

# AVALIAÇÃO DA PRECISÃO DA CLASSIFICAÇÃO DE COBERTURA DO SOLO NA IDENTIFICAÇÃO DE ÁREAS POTENCIAIS PARA O APARECIMENTO DE CRIADOUROS DE *Aedes Aegypti*

Israel Henrique Ribeiro Rios

*Universidade de São Paulo*

Maryly Weyll SantAnna

*Universidade de São Paulo*

José Alberto Quintanilha

*Universidade de São Paulo*

Carlos Henrique Grohmann

*Universidade de São Paulo*

## RESUMO

Uma ferramenta bastante utilizada atualmente é a classificação de cobertura do solo, em vários campos da ciência. Com o avanço das técnicas de sensoriamento remoto, pode-se utilizar resoluções melhores de imagens, e com isso a acurácia da classificação melhora. Este trabalho utiliza imagens do satélite World-View 3 para a classificação supervisionada de dois bairros da cidade de Campinas, São Paulo. As classes foram divididas em Árvore, Grama, Asfalto, Sombra, Cimento, Piscina/Água, Solo exposto e Telhados 1, 2 e 3. Essa classificação pode ser importante na procura de possíveis locais com focos de mosquito *Ae. Aegypti*, auxiliando os gestores e agentes comunitários de saúde. Foram averiguadas a matriz de erro, o coeficiente Kappa e a acurácia sob três diferentes algoritmos: Mínima Distância, Máxima Verossimilhança e *Random Forest*. De forma geral, o algoritmo de destaque foi o de Máxima Verossimilhança, índices Kappa classificados como muito bons. Esse tipo de procedimento pode ser utilizado para identificar áreas mais propensas ao aparecimento de ovos de mosquito do *Aedes Aegypti*, sobretudo em áreas próximas a acúmulo de água e telhados de amianto em zonas de maior vulnerabilidade social.

*Palavras-chave:* Sensoriamento remoto, classificação supervisionada, ocupação do solo.

## ABSTRACT

A widely used tool currently is land cover classification, in various fields of science. With the advancement of remote sensing techniques, better image resolutions can be used, and thus the classification accuracy improves. This work uses images from the World-View 3 satellite for the supervised classification of two neighborhoods in the city of Campinas, São Paulo. The classes were divided into Tree, Grass, Asphalt, Shade, Cement, Pool/Water, Exposed soil and Roofs 1, 2 and 3. This classification can be important when looking for possible locations with *Ae. mosquito* outbreaks. *Aegypti*, helping managers and community health agents. The error matrix, Kappa coefficient and accuracy were investigated under three different algorithms: Minimum Distance, Maximum Likelihood and Random Forest. In general, the highlighted algorithm was Maximum Likelihood, Kappa indices classified as very good. This type of procedure can be used to identify areas more prone to the appearance of *Aedes Aegypti* mosquito eggs, especially in areas close to water accumulation and asbestos roofs in areas of greater social vulnerability.

*Key words:* Remote sensing, supervised classification, land cover.

## Introdução

Técnicas de geoprocessamento e sensoriamento remoto vêm sendo utilizadas como ferramentas importantes no monitoramento da saúde populacional de centros urbanos (WU *et al.*, 2022). As informações geradas por meio de imagens de satélites contribuem para caracterização, mapeamento e identificação de agentes existentes no espaço, bem como na avaliação da extensão e intensidade da ação humana em áreas avaliadas (ZHAO *et al.*, 2022). A partir da análise de dados existentes nos espaços geográficos é possível identificar novos subsídios capazes de auxiliar no planejamento, avaliação e monitoramento espacial de ações de controle já implementadas, bem como a existência de possíveis riscos ambientais relacionados à saúde humana (DE OLIVEIRA LAGE, *et al.*, 2020). Dentre as doenças com grande impacto socioambiental encontram-se as arboviroses, as quais configuram-se como um importante problema de saúde pública global, principalmente em regiões tropicais (DONALISIO *et al.*, 2017; KAMAL *et al.*, 2018).

Entre as principais arboviroses existentes no Brasil destacam-se a dengue, zika e chikungunya, presentes em praticamente todo território nacional e com ocorrência de surtos frequentes ao longo dos anos (LIMA-CAMARA, 2016; AGUIAR *et al.*, 2018; PÉRISSÉ *et al.*, 2020). O principal vetor transmissor dessas enfermidades no país é o mosquito *Aedes aegypti* (VARIZA *et al.*, 2022; BRASIL, 2019), o qual tem sua reprodução acentuada em áreas urbanas (ALMEIDA *et al.*, 2020), visto que a presença humana, além de ser fonte de alimentação para as fêmeas (únicas transmissoras), também contribui para a existência de ambientes favoráveis à proliferação do inseto, como criadouros de oviposição (FERRAGUT *et al.*, 2022).

Tipos específicos de uso e cobertura do solo podem favorecer a proliferação de insetos, visto que áreas com grande cobertura vegetal e presença de água estão fortemente associadas com maior existência de desses vetores (LORENZ, *et al.*, 2020a). Somando-se a isso, a presença de telhados de amianto e lajes também apresentam associação positiva a maior ocorrência de mosquitos, pois a laje contribui para a existência de possíveis reservatório de larvas e os telhados de amianto estão associados às piores condições socioeconômicas (LORENZ, *et al.*, 2020b). Estudos apontam que áreas ocupadas por populações com menor nível econômico e maior densidade populacional são mais vulneráveis à ocorrência de doenças provocadas pelos arbovírus (TEICH *et al.*, 2017; DEALMEIDA e DA SILVA, 2018; CUNHA *et al.*, 2021a). Além disso, no estudo realizado por De Oliveira Lage (2020) ficou evidenciado a preferência do mosquito em depositar seus ovos em locais que aparecem com tons mais escuros em imagens satelitais e de drones, como sombras (principalmente de árvores e vegetação de maior porte), telhados e pavimentos de cimento e asfalto.

Percebe-se que a identificação e classificação de áreas urbanas favoráveis à existência de criadouros do *Ae. aegypti* configuram-se como importantes indicadores para o risco do aparecimento de arboviroses (CUNHA, 2021a). Desta forma, a compreensão do espaço urbano por meio de técnicas de geoprocessamento torna-se um fator auxiliador para o monitoramento e direcionamento de medidas de prevenção e controle dessas doenças (MOURA, 2020). O uso dessas técnicas é considerado um método rápido e com baixo custo operacional, facilitando assim a caracterização e identificação de possíveis áreas de risco para a presença do *Ae. aegypti* (CUNHA, 2021a).

A presença de criadouros está relacionada ao aumento do risco da transmissão de doenças (ABÍLIO, 2018), tornando a ocorrência de arbovirose mais frequente em regiões que apresentem maior disponibilidade de criadouros. Partindo dessa análise, este estudo tem como objetivo identificar áreas potenciais para o aparecimento de criadouros do vetor, através da classificação da cobertura dos solos em duas áreas distintas do município de Campinas, diferenciadas pelo risco de transmissão da doença. Foram utilizados três algoritmos de classificação em imagens digitais: Mínima Distância (*Minimum Distance/MD*), Máxima Verossimilhança (*Maximum Likelihood/ML*) e *Random Forest (RF)*, a fim de identificar os tipos de cobertura do solo urbano existentes na região e, após sua análise, na identificar áreas com maior e menor risco para a existência de criadouros e, conseqüentemente, de ocorrência de casos de arboviroses registrados na região.

Para a identificação e classificação das coberturas urbanas, utilizou-se uma imagem de alta resolução espacial obtida pelo sensor World-View 3 (DATA, RESOLUÇÕES ESPACIAIS E ESPECTRAIS) e a ferramenta *Semi Automatic Classification Plugin*, disponível no programa Quantum GIS (QGIS) versão 3.16.16. A exatidão das classificações foi avaliada utilizando Índice Kappa, a Matriz de Erro e os erros de comissão e de omissão obtidos em cada classificação. Espera-se com isso, resultados similares que possam indicar, com maior propriedade, aspectos fundamentais para o planejamento urbano e de saúde pública.

## *Material e métodos*

Para a realização deste estudo seguiram-se as seguintes etapas: 1) revisão da literatura sobre o uso de sensoriamento remoto na epidemiologia de arboviroses; 2) aquisição da imagem pelo satélite World-View 3 do Grupo de Pesquisa de Análise Espacial em Saúde (GEANES) da USP, projeto FAPESB; 3) adequação a partir do projeto com definição das classes de imagens e do uso dos classificadores; 4) obtenção dos parâmetros estatísticos das classificações; 5) análise dos resultados à luz do problema epidemiológico.

### *Área de estudo*

O município de Campinas está localizado na porção centro-leste do Estado de São Paulo (47°04'40" Longitude Oeste e 22°53'20" Latitude Sul), com altitude média de 680 metros acima do nível do mar, área territorial de 794,571 km<sup>2</sup> e população estimada de 1.223.237 habitantes (IBGE, 2021). As unidades administrativas escolhidas para a análise de ocupação do solo foram dois bairros, escolhendo-se um bairro de menor incidência de arboviroses e outro com maior incidência, de acordo com os dados fornecidos pelo serviço de epidemiologia municipal.

