambientais, a exemplo de Lima *júnior et al.* (2014), que propõem o uso do índice para estimativa da biomassa na Caatinga analisando a sazonalidade no bioma e Bezerra *et al.* (2019) para avaliação da cobertura vegetal em região semiárida, dentre outros.

De acordo com Cardoso *et al.* (2021), a vegetação do Bioma Caatinga possui características muito peculiares, sendo uma região que tem forte influência dos regimes pluviométricos, de tal forma que o início da produção de biomassa na vegetação ocorre de forma muito rápida, logo após as primeiras chuvas. A Caatinga é um bioma representativo do semiárido brasileiro, representando um patrimônio biológico para a região semiárida, sendo o quarto maior bioma do Brasil, considerado também uma das mais extensas regiões secas do mundo. Apresenta características peculiares como altas temperaturas e grandes variações pluviométricas que repercutem em grandes mudanças no aspecto da vegetação e em suas fases fenológicas (Giongo *et al.*, 2011). Portanto, compreender a distribuição espacial das diferentes classes vegetativas naturais desse bioma, auxilia na compreensão do mesmo e na aplicação de políticas de preservação, monitoramento e gestão deste bioma.

É importante ressaltar ainda, que a vegetação do Bioma Caatinga está inserida num processo de desertificação de regiões semiáridas como as encontradas no sertão paraibano. Estudos como o de Alves *et al.* (2017) indicam que as causas relacionadas ao avanço deste processo estão fortemente relacionadas com as ações antrópicas, ficando em segundo plano à influência das mudanças climáticas. Portanto, entender a dinâmica da cobertura do solo nessa região ajuda no controle da supressão da vegetação nativa provocada pelo avanço da urbanização que ocorre em diversos municípios da Paraíba.

## **2. OBJETIVOS**

### **2.1. Objetivo geral**

Avaliar a classificação da cobertura do solo em área de Caatinga no sertão paraibano sob diferentes metodologias.

### **2.2. Objetivos específicos**

- Classificar a cobertura do solo em área de Caatinga do sertão paraibano;
- Analisar o desempenho de metodologias de classificação e seus efeitos;
- Identificar as variáveis preditoras com maior importância na classificação da área de estudo através de cada método;
- Definir o método com melhor desempenho entre aqueles avaliados.

## **3. FUNDAMENTAÇÃO TEÓRICA**

### **3.1. Bioma Caatinga**

O Brasil possui uma imensa extensão territorial, aproximadamente 8.500.000 km<sup>2</sup>, com diversidade de clima, solos e vegetação. Em toda a extensão do país existem porções com características semelhantes e biodiversidade contínua, que define os chamados “biomas”. Um bioma é entendido como um conjunto de todos os seres vivos em uma região que apresenta uniformidade e uma história comum em sua formação. Existem seis biomas bem definidos no Brasil: Amazônia (49,29%), Cerrado (23,92%), Mata Atlântica (13,04%), Caatinga (9,92%), Pampa (2,07%) e Pantanal (1,76%) (IBGE, 2023).

A Caatinga é um bioma exclusivamente brasileiro. Etimologicamente, o termo Caatinga (Tupi-Guarani) significa floresta branca, devido sua paisagem branqueada, caracterizada pelo aparecimento dos troncos de árvores na estação seca, quando a vegetação perde quase que totalmente sua folhagem. Quase 10% do território nacional é ocupado pela Caatinga, que abrange parte dos estados de Alagoas, Bahia, Ceará, Maranhão, Minas Gerais, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe (IBGE, 2023).

A Caatinga surgiu com a extinção da floresta tropical anteriormente existente na

região, ocorrida no final da era glacial, há dez mil anos. Além disso, a Caatinga é heterogênea, sendo muito rica em biodiversidade e espécies endêmicas. A vegetação da caatinga é classificada como savana de estepe e a paisagem é bastante diversificada devido à variação acentuada de precipitação, fertilidade, tipos de solo e topografia (SILVA *et al.*, 2018).

A particularidade da cobertura vegetal da Caatinga permite a classificação da paisagem em dois tipos distintos: agreste (serras agrestes) e sertão (sertão). O primeiro envolve a transição entre a área seca e a Mata Atlântica, que é dominada por árvores arbustivas, com muitos ramos entrelaçados; e o segundo (sertão / sertão) possui vegetação mais rústica (SILVA & CRUZ, 2018).

Na Caatinga, as temperaturas médias anuais são altas (25°-29°C), com alta radiação solar. O clima é semiárido, principalmente decorrente do predomínio de massas de ar estáveis, impulsionadas pelos ventos alísios do Sudeste, que têm sua origem na ação anticiclone do Atlântico Sul, conseqüentemente, há uma baixa taxa de umidade relativa e um potencial de evapotranspiração muito alto. O solo que é derivado de diferentes tipos de rocha ainda é raso e pedregoso. As secas são cíclicas e prolongadas. A estação chuvosa é curta, começando no início do ano e com chuvas mais baixas e irregulares. A ocorrência de fenômenos catastróficos, como secas e inundações, é muito comum (SILVA *et al.*, 2018).

A agricultura é a principal atividade econômica realizada na Caatinga. Alguns projetos de irrigação para agricultura são desenvolvidos nos vales do São Francisco e no Rio Parnaíba, que são os principais rios da região (SILVA *et al.*, 2016). A região tem problemas sociais, como baixos níveis de renda e escolaridade e alta taxa de mortalidade infantil.

No semiárido, a vegetação é composta prioritariamente por arbustos que perdem as folhas nos meses mais secos ou por pastagens que secam na época de estiagem, essas são características da Caatinga presente na região Nordeste do Brasil. A vegetação da Caatinga é muito sensível às variações climáticas, principalmente com relação as alterações de disponibilidade hídrica, o que dificulta as análises de padrões, por necessitar de um período longo de dados, o qual possa representar fielmente o comportamento da Caatinga às mudanças climáticas e às ações antrópicas (ANDRADE *et al.*, 2021).

No Bioma Caatinga, as ações antrópicas ao longo dos anos vêm modificando a paisagem, intensificando o processo de substituição das áreas naturais por diversos tipos de uso do solo e reduzindo áreas florestadas. Apesar de sua grande riqueza em biodiversidade, a

Caatinga é o bioma brasileiro mais negligenciado. O uso indiscriminado do solo e dos recursos naturais, durante o processo de ocupação e exploração, degradou a Caatinga. A singularidade, o uso insustentável, a riqueza particular e a pequena extensão expressam a necessidade de estudos que busquem preservar o bioma (MATSUSHITA *et al.*, 2006; LEAL *et al.*, 2005).

Diversos fatores estão diretamente ligados a este processo, uma dessas causas é a exploração intensa de áreas com agricultura familiar e pecuária, muito comum na região semiárida do nordeste brasileiro. A maior exposição do solo faz com que os processos erosivos se intensifiquem, ocorrendo também perda da biodiversidade. O monitoramento do uso e da cobertura do solo (LULC, sigla em inglês), através de informações espaço-temporais detalhadas das modificações ocorridas na paisagem, é a principal ferramenta para o conhecimento das capacidades e limitações de uma região. (COELHO *et al.*, 2014; VANZELA *et al.*, 2010).

### **3.2. Sensoriamento remoto (SR)**

O sensoriamento remoto pode ser definido, segundo Novo (2008), como sendo a utilização simultânea de sensores, equipamentos de transmissão de dados a bordo de aeronaves, espaçonaves, ou outras plataformas, cujo objetivo é estudar eventos, fenômenos e processos que ocorrem na superfície da Terra a partir do registro e da análise das interações entre a radiação eletromagnética e as substâncias que compõem a superfície terrestre em suas mais diversas manifestações. Desta forma, as técnicas de sensoriamento remoto têm como base quatro elementos: radiância eletromagnética (REM), Fonte e alvo. A fonte varia de acordo com a aplicação. O Sol é a principal fonte para sensores ópticos, já para sensores termais e de microondas passivos, a Terra é a principal fonte. O alvo é o elemento que se pretende extrair informação e o sensor é o instrumento que coleta a REM refletida ou emitida pelo alvo (Angelis, 2011).

As técnicas de SR são utilizadas com o intuito de identificar determinadas características dos alvos em análise sem a necessidade de contato direto com os mesmos. É assim que os sensores acoplados nos satélites que orbitam o planeta Terra captam informações espectrais da superfície, registrando os níveis de radiação eletromagnética refletida e/ou emitida pelos alvos. Para que se possa compreender melhor como se viabiliza a aplicação dessas técnicas no estudo dos recursos naturais, é necessário o entendimento de pelo menos quatro parâmetros radiométricos. O primeiro deles, refere-se à irradiância. Em termos simples,

a irradiância representa a intensidade do fluxo radiante que atinge uma determinada superfície e que é proveniente de todas as direções (Jensen, 2009).

### ***3.2.1. Sensoriamento remoto da vegetação***

As plantas são fundamentalmente dependentes do seu ambiente físico para crescimento, sobrevivência e reprodução. Para entender seu comportamento, são imprescindíveis a utilização de ferramentas tanto para quantificação do meio ambiente quanto para o estudo das plantas e seu funcionamento. De fato, nos últimos anos, houve um reconhecimento cada vez maior de que para a melhoria futura nos rendimentos das culturas e a melhoria do monitoramento e manejo dos ecossistemas naturais, por exemplo em resposta às mudanças climáticas, necessita-se de uma boa compreensão dos mecanismos subjacentes à resposta da planta. Da mesma forma para o ambiente, a previsão das mudanças climáticas e dos impactos por elas causados, são criticamente dependentes do entendimento do papel da vegetação no controle de mudanças no balanço energético terrestre (Jones & Vaughan, 2010).

As técnicas de SR, no sentido da coleta de informações por um dispositivo separado do alvo, estão tornando-se cada vez mais componentes importantes entre as opções de tecnologias disponíveis para o estudo dos sistemas de vegetação e seu funcionamento. Isto é, apesar do fato de que muitas aplicações apenas fornecem estimativas indiretas das variáveis biofísicas de interesse. Além de não necessitar de contato direto com o alvo, outra vantagem é a facilidade com que as observações são facilmente extrapoladas para as escalas maiores que são de particular interesse em estudos ecológicos e climáticos. Mesmo na escala da planta, as imagens obtidas por sensores remotos podem permitir a amostragem rápida de um grande número de plantas conforme necessário, por exemplo, no cultivo de plantas e na identificação de variedades com características fisiológicas específicas. (Jones & Vaughan, 2010).

Estudos da cobertura vegetal podem envolver quatro diferentes níveis de coleta de dados: laboratório, campo, aeronave e orbital. Em laboratório, comumente são consideradas as folhas, partes de plantas ou até alguns arranjos de plantas, dos quais são coletados dados radiométricos com o objetivo de caracterizar espectralmente aspectos relacionados ao processo de interação entre a REM e a vegetação. Em campo, os dados podem ser coletados diretamente das folhas ou através de dispositivos como plataformas (móveis ou fixas), teleféricos, etc; que permitem a inserção dos sensores imediatamente acima dos dosséis vegetais segundo as mais diferentes disposições, como no caso dos VANT's.

A principal motivação dos estudos em vegetação envolvendo a aplicação das técnicas

de SR, fundamenta-se na compreensão da “aparência” que uma determinada cobertura vegetal assume em um produto de sensoriamento remoto, a qual resulta de um processo complexo que envolve muitos parâmetros e fatores ambientais. Além disso, um dossel é constituído por muitos elementos da própria vegetação, como folhas, galhos, frutos, flores, etc. (Moreira, 2007).

De todos os elementos constituintes da vegetação, a folha constitui o principal deles quando se considera o processo de interação descrito. Para um melhor entendimento das características de reflectância da REM incidente sobre uma folha é necessário o conhecimento de sua composição química, principalmente tipo e quantidade de pigmentos fotossintetizantes, e de sua morfologia interna (distribuição e quantidade de tecidos, espaços intercelulares, etc).

O comportamento espectral de uma folha é função de sua composição, morfologia e estrutura interna. Desde que as características da folha sejam geneticamente controladas, existirão, portanto, diferenças no comportamento espectral entre grupos geneticamente distintos. Levando em consideração o conceito da reflectância interna numa folha e os conhecimentos do espectro de absorção da clorofila, Willstatter e Stoll (1918), desenvolveram uma teoria sobre a trajetória da REM dentro de uma folha, a qual é válida até hoje.

Os referidos autores basearam sua teoria na estrutura interna das folhas e na reflectância potencial das superfícies. Segundo eles, a trajetória da REM se daria ao longo de vários meios, sendo estes compostos pela água, ar, membranas celulares, etc. Um mesmo feixe de radiação poderia passar, por exemplo, através de uma camada de água, que possui um índice de refração de 1,33, e em seguida atravessar um espaço preenchido com ar, que possui um índice de refração igual a 1,00. Os comprimentos de onda relativos ao ultravioleta não foram considerados, porque uma grande quantidade dessa energia é absorvida pela atmosfera e a vegetação não a utiliza. A região do espectro compreendida entre 400 nm e 2600 nm pode ser dividida em três áreas: região do visível (400 nm a 700 nm), região do infravermelho próximo (700 nm a 1300 nm) e a região do infravermelho médio (1300 nm a 2600 nm). Os principais aspectos relacionados ao comportamento espectral da folha, em cada uma destas regiões são:

- Região do visível: nesta região os pigmentos existentes nas folhas dominam a reflectância espectral. Estes pigmentos, geralmente encontrados nos cloroplastos são: clorofila (65%), carotenos (6%), e xantofilas (29%). Os valores percentuais destes pigmentos podem variar grandemente de espécie para espécie. A energia radiante interage com a estrutura foliar por absorção e por espalhamento. A energia é absorvida seletivamente pela clorofila e é convertida em calor ou fluorescência, e também convertida fotoquimicamente em energia

estocada na forma de componentes orgânicos através da fotossíntese (Jensen, 2009).

- Região do infravermelho próximo: nesta região existe uma absorção pequena da REM e considerável espalhamento interno na folha. A absorção da água é geralmente baixa e a reflectância espectral é quase constante. Gates *et al.* (1965) determinou que a reflectância espectral de folhas nessa região do espectro eletromagnético é o resultado da interação da energia incidente com a estrutura do mesófilo. Fatores externos à folha, como disponibilidade de água por exemplo, podem causar alterações na relação água-ar no mesófilo, podendo alterar a reflectância de uma folha nesta região. De maneira geral, quanto mais lacunosa for a estrutura interna foliar, maior será o espalhamento interno da radiação incidente, e conseqüentemente, maior será também a reflectância.

- Região do infravermelho médio: a absorção devido à água líquida predomina na reflectância espectral das folhas na região do infravermelho próximo. Considerando a água líquida, esta apresenta na região em torno de 2000 nm, uma reflectância geralmente pequena, sendo menor do que 10% para um ângulo de incidência de 65° e menor do que 5% para um ângulo de incidência de 20°. A água absorve consideravelmente a REM incidente na região espectral compreendida entre 1300 nm a 2000 nm. Em termos mais pontuais, a absorção da água se dá em 1100 nm, 1450 nm, 1950 nm e 2700 nm (Jensen, 2009).

Nesse sentido, quando o alvo em estudo é a vegetação, as análises tornam-se relativamente mais complexas. Segundo Houborg (2015), estudos sobre as aplicações do sensoriamento remoto para a caracterização das copas de vegetação começaram a acelerar no final da década de 1960. Como as propriedades de reflectância da terra vegetativa variam sazonalmente em relação à fenologia da vegetação, o sensoriamento remoto possibilita o monitoramento da dinâmica da vegetação, uma vez que possibilitam o estudo desde escalas macro (ecossistemas) até as microescalas (regiões específicas e áreas de interesse), dependendo do sensor utilizado na análise. (Pastor-Guzman *et al.*, 2018).

### **3.2.2. Índices de Vegetação (IVs)**

Os índices de vegetação são medidas radiométricas adimensionais, cujo objetivo é indicar a abundância relativa e a atividade da vegetação verde. Ainda de acordo com Jensen (2009), um índice ideal de vegetação, deve:

- Maximizar a sensibilidade a parâmetros biofísicos da vegetação, preferencialmente de forma linear, de maneira que esta sensibilidade seja fidedigna para uma grande

amplitude de condições da vegetação e para facilitar a validação e calibração do índice;

- Normalizar ou modelar os efeitos externos, tais como: ângulo solar, ângulo de visada do sensor, e as interferências atmosféricas, de forma a possibilitar comparações espaciais e temporais;
- Normalizar os efeitos internos, tais como: as variações do substrato abaixo do dossel, incluindo topografia, solos e diferenças quanto a vegetação senescente ou a presença de ramos lenhosos.

Em diversos estudos sobre os recursos naturais, os índices de vegetação são parâmetros fundamentais, pois possibilitam quantificar e qualificar a biomassa e fornecer informações importantes a respeito do ciclo de crescimento e senescência da vegetação, ou seja, o entendimento de sua fenologia. Estes índices nada mais são do que combinações lineares das reflectâncias em diferentes bandas do espectro, especificamente as bandas do vermelho (*Red*) e do infravermelho próximo (*Nir*), por serem mais sensíveis as respostas da vegetação (Souza, 2016; Barbosa; Kumar, 2016; Workie & Debella, 2017).

Inúmeros IVs já foram desenvolvidos para diferentes finalidades. Estudos em todo o mundo usam as características de absorção associadas com constituintes bioquímicos para avaliar diferentes tipos de vegetação. Assim, diferentes índices são utilizados no levantamento da biomassa vegetal e são otimizados para avaliar um processo de interesse (Brandão, 2008). Sendo assim, os IVs podem ser divididos em índices sensíveis à concentração do pigmento (clorofila) e índices estruturais, que podem ou não sofrer a influência das propriedades óticas do solo (Haboudane *et al.*, 2004; Brandão *et al.*, 2008).

Tucker (1979) estudou diferentes combinações lineares das bandas do verde, vermelho e infravermelho próximo para monitorar propriedades como a biomassa, conteúdo de água na folha e conteúdo de clorofila. Além disso, validou o *NDVI* que foi desenvolvido por Rouse *et al.* (1973). O *NDVI* é um IV estrutural, muito utilizado para monitoramento agrícola, devido sua forte correlação linear com o crescimento das culturas. Além disso, a razão simples entre o *Nir* e o *Red* apresenta uma alta correlação com o desenvolvimento das plantas e com seus estágios de crescimento, sendo esse um dos primeiros IVs estudados. (Birth; Mcvey, 1968).

Segundo Jensen (2009), pesquisas têm apontado, desde metade do século passado, que existe uma relação direta entre o efeito verificado no *Nir* e as variações da biomassa. Em contrapartida, tem sido notado que a resposta no visível, especificamente no *Red*, ocorre de maneira inversa em relação às variações da biomassa da vegetação. Sendo assim, quanto maior for o verdor da vegetação (nível alto de atividade fotossintética) maior será o valor da

reflectância no *Nir* e menor será no *Red*. A medida que o nível de atividade fotossintetizante diminui, os papéis se invertem e a planta passa a refletir muito no *Red* e pouco no *Nir*. Portanto, combinações lineares que envolvam essas duas faixas de comprimentos de onda, podem representar muito bem as condições da vegetação e sua relação com os ciclos hidrológicos, é o que ocorre com o *NDVI*.

Muitas pesquisas têm utilizado séries temporais de índices de vegetação para o monitoramento da vegetação, por diversos sensores a bordo de satélites, como o MODIS-Moderate Resolution Image Spectroradiometer (Cai *et al.*, 2017; Souza *et al.*, 2016; Kuplich *et al.*, 2013), os sensores TM, ETM+ e OLI dos satélites LANDSAT 5, 7 e 8, respectivamente (Silva *et al.*, 2018; Melaas *et al.*, 2013) e o AVHRR a bordo do satélite NOAA (Zhang *et al.*, 2018).

Muitos desses estudos, com foco na análise de grandes áreas, utilizam principalmente, séries temporais de índices de vegetação adquiridos por sensores de resolução moderada, como o MODIS, cuja resolução espacial é de 250 m, principalmente por ter alta resolução temporal (tempo de revisita de até um dia), como os trabalhos realizados na China (Li *et al.*, 2017; Lu *et al.*, 2016) e na África (Di Vittorio; Georgakakos, 2018).

Contudo, em áreas com heterogeneidade marcante, como ocorre na vegetação Caatinga, utilizar dados de sensores com resolução espacial moderada podem resultar em interpretações incorretas sobre a dinâmica da vegetação. Nesse sentido, dados obtidos a partir de sensores de resolução espacial maiores, como os VANT's, são ideais por fornecerem informações espaciais muito mais detalhadas (Liu *et al.*, 2017; Boschetti, 2015).

Diversas pesquisas realizadas na Caatinga obtiveram êxito na utilização de imagens de alta resolução obtidas a partir de câmeras acopladas em VANT para uma infinidade de aplicações, como no estudo de Santos *et al.* (2022) em que os autores analisaram o índice de área foliar para uma área de Caatinga em Petrolina – PE, além dos estudos de Santos *et al.* (2021) que utilizaram dados de *NDVI* obtidos a partir de VANT para verificar a precisão dos sensores acoplados aos satélites Sentinel – 2 e Landsat- 8 para captação das características estruturais da vegetação na Caatinga. Estudo aplicados à áreas de pastagem e cultivo também obtiveram sucesso com aplicação de VANT, é o caso do estudo de Sampaio *et al.* (2020) que compararam índices de vegetação gerados a partir de imagens de alta resolução obtidos a partir de VANT com os mesmos índices gerados a partir de imagens Landsat-8 para uma área de pastagem na Embrapa – SP e da pesquisa de Sá (2022) que aplicou imagens obtidas por VANT para uma área de cultivo no submédio do São Francisco em Juazeiro-BA.

### ***3.2.3. Veículo Aéreo Não Tripulado (VANT)***

O termo Veículo Aéreo Não Tripulado (VANT) é mundialmente reconhecido e inclui uma grande gama de aeronaves que são autônomas, semiautônomas ou remotamente operadas. Segundo a ABA - (Associação Brasileira de Aerodelismo), a definição para Veículo Aéreo Não Tripulado (VANT) é: “um veículo capaz de voar na atmosfera, fora do efeito de solo, que foi projetado ou modificado para não receber um piloto humano e que é operado por controle remoto ou autônomo”. De acordo com Medeiros (2007), os veículos Aéreos Não Tripulados (VANT's), popularmente conhecidos como Drones, são pequenas aeronaves, sem qualquer tipo de contato físico direto, capazes de executar diversas tarefas, tais como monitoramento, reconhecimento tático, vigilância e mapeamento, entre outras.

No Brasil, os primeiros relatos de VANT's ocorreram na década de 80, quando o Centro Técnico Aeroespacial (CTA) desenvolveu o projeto Acauã, que tinha fins militares (Medeiros, 2007). As pesquisas nessa área, no Brasil, ainda são muito recentes e os principais trabalhos com VANT's têm no escopo as aplicações civis, como vigilância policial de áreas urbanas e de fronteira, inspeções de linhas de transmissão de energia, monitoramento de atividades em áreas agrícolas, acompanhamento de safra, controle de pragas e de queimadas.

O interesse em VANT's tem crescido ao redor do mundo. Avanços recentes na tecnologia computacional, desenvolvimento de software, materiais mais leves, sistemas globais de navegação, avançados links de dados, sofisticados sensores e a miniaturização são os motivos do aumento de desenvolvimentos de VANTs (Jorge e Inamasu, 2014).

Com o auxílio do Geographic Position System (GPS), os VANT's apresentam boa acurácia em estudos ambientais, sendo muito utilizados na agricultura de precisão e no monitoramento de habitats, pois permitem a coleta de dados com alta resolução temporal. Outras vantagens apresentadas, são a sua capacidade de sobrevoar áreas em baixas altitudes (30 – 300 m) e, conseqüentemente, sem interferência das nuvens, que os capacita a capturar imagens em alta resolução espacial e a coletar dados em locais de difícil acesso (Silva *et al.*, 2014).

Imagens aéreas com alta resolução espacial e temporal contribuem para obtenção de informações em campo e, conseqüentemente, para a caracterização da área analisada em escala detalhada. Dados obtidos por VANT's estão sendo usados com enfoques diferenciados, como no monitoramento de copas de árvores em ruas de áreas urbanas (Alves, 2017), classificação

de cobertura do solo (Jiang, 2017), monitoramento agrícola (Niu, 2019), estimativa de índice de área foliar (Córcoles *et al.*, 2013), estimativa de evapotranspiração de culturas e de biomas (Sagan *et al.*, 2019, Brenner *et al.*, 2018, Park, 2018, Hoffmann *et al.*, 2016, Ortega-Farías *et al.*, 2016, Xia *et al.*, 2016), entre outros.

Muitos esforços têm sido empregados no desenvolvimento dos VANT's, permitindo que os mesmos voem em altitudes, distâncias e velocidades cada vez maiores. Diversos equipamentos, tais como estabilizadores automáticos para câmeras, câmeras com maior resolução e menor peso, transmissores de vídeo de alta qualidade e longo alcance e módulos de posicionamento global de alta precisão, agregam funcionalidade aos VANT's, tornando-os excelentes ferramentas de dados, inclusive em tempo real (Matsuo, 2011).

### **3.3. Classificação da cobertura do solo com VANT's e Random Forest**

O monitoramento do ambiente baseado em imagens aéreas obtidas por VANT's tem se tornado cada vez mais comum (Anderson E Gaston, 2013; Silveira *et al.*, 2022; Carvalho e Carvalho, 2022; Santos *et al.*, 2021; Santos *et al.*, 2022). Os dados obtidos por satélites, ainda que tenham melhorado muito em termos de resolução espacial e temporal, são menos eficientes do que os drones para captar mudanças espaciais que ocorrem em curtos espaços, na ordem dos centímetros, o que é facilmente alcançado com os dados obtidos por VANT. (Marzialetti *et al.*, 2021; Gonçalves *et al.*, 2022; Olariu, *et al.*, 2022; Kattenborn *et al.*, 2019 e Lehmann *et al.*, 2017).

Os VANT's podem ser equipados com sensores RGB, multiespectrais, hiperespectrais e ativos. Embora Olariu *et al.* (2022) apresentaram resultados satisfatórios usando apenas câmeras RGB, sensores multiespectrais e hiperespectrais também podem ser usados para mapear vegetação (Kattenborn *et al.*, 2019; Berni *et al.*, 2009; Strecha *et al.*, 2012). Porém, faltam estudos combinando sensores RGB e multiespectrais para classificação da cobertura do solo, aplicados à vegetação da Caatinga.

Pela característica heterogênea do bioma estudado (alto grau de biodiversidade) e pela alta resolução espacial obtidas com os VANT's, a combinação de sensores para a finalidade de classificação da cobertura do solo exige alta capacidade de processamento de dados. Algoritmos de aprendizado de máquina são ferramentas amplamente utilizadas na classificação de sensoriamento remoto. Eles podem modelar assinaturas de classes complexas e aceitar diferentes tipos de dados de entrada, incluindo dados sem distribuição normal. Há evidências suficientes do desempenho superior dos algoritmos de aprendizado de máquina em

relação às técnicas tradicionais, como a estimativa de máxima verossimilhança (Pal *et al.*, 2002; Pal *et al.*, 2003; Belgiu *et al.*, 2016 e Maxwell *et al.*, 2018).

Um desses algoritmos é o *Random Forest* (RF) (Breiman, 2001), ele tem apresentado bons resultados para aplicação em classificação de cobertura do solo (Ghosh e Joshi, 2014; Chen *et al.*, 2018; Andrade *et al.*, 2021). O RF é um algoritmo de aprendizado de máquina que se baseia na técnica de *ensemble*, o que significa que ele combina várias árvores de decisão para realizar a classificação.

