

Covariáveis preditivas e estratégias de amostragem para mapeamento automático do uso e ocupação na sub-região sudoeste da região metropolitana de São Paulo

Predictive variables and sampling strategies for automatic classification of landcover at southwest subregion of the metropolitan region of São Paulo

Alan Pessoa Valadares

Instituto Agronômico, Campinas, SP, Brasil

alanvaladares@ipt.br

 <https://orcid.org/0000-0002-9002-0114>

Carlos Tadeu de Carvalho Gamba

Instituto de Pesquisas Tecnológicas, São Paulo, SP, Brasil

carlosgamba@ipt.br

 <https://orcid.org/0000-0002-4139-4355>

Ricardo Marques Coelho

Instituto Agronômico, Campinas, SP, Brasil

ricardo.coelho@sp.gov.br

 <https://orcid.org/0000-0001-8508-380X>

RESUMO

Alternativamente à classificação com base em pixels, a classificação baseada em objetos tem alcançado resultados satisfatórios quando aplicada a imagens de alta resolução, para mapeamento de uso e ocupação. Contudo, em nível de semidetalhado, pouco se sabe sobre as vantagens e desvantagens com cada abordagem, utilizando imagens com resolução espacial média (~10 m). O efeito da quantidade e da frequência na distribuição de amostras para o treinamento de modelos preditivos, em cada classe de interesse, e a eficiência dos diferentes tipos de covariáveis preditivas foram pouco investigados. O objetivo deste estudo foi avaliar diferentes estratégias de amostragem para o treinamento de modelos preditivos e a eficiência das covariáveis, comparando abordagens de classificação com base em pixels e em objetos. A área de estudo foi a sub-região sudoeste da Região Metropolitana de São Paulo e a classificação supervisionada foi realizada utilizando o algoritmo *Random Forest*. Foram testados diferentes conjuntos de dados, quanto à quantidade e frequência da distribuição das amostras nas classes de uso e ocupação, para treinamento do classificador. Utilizaram-se informações espectrais da missão Sentinel-2, informações topográficas de interferometria por radar da missão SRTM, índices de vegetação e de água, além de informações sobre a variabilidade temporal e espacial de covariáveis. A qualidade do mapeamento foi melhor quando presentes no conjunto de covariáveis preditivas os índices de vegetação e de água, a variabilidade espacial e a topografia, na abordagem de classificação com base em pixels, alcançando 82,4% de acurácia e índice Kappa de 0,80. Na classificação com base em objetos, a variabilidade temporal dos índices e a morfologia dos objetos melhoraram os resultados, alcançando acurácia e índice Kappa de 78,15% e 0,75, respectivamente. O resultado da classificação é bastante sensível à quantidade e à distribuição das amostras entre as classes no treinamento. As abordagens testadas apresentaram qualidade semelhante.

Palavras-chave: Uso e ocupação; Random Forest; Classificação supervisionada; Covariáveis preditivas; Google Earth Engine.

ABSTRACT

Alternatively of pixel-based, object-based classification have reached satisfactory results when applied to high resolution images, for landcover mapping. However, at semi-detailed level (1:50,000-scale) the best way to go remains unknown. Few studies compared those approaches using medium resolution images (~ 10 m). The effect of quantity and frequency distribution of training observations at each landcover classes are poorly investigated, as well as the efficiency of different

kind of predictive variables. The objective of this study was to evaluate different sampling strategies for model training and the efficiency of predictive variables, comparing the results at pixel-based and object-based approaches. The mapping area was the southwest of metropolitan region of São Paulo, and the supervised classification was performed by Random Forest. Six databases were tested by changing quantity and frequency distribution of training observations at each landcover class. We used spectral information from Sentinel-2 mission, topographic information from SRTM, vegetation and water indexes, beyond of information about spatial-temporal variation of indexes. Maps performance was better when vegetation and water indexes, spatial variation and topographic information were added to training set data in pixel-based approach, reaching 82.44 % of accuracy and Kappa index of 0.80. At object-based approach, adding temporal variation of indexes and information about segments geometries were important for map improvement, reaching accuracy and Kappa index of 78.15 % e 0.75, respectively. The results of the maps depend of amount and frequency distribution of observation by classes used for training dataset. The both of approaches tested show similar quality. Toward to choose the better fit among classification results it is main to know the study region characteristics.