Os bairros escolhidos foram: Jardim Petrópolis, com baixo risco para casos de dengue e Vila 31 de março, onde apresenta elevado risco de dengue (CAMPINAS, 2022), dispostos na Figura 1. O critério de escolha das áreas foi a disponibilidade de imagens de satélite com boa qualidade (sem presença de nuvens) dentro da cidade; a possibilidade de delimitação dos arquivos vetoriais dos bairros; e oferta das imagens pelo grupo de pesquisa da universidade. De acordo com os avisos mensais emitidos pelo serviço de epidemiologia do município de Campinas, nos

cinco primeiros meses de 2022, o bairro de Vila 31 de Março se sobressaiu na incidência de casos de arboviroses, sendo mencionado como um dos que apresentou maior incidência de dengue, motivo pelo qual este foi escolhido para representar a área com maior risco para presença de focos do *Aedes aegypti*. Já o bairro Jardim Petrópolis não foi mencionado nas notas de alertas emitidas pela prefeitura, o que sugere baixa incidência de casos, sendo escolhido para representar a área de menor incidência e baixo risco para presença de focos do vetor (CAMPINAS, 2022).

### *Base de dados*

A classificação da ocupação do solo foi realizada a partir da imagem GlobalGeo, 30 cm de resolução espacial, World-View 3, bandas R, G, B e NIR, do mosaico de 13/05/2019, 14/06/2019 e 05/03/2020. As imagens obtidas já haviam sido processadas (ortorretificadas, mosaicadas e realçadas) em formato GeoTIFF – 8 bits, projeção UTM e *datum* SIRGAS 2000.

Os arquivos em *Shapefile* relativos aos bairros foram criados no Google Earth pelos autores a partir de informações de delimitação de bairros dos mapas do Google.

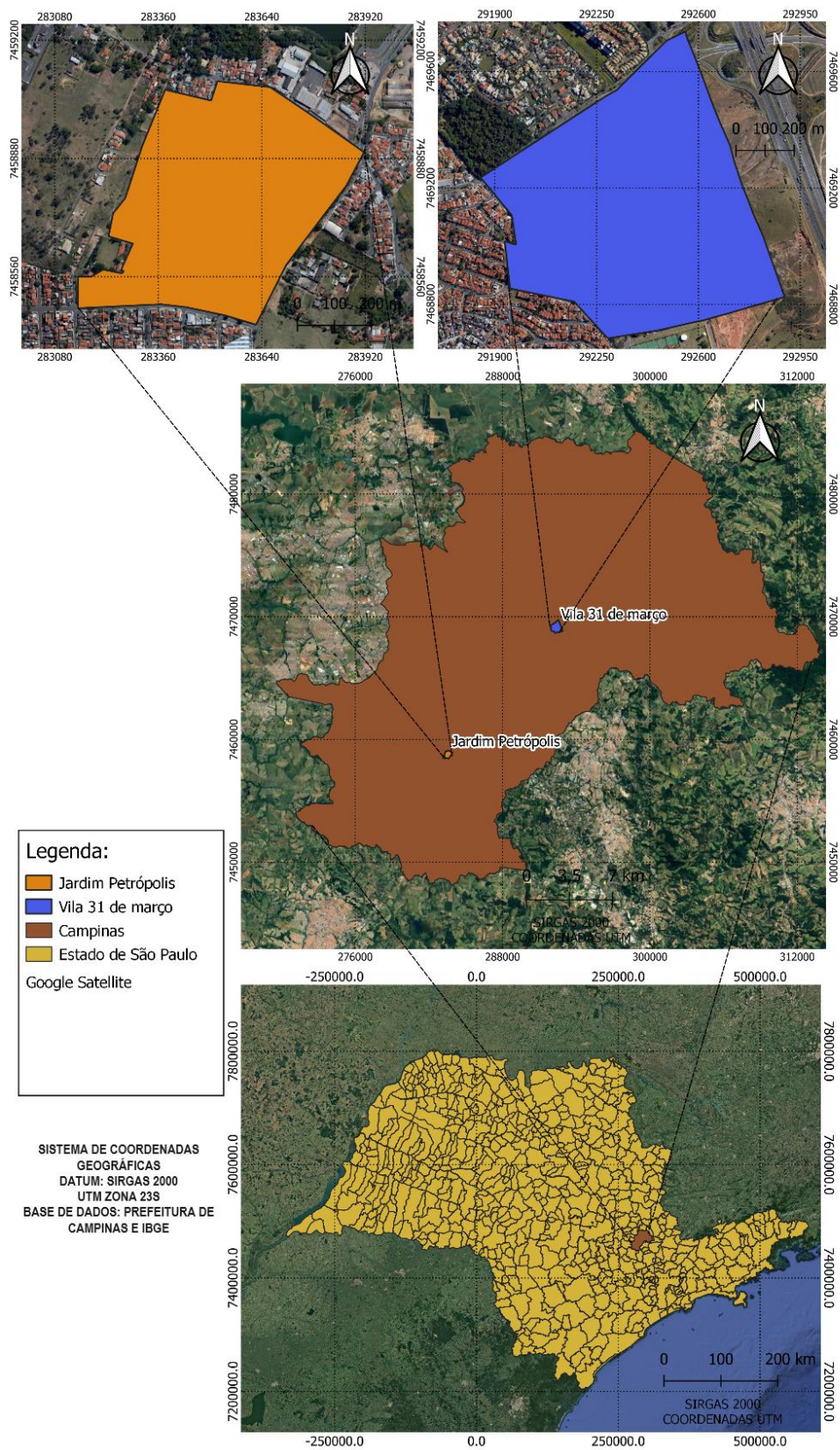
### *Fluxograma geral*

A Figura 2 mostra um fluxograma resumindo o processo utilizado neste trabalho. A partir da imagem de satélite foram utilizados dois complementos/plugins do software QGIS para classificação e posterior validação de cada uma das classificações geradas. A ferramenta *Semi Automatic Classification Plugin (SCP)*, por método de classificação supervisionada, foi utilizada para processar as classes de cobertura do solo sob três métodos distintos: Mínima Distância (*Minimum Distance/MD*), o de Máxima Verossimilhança (*Maximum Likelihood/ ML*) e o *Random Forest (RF)*. O treinamento da classificação supervisionada foi realizado pelos autores. O *Accuracy Assessment of Thematic Maps (AcATaMA)* foi utilizado para validação e cálculos estatísticos referentes às imagens classificadas.

### *Classificação da imagem*

A classificação da imagem foi executada por meio do programa QGIS 3.16.16 com a utilização do *Semi Automatic Classification Plugin*, utilizando amostragem supervisionada. A quantidade de amostras para cada classe foi definida a partir de uma estimativa visual da área de cada classe, seguindo o critério de diferentes cores – baseadas nas bandas espectrais – para cada amostra, apresentado no Quadro 1. A definição das classes foi baseada no estudo de De Oliveira Lage (2020), tendo-se um total de 96 amostras.

O número de amostras para validação utilizado para cada classe consta também no Quadro 1, estimados numa proporção de acordo com a quantidade utilizada no treinamento da classificação.