A classificação do solo utilizando o RF é uma aplicação poderosa e eficaz da aprendizagem de máquina em ambientes de sensoriamento remoto e geoprocessamento. A capacidade de determinar a cobertura do solo com precisão é fundamental para diversas áreas, incluindo planejamento urbano, monitoramento ambiental, agricultura de precisão e gestão de recursos naturais. O RF, com suas características de *ensemble* e capacidade de lidar com dados de alta dimensionalidade, se destaca como uma ferramenta valiosa para essa tarefa.

O processo de classificação do solo com o RF começa com a aquisição de imagens de sensoriamento remoto que capturam informações sobre a superfície terrestre em várias bandas espectrais. Essas imagens são posteriormente processadas para extrair características relevantes, como intensidades espectrais em diferentes bandas, índices de vegetação, texturas e outras informações que ajudam a caracterizar a cobertura do solo. O próximo passo envolve a preparação de um conjunto de treinamento, que consiste em amostras rotuladas de diferentes classes de cobertura do solo. Por exemplo, essas classes podem incluir florestas, áreas urbanas, corpos d'água, campos agrícolas e outras categorias relevantes. O RF é treinado com base nesses exemplos, aprendendo a distinguir padrões associados a cada classe.

O diferencial do algoritmo RF reside em sua capacidade de criar múltiplas árvores de decisão independentes durante o treinamento. Cada árvore é treinada em uma subamostra aleatória dos dados de treinamento, o que ajuda a reduzir o risco de *overfitting*, onde o modelo se ajustaria em excesso aos dados de treinamento. Durante a classificação, todas as árvores contribuem para a decisão final, e a classe mais votada é atribuída a cada pixel na imagem de sensoriamento remoto.

Além disso, fornece informações sobre a importância das variáveis preditoras usadas no processo de classificação, isso é importante para entender quais características ou bandas espectrais são mais relevantes para a classificação do solo, auxiliando na interpretação dos resultados e na identificação de padrões significativos.

## **4. MATERIAL E MÉTODOS**

### **4.1. Área de estudo**

A área de compreensão desta pesquisa está totalmente inserida no Semiárido do Brasil (Figura 1), nos domínios do Bioma Caatinga, que é uma FTSS com inúmeras variações paisagísticas. Segundo o IBGE (2022) ela é classificada como Savana Estépica, com área mais representativa abrangendo a Caatinga do Semiárido Nordeste. O clima predominante, conforme a classificação climática de Köppen, é do tipo Semiárido e de subtipo seco e quente (BSh), com precipitação pluvial anual variando entre os valores médios aproximados a 300,0 mm e 600,0 mm e temperatura média superior a 26°C. O período chuvoso na região de Patos – PB começa em janeiro/fevereiro e encerra-se em maio/junho, enquanto o período de estiagem (seco) é prolongado, de seis a sete meses, tendo início em maio/junho e estendendo-se até dezembro. A precipitação média a partir de uma série histórica dos últimos 10 anos (2013 a 2023) foi de aproximadamente 398 mm/ano, os dados são provenientes da estação meteorológica do Instituto Nacional de Meteorologia (INMET) de coordenadas 7°4'47'' S e 37°16'22'' O.

O município de Patos, no estado da Paraíba, segundo o IBGE (2020) tem a população estimada em 108.192 habitantes e está em constante progresso ambiental, econômico e demográfico. A área de estudo, bacia Riacho do Frango, banha a cidade de Patos e está introduzido na bacia hidrográfica Piranhas-Açu. O município é cortado pelo Rio Espinhas. O riacho atualmente é o canal do Frango, onde ao longo do seu talvegue já existe um canal artificial de alvenaria, onde ocorrem frequentes alagamentos e inundações corriqueiras (ARAUJO & QUEIROZ, 2022). A bacia do riacho do Frango percorre diversos bairros da cidade como: Belo Horizonte, Jardim Lacerda, Bela Vista, Jardim Europa, e deságua no Rio Espinharas, sendo a principal bacia hidrográfica urbana do município.

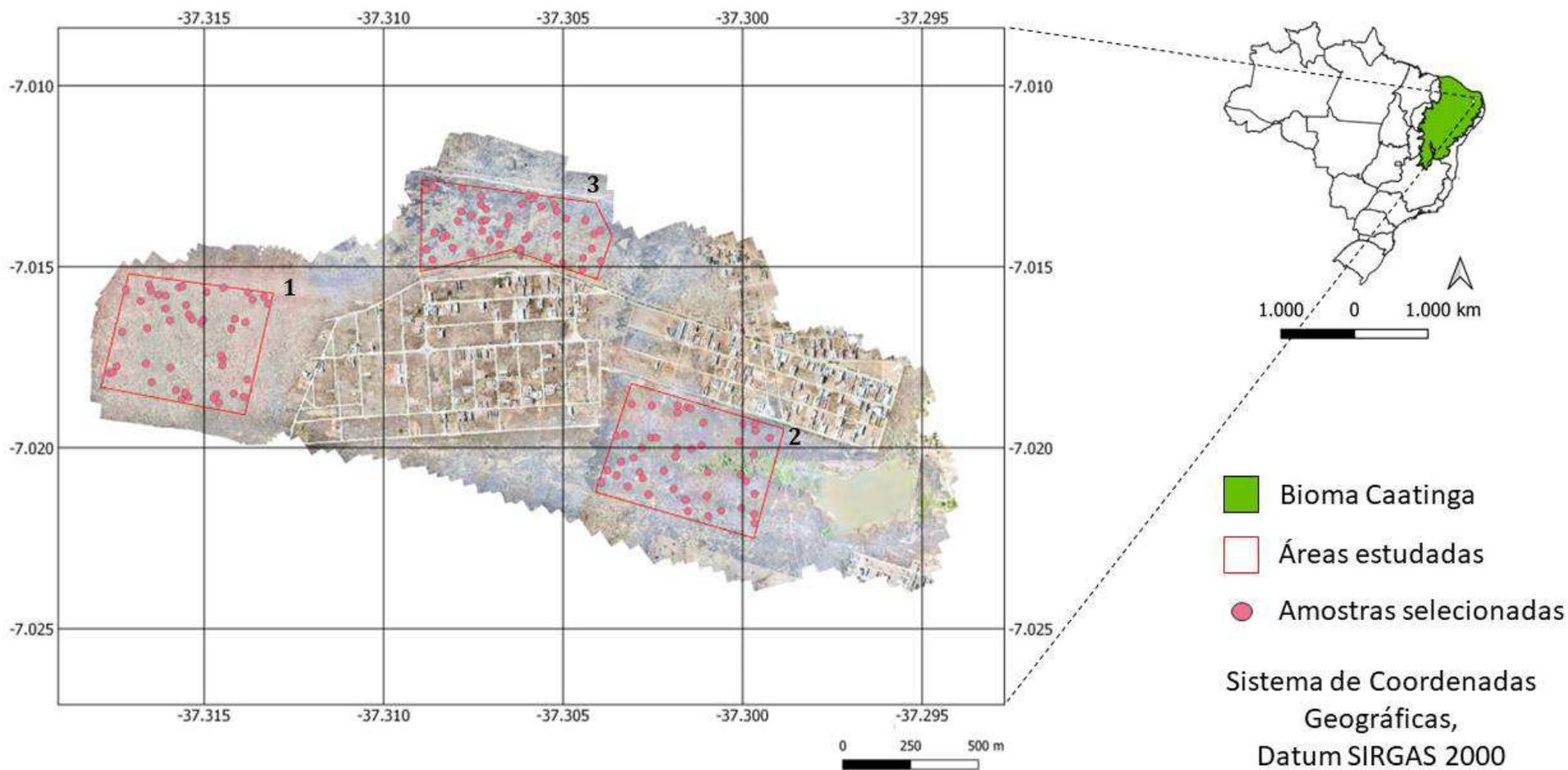


Figura 1: Localização da área de estudo: 1, 2 e 3 referem-se as áreas selecionadas.

## 4.2. Etapas metodológicas

As etapas metodológicas dessa pesquisa estão resumidas na Figura 2 e englobam: aquisição de dados, processamento de dados, definição das classes de cobertura e coleta de amostras, classificação da cobertura do solo e precisão da classificação.



Figura 2: Fluxograma representativo das etapas metodológicas para classificação da cobertura do solo na Caatinga utilizando VANT e classificador RF.

### 4.2.1 Aquisição dos dados

Neste estudo foram utilizadas imagens obtidas a partir de duas câmeras: Multiespectral Mapir Survey 3W RGN (R – red , G – Green , N – Nir (Near infrared)) acoplada ao VANT com 87° HFOV (19mm) e abertura f/2.8 e a câmera RGB (R – red , G – Green , B – Blue) modelo CMOS 1/2.3" FOV 78.8° 26 mm (formato equivalente a 35 mm) e abertura f/2.2, nativa do DJI Mavic Pro, ambas com resolução de 12 MP (4000 x 3000 pixels) e com Sistema de Posicionamento Global (GPS) embarcado. A figura 3 apresenta os equipamentos utilizados nesta pesquisa. Os voos sobre a área de estudo foram realizados entre 22/09/2023 e 22/10/2023 sempre no horário compreendido entre 12:00 e 14:00 hrs, para minimizar o efeito de sombreamento nas imagens, com condições de céu claro. O período de coleta dos dados está totalmente compreendido na estação seca, tendo o último evento de chuva ocorrido no dia 20/08/2023 para o município de Patos-PB, segundo dados da AESA (2023).

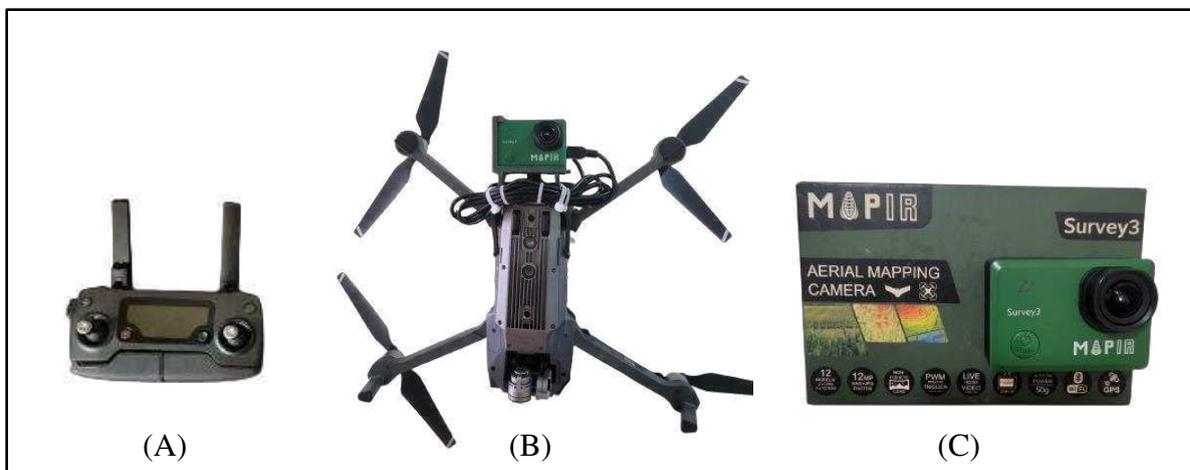


Figura 3: Equipamentos utilizados na pesquisa: Controle RC Mavic Pro (A), Drone DJI Mavic Pro com câmera e GPS acoplados (B) e Câmera multiespectral Mapir Survey 3w (C).

O percurso de cada voo foi previamente definido no *Google Earth* e depois exportado em formato *kml* para a plataforma online *Drone Deploy*, a partir da qual os outros parâmetros do planejamento de voo puderam ser definidos, os quais estão indicados na tabela 1. A partir das predefinições inseridas, o *Drone deploy* apresenta o tempo de voo, a quantidade de baterias e quantidade de imagens necessárias a serem captadas e envia as informações para o drone, que realiza o voo de forma automática, coletando as imagens de maneira a garantir as sobreposições definidas.

Tabela 1: Parâmetros do planejamento de voo no *Drone Deploy*.

Parâmetro	Valor adotado
Altura de voo	120 m
Velocidade de voo	10 m/s
Sobreposição frontal	75%
Sobreposição lateral	75%
Ângulo de inclinação do gimbal da câmera	-90°

Para a câmera Mapir survey 3W RGN foram realizados na própria câmera os ajustes de configurações para melhor desempenho na captura das fotos, os quais estão indicados na tabela 2. A Mapir foi acoplada no VANT de forma a garantir o mesmo ângulo de captura da câmera RGB.

Tabela 2: Configurações para câmera Mapir Survey 3W (RGN).

<b>Parâmetro</b>	<b>Valor adotado</b>
Intervalo de disparo	2 s
ISO	100
Equilíbrio do branco	automático
Exposição	+0,0
Velocidade do obturador	1/250 s
Formato de captura das imagens	JPEG + RAW

#### 4.2.2. *Processamento dos dados*

As imagens multiespectrais obtidas pela Mapir inicialmente foram calibradas e corrigidas espectralmente utilizando o software *Mapir Camera Control (MCC)*, fornecido pelo próprio fabricante, disponível para download na homepage da marca. No ambiente do software *MCC* foram inseridas às imagens captadas em formato *JPEG + RAW* e como saída, após o processamento, as imagens são convertidas em formato *TIFF* e guardadas em uma pasta pré-definida (geralmente a mesma pasta de origem das imagens de entrada). Já para as imagens em formato JPEG obtidas pela câmera RGB não se faz necessária nenhuma etapa de pré-processamento.

Logo após, todas as imagens aéreas foram processadas utilizando o software *Agisoft Metashape (Agisoft LLC, Russo)*, no qual seis etapas são necessárias para o processamento das imagens, que inclui o alinhamento de fotos, construção de malha, a geração da nuvem densa de pontos, o modelo digital de elevação, a geração do ortomosaico e, por último, a exportação do ortomosaico em formato *TIFF*. A quantidade de imagens utilizadas em cada etapa de processamento foi aquela referente à uma semana de coleta de dados para cada câmera, com uma média de 1500 imagens por processamento, perfazendo uma área de aproximadamente 100 hectares em cada estágio. Essa quantia foi definida de acordo com o tempo de processamento do software, de forma a otimizar a geração dos ortomosaicos, que, ao final do processamento, apresentaram uma resolução espacial de 5,5 cm/pixel para Mapir e 4,0 cm/pixel para câmera nativa do VANT.

Com os ortomosaicos prontos, foi utilizado o software *QGIS* na sua versão 3.32.3 para união dos mesmos, obtendo assim uma única imagem no formato *TIFF* representando toda a área de estudo. A partir do mosaico, foram analisadas as condições de sombreamento de nuvens na imagem e definidas áreas específicas nesta ortofoto para a classificação proposta

neste trabalho. Foram escolhidas três áreas, com a área 1 correspondendo à 16,60 ha, área 2 com 17,35 ha e a área 3 com 12,32 ha, indicadas na Figura 3. Os critérios de escolha foram: enquadramento na Bacia do Riacho do Frango, ausência de sombra de nuvens e ausência de áreas urbanizadas. As imagens que correspondem a estas áreas foram obtidas nos dias 18, 19 e 20 de setembro de 2023 (área 1), 22 e 27 de setembro de 2023 (área 2) e 04, 05 e 06 de Outubro de 2023 (área 3).

Ainda no ambiente do *QGIS*, a partir da ferramenta “extração de dados”, a ortofoto foi recortada para as áreas de interesse. Por último, foram extraídas as bandas espectrais para as áreas escolhidas, sendo 3 bandas (R, G e B) para a câmera nativa e 3 bandas (R, G, Nir) para a câmera Mapir. As bandas foram então salvas em formato *TIFF* para posterior normalização. Com as bandas de cada imagem prontas, foi utilizado um script dentro da plataforma Rstudio (Anexo A) para normalizar os valores de cada banda através do método de normalização pela média (*mean-based normalization*). Essa abordagem foi utilizada para equalizar as escalas das diferentes bandas espectrais, o que é especialmente útil para comparar ou combinar informações de várias bandas em sensoriamento remoto ou análise de imagens multiespectrais aplicadas em algoritmos de classificação como o RF. Além disso, ajuda a reduzir o efeito de diferentes níveis de iluminação ou sensibilidade do sensor em cada banda. Ainda na plataforma do *Rstudio*, com as bandas já normalizadas, foram gerados os IV's que, além das bandas normalizadas, servirão como parâmetros de entrada para o RF: *NDVI* (Equação 1), obtido com as bandas normalizadas da Mapir e do *NGRDI* (Equação 2), obtido pelas bandas normalizadas do sensor nativo do drone.

$$NDVI = \frac{Nir-Red}{Nir+Red} \quad (1)$$

$$NGRDI = \frac{Green-Red}{Green+Red} \quad (2)$$

#### **4.2.3. Definição das classes de cobertura e coleta de amostras**

A definição das classes de cobertura do solo da Caatinga seguiu a metodologia proposta por Chaves *et al.* (2008). Esses autores estudaram a vegetação da Caatinga em seus diferentes estágios de antropização, tendo como parâmetros: o tamanho, características morfológicas e graus de cobertura, como indicado na Figura 4.

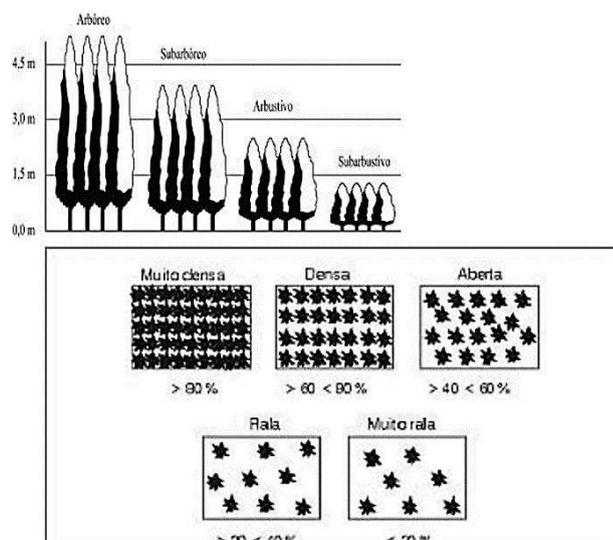


Figura 4 : Grupos e classes para a vegetação Caatinga, Fonte: Chaves et al. (2008).

Segundo os autores, utilizando combinações desses grupos e classes de acordo com a porcentagem de cada grupo na parcela analisada, 201 padrões de vegetação Caatinga poderiam ser diferenciados. Neste trabalho, a partir da análise visual de campo e das imagens, foram diferenciados apenas 3 padrões, quais sejam: Arbustiva arbórea densa, Subarbustiva muito rala e Solo exposto, que foram aplicados na coleta de amostras e na classificação. A imagem apresentada na Figura 5 exemplifica a distribuição destas classes para uma pequena parcela de solo na área de estudo.

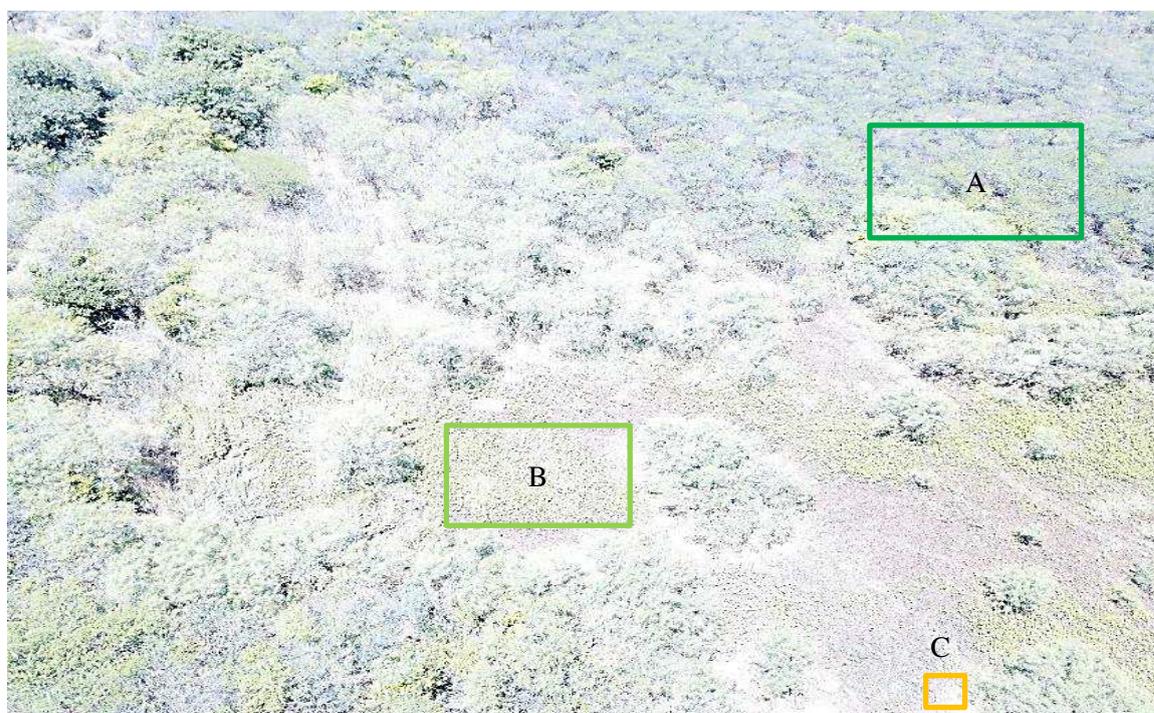


Figura 5: Classes de cobertura identificadas na área de estudo: Arbustiva arbórea densa (A), Subarbustiva muito rala (B) e Solo exposto (C).

Na coleta de amostras foram utilizados 135 pontos para identificar os diferentes agrupamentos vegetais, sendo 15 pontos para cada uma das três classes, totalizando 45 pontos em cada área avaliada. Essas amostras foram definidas dentro do ambiente do *QGIS* e exportadas de forma agrupada para cada área em um único arquivo no formato *shapefile*. Do total de pixels amostrados para cada classe (15), 70% (10) foram utilizados para o treinamento do modelo e 30% (5) para teste, definidos de forma aleatória. Os dados utilizados para treinamento e validação do modelo em cada área avaliada estão indicados no Anexo B.

#### **4.2.4. Classificação da cobertura do solo**

Nesta etapa foi utilizado um script desenvolvido dentro da plataforma *Rstudio*, utilizando o pacote *Random Forest R* (Liaw e Wiener, 2002). O RF é um classificador baseado em árvores de decisão, que estabelece regras para a tomada de decisão, que por sua vez dependem das variáveis preditoras. O algoritmo então, cria uma estrutura similar à um fluxograma, onde as condições são verificadas em cada “nó” e, dependendo do resultado se positivo ou negativo, segue por ramos diferentes, até a finalização da árvore.

Os principais parâmetros do RF, definidos pelo usuário, são o número máximo de árvores de decisão a serem geradas no fluxograma (*n tree*) e o número de variáveis usadas aleatoriamente para dividir cada nó (*m try*) da árvore (Belgiu edrăgut, 2016; Htitiou et al., 2019). Neste estudo, foram utilizadas mil (1.000) árvores de decisão e o *m try* foi definido como a raiz quadrada do número de variáveis preditoras, neste caso duas (2). Além disso, as amostras coletadas foram separadas em duas partes: treinamento e teste, que correspondem à 70% e 30% das amostras, respectivamente.

Para classificação foram utilizadas duas abordagens distintas, quais sejam:

- Método 1: classificação baseada em quatro variáveis preditoras correspondentes às bandas normalizadas do vermelho (R), verde (G), Infravermelho Próximo (Nir) e o *NDVI* de uma imagem mono-temporal de alta resolução espacial obtida a partir da câmera Mapir Survey 3W (RGN) acoplada ao VANT.
- Método 2: classificação baseada em quatro variáveis preditoras correspondentes às bandas normalizadas do vermelho (R), verde (G), azul (B) e o *NGRDI* de uma imagem mono-temporal de alta resolução espacial obtida a partir da câmera nativa do VANT DJI Mavic Pro.

#### 4.2.5. Importância das variáveis e precisão da classificação

Para avaliar a importância de cada variável para precisão da classificação no RF em cada modelo, foi utilizado o *Mean Decrease Gini* (MDG) (Corcoran *et al.*, 2013; Nitze *et al.*, 2015). Quanto maior for o valor do indicador, maior será o grau de importância da variável (Zhang *et al.*, 2019). Finalmente, foi verificado qual melhor método para classificação da área de estudo através dos índices: Acurácia Global (AG), Coeficiente Kappa (k), Precisão do produtor (PP) e Precisão do usuário (PU), obtidos a partir da matriz de confusão de cada método em cada uma das três áreas avaliadas.

O *Mean Decrease Gini* (MDG) (Nitze *et al.*, 2015) é um índice utilizado no RF para definir a importância das variáveis preditoras, e é utilizado para selecionar a melhor divisão em cada nó da árvore, calculando a impureza de cada variável no modelo (Jin *et al.*, 2018). A expressão do MDG é demonstrada pela Equação 3:

$$MDG = \sum_{i=1}^n p_k(1 - p_k) = 1 - \sum_{i=1}^n p_k^2 \quad (3)$$

Onde,  $P_k$  é a probabilidade da  $k$ -ésima categoria entre as  $n$  categorias e pode ser estimada a partir da Equação 4:

$$p_k = MDG(D) - \left[ \left( \frac{|D1|}{D} \cdot MDG(D1) \right) + \left( \frac{|D2|}{D} \cdot MDG(D2) \right) \right] \quad (4)$$

D1 e D2 são duas categorias nas quais o conjunto D é dividido de acordo com as características das classes. Portanto, quanto maior for o valor do índice MDG, maior será a importância da variável preditora avaliada e indica também a qualidade da classificação realizada.

A precisão da classificação supervisionada pode ser realizada com base em quatro indicadores de desempenho de precisão calculados a partir da matriz de confusão: Acurácia Geral (AG), Coeficiente Kappa (k), precisão do produtor (PP) e precisão do usuário (PU). A matriz de confusão é uma tabela que descreve o desempenho de um modelo de classificação em termos de suas previsões para cada classe com base nos valores de referência indicados nos dados de treinamento do modelo (Foody, 2002) (Tabela 3).

Tabela 3: Matriz de confusão genérica

Classificação	Dados de referência				Total
	Classe 1	Classe 2	Classe 3	Classe n	
Classe 1	X11	X12	X13	X1m	X1+
Classe 2	X21	X22	X23	X2m	X2+
Classe 3	X31	X32	X33	X3m	X3+
Classe m	Xm1	Xm2	Xm3	Xmm	Xm+
Total	X+1	X+2	X+3	X+m	n

O coeficiente de concordância Kappa ( $k$ ) utiliza todos os elementos da matriz de confusão em seu cálculo (Equação 5). Valores de  $k$  menores que 0 significa nenhuma concordância e próximos de 1 significa concordância perfeita, ele constitui um importante avaliador de precisão na análise de imagens.

$$k = \frac{n \sum_{i=1}^m x_{ii} - \sum_{i=1}^m x_{i+} \cdot x_{+i}}{n^2 - \sum_{i=1}^m x_{i+} \cdot x_{+i}} \quad (5)$$

Onde,  $n$  é o número total de amostras e  $m$  é número total de classes,  $x_{ii}$  é o valor da diagonal da matriz de confusão na linha e na coluna  $i$ ,  $x_{i+}$  é a soma da linha  $i$  e  $x_{+i}$  é a soma da coluna  $i$  da matriz.