Keywords: Landcover; Random Forest; Supervised classification; Predictive covariates; Google Earth Engine.

1. INTRODUÇÃO

Mapas de uso da terra são produtos cartográficos de grande relevância para o gerenciamento dos recursos naturais, permitindo análises espaço-temporais em diversas áreas do conhecimento (GÓMEZ; WHITE; WULDER, 2016). Com o monitoramento do território, é possível inferir a dinâmica dos corpos d'água, da cobertura vegetal, da urbanização e também a ocorrência de crimes ambientais, por exemplo. Dada sua importância, métodos que tornem esses mapeamentos mais rápidos, exatos e menos onerosos são altamente desejáveis. É o que se espera da utilização de métodos de classificação automática do uso da terra.

A inteligência artificial tem sido testada em diversos países para automatizar a classificação do uso da terra a partir de informações extraídas de imagens geradas por sensoriamento remoto. Contudo, ainda não há resultados conclusivos que indiquem qual o melhor caminho a seguir (KHATAMI; MOUNTRAKIS; STEHMAN, 2016). Cada estudo tem suas particularidades, que se referem ao contexto próprio de cada local e às abordagens metodológicas empregadas, tornando muitas vezes difícil uma comparação justa entre os trabalhos, mesmo considerando uma linha metodológica geral (MA; *et al.*, 2017).

Dois tipos de abordagens têm sido empregados na classificação supervisionada do uso da terra por aprendizagem de máquina: a abordagem baseada em pixels, na qual um classificador considera cada pixel individualmente como uma instância a ser classificada; e a classificação orientada a objetos (classificação por segmentação), capaz de associar as formas geométricas dos objetos às suas respectivas classes, por meio da segmentação de imagens (Object-based Image Analysis - OBIA), possibilitando que cada objeto (polígono) seja tratado como uma instância a ser classificada (HAY; CASTILLA, 2008).

Segundo Blaschke *et al.* (2014), a abordagem OBIA surge como um novo paradigma no sensoriamento remoto, uma vez que imagens de alta resolução se tornaram prontamente disponíveis. Nestas imagens, a resolução espacial dos pixels é menor que a maioria dos objetos de interesse nos mapeamentos. Entre as principais vantagens apontadas pela utilização da abordagem OBIA, estão a redução da variação espectral dentro das classes nos dados de treinamento e a remoção do efeito “sal e pimenta”, comum nos mapas gerados pela abordagem com base em pixels (LIU; XIA, 2010).

Muitos estudos comparativos entre as duas abordagens utilizaram imagens de alta (entre 1 m e 5 m) ou de muito alta resolução (< 1 m), contrastando com a baixa quantidade de trabalhos realizados em imagens com resolução espacial em torno de 30 m e ainda menor daqueles que utilizaram imagens com resolução espacial em torno de 10 m. Belgiu e Csillik (2018) compararam as duas abordagens utilizando imagens do Sentinel-2 e encontraram valores muito próximos entre os dois métodos para o mapeamento de culturas agrícolas.

Outros aspectos que precisam ser investigados são a quantidade de informações para o treinamento de modelos e a proporção de observações entre as classes para poder alcançar resultados satisfatórios. Estes aspectos apresentam sérias implicações do ponto de vista prático. Li *et al.* (2014) avaliaram o efeito da quantidade de dados para treinamento e verificaram que diferentes algoritmos de classificação apresentam diferentes quantidades mínimas de observações e uma tendência à estabilização da acurácia global a partir de certas quantidades de dados para treinamento. Além disso, a quantidade ótima de observações e a proporção entre as classes também devem ser afetadas pela diversidade de classes no contexto de cada estudo.