**Figura 1.** Município de Campinas e áreas envolvidas. Fonte: Autores, 2022. Imagens de fundo: Google Earth.

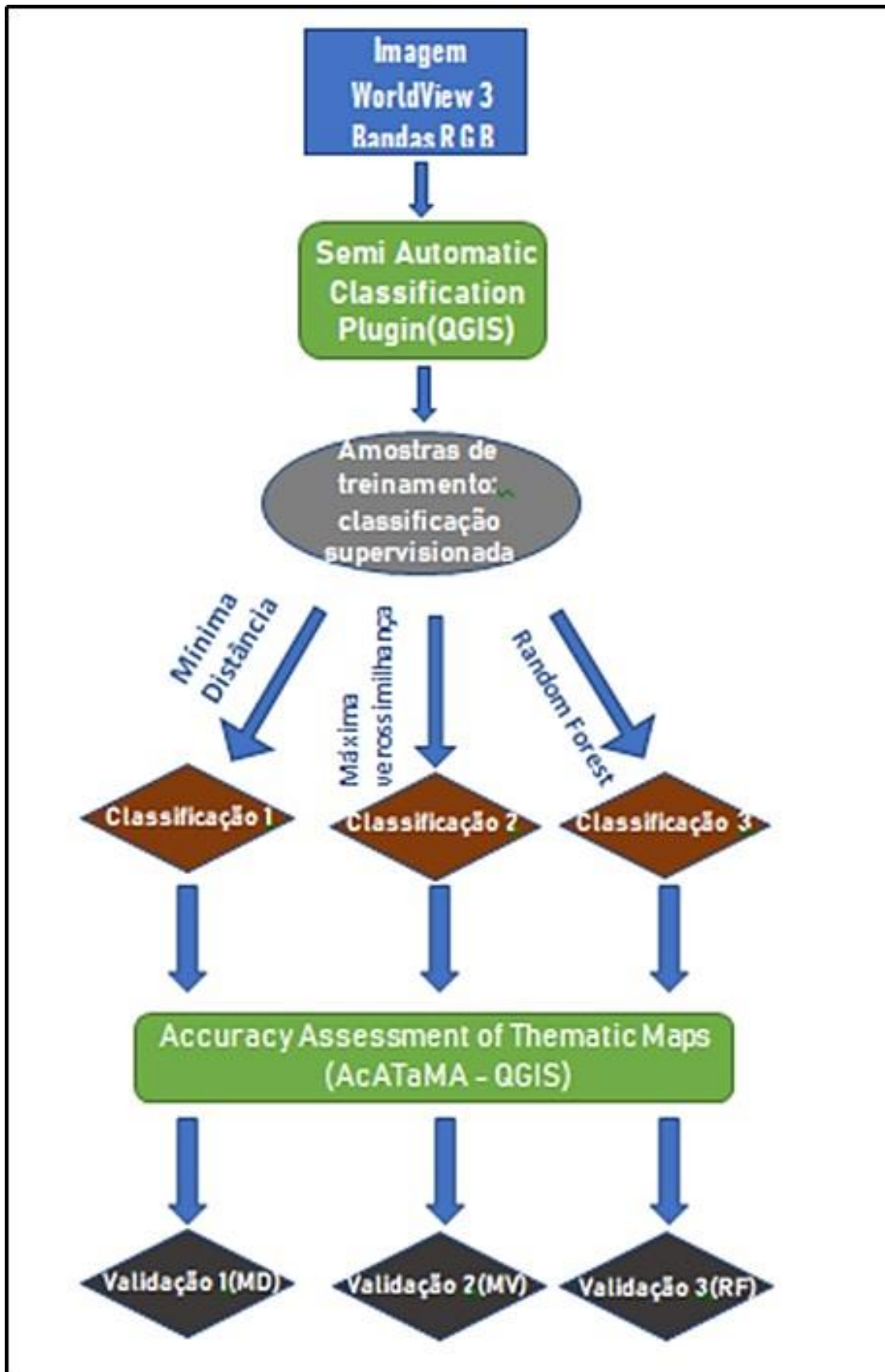












Figura 2. Fluxograma. Fonte: Autores, 2022.

Classes de cobertura do solo	Cor utilizada	Número de amostras coletadas para o treinamento	Número de amostras coletadas para a validação
Árvore		10	30
Gramma		10	30
Asfalto		12	30
Sombra		10	30
Cimento		10	30
Piscina/Água		8	30
Solo exposto		10	30
Telhado 1 telha		12	30
Telhado 2 claro		10	30
Telhado 3 escuro		10	30
Total de amostras		102	300

**Quadro 1.** Número de amostras por classe e cores utilizadas. Fonte: Autores, 2022.

O algoritmo da Distância Mínima utiliza as características espectrais das amostras de treinamento que foram escolhidas como representantes das classes de objetos selecionados, calculando a Distância Euclidiana entre os valores dos *pixels* candidatos e os valores médios de cada classe. O candidato é alocado para a classe com a Distância Euclidiana mais curta (LEE *et al.*, 2016).

A classificação por Máxima Verossimilhança/ML supervisionada realiza a classificação dos pixels com base em parâmetros estatísticos como: média, variância e matrizes de covariância das imagens, verificando a probabilidade que cada pixel tem de pertencer a cada classe definida. Já o algoritmo *Random Forest* está fundamentado no princípio de árvores de decisão, a partir da seleção aleatória de um subconjunto de amostras e suas variáveis de treinamento (CHO *et al.*, 2021; BELGIU e DRĂGUȚ, 2016). Utilizado o número de dez árvores para a classificação.

### *Validação da classificação*

Após o processamento da classificação, foi realizada a avaliação a fim de averiguar a acurácia desse procedimento. Isso se deu por meio da ferramenta (complemento do programa) AcA-TaMA, com a finalidade de obter a acurácia das classificações, comparando as classes das imagens classificadas com a imagem de satélite original. Para avaliar os resultados da classificação foi utilizada a matriz de erro/confusão (DIBS *et al.*, 2020).

Além disso, foram utilizados o índice Kappa e a Acurácia Global, pois são consideradas boas ferramentas de avaliação da qualidade dos resultados da classificação (DIBS *et al.*, 2020). Rwanda (2017) coloca que a precisão do produtor mede erros de omissão, que é uma medida de quão bem os tipos de cobertura do solo no mundo real podem ser classificados. Já a precisão do usuário mede erros de comissão, o que representa a probabilidade de um pixel classificado corresponder ao tipo de cobertura do solo da sua correspondente localização no mundo real. A matriz

de erro e o coeficiente kappa tornaram-se um meio padrão de avaliação da precisão da classificação de imagens; a classificação pelo Kappa deu-se conforme Congalton e Green (1999).

## *Resultados e discussão*

Os mapas de classificação da cobertura do solo envolvendo os três algoritmos constam nas figuras a seguir. Na Figura 3 observa-se as classes de uso/ocupação na área do bairro Jardim Petrópolis e na Figura 4 as classes na área do bairro Vila 31 de Março. Os telhados escuros foram associados ao material amianto e os telhados claros aos telhados de vidro ou PVC.

Dentro das imagens se pode observar no bairro Jardim Petrópolis uma parte central de predomínio de vegetação com moradias espalhadas nas laterais da área. Já no bairro Vila 31 de Março a área de abrangência de vegetação (grama e árvore) é bem maior, sendo que ela cobre praticamente toda a região lateral direita da área, existindo moradias na parte esquerda e sul do mapa. Percebe-se que de modo geral os algoritmos apresentam semelhanças nas classificações, tanto na área com menor vegetação (Jardim Petrópolis), como na área de maior vegetação (Vila 31 de Março).

Colaborando nessa análise, temos nas Figuras 5 e 6 as porcentagens de cada classe de cobertura do solo em cada bairro. Visualmente observado e confirmado pelos dados nos gráficos pode-se perceber que a área de Jardim Petrópolis tem uma área maior relativa à classe de asfalto devido a um número maior de vias de tráfego. A maior área de abrangência de vegetação no bairro Vila 31 de Março foi confirmada através dos gráficos. Uma maior proporção de áreas de telhados de telha foi observada também no bairro Jardim Petrópolis. Além disso, foi detectado um maior número de áreas com piscina também no bairro Jardim Petrópolis, o que é um indício da maior condição de renda da população dentro desse bairro (CARLUCCI *et al.*, 2020). Dentro do trabalho, classificou-se a área de Vila 31 de Março como alto risco de incidência, a alta proporção de áreas de vegetação dentro desse bairro pode ser um fator que colabore para a dispersão de vetores nas localidades (LORENZ, *et al.*, 2020a).

Percebe-se que embora o bairro Jardim Petrópolis tenha maior concentração de áreas mais escuras, representadas por sombra, cimento e asfalto, o que contribuiria para o aumento de vetores, segundo De Oliveira Lage *et al.* (2020), esse bairro possui baixa incidência de casos. Um fator que pode explicar essa divergência é o maior nível econômico desse bairro quando comparado ao bairro de Vila 31 de Março, corroborando a importância de fatores socioeconômicos no aumento da incidência de arboviroses (DEALMEIDA e DA SILVA, 2018; CUNHA *et al.*, 2021a).

Conforme Almeida *et al* (2007), as distinções na transmissão da dengue são determinadas pelas interações entre ser humano, os vetores e o meio ambiente, e essa interação define o contato entre mosquitos e indivíduos suscetíveis e infectados. Neste estudo de Almeida *et al* (2007), diagnosticou-se que fatores significativamente associados às áreas de maior risco foram a menor renda do chefe da família, maior densidade domiciliar e maior proporção de crianças e idosos mulheres. As áreas mais vulneráveis sob o ponto de vista socioeconômico apresentam morfologias distintas, ora com maior densidade de edificações, ora com presença maior da vegetação, principalmente em áreas de transição entre o urbano e rural. No estudo de (LORENZ, 2020a), os níveis