A acurácia geral (AG) mede a taxa global de previsões corretas feitas pelo modelo em relação ao número total de previsões, ou seja, é quociente da soma dos elementos da diagonal principal da matriz de confusão pelo número total de amostras de referência (Equação 6).

$$AG = \frac{x_{11} + x_{22} + x_{33} \dots + x_{mm}}{n} \quad (6)$$

A precisão do produtor (PP), também conhecida como sensibilidade ou Taxa de Verdadeiros Positivos, é o quociente do número total de amostras classificadas corretamente pelo número total de amostras de referência para aquela classe (Equação 7). Já a precisão do usuário (PU), também conhecida como Valor Preditivo Positivo, é o quociente entre a quantidade total de amostras classificadas corretamente da classe pelo número total de amostras classificadas naquela classe (Equação 8).

$$PP = \frac{x_{ii}}{x_{+i}} \quad (7)$$

$$PU = \frac{x_{ii}}{x_{i+}} \quad (8)$$

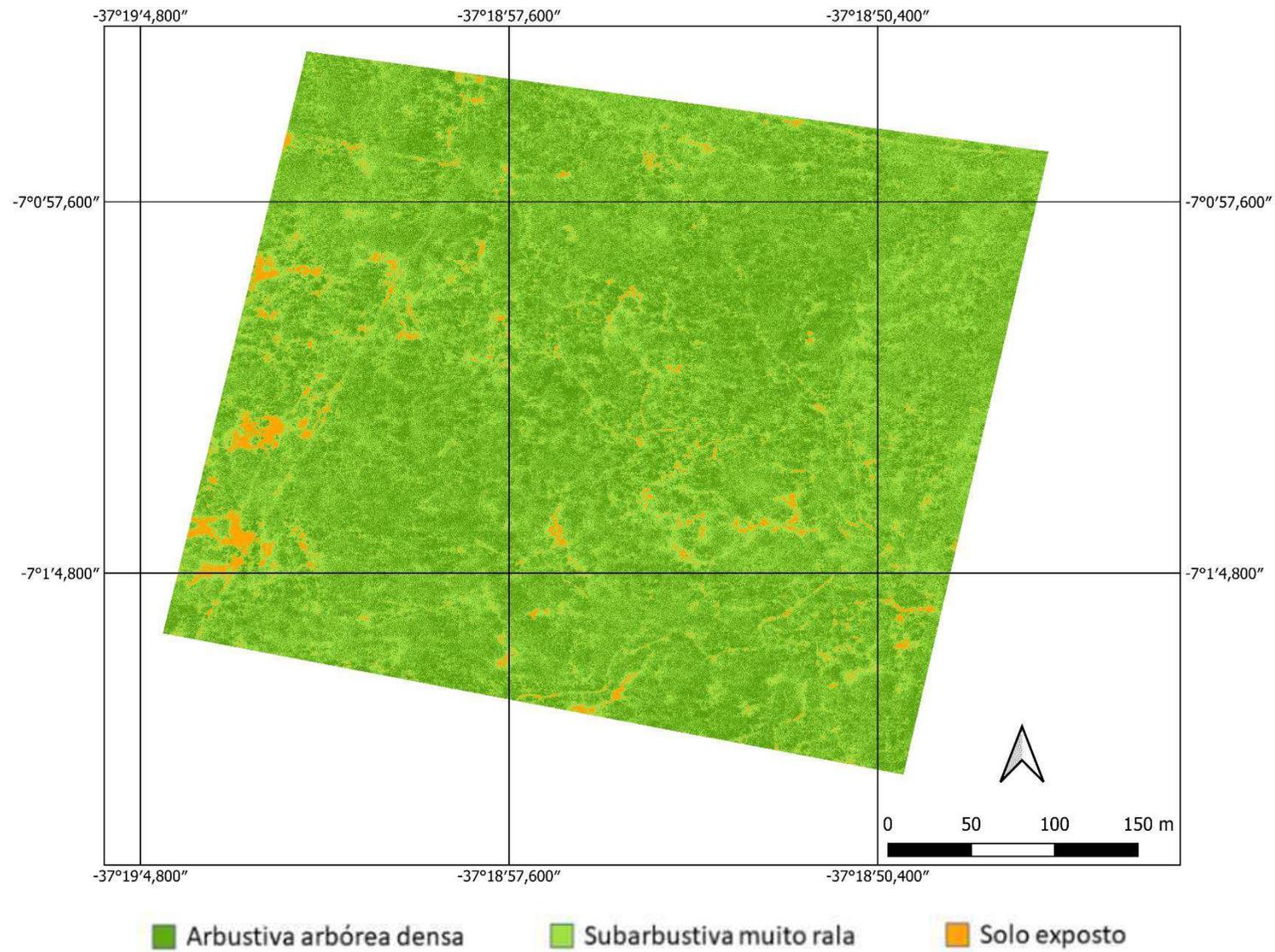
Todos os índices analisados (AG, k, PP e PU) variam de 0 % até 100 %, valores que representam o pior e o melhor desempenho do método de classificação, respectivamente.

## **5. RESULTADOS E DISCUSSÕES**

### **5.1. Classificação para área 1**

A classificação realizada para a área 1 (Figura 6A), utilizando as bandas do Nir, R, G e *NDVI* (método 1) apresentou uma melhor acurácia global (AG = 93,0%) quando comparada à classificação realizada pelo método 2 (Figura 6B), utilizando as bandas do R, G, B e *NGRDI*, cujo valor de AG foi de 40,0%. Na análise do índice Kappa, o método 1 também apresentou resultado superior ao método 2, com valores de k de 0,90 e 0,10, respectivamente.

(A)



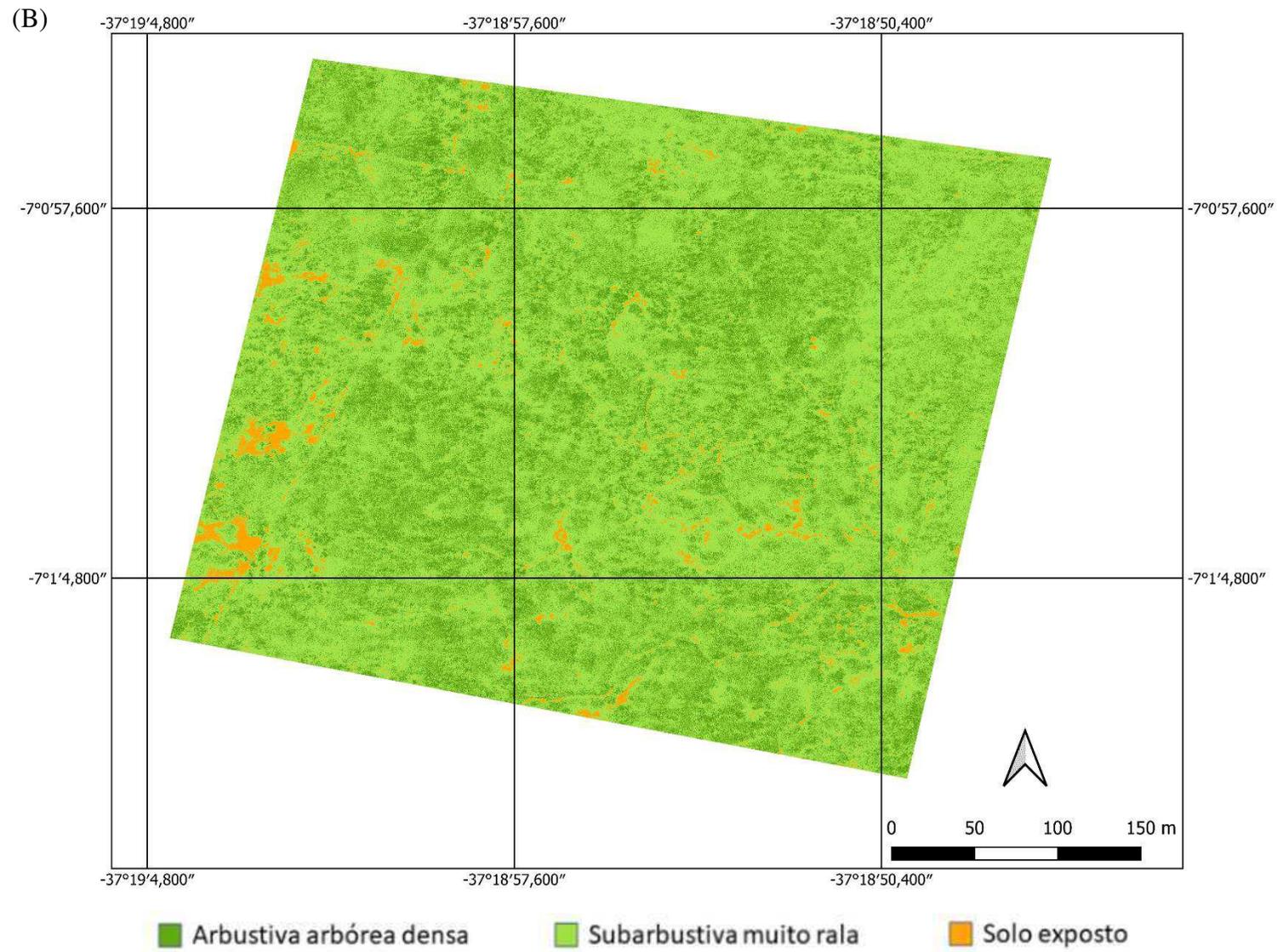


Figura 6: Classes de cobertura de solo para a área 1: método 1 (A) e método 2 (B).

As diferenças presentes nas duas classificações tornam-se ainda mais evidentes a partir da análise da percentagem de cobertura de cada classe em cada método. As figuras 7A e 7B indicam a percentagem de cobertura de cada classe para o método 1 e 2, respectivamente.

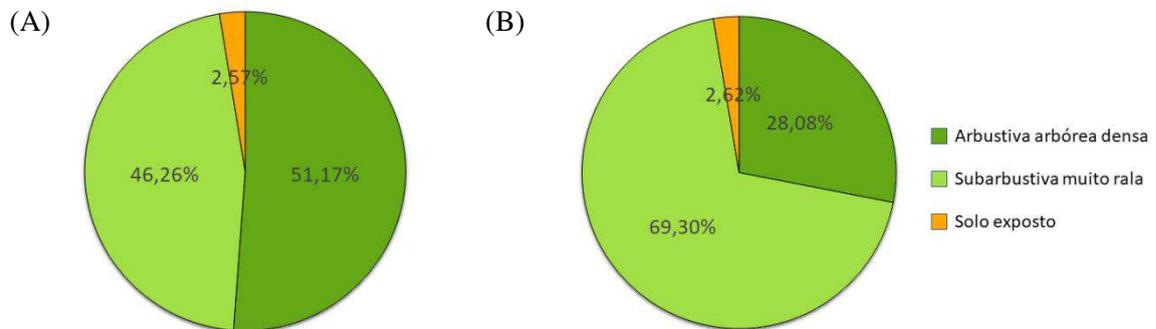


Figura 7: Distribuição percentual das classes de cobertura para a área 1: método 1 (A) e método 2 (B).

É possível notar que as percentagens de cobertura para Solo exposto foram muito próximas, indicando que nesta área os dois métodos apresentaram bons resultados para detecção desta classe, com valores de PU e PP de 100,0% para o método 1 e 80,0% para o método 2, como observa-se nas Tabelas 4A e 4B referentes à matriz de confusão nos métodos 1 e 2, respectivamente. O método 2 apresentou desempenho insatisfatório ao classificar os pixels para as classes Arbustiva arbórea densa (PU=PP=0,0%) e Subarbustiva muito rala (PP = 40,0% e PU = 29,0%), ao passo em que o método 1 apresentou valores de PU e PP iguais à 83,0% e 100,0% na classe Arbustiva arbórea densa e 100,0% e 80,0% na classe Subarbustiva muito rala, nesse caso, com desempenho muito superior ao método 2.

A partir dos gráficos apresentados nas Figuras 7A e 7B é possível perceber que a classes Arbustiva arbórea densa e Subarbustiva muito rala apresentaram as maiores diferenças entre os métodos (23,09% e 23,04%, respectivamente).

Tabela 4: Matrizes de confusão para a área 1.

(A)

Matriz de confusão - método 1					
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)
Arbustiva arbórea densa	5	1	0	6	83,0
Subarbustiva muito rala	0	4	0	4	100,0
Solo exposto	0	0	5	5	100,0
Total	5	5	5	15	-
PP (%)	100,0	80,0	100,0	AG (%)	93,0
k	0,90				

(B)

Matriz de confusão - método 2					
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)
Arbustiva arbórea densa	0	2	1	3	0,0
Subarbustiva muito rala	5	2	0	7	29,0
Solo exposto	0	1	4	5	80,0
Total	5	5	5	15	-
PP (%)	0,0	40,0	80,0	AG (%)	40,0
k	0,10				

(A) Método 1 e (B) Método 2.

Além disso, através da análise do *MDG* é possível notar que a variável preditora com maior importância no método 1 para a área 1 foi a banda do vermelho (R), seguido pelo *NDVI*. A variável menos importante foi a banda do verde (G), como mostra a figura 8A. No método 2 a variável mais importante foi o *NGRDI*, seguida pelas bandas do verde (G) e vermelho (R). A banda com menor importância para o método 2 foi a banda do azul (B), como mostra a figura 8B.

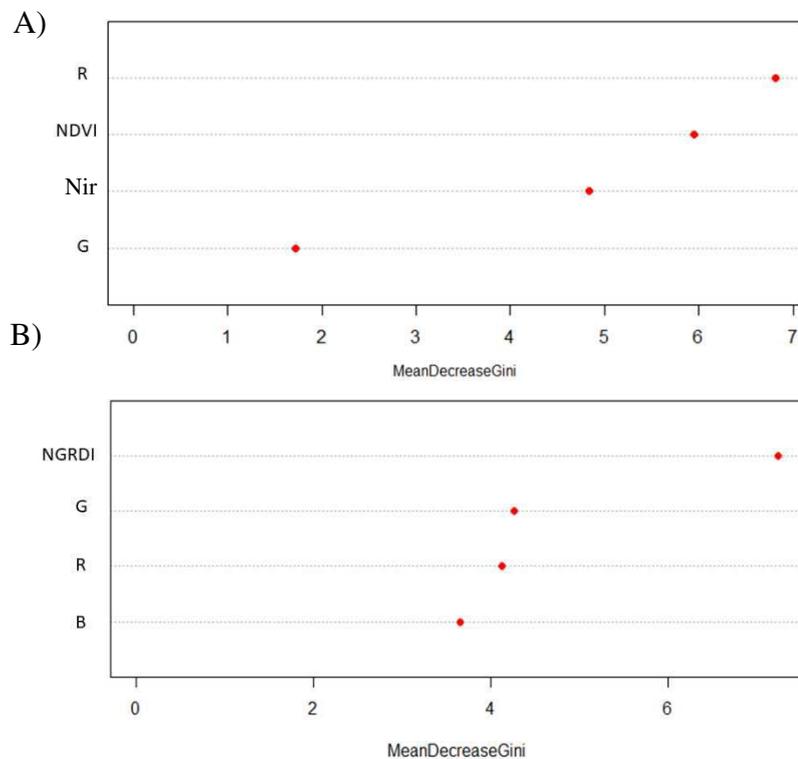


Figura 8: Importância das variáveis predictoras por *MDG* na área 1: método 1 (A) e método 2 (B).

As Figura 9A e 9B apresentam os boxplots para os dados de treinamento das bandas e

índices analisados para as 3 classes de cobertura identificadas na área 1 nos métodos 1 e 2, respectivamente. É possível notar que o *NDVI* calculado a partir das bandas R e *Nir* da câmera Mapir Survey 3W foi a melhor variável entre todas para captar as diferenças entre as classes. De forma geral todas as variáveis predictoras do método 1 se mostram mais eficazes na diferenciação entre as classes, quando se observa os *boxplots* das figuras 9A e 9B. Além disso, é possível perceber que o *NDVI* apresentou maiores valores para as classes Arbustiva arborea densa e Subarbustiva muito rala e menores valores para a classe de solo exposto, como era esperado para o índice.

Já para o método 2, o índice *NGRDI* obteve desempenho semelhante ao apresentar maiores valores para as classes Arbustiva arborea densa e Subarbustiva muito rala e menores para Solo exposto. Ainda pode-se notar através da Figura 9B, que as classes de Solo exposto e Subarbustiva muito rala apresentaram valores semelhantes para todas as variáveis, o que dificulta a diferenciação entre essas classes e consequentemente interfere no desempenho do método 2. Por fim, ao analisar as bandas do verde (G) e do vermelho (R) para as duas câmeras, é possível notar que as bandas obtidas a partir da câmera Mapir Survey 3W (método 1) apresentaram maior capacidade de diferenciação entre as classes, quando comparadas as mesmas bandas para a câmera nativa do drone DJI Mavic Pro.

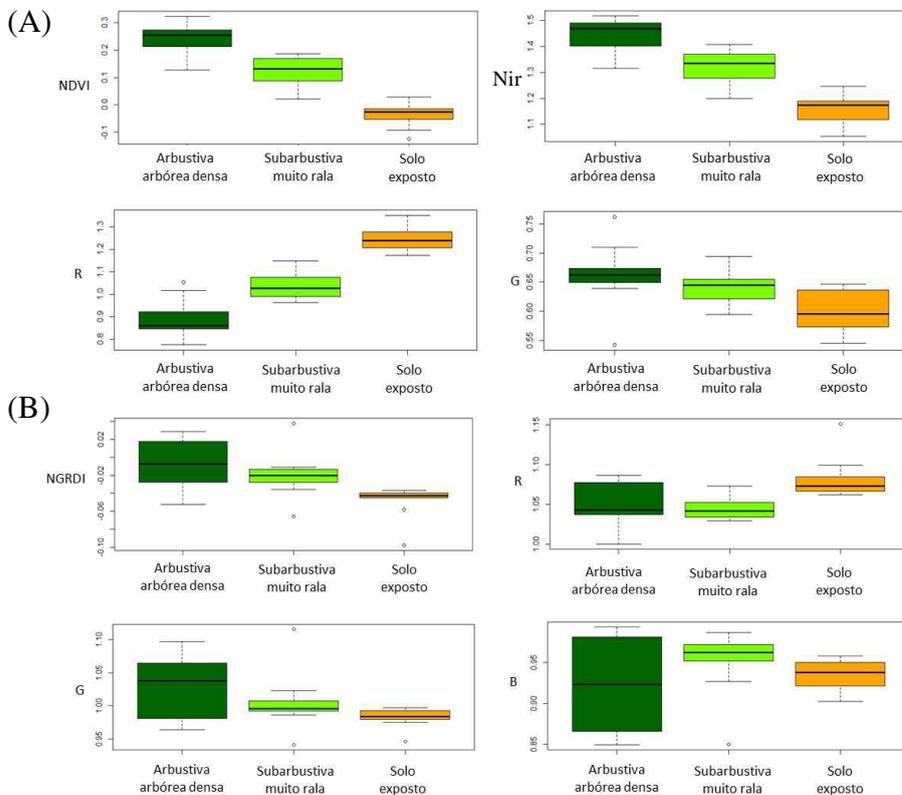
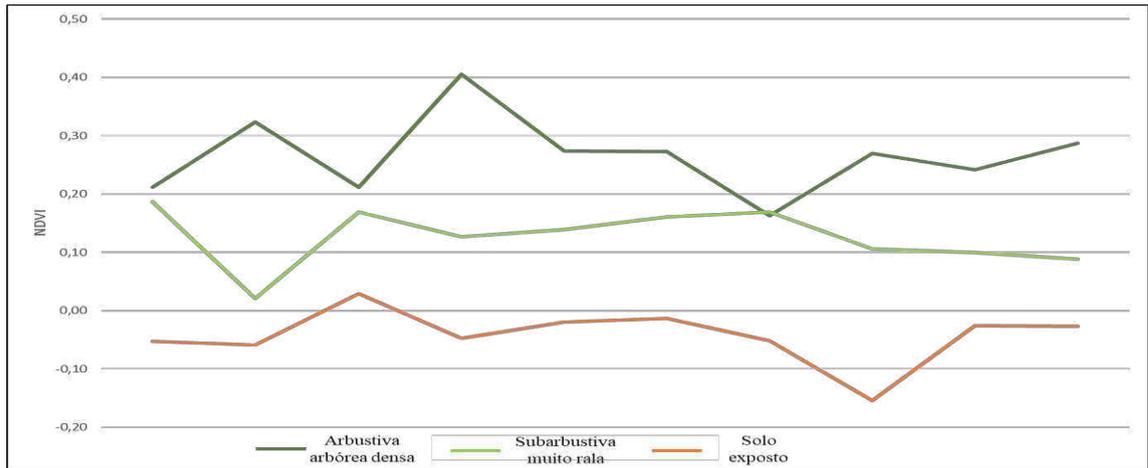


Figura 9: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 1: método 1 (A) e método 2 (B).

Ao realizar a análise dos IV's obtidos por cada câmera, pode-se notar que o NDVI apresenta melhor desempenho do que o NGRDI para diferenciação entre as classes, isso torna-se claro a partir da análise das Figuras 10A e 10B e pode ser explicado pelo fato do NDVI incorporar a banda do Infravermelho próximo (Nir) que capta informações importantes da fisiologia da vegetação não absorvidas pelas bandas Green e Red do espectro visível, as quais são utilizadas para o cálculo do NGRDI.

(A)



(B)

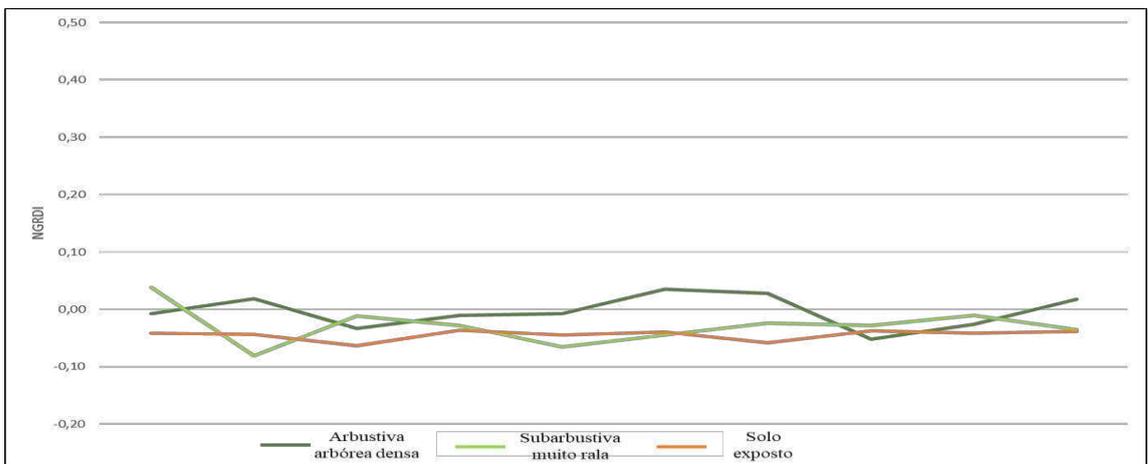
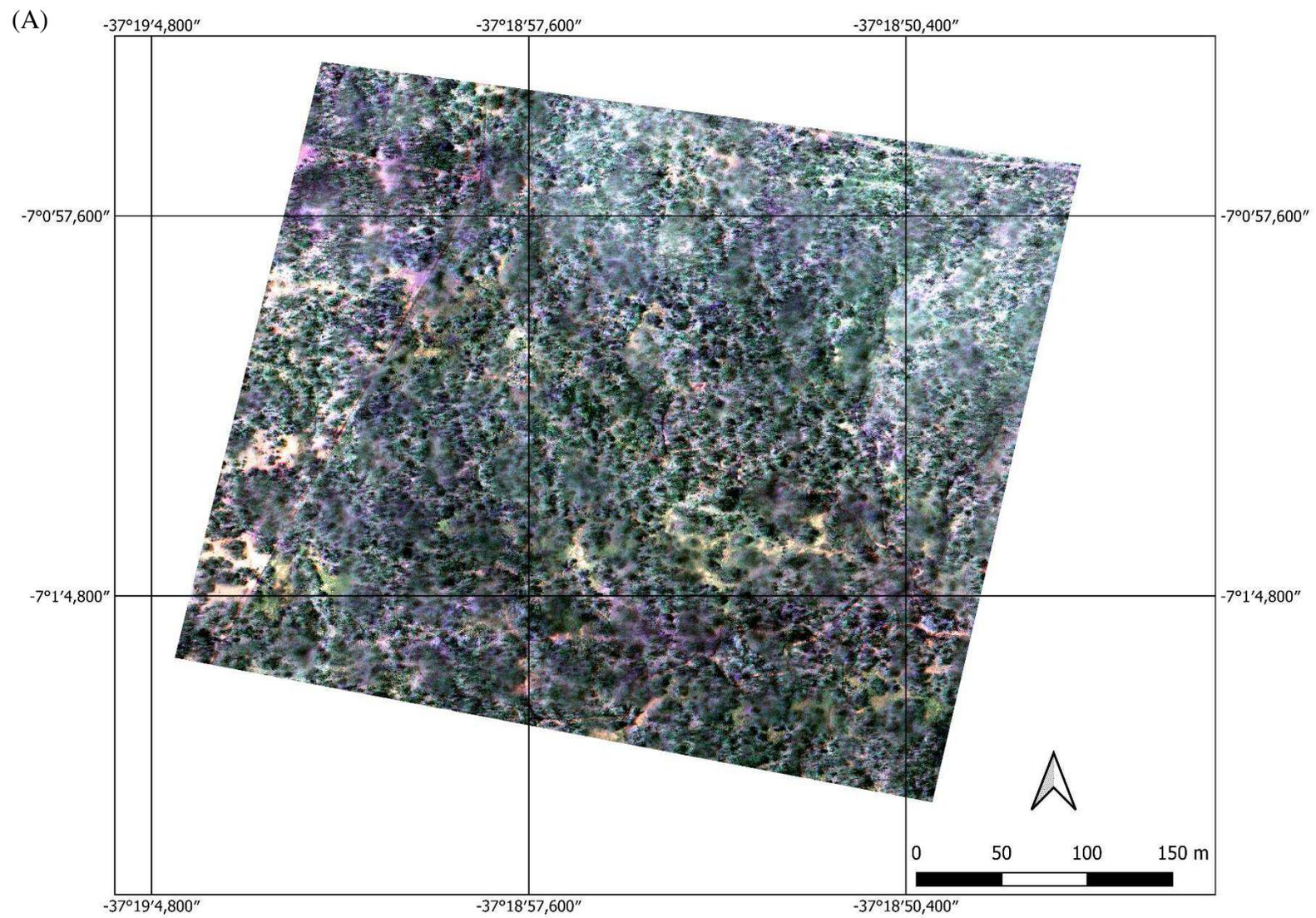


Figura 10: Comportamento dos IV's para os dados de treinamento em cada classe na área 1.

Por último, a partir de uma análise visual das Figuras 11A e 11B também é possível perceber que o método 1 foi mais preciso ao definir cada uma das classes de cobertura, já que visivelmente a classe Arbustiva arbórea densa é preponderante na área 1, o que corrobora com os resultados encontrados na classificação realizada a partir dos dados da câmera multiespectral.