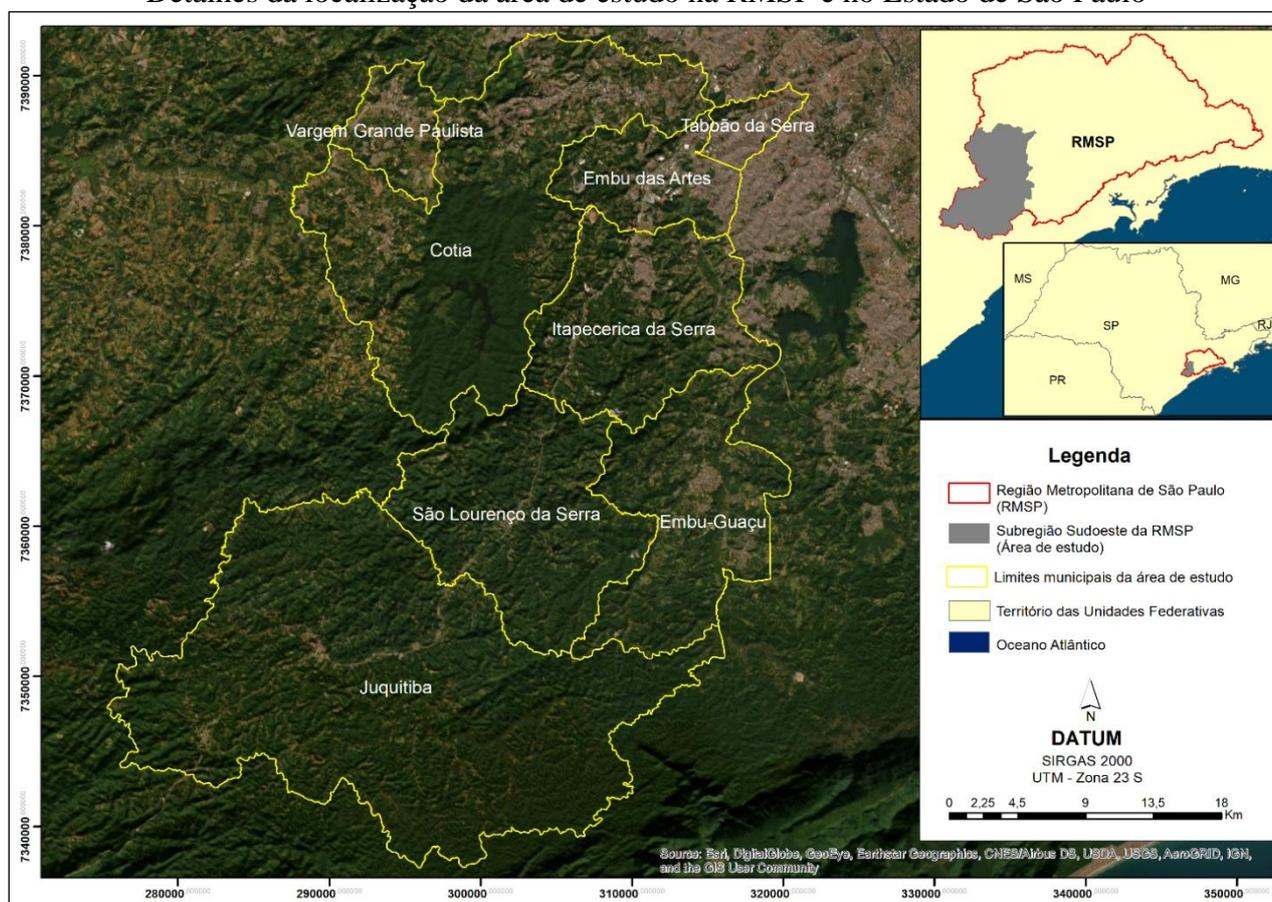
Além das informações espectrais obtidas pelos sensores, a inclusão de informações topográficas, de atributos que informam sobre a variabilidade espacial das variáveis preditivas (texturas) e sobre a variabilidade temporal de índices como o NDVI e o NDWI (GAO, 1996), tem melhorado o desempenho dos classificadores na diferenciação entre as classes (KHATAMI; MOUNTRAKIS; STEHMAN, 2016).

Neste trabalho avaliamos o efeito da quantidade e da distribuição de observações para treinamento de modelos, as características e a eficiência de variáveis preditivas, comparando as abordagens de classificação supervisionada com base em pixels e em objetos.

2. MATERIAL E MÉTODOS

A área de estudo representa a sub-região sudoeste da Região Metropolitana de São Paulo (RMSP), abrangendo os municípios de Cotia, Embu das Artes, Embu-Guaçu, Itapeverica da Serra, Juquitiba, São Lourenço da Serra, Taboão da Serra e Vargem Grande Paulista, como mostra a **figura 1**. Esta região permite avaliar o mapeamento do uso da terra em escalas de publicação próximas a 1:50.000, com área mínima mapeável de aproximadamente 10 ha.