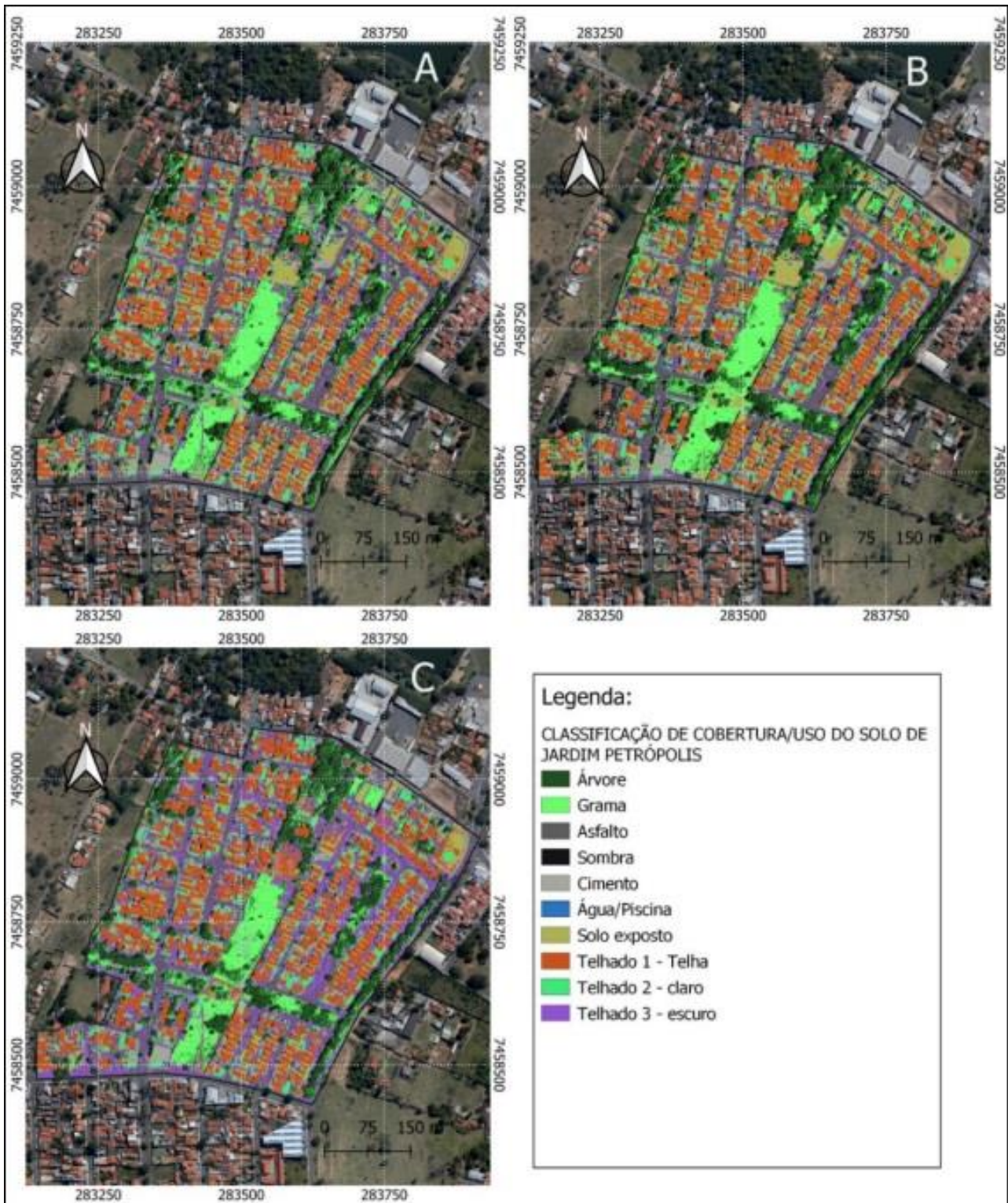
de infestação não foram associados às áreas cobertas por árvores; em que embora a região estivesse localizada em uma região tropical com ótimas condições para o ciclo de vida do mosquito, a vegetação tipicamente densa que ali cresce não é necessariamente favorável aos mosquitos que preferem se reproduzir dentro e ao redor das estruturas urbanas. Colocam ainda que os mosquitos urbanos que preferem recipientes artificiais de água para oviposição; e foi encontrado uma relação positiva em ambientes de telhado de amianto e solo exposto (que os mosquitos preferem habitar). No estudo realizado por Cunha *et al.*, (2021b) a associação dengue-verde foi modificada por fatores socioeconômicos, sendo encontrada uma associação positiva entre casos de dengue e vegetação em localidades com maior vulnerabilidade socioeconômica, e uma associação negativa em regiões com menor vulnerabilidade social, o que pode ser confirmado neste estudo.

Com relação ao tipo de telhado não foi possível perceber diferença nos tipos de telhados, visto que ocorreu proporção semelhante de tipos de telhados nas áreas avaliadas, não sendo possível perceber a influência de telhados de amianto e lage no aumento de casos de dengue, como o que foi observado no estudo de Lorenz *et al.* (2020b).

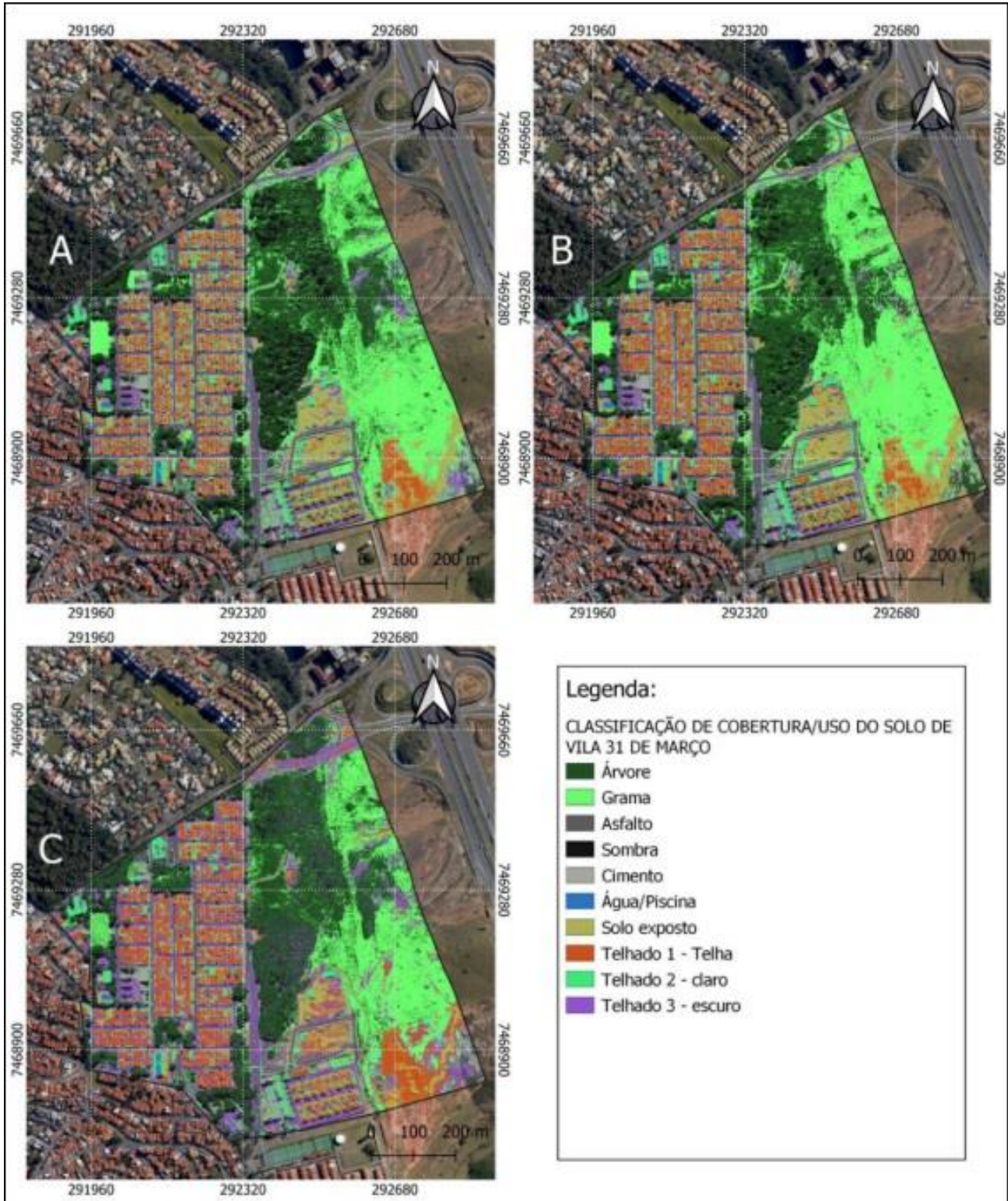
Diante dessas informações percebe-se que o algoritmo de Máxima Verossimilhança tende a ter uma área maior da classe relacionada à água e uma razoável área de sombra, o que pode ser importante na averiguação de locais com foco de dengue, já que com uma área maior tenderia a errar por excesso e não por falta.

Com isso, prosseguiu-se para a validação dos mapas de classes com a elaboração do cálculo das matrizes de erro referentes a cada bairro, acompanhada do valor das acurácias individuais de cada classe e o valor do coeficiente Kappa, conforme apresentado nas Tabelas 1 e 2.