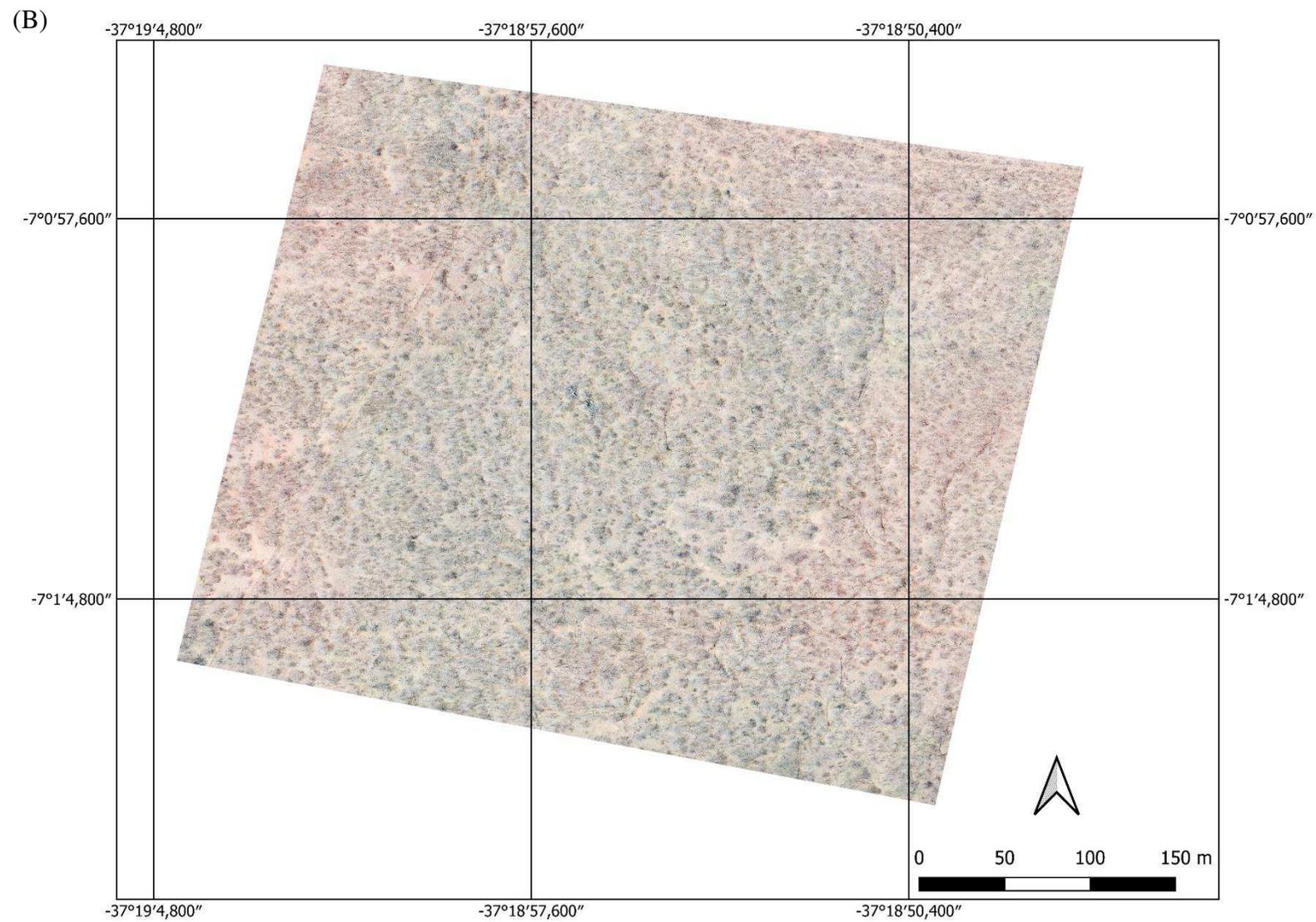
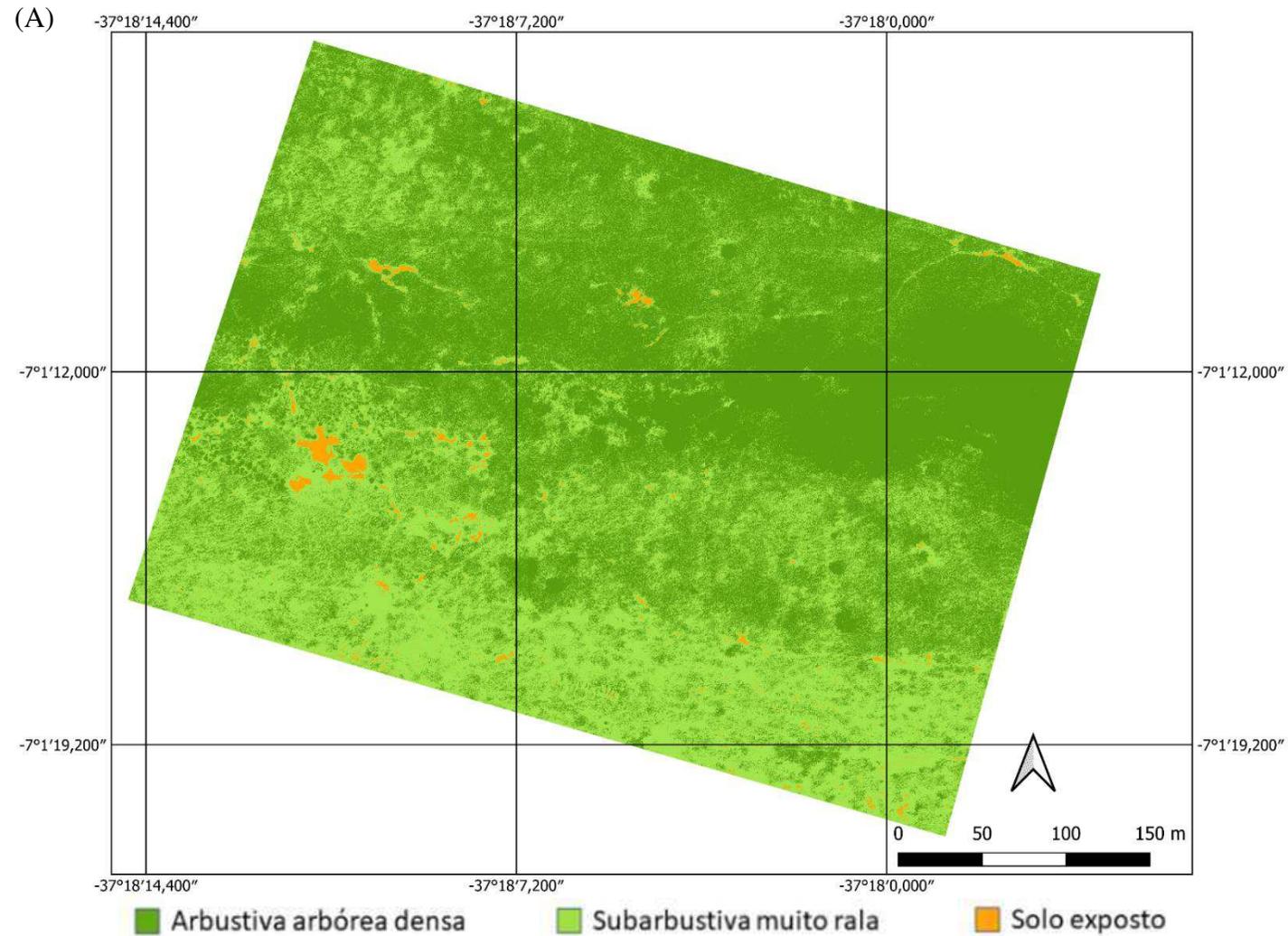


Figura 11: Composições para a área 1: RNG (A) e RGB (B).

## 5.2. Classificação para área 2

Para a área 2, a classificação pelo método 2 apresentou  $AG = 87,0\%$  e  $k = 0,80$ , valores bem próximos do método 1 que teve desempenho melhor com  $AG$  igual à  $93,0\%$  e  $k$  igual à  $0,90$ . As figuras 12A e 12B apresentam as classificações obtidas pelo método 1 e 2, respectivamente.



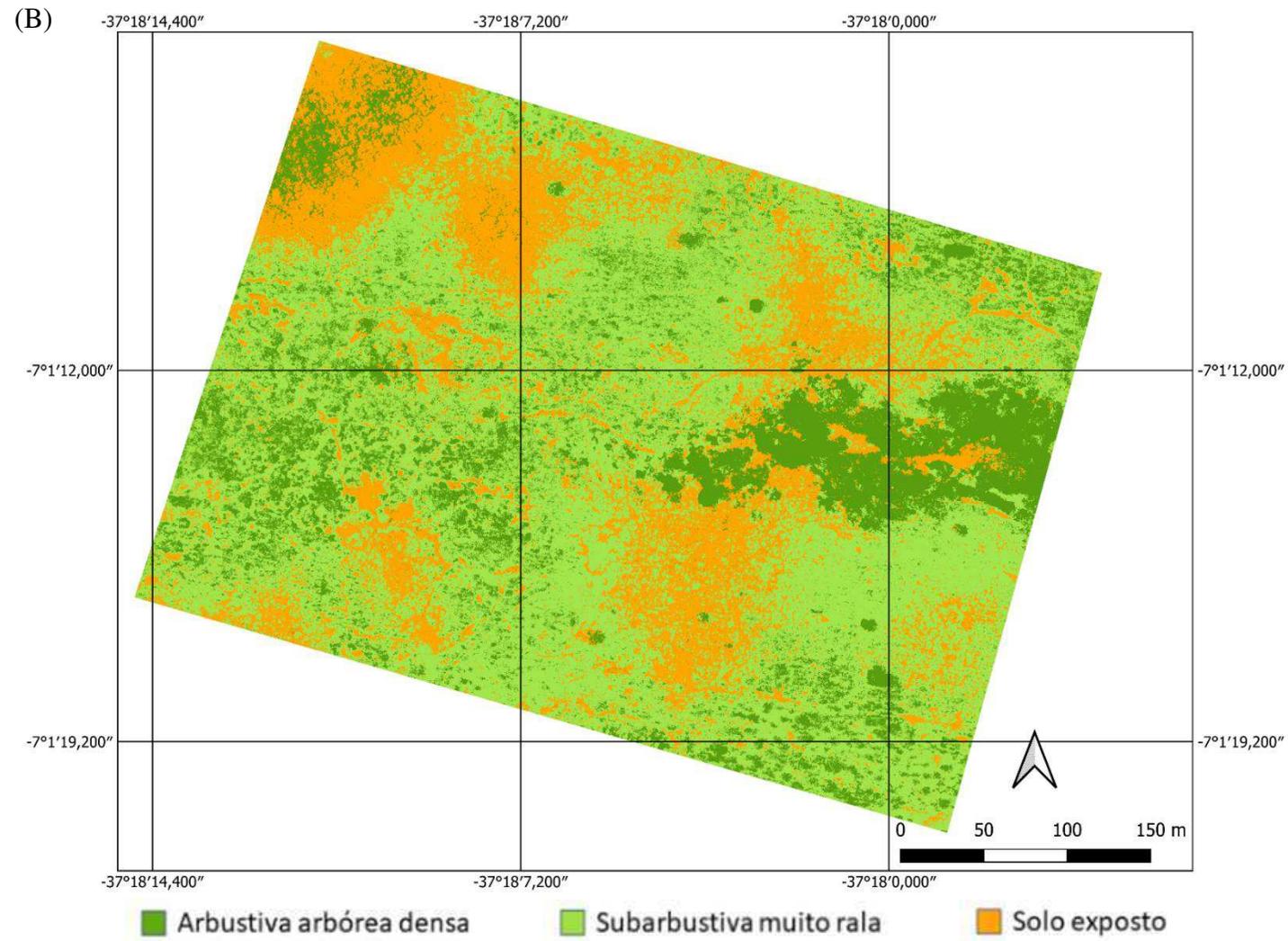


Figura 12: Classes de cobertura de solo para a área 1: método 1 (A) e método 2 (B).

Apesar do desempenho semelhante, é possível perceber que no método 2 (figura 12B), a classe de Solo exposto apresentou uma cobertura bem superior quando comparada com a classificação realizada pelo método 1 apresentada na figura 12A, o que indica uma dificuldade do método 2 na separação das classes. Além disso, houveram definições diferentes para as classes Arbustiva arbórea densa e Subarbustiva muito rala nos dois métodos. Essas diferenças podem ser melhor visualizadas através das figuras 13A e 13B.

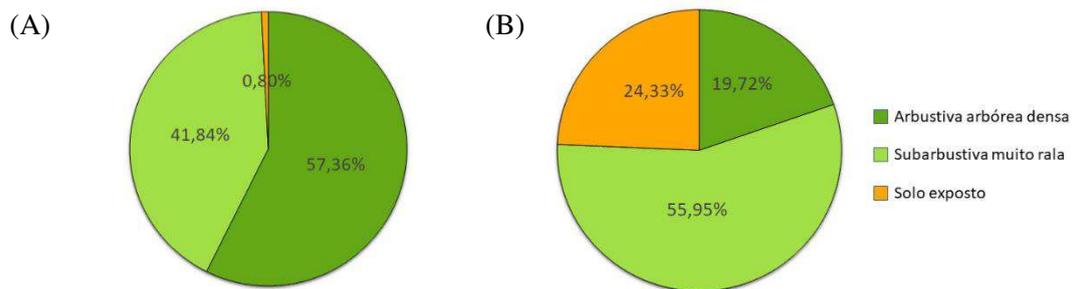


Figura 13: Distribuição percentual das classes de cobertura para a área 2: método 1 (A) e método 2 (B).

A partir dos gráficos apresentados nas figuras 13A e 13B é possível perceber que a classe Arbustiva arbórea densa foi a que apresentou maior diferença na distribuição espacial entre os métodos, seguida por Solo exposto e Subarbustiva muito rala, nessa ordem. Isso pode ser explicado pela matriz de confusão de cada método, indicada nas tabelas 5A e 5B.

Tabela 5: Matrizes de confusão para área 2.

(A)

Matriz de confusão - método 1					
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)
Arbustiva arbórea densa	5	0	0	5	100,0
Subarbustiva muito rala	0	5	0	5	100,0
Solo exposto	0	1	4	5	80,0
Total	5	6	4	15	-
PP (%)	100,0	83,0	100,0	AG (%)	93,0
k	0,90				

(B)

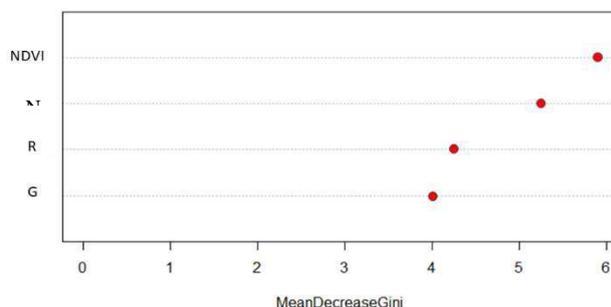
Matriz de confusão - método 2					
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)
Arbustiva arbórea densa	5	0	0	5	100,0
Subarbustiva muito rala	0	4	1	5	80,0
Solo exposto	0	1	4	5	80,0
Total	5	5	5	15	-
PP (%)	100	80	80	AG (%)	87,0
k	0,80				

(A) Método 1 e (B) Método 2.

A classe que apresentou melhor desempenho nos dois métodos foi a Arbustiva arbórea densa com valores de PP e PU de 100,0%, seguida pela classe Subarbustiva muito rala que apresentou valores de PP de 83,0% e 80,0% e valores de PU iguais à 100,0% e 80,0%, nos métodos 1 e 2, respectivamente. Para esta área a classe Solo exposto teve o pior desempenho.

Além disso, através da análise do *MDG* é possível notar que a variável preditora com maior importância no método 1 para a área 2 foi o *NDVI*, seguido pelo *Nir*. A variável menos importante foi a banda G, como mostra a figura 14A. No método 2 a variável mais importante foi a banda B, seguida pelo *NGRDI*, como indicado na figura 14B.

(A)



(B)

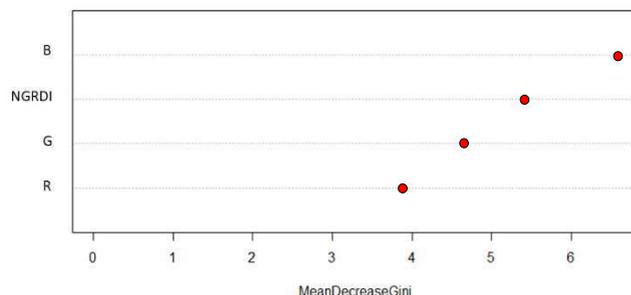


Figura 14: Importância das variáveis preditoras por *MDG* na área 2: método 1 (A) e método 2 (B).

As Figura 15A e 15B apresentam os boxplots para os dados de treinamento das bandas e índices analisados para as 3 classes de cobertura identificadas na área 2 nos métodos 1 e 2, respectivamente. É possível notar que o *NDVI* calculado a partir das bandas R e *Nir* da câmera

Mapir Survey 3W foi a melhor variável entre todas para captar as diferenças entre as classes. De forma geral todas as variáveis predictoras do método 1 se mostram mais eficazes na diferenciação entre as classes, quando se observa os *boxplots* das Figuras 15A e 15B. Além disso, é possível perceber que o *NDVI* apresentou maiores valores para as classes Arbustiva arbórea densa e Subarbustiva muito rala e menores valores para a classe de solo exposto, como era esperado para o índice. Já para o método 2, o índice *NGRDI* não obteve desempenho semelhante ao apresentar valores maiores para a classe de solo exposto e menores para as classes Arbustiva arbórea densa e Subarbustiva muito rala, o que não é esperado para este índice. Ainda pode-se notar através da Figura 15B, que as classes de Solo exposto e Subarbustiva muito rala apresentaram valores semelhantes para todas as variáveis, o que dificulta a diferenciação entre essas classes e consequentemente interfere no desempenho do método 2.

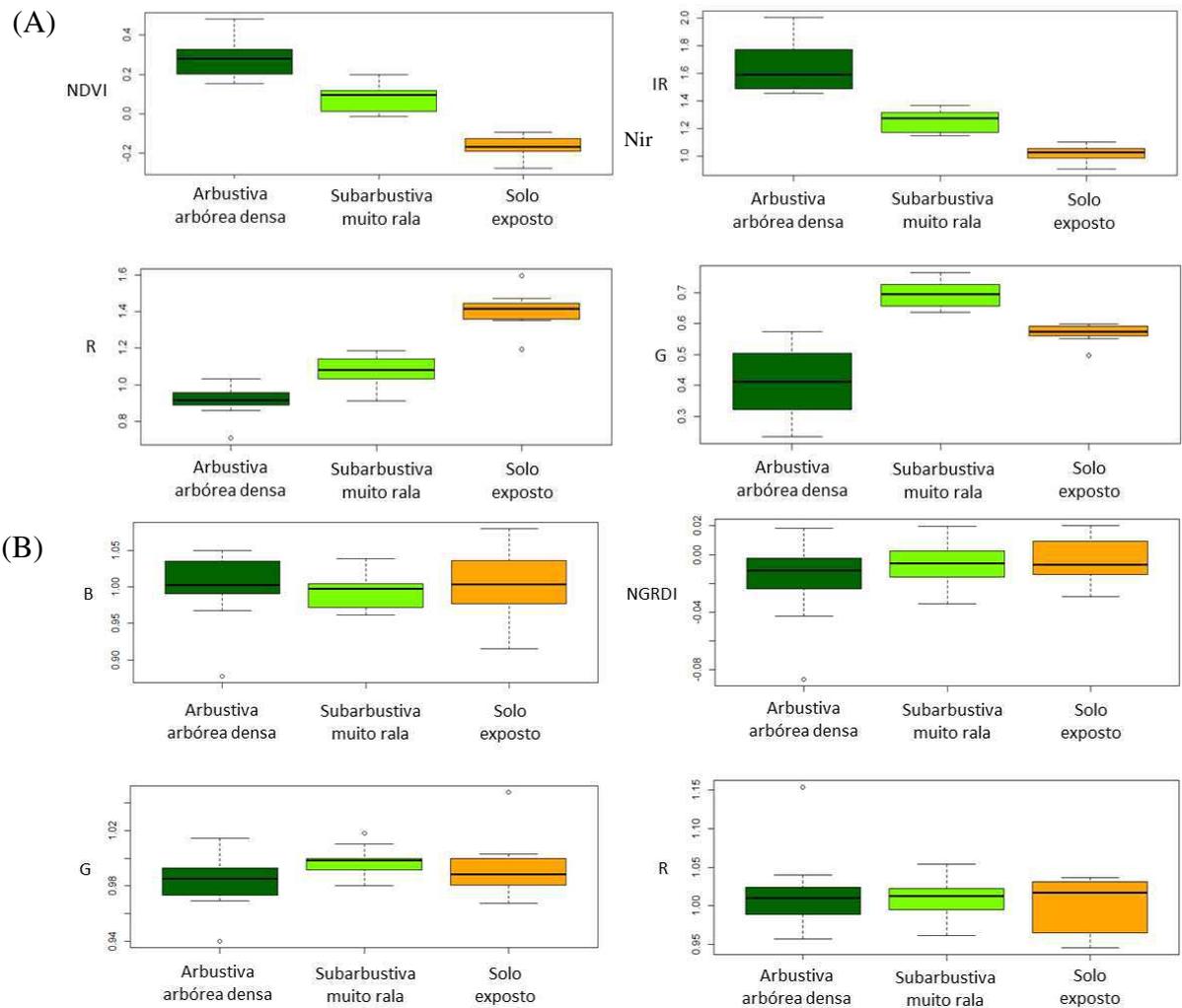
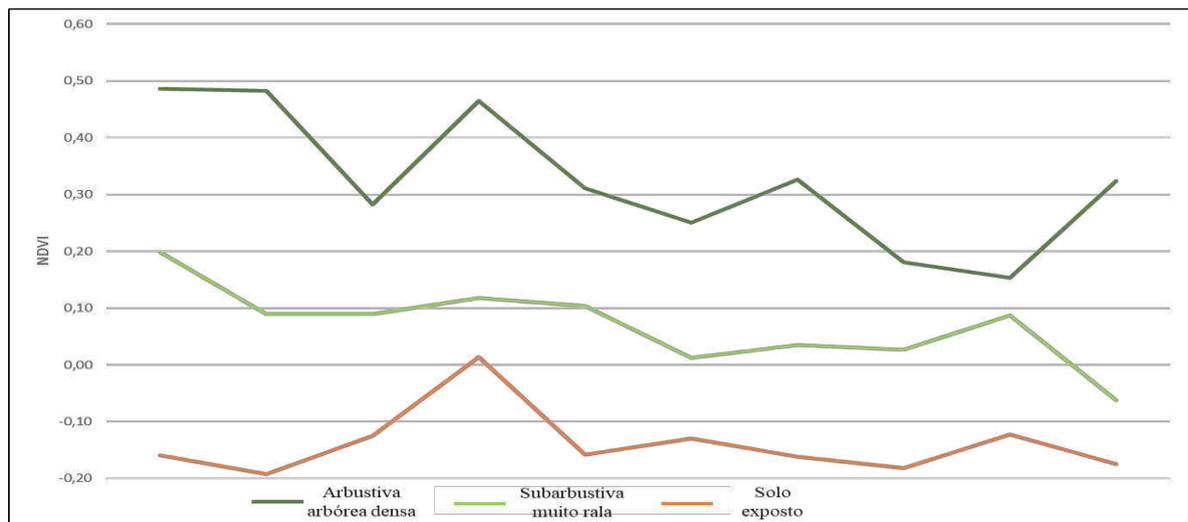


Figura 15: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 2: método 1 (A) e método 2 (B).

Da análise gráfica dos IV's apresentada nas Figuras 16A e 16B observa-se comportamento semelhante ao que ocorreu na área 1, com o NDVI apresentando maior capacidade de distinção entre as Classes. Além disso, nesta área o NDVI apresentou maiores valores na Classe Arbustiva arbórea densa quando comparado com os valores para a mesma classe na área 1, o que pode ser explicado pela proximidade da área 2 com um corpo hídrico, o que torna a vegetação naquela área mais densa e verde.

(A)



(B)

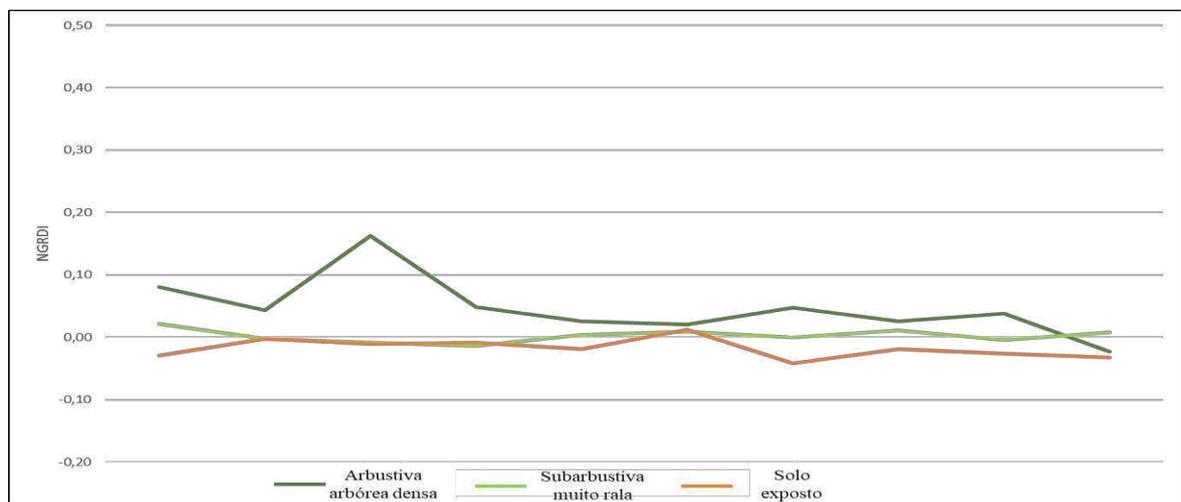
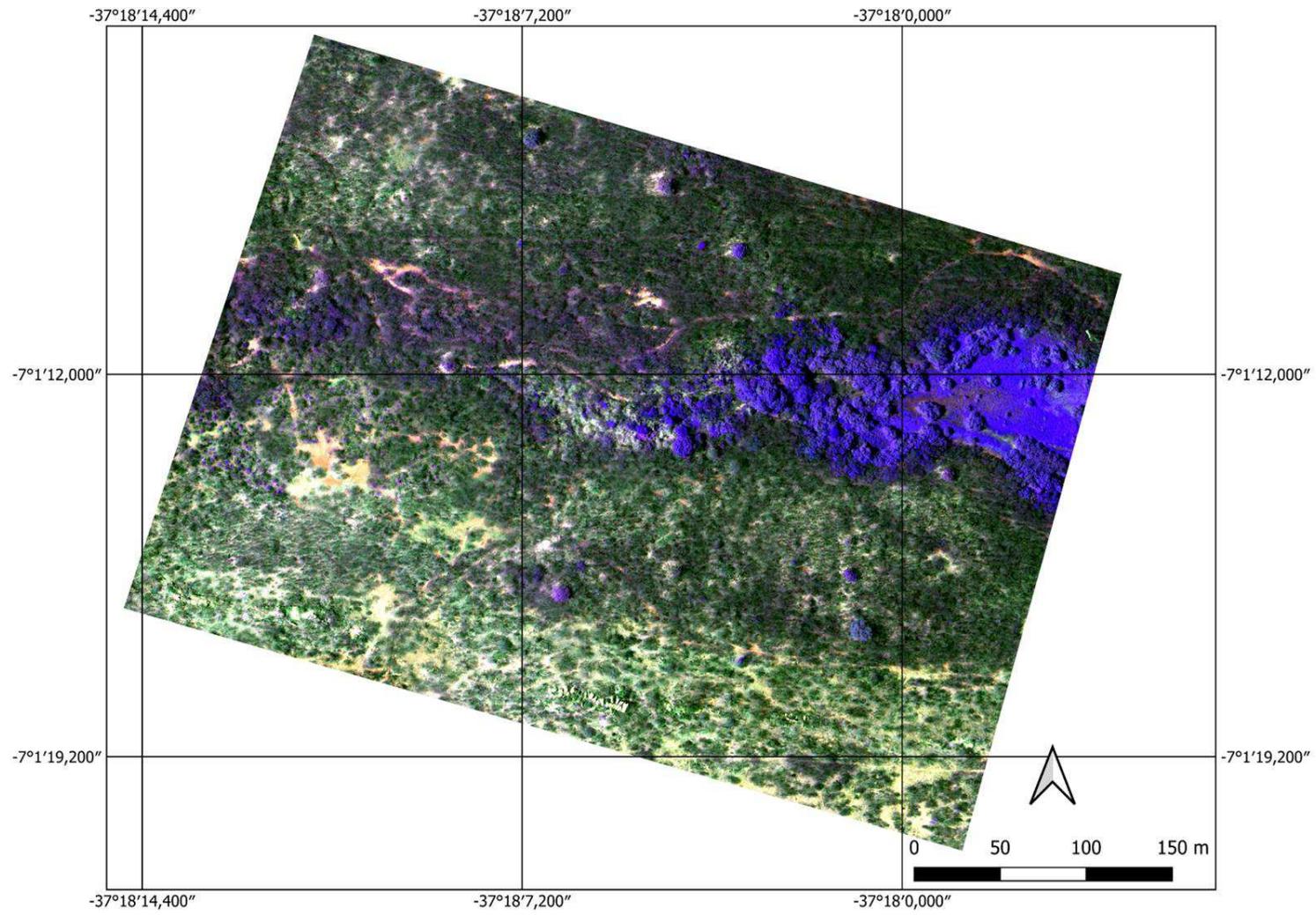


Figura 16: Comportamento dos IV's para os dados de treinamento em cada classe na área 2

Ainda, a partir das Figuras 17A e 17B, também é possível perceber que o método 1 foi mais preciso ao definir cada uma das classes de cobertura, já que visivelmente a classe Arbustiva arbórea densa é preponderante na área 2, assim como ocorre para a área 1.