Figura 1: Municípios da sub-região sudoeste da Região Metropolitana de São Paulo (RMSP). Detalhes da localização da área de estudo na RMSP e no Estado de São Paulo



Fonte: Elaborado pelos autores (2022).

Para esse nível de detalhe, optou-se por utilizar as imagens do satélite Sentinel-2, da Agência Espacial Europeia (European Space Agency - ESA). O Sentinel-2 produz imagens multispectrais com resolução espacial de 10 m nas bandas do visível e do infravermelho próximo (Near Infrared – NIR), 20 m nas bandas do infravermelho de ondas curtas (Short Wave Infrared – SWIR) e Red Edge (DRUSCH *et al.*, 2012). O **quadro 1** resume as especificações espectrais das bandas do Sentinel-2 utilizadas neste estudo.

O processamento digital das imagens, a geração das variáveis preditivas, a classificação supervisionada e a validação do mapa final foram executadas no Google Earth Engine (GEE). Trata-se de uma plataforma web que integra sistemas de informação geográfica, sensoriamento remoto e

aprendizagem de máquina (GORELICK *et al.*, 2017). O GEE dispõe de uma ampla biblioteca de dados de sensores orbitais e de informações derivadas destes sensores. Além disso, apresenta uma enorme capacidade de processamento, sendo capaz de lidar com quantidades massivas de dados, podendo gerar informações em larga escala (SHELESTOV *et al.*, 2017). A classificação supervisionada utilizando algoritmos sofisticados demanda alto poder de processamento de dados. Por permitir que os dados sejam processados em nuvem, a utilização do GEE evita a necessidade de ter um hardware de alto desempenho. O GEE permite realizar procedimentos complexos em poucos minutos (ROSA, 2018), trazendo vantagens operacionais bastante satisfatórias.

Quadro 1: Informações sobre as bandas utilizadas do Sentinel-2. (Adaptado de DRUSCH *et al.*, 2012)

Nome	Resolução	Comprimento de onda	Descrição
B2	10 m	496.6 nm (S2A) / 492.1nm (S2B)	Azul
B3	10 m	560 nm (S2A) / 559 nm (S2B)	Verde
B4	10 m	664.5 nm (S2A) / 665 nm (S2B)	Vermelho
B5	20 m	703.9 nm (S2A) / 703.8 nm (S2B)	Red Edge 1
B6	20 m	740.2 nm (S2A) / 739.1 nm (S2B)	Red Edge 2
B7	20 m	782.5 nm (S2A) / 779.7 nm (S2B)	Red Edge 3
B8	10 m	835.1 nm (S2A) / 833 nm (S2B)	NIR
B8A	20 m	864.8 nm (S2A) / 864 nm (S2B)	Red Edge 4
B11	20 m	1613.7 nm (S2A) / 1610.4 nm (S2B)	SWIR 1
B12	20 m	2202.4 nm (S2A) / 2185.7 nm (S2B)	SWIR 2
QA60	60 m		Máscara de nuvens

Fonte: Adaptado de DRUSCH *et al.* (2012).

Foi selecionada uma coleção de imagens do satélite Sentinel-2 registradas entre 1º de janeiro de 2017 e 1º de janeiro de 2018 sobre a área de estudo, com percentual de cobertura por nuvens máximo de 10%. Sobre essa coleção foi aplicado um filtro para remover as áreas próximas às bordas das imagens, utilizando faixas com 500 m de distância. Também foi aplicada uma máscara para remover eventuais nuvens e névoas presentes nas imagens, utilizando a banda QA60 do Sentinel-2. A utilização da coleção de imagens permite calcular a variabilidade temporal de variáveis preditivas, o que pode ajudar na classificação de áreas agrícolas, por exemplo.