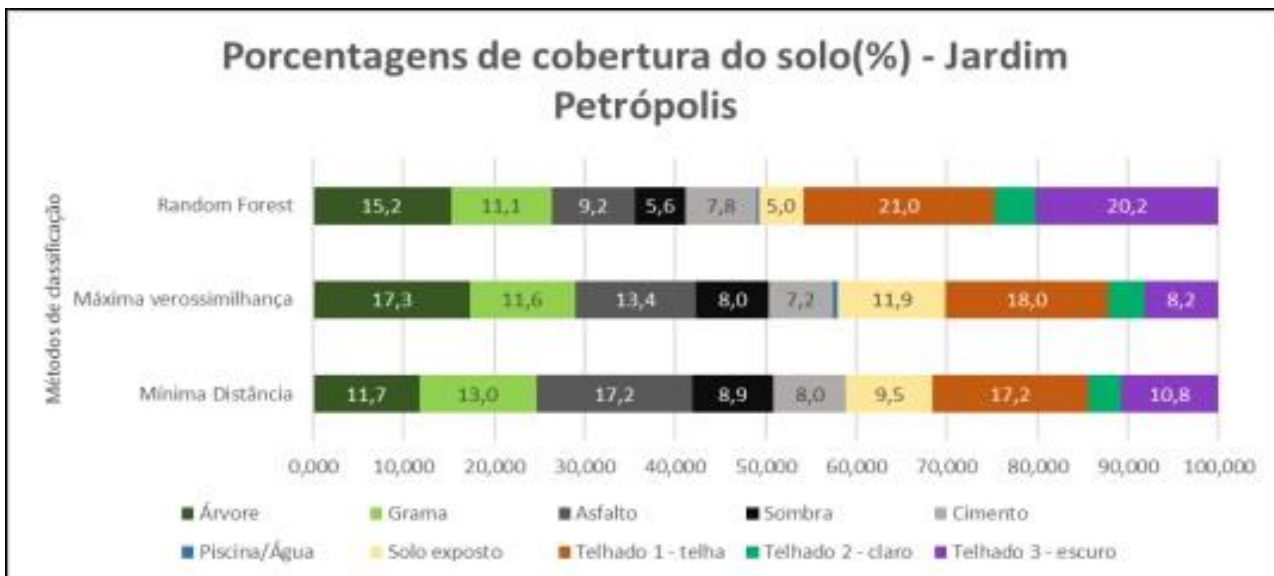
A validação envolve dois tipos de erro, conforme apresentado por Rwangi *et al.* (2017), os erros de comissão que descrevem a chance de um pixel ser associado a uma determinada classe, porém pertencer a outra classe; e o erro de omissão que se reflete no número de pontos que não estão incluídos na categoria, enquanto realmente pertencem à categoria. Ainda segundo Rwangi *et al.* (2017), a acurácia do usuário ou do consumidor refere-se ao erro de comissão (O.C), enquanto a acurácia do produtor refere-se ao erro de omissão (O.E), conforme apresentados nas Tabela 1 e Tabela 2 a seguir.



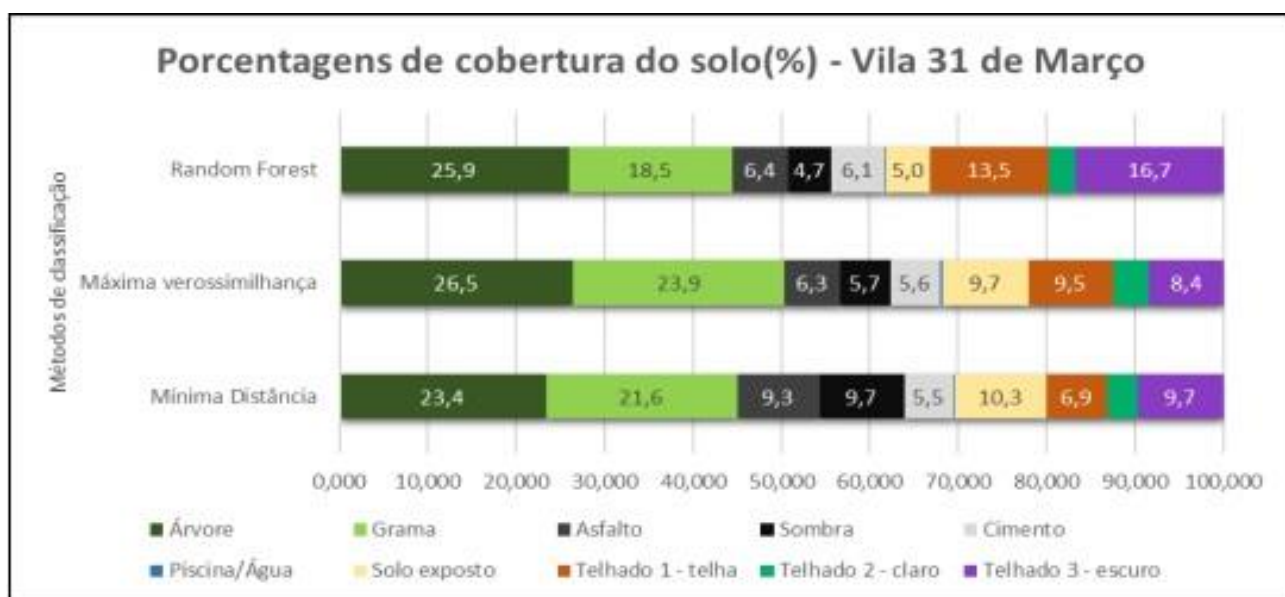
**Figura 3.** Resultado da classificação supervisionada pelos algoritmos MD (A), ML (B) e RF (C) na área de Jardim Petrópolis. Fonte: World-View 3, 2022. Imagens de fundo: Google Earth.



**Figura 4.** Resultado da classificação supervisionada pelos algoritmos MD (A), ML (B) e RF (C) na área da Vila 31 de Março.



**Figura 5.** Porcentagem de ocupação do solo segundo os diferentes algoritmos de classificação na área de Jardim Petrópolis. Fonte: Autores, 2022.



**Figura 6.** Porcentagem de ocupação do solo segundo os diferentes algoritmos de classificação na área de Vila 31 de Março. Fonte: Autores, 2022.

Mínima distância															
Classes	1 (Árvo-re)	2 (Gra-ma)	3 (Asfal-to)	4 (Som-bra)	5 (Cimen-to)	6 (Pisci-na/Água)	7 (Solo expos-to)	8 (Telha-do 1-telha)	9 (Telha-do 2-claro)	10 (Telha-do 3-escuro)	Total	E. C.	P(B )	P(A)XP(B)	User accuracy
1	21	2	0	6	0	0	0	1	0	0	30	9	0,1	0,01	0,70
2	6	15	1	2	0	0	5	0	0	1	30	15	0,1	0,00866667	0,50
3	1	4	18	1	0	0	3	1	0	2	30	12	0,1	0,009	0,60
4	1	1	0	27	0	0	1	0	0	0	30	3	0,1	0,015	0,90
5	0	0	0	0	20	0	3	0	6	1	30	10	0,1	0,00666667	0,67
6	0	0	0	0	0	28	0	0	2	0	30	2	0,1	0,00966667	0,93
7	0	0	0	0	0	0	20	7	3	0	30	23	0,1	0,012	0,67
8	0	0	0	0	0	0	1	29	0	0	30	1	0,1	0,01266667	0,97
9	0	0	0	0	0	1	0	0	29	0	30	1	0,1	0,01366667	0,97
10	1	4	8	9	0	0	3	0	1	4	30	26	0,1	0,00266667	0,13
total	30	26	27	45	20	29	36	38	41	8	300	102		0,1	
P(A)	0,10	0,09	0,09	0,15	0,07	0,10	0,12	0,13	0,14	0,03					KAPPA
E.O.	9	11	9	18	0	1	16	9	12	4	89				0,670
Produce-accuracy	0,7	0,57692	0,666667	0,6	1	0,96552	0,555556	0,763158	0,707317	0,5				K1	0,703
Máxima verossimilhança															
Classes	1 (Árvo-re)	2 (Gra-ma)	3 (Asfal-to)	4 (Som-bra)	5 (Cimen-to)	6 (Pisci-na/Água)	7 (Solo expos-to)	8 (Telha-do 1-telha)	9 (Telha-do 2-claro)	10 (Telha-do 3-escuro)	Total	E. C.	P(B )	P(A)XP(B)	User accuracy
1	25	3	0	2	0	0	0	0	0	0	30	5	0,1	0,01133333	0,83
2	5	23	0	0	0	0	2	0	0	0	30	7	0,1	0,01033333	0,77
3	0	1	23	0	2	0	0	0	0	4	30	7	0,1	0,00966667	0,77
4	1	0	0	28	0	0	0	0	0	1	30	2	0,1	0,013	0,93
5	0	0	1	0	20	0	5	0	4	0	30	10	0,1	0,00866667	0,67
6	0	3	1	0	2	17	0	0	7	0	30	13	0,1	0,006	0,57
7	0	0	0	0	0	0	15	14	1	0	30	16	0,1	0,00766667	0,50
8	0	0	0	0	0	0	0	30	0	0	30	0	0,1	0,01466667	1,00
9	0	0	0	0	1	1	0	0	28	0	30	2	0,1	0,01333333	0,93
10	3	1	4	9	1	0	1	0	0	11	30	19	0,1	0,00533333	0,37
total	34	31	29	39	26	18	23	44	40	16	300	81		0,1	
P(A)	0,11	0,10	0,10	0,13	0,09	0,06	0,08	0,15	0,13	0,05					KAPPA
E.O.	9	8	6	11	6	1	8	14	12	5	80				0,703
Produce-accuracy	0,74	0,74	0,79	0,72	0,77	0,94	0,65	0,68	0,70	0,69				K1	0,733
Random Forest															
Classes	1 (Árvo-re)	2 (Gra-ma)	3 (Asfal-to)	4 (Som-bra)	5 (Cimen-to)	6 (Pisci-na/Água)	7 (Solo expos-to)	8 (Telha-do 1-telha)	9 (Telha-do 2-claro)	10 (Telha-do 3-escuro)	Total	E. C.	P(B )	P(A)XP(B)	User accuracy
1	21	7	0	2	0	0	0	0	0	0	30	9	0,1	0,01066667	0,70
2	2	17	1	3	1	0	5	0	0	1	30	13	0,1	0,00966667	0,57
3	0	2	22	0	1	0	1	0	0	4	30	8	0,1	0,00866667	0,73
4	2	0	0	28	0	0	0	0	0	0	30	2	0,1	0,01333333	0,93