(A)



(B)

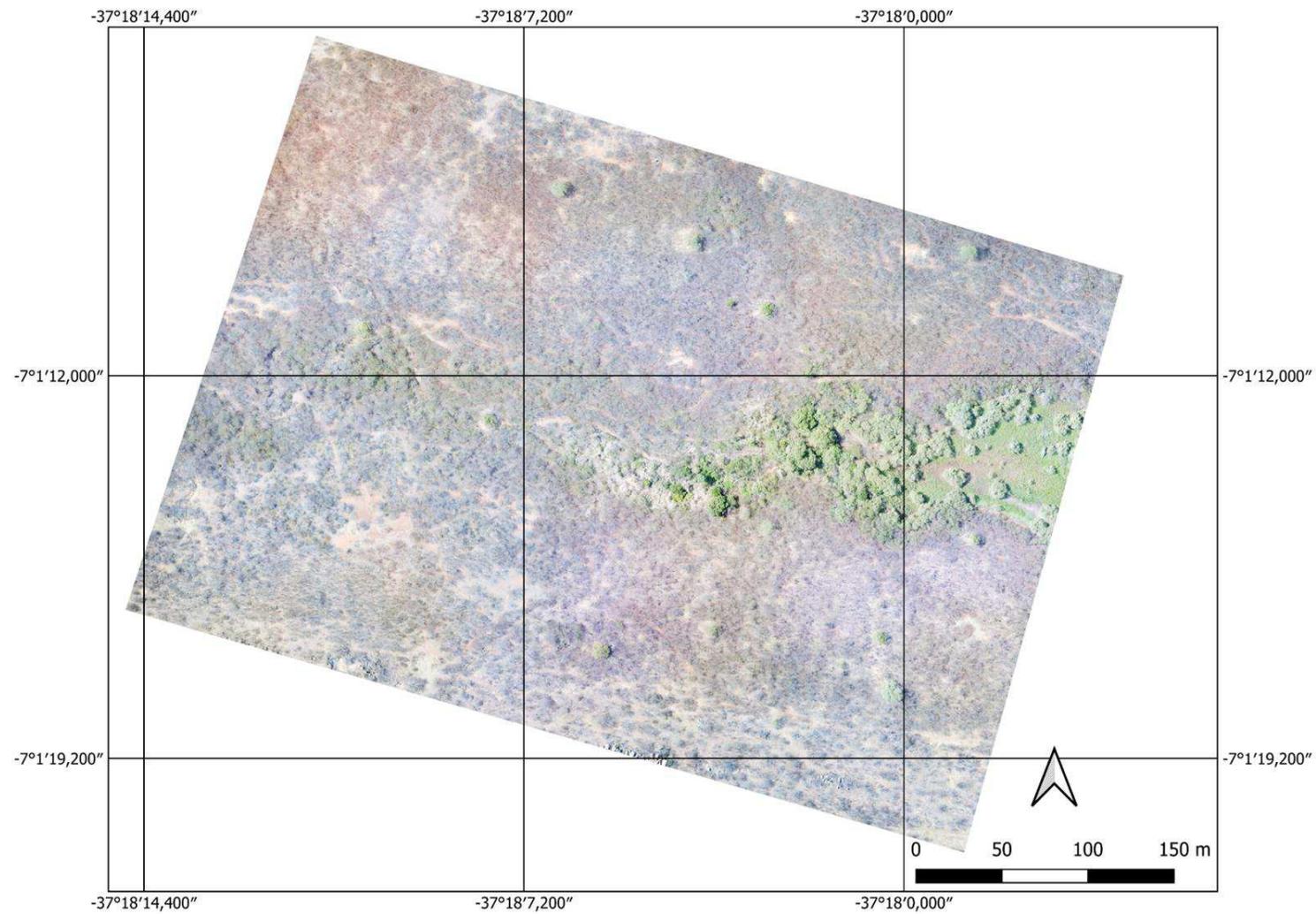
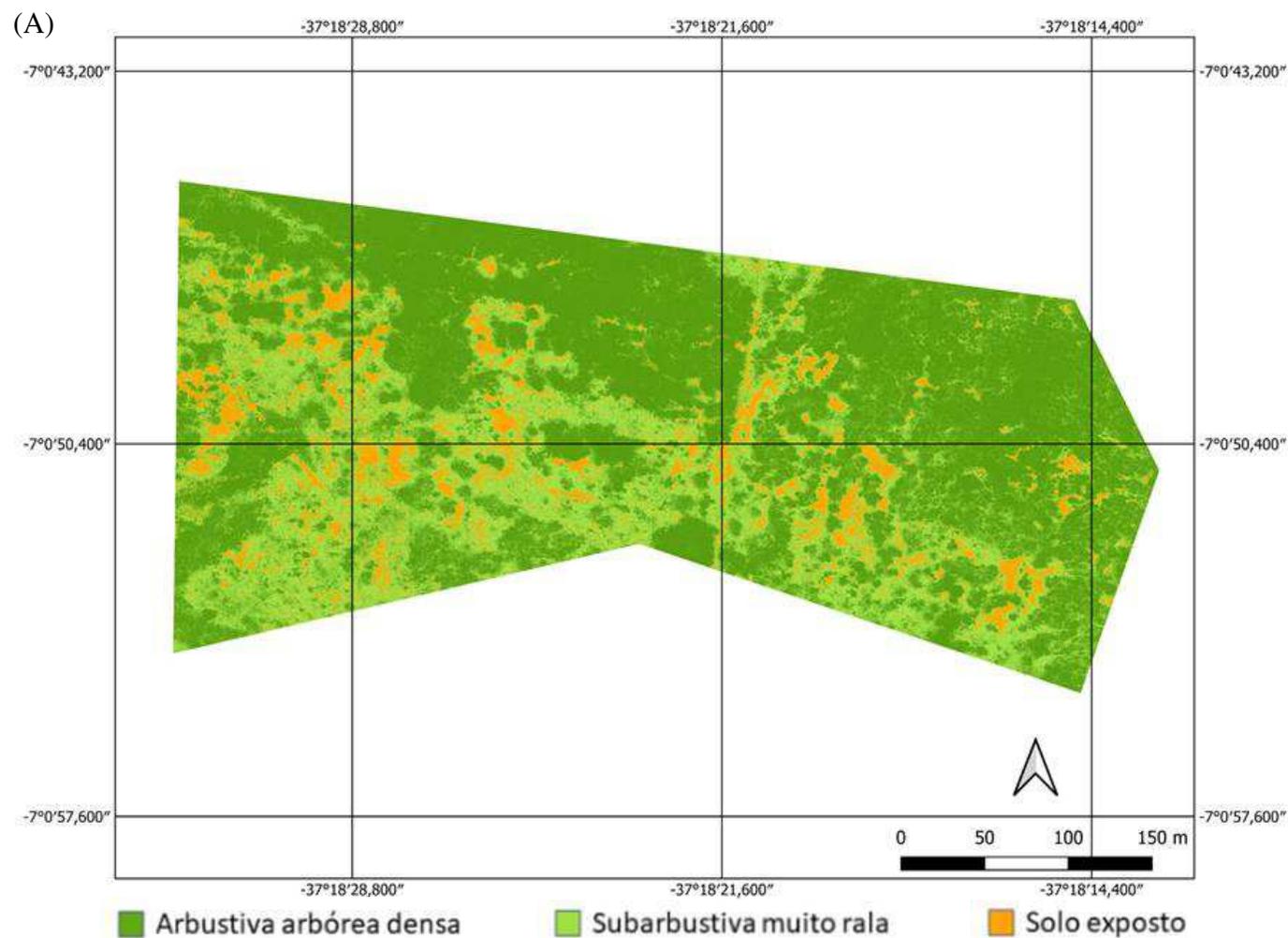


Figura 17: Composições para a área 2: RNG (A) e RGB (B).

### 5.3. Classificação para área 3

Na área 3, a classificação pelo método 1 apresentou  $AG = 100,0\%$  e  $k = 1,00$ , melhores resultados entre as áreas e métodos avaliados. Já o método 2 para esta área apresentou  $AG = 67,0\%$  e  $k = 0,50$ , tendo o segundo pior desempenho do método para as três áreas avaliadas, ficando atrás apenas da área 1. Com a diferença no desempenho nos dois métodos era esperado que a distribuição espacial das classes fosse bastante distante entre eles, o que pode ser notado nas Figuras 18A e 18B.



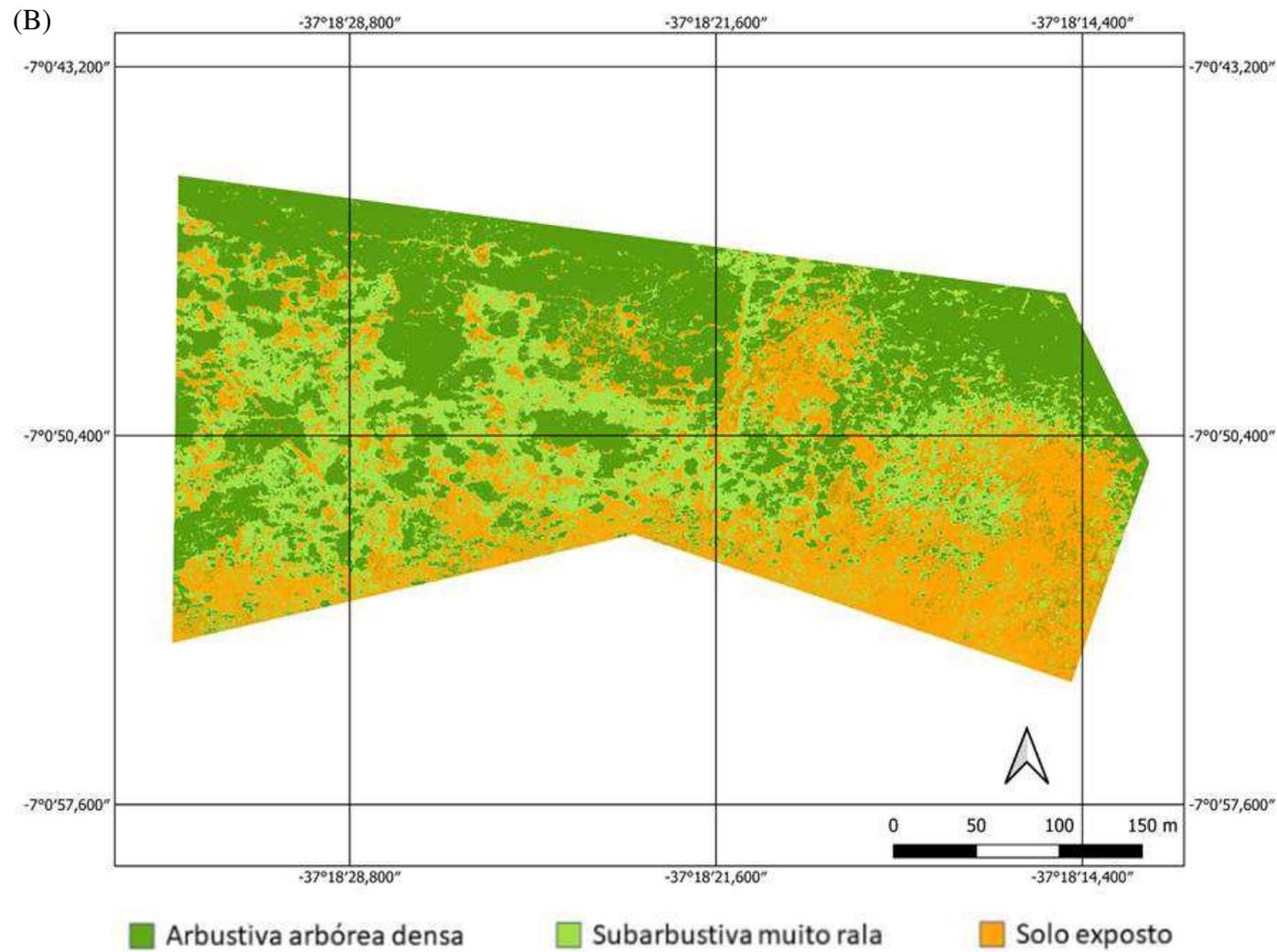


Figura 18: Classes de cobertura de solo para a área 3: método 1 (A) e método 2 (B).

A classe que se apresentou de forma mais semelhante nos dois métodos, foi a Subarbustiva muito rala, enquanto as classes Arbustiva arbórea densa e Solo exposto obtiveram distribuição espacial bastante distinta. Essas diferenças podem ser melhor visualizadas através das figuras 19A e 19B.

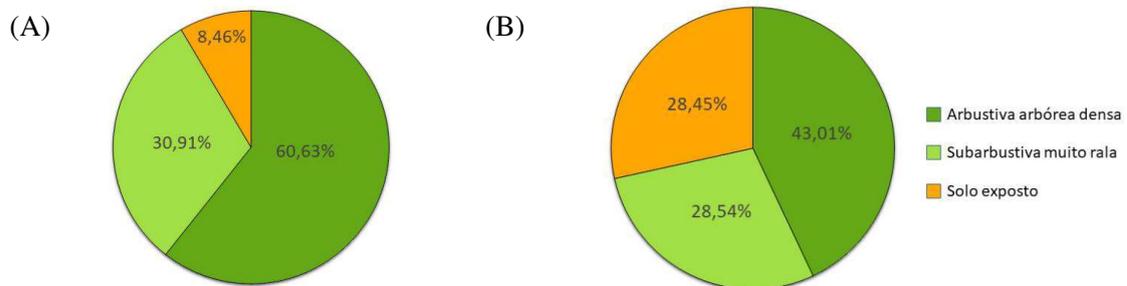


Figura 19: Distribuição percentual das classes de cobertura para a área 2: método 1 (A) e método 2 (B).

A partir dos gráficos apresentados nas figuras 19A e 19B é possível perceber que a classe Solo exposto foi a que apresentou maior diferença entre os métodos, seguida por Arbustiva arbórea densa e Subarbustiva muito rala, nessa ordem. Isso pode ser explicado pela matriz de confusão dos dados de treinamento para as duas classes, indicadas nas tabelas 6A e 6B.

Tabela 6: Matrizes de confusão para área 3.

(A) Matriz de confusão - método 1						
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)	
Arbustiva arbórea densa	5	0	0	5	100,0	
Subarbustiva muito rala	0	5	0	5	100,0	
Solo exposto	0	0	5	5	100,0	
Total	5	5	5	15	-	
PP (%)	100,0	100,0	100,0	AG (%)	100,0	
k	1,0					

(B) Matriz de confusão - método 2						
Classes	Arbustiva arbórea densa	Subarbustiva muito rala	Solo exposto	Total	PU (%)	
Arbustiva arbórea densa	5	0	1	6	83,0	
Subarbustiva muito rala	0	3	2	5	60,0	
Solo exposto	0	2	2	4	50,0	
Total	5	5	5	15	-	
PP (%)	100,0	60,0	40,0	AG (%)	67	
k	0,50					

(A) Método 1 e (B) Método 2.

Nesta área, o método 2 apresentou valores de PP e PU bem abaixo do método 1 para as classes Subarrestiva muito rala e Solo exposto, sendo a última de pior desempenho com valores de PP e PU de 40,0% e 50,0%, respectivamente. Esse resultado é condicente com o apresentado na área 2. O método 1 teve o melhor desempenho geral com valores de PP e PU iguais à 100,0% para todas as classes.

Além disso, através da análise do *MDG* é possível notar que a variável preditora com maior importância no método 1 para a área 3 foi o *NDVI*, seguido pelo *Nir*, mesmo resultado obtido na área 2. A variável menos importante foi a banda *G*, como mostra a figura 20A. No método 2 a variável mais importante foi a banda *R*, seguida pelo *B*, como indicado na figura 20B.

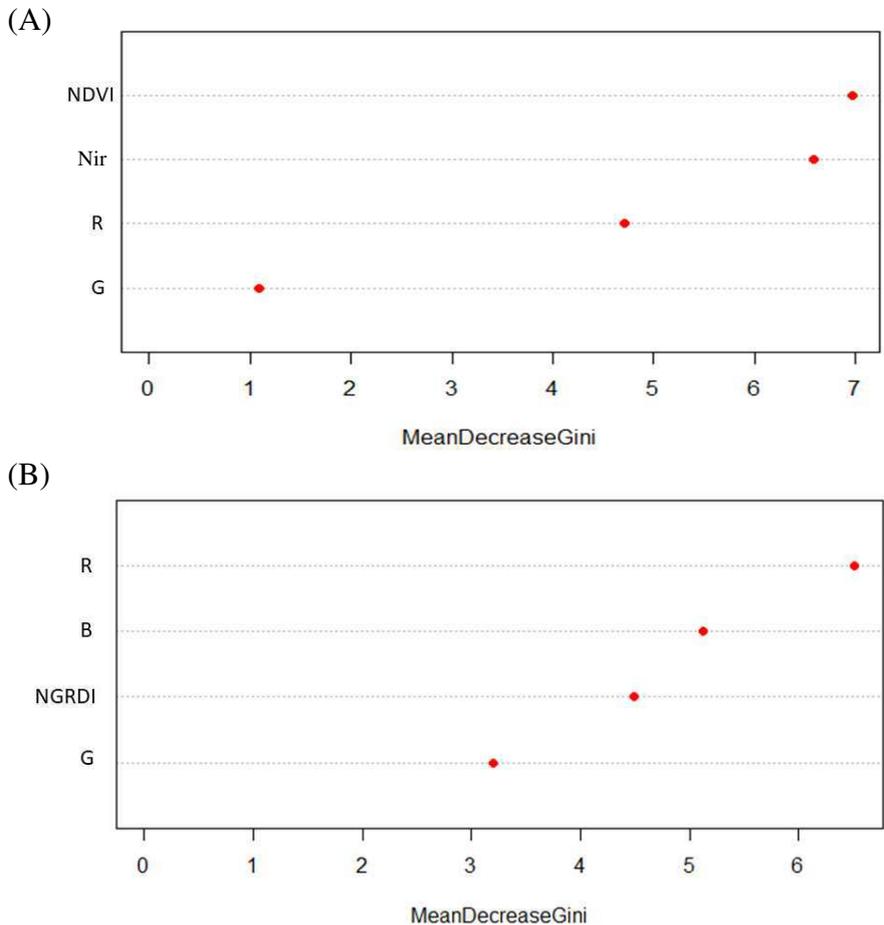


Figura 20: Importância das variáveis predictoras por *MDG* na área 3: método 1 (A) e método 2 (B).

As Figura 21A e 21B apresentam os *boxplots* para os dados de treinamento das bandas e índices analisados para as 3 classes de cobertura identificadas na área 3 nos método 1 e 2, respectivamente. É possível notar que o *NDVI* foi a melhor variável entre todas para captar as diferenças entre as classes. De forma geral todas as variáveis predictoras do método 1 se

mostram mais eficazes na diferenciação entre as classes, quando se observa os *boxplots* das figuras 21A e 21B. Além disso, é possível perceber que o *NDVI* apresentou maiores valores para as classes Arbustiva arborea densa e Subarbustiva muito rala e menores valores para a classe de solo exposto, como era esperado para o índice. Já para o método 2, o índice *NGRDI* não obteve desempenho semelhante ao apresentar valores maiores para a classe de solo exposto e menores para as classes Arbustiva arborea densa e Subarbustiva muito rala, o que não é esperado para este índice. Ainda pode-se notar através da Figura 21B, que as classes de Solo exposto e Subarbustiva muito rala apresentaram valores semelhantes para todas as variáveis, o que dificulta a diferenciação entre essas classes e consequentemente interfere no desempenho do método 2.

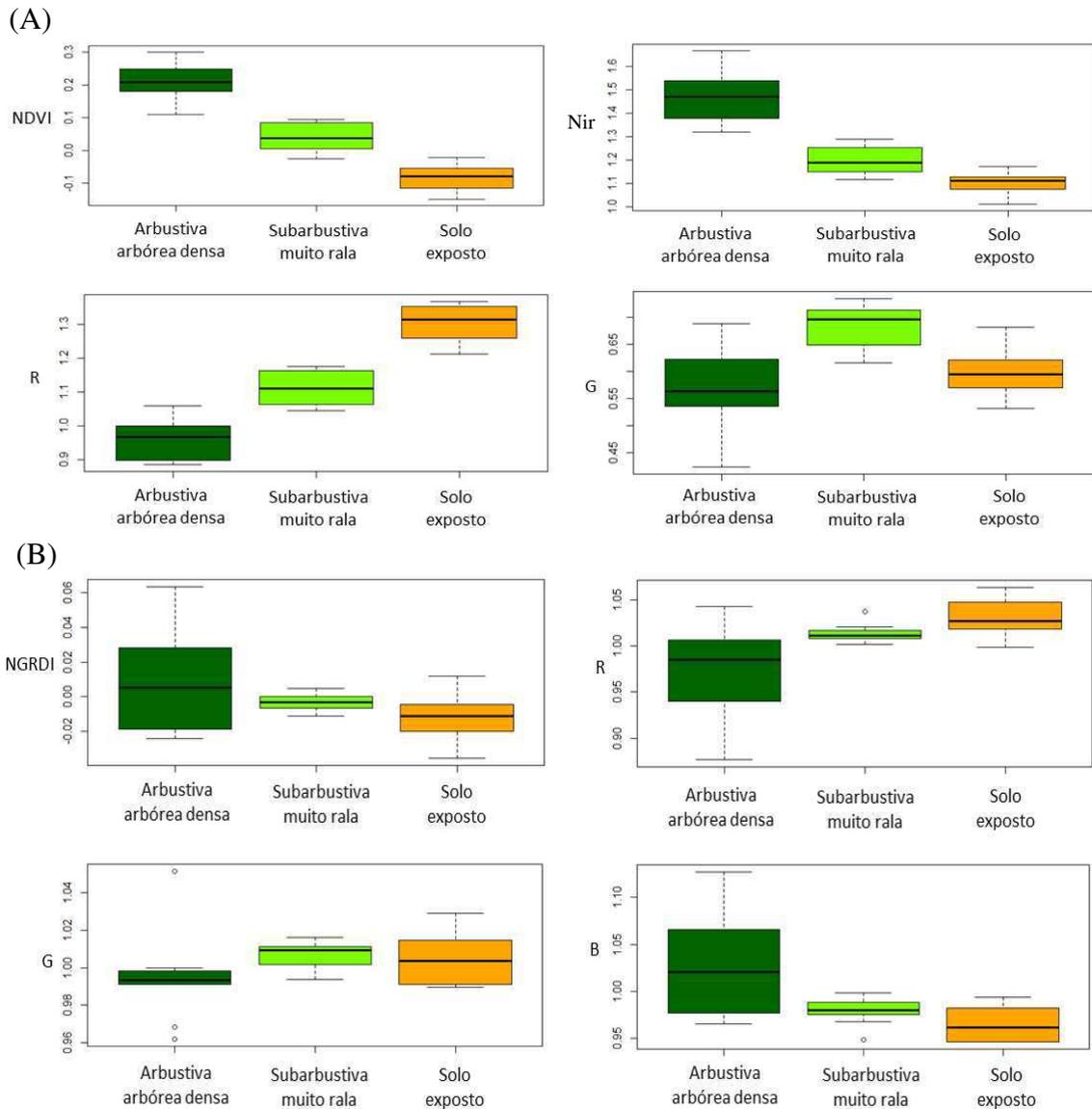
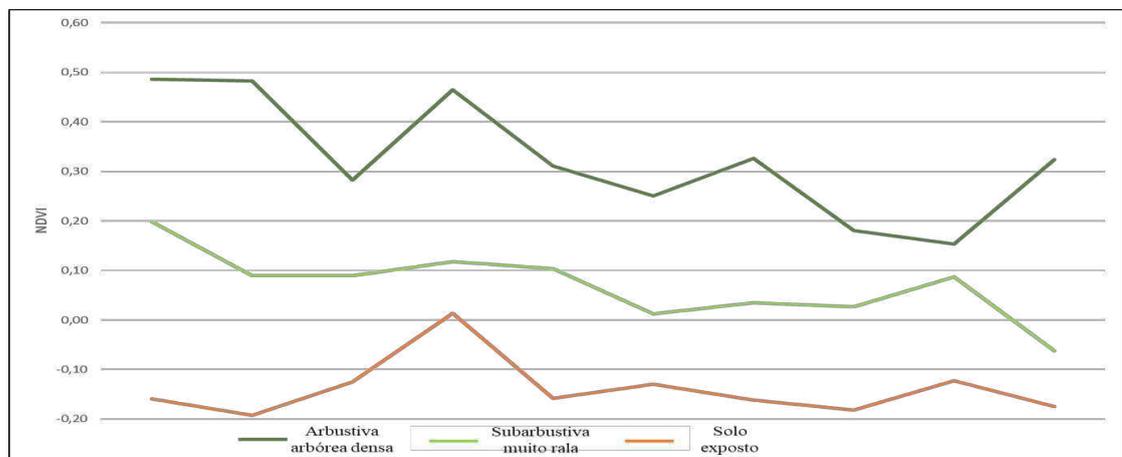


Figura 21: Variação dos valores dos parâmetros avaliados em cada classe nos dados de treinamento do modelo para a área 3: método 1 (A) e método 2 (B).

Em relação ao comportamento dos IV's observa-se, através das Figuras 22A e 22B, que o NDVI apresentou maior capacidade de distinção entre as Classes. Além disso, a área 3 apresentou os menores valores de NDVI entre as áreas avaliadas, o que pode ser explicado pela maior porcentagem da classe Solo exposto nesta área, já que para essa classe este é um comportamento esperado para o índice. As classes Arbustiva arbórea densa e Subarbustiva muito rala também apresentaram os menores valores entre todas as áreas, o que indica que o vigor vegetativo nesta área é pequeno, que pode estar sendo influenciado pelo avanço da mancha urbana nesta área.

(A)



(B)

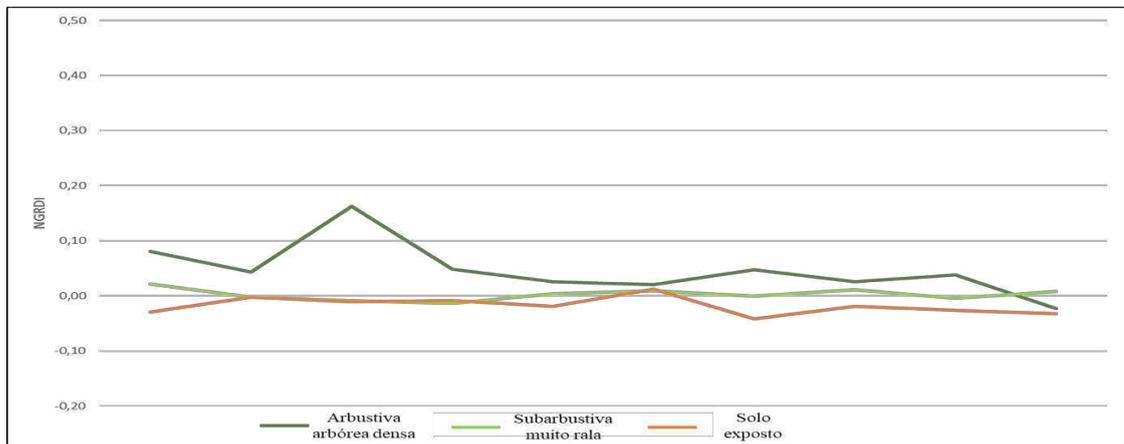
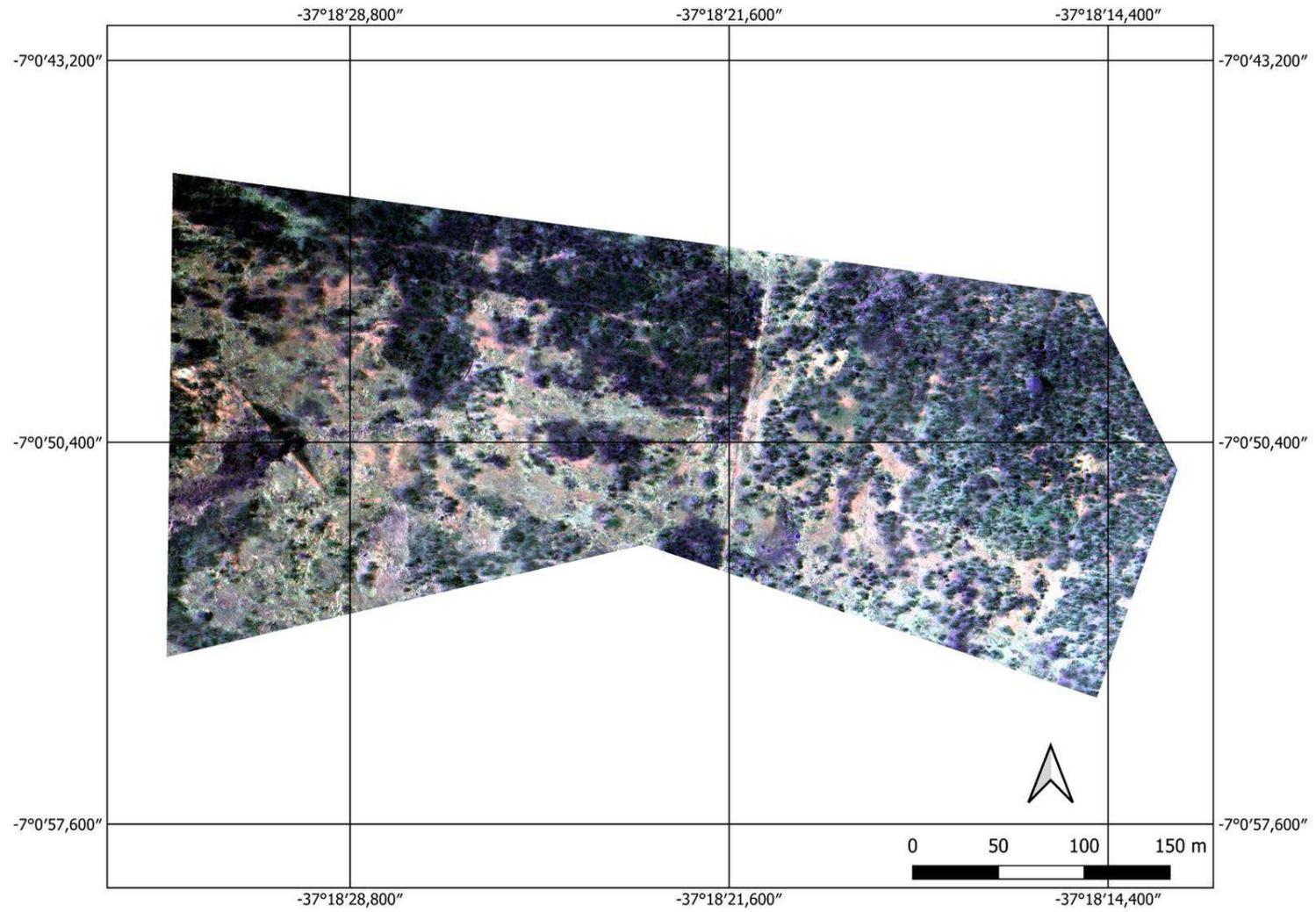


Figura 22: Comportamento dos IV's para os dados de treinamento em cada classe na área 3.

Por último, realizando uma análise visual das figuras 23A e 23B, também é possível perceber que o método 1 foi mais preciso ao definir cada uma das classes de cobertura, já que visivelmente a classe Arbustiva arbórea densa é preponderante, assim como ocorreu nas áreas 1 e 2.

(A)



(B)

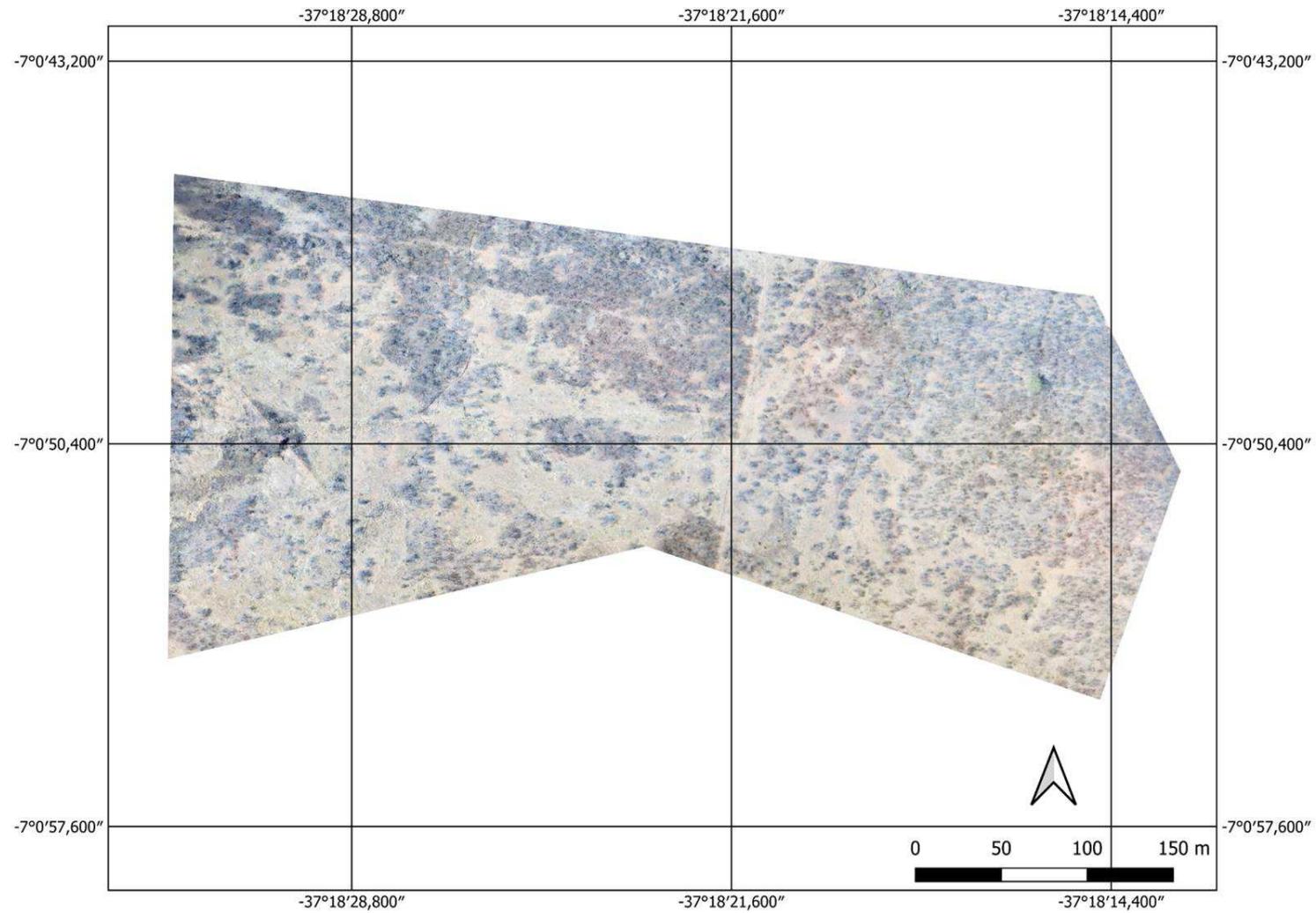


Figura 23: Composições para a área 3: RNG (A) e RGB (B).

#### 5.4. Análise comparativa do desempenho dos métodos avaliados

Ao avaliar as três áreas, este trabalho indica que o método 1 obteve melhor desempenho na classificação para a Caatinga, isto pode ser validado pelos índices AG e k apresentados pelos dois métodos, os quais estão indicados na Figura 24. De forma geral o método 1 apresentou melhor desempenho ao avaliar as variações dos índices AG e k nas três áreas, sendo a diferença para o método 2 maior nas áreas 1 e 3, onde o método 1 obteve AG de 93,0% e 100,0%, respectivamente.

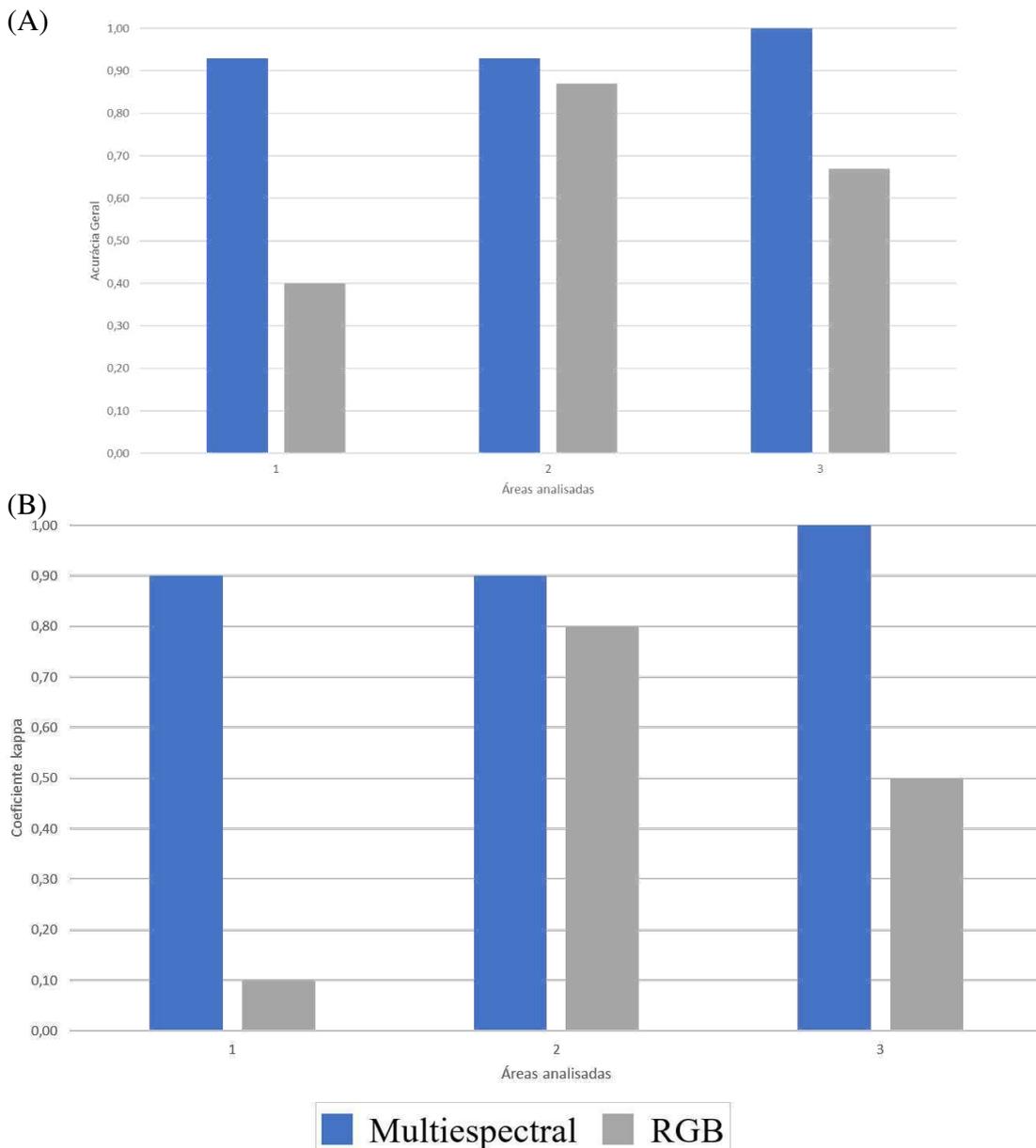


Figura 24: Índices de precisão da classificação: Acurácia Geral – AG (A) e Coeficiente kappa – k (B).

No estudo de Andrade *et al.* (2021), os autores classificaram uma área de Caatinga utilizando o *Random Forest* tendo como parâmetros de entrada dados multiespectrais multi-temporais e monotemporais obtidos do satélite *LANDSAT* cuja resolução espacial é de 30 metros. Na classificação aplicando dados multiespectrais mono-temporais os autores verificaram uma Acurácia Global de 81,40% e coeficiente Kappa de 0,78, valores inferiores aos encontrados no método 1 deste trabalho, com valores médios de AG = 95,33% e k = 0,93. A diferença nesses valores pode está associada à muitos fatores definidos em cada estudo, como: quantidade de classes, quantidade de amostras, resolução espacial dos dados, entre outros. Contudo, indica uma tendência de melhoria nos resultados ao utilizar dados com alta resolução espacial, necessitando de mais estudos para a confirmação desta tendência.

Além disso, diversos estudos que aplicaram o *Random Forest* para classificação de áreas vegetativas e de cultivo obtiveram resultados semelhantes ao desta pesquisa ao utilizar dados multiespectrais de alta resolução espacial obtidos por VANT. Por exemplo, no estudo de Ma *et al.* (2021) os autores avaliaram a eficiência de dois classificadores (*Random Forest* e *Support Vector Machine -SVM*) na distinção de cultivos em diferentes áreas agrícolas e observaram que em áreas com maior heterogeneidade (alta complexidade) os classificadores apresentavam redução nos índices de precisão (Kappa e Acurácia Global) em relação à áreas de cultivo mais homogêneas (baixa complexidade). Contudo, os valores obtidos em áreas de alta complexidade são considerados pelos autores como altos, com AG médio de 88,99% para o classificador *Random Forest*. Resultado semelhante foi encontrado no estudo de Alzhanov; Nugumanova (2024) onde os autores avaliaram a precisão do *Random Forest* modelado a partir de dados de alta resolução obtidos por VANT em áreas de cultivo e obtiveram valores de AG de 74,0% e Kappa de 75,0%. O que corrobora com os resultados encontrados nesta pesquisa.

Já na pesquisa de Pablo *et al.* (2018) os autores avaliaram a precisão de índices de vegetação obtidos a partir de câmeras RGB e multiespectrais acopladas em VANT para avaliar a Fração Vegetativa (FV) em cultivo de milho e identificaram que o índice que apresentou melhor desempenho foi o GNDVI (NDVI verde). Resultado que corrobora com este estudo no sentido de que a banda do Infravermelho presente nestes índices melhora a obtenção de informações a respeito da vegetação de forma geral.

Atualmente um dos produtos disponíveis gratuitamente para classes de cobertura do solo no Brasil é o Mapbiomas (Souza Jr. *et al.*, 2020). O projeto conta com um painel informativo que avalia a precisão obtida pelos estudos por ele realizados. O índice AG para todo o território brasileiro é de 91,30 %, e na região da Caatinga, onde este estudo foi realizado, é de 81,80% (Espindola *et al.*, 2021; Alencar *et al.*, 2020). A validação por comparação de um projeto consolidado a nível nacional e internacional demonstra que os resultados obtidos neste estudo são válidos, uma vez que os valores aqui encontrados foram superiores aos apresentados no Mapbiomas para a área de Caatinga.

## 6. CONCLUSÕES

Ao analisar todas as áreas ficou evidente que o método 1 testado neste trabalho obteve bom desempenho na classificação da cobertura do solo para a Caatinga com AG média de 95,33% e k médio de 0,93, diferentemente do método 2 que para a maioria das áreas não apresentou bons resultados, com AG média de 64,67% e k médio de 0,47. Isso indica que o uso de câmeras multiespectrais, como é o caso da Mapir Survey 3W, que possam fornecer bons índices como o NDVI, são excelentes ferramentas na análise vegetativa e consequentemente na classificação da cobertura do solo.

De forma geral, é possível afirmar que o método 1 mostrou-se eficiente nas áreas analisadas apresentando bons resultados, o que coloca o Random Forest como uma excelente ferramenta nos processos de classificação de cobertura do solo. Contudo, são necessários mais testes em áreas de Caatinga que possuam uma quantidade maior de padrões vegetativos para garantir a eficácia do modelo para qualquer porção de Caatinga. Além disso, outras variáveis podem ser inseridas na predição do modelo RF, como outras bandas espectrais, a exemplo da banda termal, e/ou a altura da copa das árvores obtida a partir de um Modelo Digital de Superfície (MDS) e de um Modelo Digital de Elevação (MDT), sabendo que o Random Forest tem a capacidade de receber dados de fontes diversas e com propriedades distintas, o que é uma das importantes vantagens desse classificador.

A qualidade dos dados alcançada com o uso de sensores acoplados em VANT é muitas vezes superior àquela conseguida com dados orbitais captados por satélites, o que torna essa abordagem vantajosa para obtenção precisa de informações sobre a cobertura do solo. Além disso, o mapeamento realizado com o VANT possibilita também alta precisão para aplicações em diversas áreas de Engenharia, como levantamento topográfico, estudos batimétricos de reservatórios, análise do fluxo de vazões em bacias hidrográficas, entre outros. Por fim, conclui-se que o uso de VANT com a câmera Mapir Survey 3W embarcada é uma ferramenta eficiente para classificação da cobertura de solo na Caatinga, com alta capacidade de distinção entre os padrões vegetativos, propiciada principalmente por sua característica multiespectral.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

**AESA.** Agência Executiva de Gestão das Águas do Estado da Paraíba. Disponível em: <<http://www.aesa.pb.gov.br/aesa-website/meteorologia-chuvas/>>. Acesso em 30 de novembro de 2023.

ALHASSAN, V.; HENRY, C.; RAMANNA, S.; STORIE, C. A deep learning framework for landuse/landcover mapping and analysis using multispectral satellite imagery. **Neural Computing and Applications**, p.116, 2019.

ALVES, B.L.T.; AZEVEDO, de V.P.; DOS SANTOS, C.A.C. Influence of climate variability on land degradation (desertification) in the watershed of the upper Paraíba River. **Theoretical and Applied Climatology**, v.127, 741-751, 2017.

ALZHANOV, A.; NUGUMANOVA, A. Crop classification using UAV multispectral images with gray-level co-occurrence matrix features. **Procedia Computer Science**, v.231, p.734-739, 2024.

ANDERSON, K.; GASTON, K.J. Lightweight Unmanned Aerial Vehicles Will Revolutionize Spatial Ecology. **Frontiers in Ecology and the Environment**, v.11, 138–146, 2013.

ANDRADE, J.; CUNHA, J.; SILVA, J.; RUFINO, I.; GALVÃO, C. Evaluating single and multi-date Landsat classifications of land-cover in a seasonally dry tropical forest. **Remote Sensing**, p.100-515, 2021.

ANGELIS, C. F. Princípios de Sensoriamento Remoto. **Instituto Nacional de Pesquisas Espaciais, Divisão de satélites e Sistemas Ambientais – INPE-DAS**, 2011.

ARAUJO, L de K.D.; QUEIROZ, F.M.M. CARACTERIZAÇÃO GEOMORFOLÓGICA DE UMA BACIA URBANA NO MUNICÍPIO DE PATOS/PB USANDO SIGS. *In: SIMPÓSIO DE HIDRÁULICA E RECURSOS HÍDRICOS DOS PAÍSES DE LÍNGUA PORTUGUESA*, 15., 2022, Caruaru-PE. Anais [...].

BARBOSA, A.H.; KUMAR, L.V.T. Influence of rainfall variability on the vegetation dynamics over Northeastern Brazil. **Journal of Arid Environments**, v.124, p.377-387, 2016.

BELGIU, M. & DRĂGUT, L. Random forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, 114, 24-31, 2016.

BERNI, J.; ZARCO-TEJADA, P.J.; SUAREZ, L.; FERERES, E. Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle. **IEEE Transactions on Geoscience and Remote Sensing**, v.47, p.722–738, 2009.

BEUCHLE, R.; GRECCHI, C. R.; SHIMABUKURO, E.Y.; SELIGER, R.; EVA, D.H.; SANO, E.; ACHARD, F. Land cover changes in the Brazilian Cerrado and Caatinga biomes

from 1990 to 2010 based on a systematic remote sensing sampling approach. **Applied Geography**, v.58, p.116-127, 2015.

BEZERRA, U. A.; OLIVEIRA, L. M. M.; CANDEIAS, A. L. B.; SILVA, B. B.; LEITE, A. C. L. D. S.; SILVA, L. T. M. D. S. Comparativo do Índice de Vegetação de Diferença Normalizada (NDVI) entre os Sensores OLI-Satélite Landsat-8 e MSI-Satélite Sentinel-2 em Região Semiárida. **Anuário do Instituto de Geociências**, v.41, p.167-177, 2019.

BIRTH, G. S.; MCVEY, G. Measuring the colour of growing turf with a reflectance spectrophotometer. **Agronomy Journal**, v. 60, n. 6, p. 640-643, 1968.

BOSCHETTI, L.; ROY, D. P.; JUSTICE, C. O.; HUMBER, M. L. MODIS – Landsat fusion for large area 30 m burned area mapping. **Remote Sensing of Environment**, v. 161, p. 27–42, 2015.

BRANDÃO, Z. N. Agricultura de precisão para gerenciamento do algodão. In: AZEVÊDO, D. M. P.; BELTRÃO, N. E. M. **O agronegócio do algodão no Brasil. Brasília: Embrapa Informação Tecnológica**, 1309 p. cap. 20. v. 2, 2008.

BREIMAN, L. Random Forests. **Machine Learning**, v.45, p.5–32, 2001.

BRENNER, C.; ZEEMAN, M.; BERNHARDT, M.; SCHULZ, K. Estimation of Evapotranspiration of Temperate Grassland Based on High-Resolution Thermal and Visible Range Imagery from Unmanned Aerial Systems. **International Journal of Remote Sensing**, v.39, p.5141–5174, 2018.

BRITO, P. V. S.; SANTOS, T. O.; MORAIS, Y. B. C. B. Análise da sazonalidade da vegetação de caatinga do município de Petrolina (Pernambuco) a partir de imagens do satélite Landsat 8–OLI. **Revista Brasileira de Sensoriamento Remoto**, v.1, 29 – 37, 2020.

CAI, Z.; JÖNSSON, P.; JIN, H.; EKLUNDH, L. Performance of smoothing methods for reconstructing NDVI time-series and estimating vegetation phenology from MODIS data. **Remote Sensing**, v.9, p.1271, 2017.

CARDOSO, P. V.; SILVA SEABRA, V.; XAVIER, R. A.; MORAIS RODRIGUES, E. Mapeamento de Áreas de Caatinga Através do Random Forrest: Estudo de caso na Baía do Rio Taperoá. **Revista Georaguaia**, v.11, p.55-68, 2021.

CARVALHO, R DA S.; CARVALHO, J.E.B. Processamento digital e aprendizado de máquina de imagens aéreas obtidas por aeronaves remotamente pilotadas (RPA) para estimar percentual de cobertura do solo por gramíneas perenes em pomar de citros. **Cruz das Almas, BA : Embrapa Mandioca e Fruticultura**, 2022.

CHAVES, I. D. B.; LOPES, V. L.; FOLLIOTT, P. F.; & PAES-SILVA, A. P. Uma classificação morfoestrutural para descrição e avaliação da biomassa da vegetação da caatinga. **Revista Caatinga**, v.21(2), p.204-213, 2008.

CHEN, Y.; LU, D.; MORAN, E.; BATISTELLA, M.; DUTRA, L.V.; SANCHES, I.D.; SILVA, R.F.B.; HUANG, J.; LUIZ, A.J.B.; OLIVEIRA, M.A.F. Mapping croplands,

cropping patterns, and crop types using MODIS time-series data. **International Journal of Applied Earth Observation and Geoinformation**, v.69, p.133–147, 2018.

COELHO, V. H. R.; MONTENEGRO, M.G.L.S.; ALMEIDA, das N.C.; LIMA, de R.V.E.; NETO, R.A.; MOURA, de S.S.G. Dinâmica do uso e ocupação do solo em uma bacia hidrográfica do semiárido brasileiro. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.18 (1), p.64-72, 2014.

CÓRCOLES, J. I.; ORTEGA, J. F.; HERNANDEZ, D. ; MORENO, M. A. Estimation of leaf area index in onion (*Allium cepa* L.) using an unmanned aerial vehicle. **Biosystems Engineering**, v.115(1), 2013.

CORCORAN, J.; KNIGHT, J.; GALLANT, A. Influence of MultiSource and MultiTemporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota. **Remote Sensing**, v.5(7), p.3212–3238, 2013.

DI VITTORIO, A.C.; GEORGAKAKOS, P.A. Land cover classification and wetland inundation mapping using MODIS. **Remote Sensing of Environment**, v.204, p.1-17, 2018.

ESPINDOLA, de G.M.; FIGUEREDO, de S.E.; JÚNIOR, P.P.; REIS FILHO, A.A. Cropland expansion as a driver of land-use change: The case of Cerrado-Caatinga transition zone in Brazil. *Environment, Development and Sustainability*, v.23, p.17146–17160, 2021.

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v.80(1), p.185-201, 2002.

GATES, D.M.; KEEGAN, J.H.; SCHLETER, C.J.; WEIDNER, R.V. Spectral properties of plants. **Applied Optics**, v.4(1), p.11-20, 1965.

GHOSH, A.; JOSHI, P.K. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. **International Journal of Applied Earth Observation and Geoinformation**, v.26, p.298–311, 2014.

GIONGO, V.; CUNHA, T. J. F.; MENDES, A. S. M.; GAVA, C. A. T. Carbono no Sistema Solo-Planta no Semiárido. **Revista Brasileira de Geografia Física**, v.6, p.1233 – 1253, 2011.