5	0	0	0	1	17	0	6	1	5	0	30	13	0,1	0,007666 67	0,57
6	0	1	0	1	1	16	0	0	11	0	30	14	0,1	0,005333 33	0,53
7	0	0	0	1	0	0	16	11	2	0	30	19	0,1	0,009666 67	0,53
8	0	0	0	1	0	0	0	29	0	0	30	1	0,1	0,014333 33	0,97
9	0	0	0	0	2	0	0	0	28	0	30	2	0,1	0,015333 33	0,93
10	7	2	3	3	1	0	1	2	0	11	30	19	0,1	0,005333 33	0,37
total	32	29	26	40	23	16	29	43	46	16	300	10 0		0,1	
P(A)	0,11	0,10	0,09	0,13	0,08	0,05	0,10	0,14	0,15	0,05					KAPPA
E.O.	11	12	4	12	6	0	13	14	18	5	95				0,648
Pro- ducer accu- racy	0,66	0,59	0,85	0,70	0,74	1,00	0,55	0,67	0,61	0,69				K1	0,683

**Tabela 1.** Tabelas unificadas da Matriz de Erro por diferentes algoritmos de classificação da área de Jardim Petrópolis. Fonte: Autores, 2022.

Classes	1 (Árvore)	2 (Gramma)	3 (Asfalto)	4 (Sombra)	5 (Cimento)	6 (Piscina/Água)	7 (Solo exposto)	8 (Telhado 1-telha)	9 (Telhado 2-claro)	10 (Telhado 3-escuro)	Total	E.C.	P(B)	P(A)XP(B)	User accuracy
1	15	10	1	1	0	0	3	0	0	0	30	15	0,10	0,01	0,67
2	1	29	0	0	0	0	0	0	0	0	30	1	0,10	0,02	0,83
3	0	7	13	0	5	0	4	0	0	1	30	17	0,10	0,01	1,00
4	2	0	0	28	0	0	0	0	0	0	30	2	0,10	0,01	1,00
5	0	0	0	0	18	0	5	0	7	0	30	12	0,10	0,01	1,00
6	0	0	0	0	0	7	0	0	23	0	30	23	0,10	0,00	0,75
7	0	0	0	0	0	0	14	16	0	0	30	16	0,10	0,01	0,17
8	0	0	0	0	0	0	1	29	0	0	30	1	0,10	0,02	1,00
9	0	0	0	0	1	0	1	2	26	0	30	4	0,10	0,02	1,00
10	5	3	2	5	1	0	12	0	0	2	30	28	0,10	0,00	0,25
total	23	49	16	34	25	7	40	47	56	3	300	119		0,10	
P(A)	0,08	0,16	0,05	0,11	0,08	0,02	0,13	0,16	0,19	0,01					KAPPA
E.O.	8	20	3	6	7	0	26	18	30	1	119				0,56
Producer accuracy	0,80	1,00	0,75	0,67	1,00	1,00	1,00	0,60	0,67	1,00				K1	0,603
Classes	1 (Árvore)	2 (Gramma)	3 (Asfalto)	4 (Sombra)	5 (Cimento)	6 (Piscina/Água)	7 (Solo exposto)	8 (Telhado 1-telha)	9 (Telhado 2-claro)	10 (Telhado 3-escuro)	Total	E.C.	P(B)	P(A)XP(B)	User accuracy
1	23	5	0	2	0	0	0	0	0	0	30	7	0,10	0,01	0,83
2	1	27	0	0	0	0	2	0	0	0	30	3	0,10	0,01	0,83
3	0	0	21	2	1	0	2	0	1	3	30	9	0,10	0,01	0,83
4	1	0	0	29	0	0	0	0	0	0	30	1	0,10	0,01	1,00
5	0	0	0	1	18	0	4	0	7	0	30	12	0,10	0,01	1,00
6	0	1	1	2	2	10	0	0	14	0	30	20	0,10	0,00	0,00
7	0	1	0	1	0	0	24	3	1	0	30	6	0,10	0,02	0,83
8	0	0	0	2	0	0	4	24	0	0	30	6	0,10	0,01	1,00
9	0	1	0	0	1	0	2	0	26	0	30	4	0,10	0,02	0,75
10	5	3	8	5	0	0	7	1	0	1	30	29	0,10	0,00	0,50
total	30	38	30	44	22	10	45	28	49	4	300	97		0,10	
P(A)	0,10	0,13	0,10	0,15	0,07	0,03	0,15	0,09	0,16	0,01					KAPPA
E.O.	7	11	9	15	4	0	21	4	23	3	97				0,641
Producer accuracy	0,71	1,00	0,83	0,86	1,00	-	0,71	1,00	0,38	1,00				K1	0,677
Classes	1 (Árvore)	2 (Gramma)	3 (Asfalto)	4 (Sombra)	5 (Cimento)	6 (Piscina/Água)	7 (Solo exposto)	8 (Telhado 1-telha)	9 (Telhado 2-claro)	10 (Telhado 3-escuro)	Total	E.C.	P(B)	P(A)XP(B)	User accuracy
1	21	6	0	2	0	0	1	0	0	0	30	9	0,10	0,01	1,00

2	1	26	0	0	0	0	3	0	0	0	30	4	0,10	0,01	0,83
3	0	4	16	3	1	0	3	0	0	3	30	14	0,10	0,01	0,83
4	0	1	0	28	0	0	1	0	0	0	30	2	0,10	0,01	1,00
5	0	0	0	1	14	0	7	0	8	0	30	16	0,10	0,01	0,80
6	0	0	0	0	0	10	0	0	20	0	30	20	0,10	0,00	0,50
7	0	1	0	0	1	0	22	6	0	0	30	8	0,10	0,02	0,67
8	0	0	0	0	0	0	5	25	0	0	30	5	0,10	0,01	1,00
9	0	0	0	0	0	0	2	0	28	0	30	2	0,10	0,02	1,00
10	3	6	8	4	1	0	7	0	0	1	30	29	0,10	0,00	0,75
total	25	44	24	38	17	10	51	31	56	4	300	109		0,10	
P(A)	0,08	0,15	0,08	0,13	0,06	0,03	0,17	0,10	0,19	0,01					KAPPA
E.O.	4	18	8	10	3	0	29	6	28	3	109				0,6
Producer accuracy	0,86	1,00	1,00	0,86	0,80	1,00	0,80	0,75	0,67	1,00				K1	0,637

**Tabela 2.** Tabelas unificadas da Matriz de Erro por diferentes algoritmos de classificação da área da Vila 31 de Março. Fonte: Autores, 2022.

Nas Matrizes de Erro das localidades os erros de omissão e comissão estão descritos por classes. Estudos apontam que as classes que envolvem vegetação e são mais escurecidas apresentam uma associação positiva com a presença do vetor (LORENZ, *et al.*, 2020a e 2020b; DE OLIVEIRA LAGE, 2020). Diante disso, foram destacadas as classes que envolvem vegetação e cor escura na avaliação do erro de omissão e erro de comissão. Percebe-se que as classes grama e sombra apresentaram menos erros de comissão nos três classificadores; as classes Telhado 1 e Telhado 2 também tiveram poucos erros de comissão. Já Telhado 3 e Piscina tiveram pouquíssimos erros de omissão. De forma geral, as classes que tiveram maior frequência de erros de comissão foram a Água/Piscina e Telhado 3. Com relação aos erros de omissão, as classes de Solo Exposto e Telhado 2 se sobressaíram. Observando-se a dificuldade de diagnosticar corretamente os pixels relativos ao Telhado 3, que se confundiram muitas vezes com diversas classes. A confusão da classe de Água/Piscina com Telhado 2 (claro) também é nítida na observação das matrizes. Uma constatação é que os telhados claros podem entrar junto com a classe do cimento, já que quando se olha imagem do *Google Earth* envolvendo a área, esse espectro de cores estão bastante próximos (acinzentadas), o que na imagem coletada do *World View* não se observou com tal nitidez (verde azulado e acinzentado).