GIOVOS, R.; TASSOPOULOS, D.; KALIVAS, D.; LOUGKOS, N.; PRIOVOLOU, A. Remote Sensing Vegetation Indices in Viticulture: A Critical Review. **Agriculture**, v.11, p.457, 2021.

GOMES, A.; QUEIROZ de M.D.; VALENTE, M.S.D.; PINTO, F. Comparing a single-sensor camera with a multisensor camera for monitoring coffee crop using Unmanned Aerial Vehicles. **Engenharia Agrícola**, v.41, p.87-97, 2021.

GONÇALVES, V.P.; RIBEIRO, E.A.W.; IMAI, N.N. Mapping Areas Invaded by *Pinus* sp. from Geographic Object-Based Image Analysis (GEOBIA) Applied on RPAS (Drone) Color Images. **Remote Sensing**, v.14, p.2805, 2022.

HABOUDANE, D.; MILLER, R. J.; PATTEY, E.; ZARCO-TEJADA, J. P.; STRACHAN, B. I. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. **Remote Sensing of Environment**, v. 90, n. 3, p. 337-352, 2004.

HOFFMANN, H.; NIETO, H.; JENSEN, R. Estimating Evaporation with Thermal UAV Data and Two-Source Energy Balance Models. **Hydrology and Earth System Sciences**, v.20, p.697–713, 2016.

HOUBORG, R.; FISHER, B.J.; SKIDMORE, K.A. Advances in remote sensing of vegetation function and traits. **International Journal Applied Earth Observations and Geoinformation**, v.43, p. 1-6, 2015.

HTITIOU, A.; BOUDHAR, A.; LEBRINI, A.; HADRIA, Y.; LIONBOUI, R.; MANSOURI, R. TYCHON, L.; BENABDELOUAHAB, B. T. The Performance of Random Forest Classification Based on Phenological Metrics Derived from Sentinel2 and Landsat 8 to Map Crop Cover in an Irrigated Semiarid Region. **Remote Sensing in Earth Systems Sciences**, v.2(4), p. 208-224, 2019.

IBGE. **Manual Técnico da Vegetação Brasileira**. Rio de Janeiro-RJ. Brasil, 271p, 2022.

IBGE – **Instituto Brasileiro de Geografia e Estatística**. Disponível em: <http://www.ibge.gov.br/>. Acesso em 15 de julho de 2023.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres (Segunda edição)**, São José dos Campos, São Paulo, Brasil: Parêntese, 2009.

JIA, K.; LIANG, S.; WEI, X.; YAO, Y.; SU, Y.; JIANG, B.; WANG, X.. Land cover classification of landsat data with phenological features extracted from time series MODIS NDVI data. **Remote Sensing**, v.6, p.11518-11532, 2014.

JIN, Y.; LIU, X.; CHEN, Y.; LIANG, X. Landcover mapping using Random Forest classification and incorporating NDVI timeseries and texture: a case study of central Shandong. **International Journal of Remote Sensing**, v.39(23), p.8703–8723, 2018.

JONES, G.H.; VAUGHAN, A.R. Remote sensing of vegetation: Principles, Techniques and Applications. **New York, EUA: OXFORD UNIVERSITY PRESS**, 2010.

JORGE, L. A. C.; INAMASU, R. Y. Uso de veículos aéreos não tripulados (VANT) em agricultura de precisão. IN: BERNARDI, A. C. C.; NAIME, J. M.; RESENDE, A. V.; BASSOI, L. H.; INAMASU, R. Y. (ED.). **Agricultura de precisão: resultados de um novo olhar**. Brasília, DF: Embrapa, p. 109-134, 2014.

KARNIELI, A.; GABAI, A.; ICHOKU, C.; ZAADY, E.; SHACHAK, M. Temporal dynamics of soil and vegetation spectral responses in a semiarid environment. **International Journal of Remote Sensing**, v.23(19), p.4073–4087, 2002.

KATTENBORN, T.; LOPATIN, T.; FÖRSTER, J.; BRAUN, M.; FASSNACHT, A.F. UAV Data as Alternative to Field Sampling to Map Woody Invasive Species Based on Combined Sentinel-1 and Sentinel-2 Data. **Remote Sensing of Environment**, v.227, p.61–73, 2019.

KUPLICH, T. M.; MOREIRA, A.; FONTANA, D. C. Série temporal de índice de vegetação sobre diferentes tipologias vegetais no Rio Grande do Sul. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.17(10), p.1123-2013, 2013.

LEAL, I.R.; SILVA, J.M.C.da.; TABARELLI, M.; LACHER JÚNIOR, E.T. Changing the course of biodiversity conservation in the Caatinga of Northeastern Brazil. **Conservation Biology**, v.19, p.701-706, 2005.

LEHMANN, R.K.; TORSTEN, P.; ZILLER, R. S.; JAN, T.; HERINGER, G.; MEIRANETO, J. A. A., BUTTSCHARDT, T.K. Open-Source Processing and Analysis of Aerial Imagery Acquired with a Low-Cost Unmanned Aerial System to Support Invasive Plant Management. **Frontiers in Environmental Science**, v.5, 44, 2017.

LI, G.; ZHANG, F.; JING, Y.; LIU, Y.; SUN, G. Response of evapotranspiration to changes in land use and land cover and climate in China during 2001–2013. **Science of the Total Environment**, v.596–597, p. 256-265, 2017.

LIAW, A.; WIENER, M. Classification and regression by Random Forest. **R news**, v.2(3), p.18-22, 2002.

LIMA JÚNIOR, de C.; ACCIOLY, O. de J.; GIONGO, V.; LIMA, A.F.L.R.; SAMPAIO, B.S.V.E.; MENEZES, C.S.R. Estimativa de biomassa lenhosa da caatinga com uso de equações alométricas e índice de vegetação. **Scientia Forestalis**, v. 42, n. 102, p. 289-298, 2014.

LIMA, N. A. S.; OLIVEIRA, L. M. M.; XAVIER, M. L. P.; SILVA JÚNIOR, U. J.; NASCIMENTO, E. F. Spectral Indices for remote sensing with the LANDSAT-5 and SENTINEL-2 satellites around the Poço Cruz reservoir-PE. **Journal of Hyperspectral Remote Sensing**, v.9, p.299-309, 2019.

LIU, R.; SHANG, R.; LIU, Y.; LU, X. Global evaluation of gap-filling approaches for seasonal NDVI with considering vegetation growth trajectory, protection of key point, noise resistance and curve stability. **Remote Sensing of Environment**, v.189, p.164-179, 2017.

LU, M.; Chen, J.; Tang, H.; Rao, Y.; Yang, P.; Wu, W. Land cover change detection by integrating object-based data blending model of Landsat and MODIS. **Remote Sensing of Environment**, v.184, p.374-386, 2016.

MA, Q.; HAN, W.; HUANG, S.; DONG, S.; LI, G.; CHEN, H. Distinguishing Planting Structures of Different Complexity from UAV Multispectral Images. **Sensors**, v.21 (6), 2021.

MAHDIANPARI, M.; SALEHI, B.; REZAEI, M.; MOHAMMADIMANESH, F.; ZHANG, Y. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. **Remote Sensing**, v.10(7), p.11-19, 2018.

PABLO, M de J.M.; SANCHEZ, G.A.; JIMENEZ, J.I.S.; CAPURATA, O.E.R.; BUSTAMANTE, O.W. Estimation of vegetation fraction using RGB and multispectral images from UAV. **International Journal of Remote Sensing**, v.40 (2), p.1 – 19, 2018.

MARZIALETTI, F.; FRATE, L.; DE SIMONE, W.; FRATTAROLI, A.R.; ACOSTA, A.T.R.; CARRANZA, M.L. Unmanned Aerial Vehicle (UAV)-Based Mapping of Acacia Saligna Invasion in the Mediterranean Coast. **Remote Sensing**, v.13, p.33-61, 2021.

MATSUO, C.A.S. **Projeto de um veículo aéreo não tripulado para pulverização aeroagrícola**. 2011. 267p. Dissertação (Mestrado em Engenharia Mecânica) – Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, 2011.

MATSUSHITA, B.; XU, M.; FUKUSHIMA, T. Characterizing the changes in landscape structure in the Lake Kasumigaura, Japan using a high-quality GIS dataset. **Landscape and Urban Planning**, v.78, p.241-250, 2006.

MAXWELL, A.E.; WARNER, T.A.; FANG, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. **International Journal of Remote Sensing**, v.39, p.2784–2817, 2018.

MEDEIROS, F. A. **Desenvolvimento de um veículo aéreo não tripulado para aplicação em agricultura de precisão**. 2007. 102p. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Federal de Santa Maria, Santa Maria, 2007.

MELAAS, E. K.; FRIEDL, M. A.; ZHU, Z. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. **Remote Sensing of Environment**, v.132, p.176-185, 2013.

MENDOZA, F.M.; GRANADOS, L.E.; GENELETTI, D.; PÉREZ-SALICRUP, R.D.; SALINAS, V. Analysing land cover and land use change processes at watershed level: A multitemporal study in the Lake Cuitzeo Watershed, Mexico (1975–2003). **Applied Geography**, v.31, p.237-250, 2011.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. Universidade Federal de Viçosa, 3ª edição, Ed. UFV, 315 p., 2007.

NITZE, I.; BARRETT, B.; CAWKWELL, F. Temporal optimisation of image acquisition for land cover classification with Random Forest and MODIS timeseries. **International Journal of Applied Earth Observation and Geoinformation**, v.34, 136–146, 2015.

NOVO, E. M. L. M. **Sensoriamento Remoto: Princípios e Aplicações**. São Paulo, Terceira Edição. Editora Edgard Blucher, 2008.

OLARIU, H.G.; MALAMBO, L.; POPESCU, S.C.; VIRGIL, C.; WILCOX, B.P. Woody Plant Encroachment: Evaluating Methodologies for Semiarid Woody Species Classification from Drone Images. **Remote Sensing**, v.14, p.16-65, 2022.

ORTEGA-FARÍAS, S.; ORTEGA-SALAZAR, S.; POBLETE, T.; KILIC, A.; ALLEN, R.; POBLETE-ECHEVERRÍA, C.; AHUMADA-ORELLANA, L.; ZUÑIGA, M.;

SEPÚLVEDA, D. Estimation of Energy Balance Components over a Drip-Irrigated Olive Orchard Using Thermal and Multispectral Cameras Placed on a Helicopter-Based Unmanned Aerial Vehicle (UAV). **Remote Sensing**, v.8, p.6-38, 2016.

PAL, M.; MATHER, P.M. An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification. **Remote Sensing of Environment**, v.86, 554–565, 2003.

PAL, M.; MATHER, P.M. Support Vector Machines for Classification in Remote Sensing. **International Journal of Remote Sensing**, v.26, p.1007–1011, 2005.

PARK, S.; RYU, D.; FUENTES, S.; CHUNG, H.; O'CONNELL, M.; KIM, J. Mapping Very-High-Resolution Evapotranspiration from Unmanned Aerial Vehicle (UAV) Imagery. **International Journal of Geo-Information**, v.10, p.2-11, 2003.

PASTOR-GUZMANA, J.; DASHA, J.; ATKINSON, M. P. Remote sensing of mangrove forest phenology and its environmental drivers. **Remote Sensing of Environment**, v. 205, p. 71-84, 2018.

ROUSE, J. W. JR. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. **Texas: Texas A&M University**, 1974.

**RSTUDIO: integrated development for R**. Boston: PBC, 2020. Disponível em: <http://www.rstudio.com/> Acesso em: 21 out. 2021.

SÁ, S DE A.C. **Sensoriamento remoto aplicado em cultivos no submédio São Francisco por imageamento aéreo com drones**. 2022. 97p. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Federal do Vale do São Francisco, Campus Juazeiro-BA, 2022.

SAGAN, V. M.; MAITINIYAZI, S.; EBLIMIT, P.; PETERSON, K.; HARTLING, K.; ESPOSITO, S.; KHANAL, F.; NEWCOMB, K.; PAULI, M.; WARD, W.; FRITSCHI, R.; SHAKOOR, F.; MOCKLER, N. T. UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue Pro R 640, and thermomap cameras. **Remote Sensing**, v.11 (3), 2019.

SAMPAIO, H. S.; BOURSCHEIDT, V.; SARRACINI, L. H.; JORGE, L. A. de C.; BETTIOL, G. M.; BERNARDI, A. C. de C. Comparação entre Índices de Vegetação obtidos por imagens aéreas com Veículo Aéreo não Tripulado (VANT) e Satélite. **Revista Brasileira de Engenharia de Biosistemas**, v.14, p.111-124, 2020.

SANTOS, C. V. B. Dos.; MOURA, M. S. B. De.; CARVALHO, H. F. de S.; GALVINCIO, J. D.; MIRANDA, R. de Q.; NISHIWAKI, A. A. M.; MONTENEGRO, S. M. G. L.. Avaliação do índice de área foliar e índice de área da planta em floresta seca utilizando modelos simplificados em imagens de alta resolução com o uso de VANT. **Journal of Hyperspectral Remote Sensing**, v12 (3), p.109-123, 2022.

SANTOS, C. V. B. dos; MOURA, M. S. B. de; GALVÍNCIO, J. D.; CARVALHO, H. F. de S.; MIRANDA, R. de Q.; MONTENEGRO, S. M. G. L. Comparação de imagens multiespectrais utilizando Satélites e VANT para a análise de mudanças estruturais em área de Floresta Seca. **Revista Brasileira de Geografia Física** v.14(6), p.2510-2522, 2021.

SCHULZ, C.; KOCH, R.; CIERJACKS, A. KLEINSCHMIT, B. Land change and loss of landscape diversity at the Caatinga phytogeographical domain – Analysis of pattern-process relationships with MODIS land cover products (2001–2012). **Journal of Arid Environments**, v.136, p.54-74, 2017.

SHIRATSUCHI, L. S.; BRANDAO, Z. N.; VICENTE, L. E.; VICTORIA, D. de C.; DUCATI, J. R.; OLIVEIRA, R. P. de; VILELA, M. de F. Sensoriamento Remoto: conceitos básicos e aplicações na Agricultura de Precisão. In: **Agricultura de precisão: resultados de um novo olhar**, 2014.

SHIROMA, G. M.; GAMARRA, R. M.; GUIRRA, A. P. M.; MOTTA, J. S.; DE FIGUEIREDO, H. R.; ENCINA, C. C. C.; PARANHOS FILHO, A. C. Comparação entre Índices de Vegetação de Diferentes Sensores na Identificação de Fitofisionomias do Cerrado. **Anuário do Instituto de Geociências**, v.43, p.18-32, 2020.

SILVA, C. R.; MACHADO, S. L. D.; ARAÚJO, A. A.; ABREU JUNIOR, C. A. M. Analysis of the phenology dynamics of Brazilian caatinga species with NDVI time series. **CERNE**, v.24 (1), p.48-58, 2018.

SILVA, S.V.D.; CRUZ, M.B.C. Tipologias de Caatinga: Uma Revisão em Apoio a Mapeamentos Através de Sensoriamento Remoto Orbital e GEOBIA. **Revista do departamento de Geografia**, v.35, USP, 2018.

SILVA, S.R.; MEDEIROS, M.B. de; LIMA, V.V.F.; PEIXOTO, M.R.; AONA, L.Y.S. Patterns of Cactaceae Species Distribution in A Protected Area in the Semiarid Caatinga Biome of North-Eastern Brazil. **Edinburgh Journal of Botany**, v.73, p.157-170, 2016.

SILVEIRA, M.A.; ANDRADE, P.;FILHO, C.; BREMM, M.; SANTOS, C.F. Aprendizado de máquina com base na resposta espectral de imagens aéreas obtidas por ANTs e aplicado no manejo de pastagens. **Bagé: Embrapa Pecuária Sul**, 2022.

SOUZA *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(17), p.27-35, 2020.

SOUZA, R.; FENG, X.; ANTONINO, A. MONTENEGRO, S.; SOUZA, E.; PORPORATO, A. Vegetation response to rainfall seasonality and interannual variability in tropical dry forests. **Hydrological processes**, v.30(20), p.3583-3595, 2016.

STRECHA, C.; FLETCHER, A.; LECHNER, A.; ERSKINE, P.; FUA, P. Developing Species Specific Vegetation Maps Using Multi-Spectral Hyperspatial Imagery from Unmanned Aerial Vehicles. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v.3, p.311–316, 2012.

TUCKER, C. J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v. 8, n. 2, p. 127-150, 1979.

VANZELA, L. S.; HERNANDEZ, F. B. T.; FRANCO, R. A. M. Influência do uso e

ocupação do solo nos recursos hídricos do Córrego Três Barras, Marinópolis. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.14 (1), p.55-64, 2010.

WASKE, B.; VAN DER LINDEN, S. Classifying multilevel imagery from SAR and optical sensors by decision fusion. **IEEE Trans. Geosci. Remote Sens.**, v.46, p.1457–1466, 2008.

WILLSTATTER, R.; STOOLL, A. Untersuchungen über die assimilation der kohlen-säure. **Springer**, Berlin, 1918.

WORKIE, G.T.; DEBELLA, J.H. Climate change and its effects on vegetation phenology across ecoregions of Ethiopia. **Global Ecology and Conservation**, v.13, p.1-13, 2017.

XIA, T.; KUSTAS, W. P.; ANDERSON, M. C.; ALFIERI, J. G.; GAO, F.; MCKEE, L.; PRUEGER, J. H.; GELI, H. M. E.; NEALE, C. M. U.; SANCHEZ, L.; MAR ALSINA, M.; WANG, Z. Mapping Evapotranspiration with High-Resolution Aircraft Imagery over Vineyards Using One- and Two-Source Modeling Schemes. **Hydrology and Earth System Sciences**, v.20, p.1523–1545, 2016.

ZHANG, Q.; KONG, D.; SHI, P.; SINGH, P.V.; SUN, P. Vegetation phenology on the Qinghai-Tibetan Plateau and its response to climate change (1982–2013). **Agricultural and Forest Meteorology**, v.248, p.408-417, 2018.

ZHANG, Z., YANG, Z., REN, W., & WEN, G. Random forestbased realtime defect detection of Al alloy in robotic arc welding using optical spectrum. **Journal of Manufacturing Processes**, v.42, p.51–59, 2019.

## ANEXO A – *SCRIPT PARA MODELO RANDOM FOREST*

Neste anexo é apresentado o *script* desenvolvido na plataforma Rstudio e utilizado nesta pesquisa para rodar os modelos *Random Forest* em cada área analisada através dos métodos 1 e 2.

### A.1. Script em R para classificação da cobertura do solo através de modelo RF

```
# Método 1

# Carregando pacotes

library(terra)
library(rstac)
library(raster)
library(rgdal)
library(dplyr)
library(randomForest)
library(caret)
library(caTools)
library(mapview)
library(ExpImage)
library(EBImage)

# Carregando um recorte da área de estudo

# X = área analisada, possíveis valores neste script = 1,2 ou 3.

shp=readOGR("D:/Doutorado/TESE/SHAPEFILES/RECORTE_AREA_DE_ESTUDO_X.
shp")

# Carregando imagem a ser recortada

imagem=stack("D:/Doutorado/TESE/RASTERS/IR_X_MAPIR.tif","D:/Doutorado/TESE/
RASTERS/RED_X_MAPIR.tif","D:/Doutorado/TESE/RASTERS/GREEN_X_MAPIR.tif"
)

# Recortando a imagem

imagemrecorte = mask(imagem,shp)

# Plotando composição RNG da imagem Mapir

plotRGB(imagemrecorte, r = 2, g = 1, b = 3, stretch="lin")

#Normalizando as bandas

IR_Mapir = (imagemrecorte$IR_X_MAPIR/ (mean(imagemrecorte$IR_X_MAPIR,
```

```

imagemrecorte$RED_X_MAPIR, imagemrecorte$GREEN_X_MAPIR)))
RED_Mapir = (imagemrecorte$RED_X_MAPIR/ (mean(imagemrecorte$IR_X_MAPIR,
imagemrecorte$RED_X_MAPIR, imagemrecorte$GREEN_X_MAPIR)))
GREEN_Mapir = (imagemrecorte$GREEN_X_MAPIR/
(mean(imagemrecorte$IR_X_MAPIR, imagemrecorte$RED_X_MAPIR,
imagemrecorte$GREEN_X_MAPIR)))
img_norm = stack(IR_Mapir, RED_Mapir, GREEN_Mapir)
names(img_norm) = c("IR_X_MAPIR", "RED_X_MAPIR", "GREEN_X_MAPIR" )

# Calculando NDVI

img_norm$NDVI = (((img_norm$IR_X_MAPIR)-(img_norm$RED_X_MAPIR)) /
((img_norm$IR_X_MAPIR) + (img_norm$RED_X_MAPIR)))

# Plotando a imagem NDVI

plot(img_norm$NDVI)

# Carregando o shapfile das amostras - PONTOS

amst = readOGR("D:/Doutorado/TESE/SHAPEFILES/AMOSTRAS_X.shp")

# Extrairdo valores das bandas na amostra

amst.ext = extract(img_norm, amst)

# Definindo os valores extraídos como uma tabela de dados

amst.ext = data.frame(amst.ext)

# Visualizando a tabela de dados

View(amst.ext)

# Atribuindo a coluna "classe" como um fator

amst.ext$classe = as.factor(amst$classe)

# Separando valores de treinamento e de validação

posicoes = sample.split(Y= amst.ext$classe, SplitRatio = 0.7)

treinamento = amst.ext[posicoes,]

validacao = amst.ext[posicoes == FALSE, ]

```

```

# Exportando os dados de treinamento e validação

Write.csv (treinamento, file = "treinamento_área_X_método_1.csv")

Write.csv (validacao, file = "validação_área_X_método_1.csv")

# Rodando o modelo para os dados de treinamento

RF.model = randomForest(classe~., data = treinamento, ntree = 1000, mtry = 2)

print(RF.model)

# Validação do modelo criado

pred.RF = predict(RF.model, validacao)

print(cm.RF)

# Predição com o modelo RF para toda a imagem

inicio = Sys.time()
set.seed(371)
classe.RF = predict(img_norm, RF.model)
fim = Sys.time()
fim - inicio

# Visualizando o mapa de classificação gerado

mapview(classe.RF, col.regions = list("orange", "darkgreen", "green"))

# Salvando em tiff. o mapa de classificação

writeRaster(classe.RF, filename = "classificacao_area_X_Mapir.tif")

# Estatísticas da classificação

# Matriz de confusão para o modelo criado

cm.RF = confusionMatrix(pred.RF, validacao$classe)

#Variável mais importante - MDG

varImpPlot(RF.model, pch = 19, col = "red")

#Boxplot para os dados de treinamento

boxplot(treinamento$NDVI~treinamento$classe, col = c("darkgreen", "lawngreen",

```

```

"orange"))
boxplot(treinamento$IR_X_MAPIR~treinamento$classe, col = c("darkgreen",
"lawngreen", "orange"))
boxplot(treinamento$RED_X_MAPIR~treinamento$classe, col = c("darkgreen",
"lawngreen", "orange"))
boxplot(treinamento$GREEN_X_MAPIR~treinamento$classe, col = c("darkgreen",
"lawngreen", "orange"))

# Método 2

# Carregando um recorte da área de estudo

shp=readOGR("D:/Doutorado/TESE/SHAPEFILES/RECORTE_AREA_DE_ESTUDO_X.
shp")

# Carregando imagem a ser recortada

imagem =
stack("D:/Doutorado/TESE/RASTERS/RED_X.tif", "D:/Doutorado/TESE/RASTERS/GRE
EN_X.tif", "D:/Doutorado/TESE/RASTERS/BLUE_X.tif")

# Recortando a imagem
imagemrecorte = mask(imagem,shp)

# Plotando composição RNG da imagem Mapir

plotRGB(imagemrecorte, r = 1, g = 2, b = 3, stretch="lin")

# Normalizando as bandas

RED_mavic = (imagemrecorte$RED_X/ (mean(imagemrecorte$RED_X,
imagemrecorte$GREEN_X, imagemrecorte$BLUE_X)))
GREEN_mavic = (imagemrecorte$GREEN_X/ (mean(imagemrecorte$RED_X,
imagemrecorte$GREEN_X, imagemrecorte$BLUE_X)))
BLUE_mavic = (imagemrecorte$BLUE_X/ (mean(imagemrecorte$RED_X,
imagemrecorte$GREEN_X, imagemrecorte$BLUE_X)))
img_norm = stack(RED_mavic, GREEN_mavic, BLUE_mavic)
names(img_norm) = c("RED_X", "GREEN_X", "BLUE_X" )

# Calculando NGRDI

img_norm$NGRDI = (((imagemrecorte$GREEN_X)-(imagemrecorte$RED_X)) /
((imagemrecorte$GREEN_X)+(imagemrecorte$RED_X)))

# Plotando a imagem NGRDI

plot(img_norm$NGRDI)

```

```

# Carregando o shapfile das amostras

amst = readOGR("D:/Doutorado/TESE/SHAPEFILES/AMOSTRAS_X.shp")

# Extrair valores das bandas na amostra

amst.ext = extract(img_norm, amst)

# Definindo os valores extraídos como uma tabela de dados

amst.ext = data.frame(amst.ext)

# Visualizando a tabela de dados

View(amst.ext)

# Atribuindo a coluna "classe" como um fator

amst.ext$classe = as.factor(amst$classe)

# Separando valores de treinamento e de validação

posicoes = sample.split(Y= amst.ext$classe, SplitRatio = 0.7)

treinamento = amst.ext[posicoes,]

validacao = amst.ext[posicoes == FALSE, ]

# Exportando os dados de treinamento e validação

Write.csv (treinamento, file = "treinamento_área_X_método_2.csv")

Write.csv (validacao, file = "validação_área_X_método_2.csv")

# Rodando o modelo para os dados de treinamento

RF.model_mavic = randomForest(classe~., data = treinamento, ntree = 1000, mtry = 2)

print(RF.model_mavic)

# Validação do modelo criado

pred.RF_mavic = predict(RF.model_mavic, validacao)

print(cm.RF_mavic)

# Predição com o modelo RF para toda a imagem

```

```

inicio = Sys.time()
set.seed(371)
classe.RF = predict(img_norm, RF.model_mavic)
fim = Sys.time()
fim - inicio

# Visualizando o mapa de classificação gerado

mapview(classe.RF, col.regions = list("orange", "darkgreen", "green"))

# Salvando em tiff. o mapa de classificação

writeRaster(classe.RF, filename = "classificacao_area_X_mavic.tif")

# Estatísticas da classificação

# Matriz de confusão para o modelo criado

cm.RF_mavic = confusionMatrix(pred.RF_mavic, validacao$classe)

# Variável mais importante - MDG

varImpPlot(RF.model_mavic, pch = 19, col = "red")

# Boxplot para os dados de treinamento

boxplot(treinamento$NGRDI~treinamento$classe, col = c("darkgreen", "lawngreen",
"orange"))
boxplot(treinamento$RED_X~treinamento$classe, col = c("darkgreen", "lawngreen",
"orange"))
boxplot(treinamento$GREEN_X~treinamento$classe, col = c("darkgreen", "lawngreen",
"orange"))
boxplot(treinamento$BLUE_X~treinamento$classe, col = c("darkgreen", "lawngreen",
"orange"))

```

## ANEXO B – DADOS DE TREINAMENTO E VALIDAÇÃO

Neste anexo são apresentados os valores de cada banda e índices extraídos das imagens e utilizados como dados de treinamento e validação do modelo Random Forest em cada método para cada área e escolhidos de forma aleatória. As amostras estão indicadas nas Tabelas B.1 a B.12.