Com base nas às acurácias individuais, dos dados gerados nas matrizes de erro, observou-se uma dificuldade geral na validação da classe Piscina/Água, confundindo-se bastante com telhados claros (onde muitos desses também tinham cor azul). Outro conflito encontrado foi a distinção entre o Telhado 3 (escuro) com Asfalto e Sombra, já que a cor do amianto seria um tom de cinza intermediário entre essas duas classes. A classificação do Telhado 2 – claros foi bastante eficaz, assim como a da Sombra, com erros de comissão iguais a zero para todos os algoritmos de classificação.

O valor dos coeficientes Kappa foi muito bom em quase todos os métodos algoritmos avaliados, sendo entre 0,6 e 0,7 segundo estudo de Freitas *et al* (2023), denotando uma concordância



satisfatória com a área de referência, conforme apresentado no Quadro 2 abaixo. Na área de Jardim Petrópolis e também no bairro Vila 31 de março o maior índice Kappa foi alcançado através do algoritmo de Máxima Verossimilhança devido aos menores erros de comissão e omissão do método no caso desta área.

Índices Kappa		
Método de classificação	Jardim Petrópolis	Vila 31 de Março
Mínima Distância	0,670	0,559
Máxima Verossimilhança	0,703	0,641
<i>Random Forest</i>	0,648	0,597

**Quadro 2.** Índices Kappa dos algoritmos de classificação. Fonte: Autores, 2022.

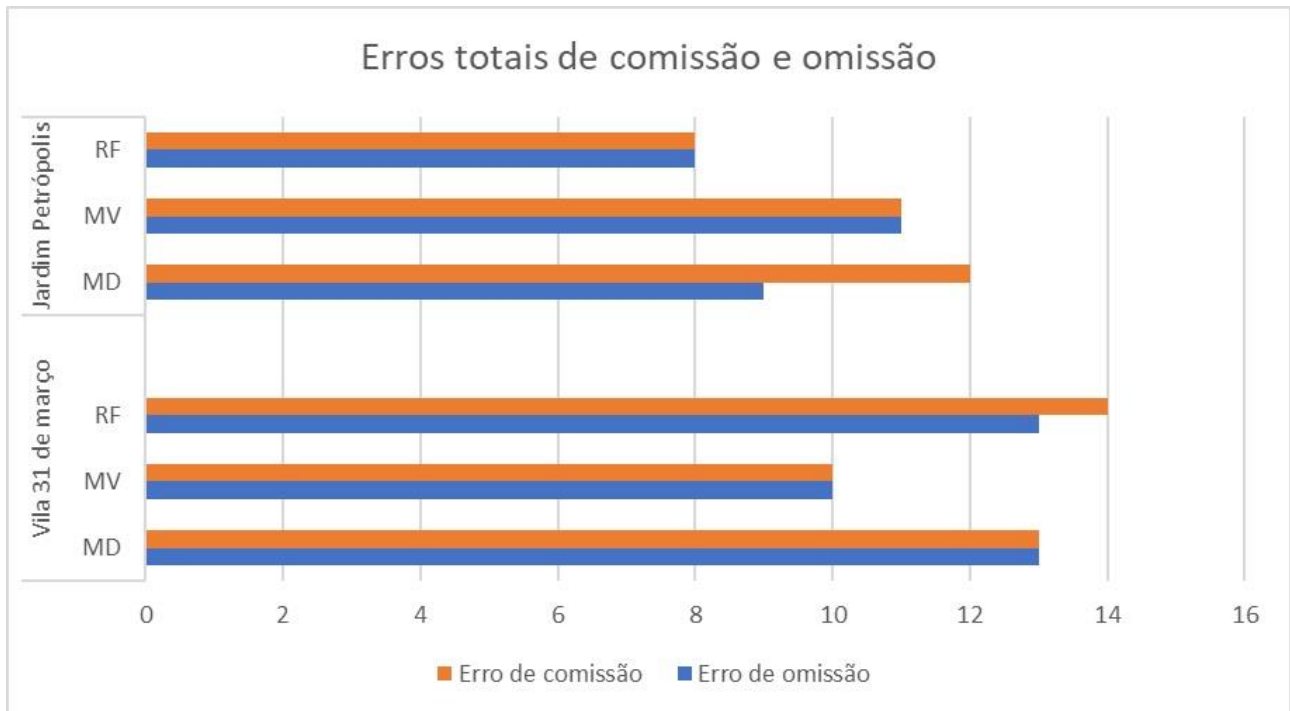
Tem-se ainda no Quadro 3 as acurácias globais/gerais calculadas para cada método de classificação. Para ambas as áreas o método da Máxima Verossimilhança teve melhor desempenho de acurácia global, com mais de 65% deste parâmetro em ambas as áreas. Considerando os valores tanto do Kappa quanto da acurácia global, percebe-se a razoável eficiência da classificação da ocupação do solo dos três métodos averiguados, em que estes têm bom desempenho na comparação com a imagem de referência.

Acurácias globais (%)		
Métodos de classificação	Jardim Petrópolis	Vila 31 de Março
Mínima distância	70,3	60,3
Máxima Verossimilhança	73,3	67,7
<i>Random Forest</i>	68,3	63,7

**Quadro 3.** Acurácias globais das classificações. Fonte: Autores, 2022.

O desempenho de cada classificação pode ser melhor demonstrado por meio dos erros de comissão e omissão de cada método em cada área averiguada assim como realizado por DUARTE e SILVA (2019), apresentados na Figura 7 abaixo. De modo específico os erros de omissão individuais correspondentes à classe de sombra são zero, indicando um bom uso de todas as três classes; já com relação a classe de Piscina/Água houve dificuldade de adequação à imagem de referência devido aos conflitos com outras classes. Os erros de comissão da classe sombra são bastante baixos.

Ao ser analisado os erros de comissão e omissão percebe-se que o Máxima Verossimilhança exibiu erros acumulados de omissão e comissão semelhantes no bairro de Jardim Petrópolis. Além disso, foi o algoritmo que apresentou o menor número de erros em ambos os bairros envolvidos. O RF apresentou um maior número de erros de comissão no bairro Vila 31 de Março, enquanto que no Jardim Petrópolis foi o algoritmo de Mínima Distância. De forma geral, com relação ao número de erros acumulados (Comissão e Omissão), os classificadores de MD e RF ficaram bastante próximos quando comparados com o MV.



**Figura 7.** Erros de comissão e omissão. Fonte: Autores, 2022.

Desta forma, foi demonstrado que os três classificadores utilizados apresentam comportamentos distintos quando utilizados em áreas diferentes, mas bons desempenhos na validação. Na área com alto risco para arbovirose, o classificador MV exibiu o maior Kappa (0,40) e menores erros acumulados de comissão e omissão. Em áreas com baixo risco, o MV foi também o que apresentou melhor Kappa (0,26) e menos erros acumulados de omissão e comissão, demonstrando a importância do treinamento com diferentes algoritmos, como etapa de decisão na escolha do algoritmo antes da realização do estudo.

### Conclusão/Considerações finais

A partir da classificação da cobertura do solo por meio de três métodos de classificação, verificou-se o desempenho de cada classificação. Com base nos resultados obtidos observou-se que os métodos de classificação de ocupação do solo aplicado apresentaram bons desempenhos, alcançando precisões gerais (globais) acima de 70% – e coeficientes Kappa acima de 20% – com exceção de dois algoritmos no Vila 31 de março – indicando que a classificação pelos três métodos teve razoáveis desempenhos dentro do contexto. O método de Máxima Verossimilhança se sobressaiu na performance do índice de acurácia global na classificação da ocupação/cobertura do solo, com valores de 0,73 e 0,68 e teve índices Kappa de 0,70 e 0,64. O método de Mínima Distância teve pior desempenho de acurácia global no caso do bairro Vila 31 de março (0,60) e baixos valores de Kappa (0,67 e 0,56) – e o *Random Forest* teve dois coeficientes Kappa muito bons nas duas áreas (0,65 e 0,60) e menor acurácia global no contexto do Jardim Petrópolis (0,68). Os resultados obtidos demonstram que é possível realizar o mapeamento da cobertura do

solo com boa qualidade utilizando dados de sensoriamento remoto e tecnologia de uso livre (QGIS).

Este trabalho calculou as porcentagens de classe de ocupação do solo correspondente a cada área envolvida – em que o bairro caracterizado como de maior risco de incidência de dengue se destacou por uma maior abrangência de grama e árvores. As informações obtidas por meio do uso destas ferramentas dão subsídio a órgãos de planejamento territorial e ambiental, além de órgãos de saúde no contexto da pré- observação de possíveis áreas com foco de mosquitos relacionadas a áreas de próximas à piscina e telhados escuros de amianto em zonas de maior vulnerabilidade social.

---

*Submetido em 29 de março de 2023.*

*Aceito para publicação em 3 de janeiro de 2024.*

## Referências

- ABÍLIO, Ana Paula *et al.* Distribution and breeding sites of *Aedes aegypti* and *Aedes albopictus* in 32 urban/peri-urban districts of Mozambique: implication for assessing the risk of arbovirus outbreaks. **PLoS neglected tropical diseases**, v. 12, n. 9, p. e0006692, 2018.
- AGUIAR, Breno S. *et al.* Potential risks of Zika and chikungunya outbreaks in Brazil: A modeling study. **International Journal of Infectious Diseases**, v. 70, p. 20-29, 2018.
- ALMEIDA, Lorena Sampaio; COTA, Ana Lídia Soares; RODRIGUES, Diego Freitas. Saneamento, Arboviroses e Determinantes Ambientais: impactos na saúde urbana. **Ciência & Saúde Coletiva**, v. 25, p. 3857-3868, 2020.
- ALMEIDA, MC de Mattos, Caiaffa WT, Assunção RM, Proietti FA. Spatial vulnerability to dengue in a Brazilian urban area during a 7-year surveillance. **Journal of Urban Health**. 2007 May;84(3):334-45. DOI: 10.1007/s11524-006-9154-2. PMID: 17243024; PMCID: PMC2231833.
- BELGIU, Mariana; DRĂGUȚ, Lucian. Random forest in remote sensing: A review of applications and future directions. **ISPRS journal of photogrammetry and remote sensing**, v. 114, p. 24-31, 2016.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. **Guia de Vigilância em Saúde**: volume único [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços. – 3ª. ed. – Brasília: Ministério da Saúde, 2019.
- CAMPINAS. DEVISA. **INFORME EPIDEMIOLÓGICO ARBOVIROSES**. 2022. Disponível em: <[https://dengue.campinas.sp.gov.br/sites/dengue.campinas.sp.gov.br/files/boletins\\_arbovirose/Informe%20Epidemiol%C3%B3gico%20Arboviroses%20Campinas%20SP%20-%202022.pdf](https://dengue.campinas.sp.gov.br/sites/dengue.campinas.sp.gov.br/files/boletins_arbovirose/Informe%20Epidemiol%C3%B3gico%20Arboviroses%20Campinas%20SP%20-%202022.pdf)>. Acessado em 03 out. 2023.
- CAMPINAS. Prefeitura de Campinas. **Relatório de informações sociais do município de campinas**. Disponível em: <[https://smcais-vis.campinas.sp.gov.br/sites/smcais-vis.campinas.sp.gov.br/files/arquivos/relatorio\\_de\\_informacoes\\_sociais\\_campinas\\_-\\_2016\\_0.pdf](https://smcais-vis.campinas.sp.gov.br/sites/smcais-vis.campinas.sp.gov.br/files/arquivos/relatorio_de_informacoes_sociais_campinas_-_2016_0.pdf)>. Acessado em 18 mai. 2022.
- CAMPINAS. Prefeitura Municipal de Campinas. **Informativo mostra 17 novas áreas com risco de transmissão de dengue**. 2022. Disponível em: <<https://portal.campinas.sp.gov.br/noticia/43990>>. Acessado em 21 out. 2022.
- CARLUCCI, Margherita *et al.* Socio-spatial disparities and the Crisis: Swimming Pools as a Proxy of Class Segregation in Athens. **Social Indicators Research**, p. 1-25, 2020.
- CHO, David Fernando *et al.* Desempenho do Algoritmo de Classificação de Imagens *Random Forest* para Mapeamento do Uso e Cobertura do Solo no Cerrado Brasileiro. Anuário do Instituto de Geociências, v. 44, UFRJ, 2021. [https://doi.org/10.11137/1982-3908\\_2021\\_44\\_37979](https://doi.org/10.11137/1982-3908_2021_44_37979).

- CUNHA, Higor Souza *et al.* Water tank and swimming pool detection based on remote sensing and deep learning: Relationship with socioeconomic level and applications in dengue control. **Plos one**, v. 16, n. 12, p. e0258681, 2021a.
- CUNHA, Maria da Consolação Magalhães *et al.* Disentangling associations between vegetation greenness and dengue in a Latin American city: Findings and challenges. **Landscape and urban planning**, v. 216, p. 104255, 2021b.
- DE OLIVEIRA LAGE, Mariana *et al.* Ovipositional Reproduction of the Dengue Vector for Identifying High-Risk Urban Areas. **EcoHealth**, v. 19, n. 1, p. 85-98, 2022.
- DEALMEIDA, Caio Américo Pereira; DA SILVA, Richarde Marques. Análise da ocorrência dos casos de dengue e sua relação com as condições socioambientais em espaços urbanos: os casos de João Pessoa, Cabedelo e Bayeux, no estado da Paraíba– Brasil. **Hygeia: Revista Brasileira de Geografia Médica e da Saúde**, v. 14, n. 27, p. 56, 2018.
- DIBS, Hayder *et al.* An optimal approach for land-use/land-cover mapping by integration and fusion of multispectral landsat OLI images: case study in Baghdad, Iraq. **Water, Air, & Soil Pollution**, v. 231, n. 9, p. 1-15, 2020.
- DONALISIO, Maria Rita; FREITAS, André Ricardo Ribas; ZUBEN, Andrea Paula Bruno Von. Arboviroses emergentes no Brasil: desafios para a clínica e implicações para a saúde pública. **Revista de saúde pública**, v. 51, 2017.
- DUARTE, Miqueias Lima; DA SILVA, Tatiana Acácio. Avaliação do desempenho de três algoritmos na classificação de uso do solo a partir de geotecnologias gratuitas. **Revista de estudos Ambientais**, v. 21, n. 1, p. 6-16, 2019.
- FERRAGUTI, Martina; MAGALLANES, Sergio; IBÁÑEZ-JUSTICIA, Adolfo. Implication of human landscape transformation on mosquito populations. In: GUTIÉRREZ-LOPEZ, R., LOGAN, J. G., PUENTE J. M. (Eds.). **Ecology and Control of Vector-borne Diseases. Wageningen Academic Publishers**, 2022.
- FREITAS, Leonardo Cristiano da Silva *et al.* Avaliação da acurácia das classes de uso e cobertura da terra do Mapbiomas (coleção 61) para o município de Carpina (PE). **Caderno Prudentino de Geografia**, Presidente Prudente, n. 45, v. 2, p. 38-52, mai-ago 2023.
- IBGE – Instituto Brasileiro de Geografia e Estatística. **Cidades**. Disponível em: <<https://cidades.ibge.gov.br/brasil/sp/campinas/panorama>>. Acesso em: 01 mai 2022.
- KAMAL, Mahmoud *et al.* Mapping the global potential distributions of two arboviral vectors *Aedes aegypti* and *Ae. albopictus* under changing climate. **Plos one**, v. 13, n. 12, p. e0210122, 2018.
- LEE, Janice Ser Huay *et al.* Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. **Remote Sensing Applications: Society and Environment**, v. 4, p. 219-224, 2016.
- LIMA-CAMARA, Tamara Nunes. Arboviroses emergentes e novos desafios para a saúde pública no Brasil. **Revista de Saúde Pública**, v. 50, 2016.
- LORENZ, Camila *et al.* Predicting *Aedes aegypti* infestation using landscape and thermal features. **Scientific reports**, v. 10, n. 1, p. 1-11, 2020a.

LORENZ, Camila *et al.* Remote sensing for risk mapping of *Aedes aegypti* infestations: Is this a practical task? **Acta Tropica**, v. 205, p. 105398, 2020b.

MOURA, Márcia Cristina Bernardo de Melo *et al.* Spatio-temporal dynamics of *Aedes aegypti* and *Aedes albopictus* oviposition in an urban area of northeastern Brazil. **Tropical Medicine & International Health**, v. 25, n. 12, p. 1510-1521, 2020.

PÉRISSÉ, André Reynaldo Santos *et al.* Zika, dengue and chikungunya population prevalence in Rio de Janeiro city, Brazil, and the importance of seroprevalence studies to estimate the real number of infected individuals. **Plos one**, v. 15, n. 12, p. e0243239, 2020.

RWANGA, Sophia S. *et al.* Accuracy assessment of land use/land cover classification using remote sensing and GIS. **International Journal of Geosciences**, v. 8, n. 04, p. 611, 2017.

TEICH, Vanessa; ARINELLI, Roberta; FAHHAM, Lucas. *Aedes aegypti* e sociedade: o impacto econômico das arboviroses no Brasil. **JBES: Brazilian Journal of Health Economics/Jornal Brasileiro de Economia da Saúde**, v. 9, n. 3, 2017.

VARIZA, Paula Fassicolo *et al.* Distribuição espaço-temporal atualizada de *Aedes* (*Stegomyia*) *albopictus* no Brasil. **Acta Tropica**, v. 232, p. 106511, 2022.

WU, Sijia *et al.* Neglected urban villages in current vector surveillance system: evidences in Guangzhou, China. **International journal of environmental research and public health**, v. 17, n. 1, p. 2, 2020.

ZHAO, Min *et al.* Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives. **Remote Sensing**, v. 11, n. 17, p. 1971, 2019.