**Tabela B.1 – Amostras de treinamento na área 1 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
17	-37,31662	-7,01759	1,42	0,92	0,66	0,21	Arbustiva arbórea densa
19	-37,31450	-7,01678	1,52	0,78	0,71	0,32	Arbustiva arbórea densa
20	-37,31621	-7,01718	1,43	0,93	0,64	0,21	Arbustiva arbórea densa
21	-37,31439	-7,01657	1,60	0,68	0,73	0,41	Arbustiva arbórea densa
23	-37,31494	-7,01877	1,48	0,85	0,67	0,27	Arbustiva arbórea densa
25	-37,31594	-7,01665	1,48	0,85	0,67	0,27	Arbustiva arbórea densa
26	-37,31446	-7,01734	1,34	0,97	0,69	0,16	Arbustiva arbórea densa
27	-37,31474	-7,01765	1,46	0,84	0,70	0,27	Arbustiva arbórea densa
29	-37,31489	-7,01842	1,46	0,89	0,65	0,24	Arbustiva arbórea densa
30	-37,31719	-7,01811	1,52	0,84	0,64	0,29	Arbustiva arbórea densa
31	-37,31706	-7,01792	1,41	0,96	0,63	0,19	Subarbustiva muito rala
32	-37,31739	-7,01807	1,20	1,15	0,65	0,02	Subarbustiva muito rala
33	-37,31484	-7,01765	1,37	0,97	0,66	0,17	Subarbustiva muito rala
37	-37,31610	-7,01666	1,34	1,04	0,62	0,13	Subarbustiva muito rala
38	-37,31713	-7,01598	1,42	1,07	0,51	0,14	Subarbustiva muito rala
39	-37,31678	-7,01723	1,37	0,99	0,64	0,16	Subarbustiva muito rala
41	-37,31576	-7,01560	1,40	0,99	0,61	0,17	Subarbustiva muito rala
43	-37,31504	-7,01755	1,33	1,08	0,60	0,11	Subarbustiva muito rala
44	-37,31349	-7,01610	1,30	1,06	0,63	0,10	Subarbustiva muito rala

45	-37,31584	-7,01582	1,28	1,07	0,65	0,09	Subarbustiva muito rala
1	-37,31726	-7,01721	1,12	1,24	0,64	-0,05	Solo exposto
5	-37,31685	-7,01571	1,15	1,29	0,57	-0,06	Solo exposto
6	-37,31440	-7,01778	1,25	1,17	0,58	0,03	Solo exposto
8	-37,31389	-7,01839	1,16	1,28	0,56	-0,05	Solo exposto
9	-37,31372	-7,01820	1,19	1,24	0,57	-0,02	Solo exposto
10	-37,31663	-7,01634	1,19	1,22	0,60	-0,01	Solo exposto
11	-37,31628	-7,01580	1,18	1,31	0,51	-0,05	Solo exposto
13	-37,31444	-7,01558	1,00	1,37	0,63	-0,15	Solo exposto
14	-37,31507	-7,01688	1,17	1,23	0,59	-0,03	Solo exposto
15	-37,31745	-7,01783	1,15	1,21	0,65	-0,03	Solo exposto

**Tabela B.2 – Amostras de validação na área 1 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
16	-37,31570	-7,01582	1,38	0,86	0,76	0,23	Arbustiva arbórea densa
18	-37,31403	-7,01892	1,32	1,02	0,67	0,13	Arbustiva arbórea densa
22	-37,31599	-7,01810	1,50	0,89	0,61	0,26	Arbustiva arbórea densa
24	-37,31639	-7,01543	1,49	0,86	0,65	0,27	Arbustiva arbórea densa
28	-37,31502	-7,01855	1,40	1,05	0,54	0,14	Arbustiva arbórea densa
34	-37,31431	-7,01656	1,20	1,10	0,69	0,04	Subarbustiva muito rala
35	-37,31460	-7,01581	1,22	1,10	0,68	0,05	Subarbustiva muito rala
36	-37,31435	-7,01579	1,33	1,01	0,65	0,13	Subarbustiva muito rala
40	-37,31495	-7,16976	1,25	1,07	0,68	0,08	Subarbustiva muito rala
42	-37,31690	-7,01836	1,33	1,06	0,62	0,11	Subarbustiva muito rala
2	-37,31573	-7,01777	1,05	1,35	0,59	-0,12	Solo exposto
3	-37,31524	-7,01579	1,20	1,26	0,55	-0,02	Solo exposto

4	-37,31571	-7,01586	1,07	1,29	0,64	-0,09	Solo exposto
7	-37,31542	-7,01866	1,14	1,30	0,56	-0,07	Solo exposto
12	-37,31616	-7,01534	1,19	1,20	0,61	-0,01	Solo exposto

**Tabela B.3 – Amostras de treinamento na área 1 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
16	-37,31570	-7,01582	1,02	1,00	0,98	-0,01	Arbustiva arbórea densa
18	-37,31403	-7,01892	1,00	1,04	0,96	0,02	Arbustiva arbórea densa
19	-37,31450	-7,01678	1,11	1,04	0,86	-0,03	Arbustiva arbórea densa
20	-37,31621	-7,01718	1,09	1,06	0,85	-0,01	Arbustiva arbórea densa
21	-37,31439	-7,01657	1,08	1,06	0,86	-0,01	Arbustiva arbórea densa
23	-37,31494	-7,01877	1,04	1,12	0,84	0,04	Arbustiva arbórea densa
24	-37,31639	-7,01543	1,04	1,10	0,87	0,03	Arbustiva arbórea densa
25	-37,31594	-7,01665	1,08	0,98	0,94	-0,05	Arbustiva arbórea densa
27	-37,31474	-7,01765	1,05	0,99	0,96	-0,03	Arbustiva arbórea densa
29	-37,31489	-7,01842	1,04	1,08	0,87	0,02	Arbustiva arbórea densa
31	-37,31706	-7,01792	1,03	1,12	0,85	0,04	Subarbustiva muito rala
32	-37,31739	-7,01807	1,11	0,94	0,95	-0,08	Subarbustiva muito rala
33	-37,31484	-7,01765	1,03	1,01	0,96	-0,01	Subarbustiva muito rala
35	-37,31460	-7,01581	1,05	1,00	0,95	-0,03	Subarbustiva muito rala
36	-37,31435	-7,01579	1,07	0,94	0,99	-0,07	Subarbustiva muito rala
37	-37,31610	-7,01666	1,06	0,97	0,97	-0,04	Subarbustiva muito rala
39	-37,31678	-7,01723	1,05	1,00	0,96	-0,02	Subarbustiva muito rala
40	-37,31495	-7,16976	1,05	0,99	0,97	-0,03	Subarbustiva muito rala
42	-37,31690	-7,01836	1,03	1,01	0,95	-0,01	Subarbustiva muito rala

43	-37,31504	-7,01755	1,06	0,99	0,95	-0,04	Subarbustiva muito rala
1	-37,31726	-7,01721	1,08	1,00	0,92	-0,04	Solo exposto
2	-37,31573	-7,01777	1,08	0,99	0,92	-0,04	Solo exposto
4	-37,31571	-7,01586	1,11	0,97	0,92	-0,06	Solo exposto
5	-37,31685	-7,01571	1,07	0,99	0,94	-0,04	Solo exposto
6	-37,31440	-7,01778	1,07	0,98	0,96	-0,04	Solo exposto
9	-37,31372	-7,01820	1,06	0,98	0,96	-0,04	Solo exposto
10	-37,31663	-7,01634	1,10	0,98	0,92	-0,06	Solo exposto
13	-37,31444	-7,01558	1,07	1,00	0,93	-0,04	Solo exposto
14	-37,31507	-7,01688	1,08	0,99	0,93	-0,04	Solo exposto
15	-37,31745	-7,01783	1,07	0,99	0,94	-0,04	Solo exposto

**Tabela B.4 – Amostras de validação na área 1 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
17	-37,31662	-7,01759	1,04	0,98	0,98	-0,03	Arbustiva arbórea densa
22	-37,31599	-7,01810	1,04	0,96	0,99	-0,04	Arbustiva arbórea densa
26	-37,31446	-7,01734	1,09	0,98	0,94	-0,05	Arbustiva arbórea densa
28	-37,31502	-7,01855	1,06	1,06	0,88	0,00	Arbustiva arbórea densa
30	-37,31719	-7,01811	1,05	1,04	0,91	-0,01	Arbustiva arbórea densa
34	-37,31431	-7,01656	1,04	1,00	0,96	-0,02	Subarbustiva muito rala
38	-37,31713	-7,01598	1,03	0,99	0,98	-0,02	Subarbustiva muito rala
41	-37,31576	-7,01560	1,03	0,99	0,97	-0,02	Subarbustiva muito rala
44	-37,31349	-7,01610	1,05	1,02	0,93	-0,01	Subarbustiva muito rala
45	-37,31584	-7,01582	1,04	1,00	0,96	-0,02	Subarbustiva muito rala

3	-37,31524	-7,01579	1,15	0,95	0,90	-0,10	Solo exposto
7	-37,31542	-7,01866	1,07	0,98	0,95	-0,04	Solo exposto
8	-37,31389	-7,01839	1,09	0,96	0,95	-0,06	Solo exposto
11	-37,31628	-7,01580	1,07	0,98	0,94	-0,04	Solo exposto
12	-37,31616	-7,01534	1,08	0,99	0,93	-0,05	Solo exposto

**Tabela B.5 – Amostras de treinamento na área 2 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
16	-37,30114	-7,02037	2,06	0,71	0,23	0,49	Arbustiva arbórea densa
17	-37,30057	-7,01999	2,01	0,70	0,29	0,48	Arbustiva arbórea densa
19	-37,29951	-7,01976	1,72	0,96	0,33	0,28	Arbustiva arbórea densa
20	-37,30085	-7,01936	2,00	0,73	0,28	0,47	Arbustiva arbórea densa
21	-37,30217	-7,01881	1,64	0,86	0,51	0,31	Arbustiva arbórea densa
22	-37,30331	-7,02089	1,54	0,93	0,53	0,25	Arbustiva arbórea densa
23	-37,29941	-7,02094	1,79	0,91	0,30	0,33	Arbustiva arbórea densa
24	-37,30200	-7,02052	1,49	1,03	0,48	0,18	Arbustiva arbórea densa
26	-37,30367	-7,02010	1,45	1,07	0,48	0,15	Arbustiva arbórea densa
27	-37,30178	-7,02115	1,77	0,91	0,32	0,32	Arbustiva arbórea densa
31	-37,30238	-7,02080	1,37	0,91	0,72	0,20	Subarbustiva muito rala
34	-37,30111	-7,02189	1,28	1,07	0,66	0,09	Subarbustiva muito rala
35	-37,30263	-7,01885	1,29	1,07	0,64	0,09	Subarbustiva muito rala
36	-37,30297	-7,02059	1,32	1,04	0,64	0,12	Subarbustiva muito rala
37	-37,30370	-7,02105	1,26	1,02	0,71	0,10	Subarbustiva muito rala
39	-37,30315	-7,02054	1,15	1,12	0,73	0,01	Subarbustiva muito rala

41	-37,30080	-7,02157	1,20	1,12	0,68	0,04	Subarbustiva muito rala
42	-37,29990	-7,02179	1,20	1,14	0,66	0,03	Subarbustiva muito rala
43	-37,30355	-7,02129	1,23	1,04	0,73	0,09	Subarbustiva muito rala
45	-37,29985	-7,02209	1,05	1,19	0,77	-0,06	Subarbustiva muito rala
1	-37,30307	-7,02042	1,03	1,42	0,55	-0,16	Solo exposto
3	-37,30252	-7,02111	0,97	1,43	0,60	-0,19	Solo exposto
7	-37,30005	-7,02154	1,05	1,35	0,59	-0,12	Solo exposto
9	-37,30128	-7,01896	1,23	1,19	0,58	0,01	Solo exposto
10	-37,30217	-7,02036	1,01	1,39	0,59	-0,16	Solo exposto
11	-37,30208	-7,02153	1,06	1,38	0,56	-0,13	Solo exposto
12	-37,29933	-7,01940	1,03	1,43	0,53	-0,16	Solo exposto
13	-37,29898	-7,01961	0,99	1,43	0,57	-0,18	Solo exposto
14	-37,30116	-7,02067	1,11	1,42	0,46	-0,12	Solo exposto
15	-37,30319	-7,02061	1,03	1,46	0,51	-0,17	Solo exposto

**Tabela B.6 – Amostras de validação na área 2 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
18	-37,30301	-7,01974	1,76	0,89	0,35	0,33	Arbustiva arbórea densa
25	-37,30179	-7,01947	1,47	0,96	0,57	0,21	Arbustiva arbórea densa
28	-37,30028	-7,01944	1,66	0,79	0,55	0,36	Arbustiva arbórea densa
29	-37,30317	-7,01863	1,52	1,01	0,47	0,20	Arbustiva arbórea densa
30	-37,30051	-7,02044	1,60	0,94	0,46	0,26	Arbustiva arbórea densa
32	-37,30252	-7,02086	1,31	1,04	0,65	0,12	Subarbustiva muito rala
33	-37,30277	-7,02122	1,17	1,17	0,66	0,00	Subarbustiva muito rala

38	-37,30161	-7,02183	1,24	1,03	0,72	0,09	Subarbustiva muito rala
40	-37,30195	-7,02136	1,15	1,19	0,66	-0,01	Subarbustiva muito rala
44	-37,30346	-7,02132	1,30	1,03	0,67	0,11	Subarbustiva muito rala
2	-37,30077	-7,02145	0,95	1,47	0,58	-0,21	Solo exposto
4	-37,30276	-7,01942	1,09	1,36	0,55	-0,11	Solo exposto
5	-37,30130	-7,01962	1,10	1,33	0,57	-0,09	Solo exposto
6	-37,30221	-7,02090	0,91	1,60	0,50	-0,28	Solo exposto
8	-37,30051	-7,02102	0,98	1,45	0,57	-0,19	Solo exposto

**Tabela B.7 – Amostras de treinamento na área 2 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
16	-37,30114	-7,02037	1,00	1,17	0,83	0,08	Arbustiva arbórea densa
18	-37,30301	-7,01974	0,96	1,05	0,99	0,04	Arbustiva arbórea densa
19	-37,29951	-7,01976	0,90	1,25	0,84	0,16	Arbustiva arbórea densa
20	-37,30085	-7,01936	0,91	1,00	1,10	0,05	Arbustiva arbórea densa
21	-37,30217	-7,01881	1,01	1,06	0,93	0,03	Arbustiva arbórea densa
22	-37,30331	-7,02089	1,04	1,08	0,88	0,02	Arbustiva arbórea densa
23	-37,29941	-7,02094	0,98	1,07	0,95	0,05	Arbustiva arbórea densa
25	-37,30179	-7,01947	0,99	1,04	0,97	0,03	Arbustiva arbórea densa
27	-37,30178	-7,02115	0,98	1,06	0,96	0,04	Arbustiva arbórea densa
28	-37,30028	-7,01944	1,03	0,99	0,98	-0,02	Arbustiva arbórea densa
32	-37,30252	-7,02086	0,97	1,01	1,02	0,02	Subarbustiva muito rala
34	-37,30111	-7,02189	1,01	1,01	0,98	0,00	Subarbustiva muito rala
35	-37,30263	-7,01885	1,02	1,00	0,98	-0,01	Subarbustiva muito rala

36	-37,30297	-7,02059	1,02	1,00	0,98	-0,01	Subarbustiva muito rala
38	-37,30161	-7,02183	0,99	1,00	1,01	0,00	Subarbustiva muito rala
39	-37,30315	-7,02054	0,98	1,00	1,01	0,01	Subarbustiva muito rala
40	-37,30195	-7,02136	1,00	1,00	1,00	0,00	Subarbustiva muito rala
41	-37,30080	-7,02157	0,99	1,01	1,00	0,01	Subarbustiva muito rala
44	-37,30346	-7,02132	1,00	0,99	1,02	0,00	Subarbustiva muito rala
45	-37,29985	-7,02209	0,97	0,99	1,04	0,01	Subarbustiva muito rala
2	-37,30077	-7,02145	1,04	0,98	0,97	-0,03	Solo exposto
3	-37,30252	-7,02111	1,01	1,00	0,99	0,00	Solo exposto
4	-37,30276	-7,01942	1,03	1,01	0,97	-0,01	Solo exposto
5	-37,30130	-7,01962	1,02	1,01	0,97	-0,01	Solo exposto
6	-37,30221	-7,02090	1,04	1,00	0,96	-0,02	Solo exposto
8	-37,30051	-7,02102	0,98	1,00	1,02	0,01	Solo exposto
11	-37,30208	-7,02153	1,08	0,99	0,93	-0,04	Solo exposto
12	-37,29933	-7,01940	1,03	0,99	0,98	-0,02	Solo exposto
14	-37,30116	-7,02067	1,04	0,99	0,97	-0,03	Solo exposto
15	-37,30319	-7,02061	1,06	0,99	0,95	-0,03	Solo exposto

**Tabela B.8 – Amostras de validação na área 2 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
17	-37,30057	-7,01999	0,96	1,07	0,97	0,05	Arbustiva arbórea densa
24	-37,30200	-7,02052	0,97	1,07	0,96	0,05	Arbustiva arbórea densa
26	-37,30367	-7,02010	0,98	1,07	0,95	0,04	Arbustiva arbórea densa
29	-37,30317	-7,01863	0,97	1,00	1,03	0,01	Arbustiva arbórea densa

30	-37,30051	-7,02044	1,12	0,98	0,90	-0,07	Arbustiva arbórea densa
31	-37,30238	-7,02080	1,01	1,00	0,98	0,00	Subarbustiva muito rala
33	-37,30277	-7,02122	1,00	0,99	1,01	-0,01	Subarbustiva muito rala
37	-37,30370	-7,02105	1,01	1,01	0,98	0,00	Subarbustiva muito rala
42	-37,29990	-7,02179	0,98	0,99	1,04	0,01	Subarbustiva muito rala
43	-37,30355	-7,02129	0,98	1,01	1,01	0,01	Subarbustiva muito rala
1	-37,30307	-7,02042	1,05	0,99	0,97	-0,03	Solo exposto
7	-37,30005	-7,02154	1,04	0,98	0,98	-0,03	Solo exposto
9	-37,30128	-7,01896	1,06	0,99	0,95	-0,04	Solo exposto
10	-37,30217	-7,02036	1,02	1,00	0,98	-0,01	Solo exposto
13	-37,29898	-7,01961	1,05	0,99	0,96	-0,03	Solo exposto

**Tabela B.9 – Amostras de treinamento na área 3 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
16	-37,30438	-7,01370	1,47	1,00	0,54	0,19	Arbustiva arbórea densa
17	-37,30615	-7,01326	1,67	0,89	0,44	0,30	Arbustiva arbórea densa
19	-37,30397	-7,01400	1,54	0,90	0,57	0,26	Arbustiva arbórea densa
20	-37,30524	-7,01327	1,35	1,06	0,59	0,12	Arbustiva arbórea densa
22	-37,30756	-7,01359	1,45	0,99	0,55	0,19	Arbustiva arbórea densa
24	-37,30756	-7,01461	1,38	1,03	0,59	0,14	Arbustiva arbórea densa
25	-37,30780	-7,01285	1,59	0,99	0,42	0,23	Arbustiva arbórea densa
28	-37,30714	-7,01341	1,36	0,95	0,69	0,18	Arbustiva arbórea densa
29	-37,30420	-7,01450	1,82	0,73	0,45	0,43	Arbustiva arbórea densa
30	-37,30517	-7,01346	1,47	0,94	0,59	0,22	Arbustiva arbórea densa

31	-37,30500	-7,01489	1,26	1,09	0,66	0,07	Subarbustiva muito rala
32	-37,30446	-7,01506	1,25	1,06	0,69	0,09	Subarbustiva muito rala
33	-37,30656	-7,01383	1,18	1,11	0,71	0,03	Subarbustiva muito rala
35	-37,30715	-7,01371	1,12	1,17	0,71	-0,03	Subarbustiva muito rala
37	-37,30586	-7,01312	1,29	1,06	0,65	0,10	Subarbustiva muito rala
39	-37,30703	-7,01426	1,19	1,14	0,66	0,02	Subarbustiva muito rala
40	-37,30676	-7,01439	1,24	1,14	0,62	0,04	Subarbustiva muito rala
41	-37,30726	-7,01400	1,14	1,16	0,70	-0,01	Subarbustiva muito rala
43	-37,30793	-7,01373	1,19	1,09	0,71	0,04	Subarbustiva muito rala
44	-37,30605	-7,01424	1,15	1,11	0,74	0,02	Subarbustiva muito rala
3	-37,30444	-7,01475	1,09	1,26	0,66	-0,07	Solo exposto
4	-37,30824	-7,01414	1,01	1,42	0,57	-0,17	Solo exposto
7	-37,30881	-7,01287	1,12	1,35	0,53	-0,10	Solo exposto
8	-37,30808	-7,01447	0,97	1,39	0,64	-0,18	Solo exposto
9	-37,30677	-7,01410	1,05	1,36	0,59	-0,13	Solo exposto
10	-37,30414	-7,01408	1,01	1,37	0,62	-0,15	Solo exposto
11	-37,30394	-7,01484	1,11	1,21	0,68	-0,04	Solo exposto
12	-37,30543	-7,01474	0,99	1,41	0,59	-0,17	Solo exposto
14	-37,30728	-7,01304	1,05	1,34	0,61	-0,12	Solo exposto
15	-37,30785	-7,01346	1,07	1,35	0,57	-0,11	Solo exposto

**Tabela B.10 – Amostras de validação na área 3 pelo método 1**

Ponto	Latitude	Longitude	Nir	Red	Green	NDVI	Classe
18	-37,30621	-7,01451	1,32	1,06	0,62	0,11	Arbustiva arbórea densa
21	-37,30881	-7,01452	1,29	1,13	0,58	0,07	Arbustiva arbórea densa

23	-37,30651	-7,01361	1,50	0,95	0,54	0,22	Arbustiva arbórea densa
26	-37,30858	-7,01405	1,46	0,98	0,56	0,20	Arbustiva arbórea densa
27	-37,30862	-7,01279	1,47	0,88	0,64	0,25	Arbustiva arbórea densa
34	-37,30836	-7,01419	1,19	1,14	0,68	0,02	Subarbustiva muito rala
36	-37,30561	-7,01333	1,22	1,07	0,71	0,06	Subarbustiva muito rala
38	-37,30864	-7,01480	1,29	1,13	0,58	0,06	Subarbustiva muito rala
41	-37,30726	-7,01400	1,14	1,16	0,70	-0,01	Subarbustiva muito rala
45	-37,30596	-7,01301	1,26	1,04	0,69	0,09	Subarbustiva muito rala
1	-37,30517	-7,01412	1,17	1,23	0,60	-0,02	Solo exposto
2	-37,30596	-7,01413	1,10	1,36	0,53	-0,11	Solo exposto
5	-37,30726	-7,01375	1,13	1,31	0,56	-0,08	Solo exposto
6	-37,30579	-7,01303	1,13	1,26	0,61	-0,05	Solo exposto
13	-37,30492	-7,01366	1,11	1,32	0,57	-0,08	Solo exposto

**Tabela B.11 – Amostras de treinamento na área 3 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
16	-37,30438	-7,01370	0,97	1,05	0,98	0,04	Arbustiva arbórea densa
18	-37,30621	-7,01451	1,04	1,00	0,97	-0,02	Arbustiva arbórea densa
19	-37,30397	-7,01400	0,99	0,99	1,02	0,00	Arbustiva arbórea densa
20	-37,30524	-7,01327	0,94	0,99	1,07	0,03	Arbustiva arbórea densa
24	-37,30756	-7,01461	1,01	0,97	1,03	-0,02	Arbustiva arbórea densa
25	-37,30780	-7,01285	0,92	0,96	1,11	0,02	Arbustiva arbórea densa
26	-37,30858	-7,01405	0,99	1,00	1,01	0,01	Arbustiva arbórea densa
28	-37,30714	-7,01341	0,88	1,00	1,13	0,06	Arbustiva arbórea densa

29	-37,30420	-7,01450	1,04	0,99	0,97	-0,02	Arbustiva arbórea densa
30	-37,30517	-7,01346	0,98	0,99	1,03	0,01	Arbustiva arbórea densa
31	-37,30500	-7,01489	1,04	1,01	0,95	-0,01	Subarbustiva muito rala
32	-37,30446	-7,01506	1,06	1,02	0,93	-0,02	Subarbustiva muito rala
33	-37,30656	-7,01383	1,02	1,00	0,98	-0,01	Subarbustiva muito rala
34	-37,30836	-7,01419	1,00	1,00	1,00	0,00	Subarbustiva muito rala
35	-37,30715	-7,01371	1,00	1,01	0,99	0,00	Subarbustiva muito rala
36	-37,30561	-7,01333	1,04	1,00	0,96	-0,02	Subarbustiva muito rala
40	-37,30676	-7,01439	1,01	1,01	0,98	0,00	Subarbustiva muito rala
41	-37,30726	-7,01400	1,01	1,01	0,98	0,00	Subarbustiva muito rala
42	-37,30728	-7,01328	1,01	1,01	0,98	0,00	Subarbustiva muito rala
44	-37,30605	-7,01424	1,02	1,00	0,98	-0,01	Subarbustiva muito rala
1	-37,30517	-7,01412	1,05	1,01	0,95	-0,02	Solo exposto
2	-37,30596	-7,01413	1,03	0,99	0,98	-0,02	Solo exposto
3	-37,30444	-7,01475	1,06	1,00	0,94	-0,03	Solo exposto
5	-37,30726	-7,01375	1,00	1,01	0,99	0,00	Solo exposto
6	-37,30579	-7,01303	1,01	1,01	0,98	0,00	Solo exposto
9	-37,30677	-7,01410	1,02	1,00	0,98	-0,01	Solo exposto
10	-37,30414	-7,01408	1,06	0,99	0,95	-0,04	Solo exposto
12	-37,30543	-7,01474	1,04	1,01	0,95	-0,01	Solo exposto
13	-37,30492	-7,01366	1,03	0,99	0,97	-0,02	Solo exposto
15	-37,30785	-7,01346	1,02	1,00	0,98	-0,01	Solo exposto

---

**Tabela B.12 – Amostras de validação na área 3 pelo método 2**

Ponto	Latitude	Longitude	Red	Green	Blue	NGRDI	Classe
17	-37,30615	-7,01326	0,91	1,03	1,06	0,06	Arbustiva arbórea densa
21	-37,30881	-7,01452	0,98	0,99	1,03	0,00	Arbustiva arbórea densa
22	-37,30756	-7,01359	0,58	1,00	1,42	0,27	Arbustiva arbórea densa
23	-37,30651	-7,01361	0,97	0,98	1,05	0,01	Arbustiva arbórea densa
27	-37,30862	-7,01279	0,93	0,97	1,11	0,02	Arbustiva arbórea densa
37	-37,30586	-7,01312	1,01	0,99	1,00	-0,01	Subarbustiva muito rala
38	-37,30864	-7,01480	1,05	1,01	0,93	-0,02	Subarbustiva muito rala
39	-37,30703	-7,01426	1,02	1,01	0,97	0,00	Subarbustiva muito rala
43	-37,30793	-7,01373	1,01	1,02	0,98	0,00	Subarbustiva muito rala
45	-37,30596	-7,01301	1,01	1,00	1,00	-0,01	Subarbustiva muito rala
4	-37,30824	-7,01414	1,05	0,99	0,96	-0,03	Solo exposto
7	-37,30881	-7,01287	1,03	1,02	0,95	0,00	Solo exposto
8	-37,30808	-7,01447	1,00	1,03	0,97	0,01	Solo exposto
11	-37,30394	-7,01484	1,03	1,00	0,97	-0,02	Solo exposto
14	-37,30728	-7,01304	1,02	1,00	0,97	-0,01	Solo exposto