Além dos dados de reflectância, foi calculada a diferença normalizada entre as bandas do infravermelho próximo (NIR – Near Infrared) e do vermelho (NDVI – Normalized Difference Vegetation Index) e entre as bandas do NIR e do verde (NDWI – Normalized Difference Water Index) (GAO, 1996), para cada imagem da coleção. As amplitudes anuais do NDVI e do NDWI foram calculadas pela diferença entre os valores máximo e mínimo desses índices, em cada pixel da área de estudo. Utilizamos o NDWI juntamente com o NDVI na expectativa de conseguir melhores resultados na diferenciação entre áreas úmidas com vegetação herbáceo-arbustiva (classe “Capoeira”) e outras áreas com predominância de vegetação herbácea não úmida (classe “Pastagem”).

Para representar os valores presentes em todas as imagens da coleção (jan/2017 a jan/2018), foram calculadas as medianas do NDVI, do NDWI, das bandas do visível (B2, B3 e B4), do NIR (B8), das Red Edge (B5, B6, B7 e B8A) e do SWIR (B11 e B12), de cada pixel nas imagens na coleção.

Para caracterizar a variabilidade espacial do NDVI e do NDWI, foram calculadas as entropias (texturas) das medianas dos índices. Este cálculo foi realizado para áreas quadradas, com 15 pixels de comprimento de aresta (Kernel de 15 px).

Como fonte de atributos topográficos, foram utilizados dados de declividade calculados a partir do modelo digital de elevação SRTM (Shuttle Radar Topography Mission) (FARR *et al.*, 2007) com resolução espacial de 30 m, disponível na biblioteca do GEE.

O arquivo com as variáveis preditivas foi gerado com resolução espacial de 10 m. As

informações armazenadas nos pixels corresponderam a dez variáveis representando as medianas de cada banda do Sentinel-2 (visível, NIR, Red Edge e SWIR), seis variáveis representando as medianas, as amplitudes e as entropias dos índices calculados (NDVI e NDWI) e uma variável topográfica (declividade), totalizando 17 variáveis preditivas para as classes de uso da terra. A nomenclatura das variáveis, sua origem, o tipo de cálculo efetuado e o tipo de informação que cada uma carrega é apresentado no **quadro 2**.

Quadro 2: Variáveis preditivas de classes de uso

Variáveis	Origem	Tipo de Informação	Cálculo
B2, B3 e B4	Sentinel-2	Espectro visível	Mediana
B5, B6, B7 e B8A	Sentinel-2	Espectro Red Edge	Mediana
B8	Sentinel-2	Espectro NIR	Mediana
B11 e B12	Sentinel-2	Espectro SWIR	Mediana
NDVI e NDWI	Calculado	Diferença normalizada	Mediana
NDVI_AMP e NDWI_AMP	Calculado	Variabilidade temporal	Diferença (máx-mín)
NDVI_TX e NDWI_TX	Calculado	Variabilidade espacial	Mediana/Entropia
Slope	SRTM	Declividade	-

Fonte: Elaborado pelos autores (2022).

As amostragens foram realizadas em duas etapas, sendo uma para treinamento dos modelos preditivos e outra para validação dos mapas de uso. A primeira amostragem foi feita com a vetorização de polígonos e designação das classes de uso por fotointerpretação de composições coloridas (cor verdadeira e falsa cor) das medianas dos valores de reflectância apresentados na coleção de imagens do Sentinel-2 dentro do período avaliado. As classes de uso da terra e suas respectivas descrições são apresentadas no **quadro 3**.

Quadro 3: Descrição das classes de uso e ocupação utilizadas no presente estudo

Classe de Uso	Descrição
Mata	Vegetação natural predominantemente arbórea (ex.: florestas, bosques e parques).
Pastagem	Vegetação predominantemente herbácea (ex.: pastagens e gramados).
Capoeira	Vegetação predominantemente arbustiva (ex.: várzeas, área com regeneração natural em fase inicial).
Urbano	Áreas urbanas em diferentes níveis de adensamento, com predomínio de uso não industrial (ex.: estradas, casas e prédios).
Agricultura	Áreas agrícolas diversas, com predomínio de lavouras anuais (ex.: horticultura, milho e cana).
Solo exposto	Solo desnudo não impermeabilizado (ex.: mineração e terraplanagem).
Indústria	Galpões de grande porte, industriais ou não (ex.: hipermercados, centros de distribuição).
Silvicultura	Reflorestamento para produções florestais diversas (ex.: eucalipto).
Corpos d'água	Espelho d'água com dimensões mapeáveis (ex.: rios, lagos e represas).

Fonte: Elaborado pelos autores (2022).

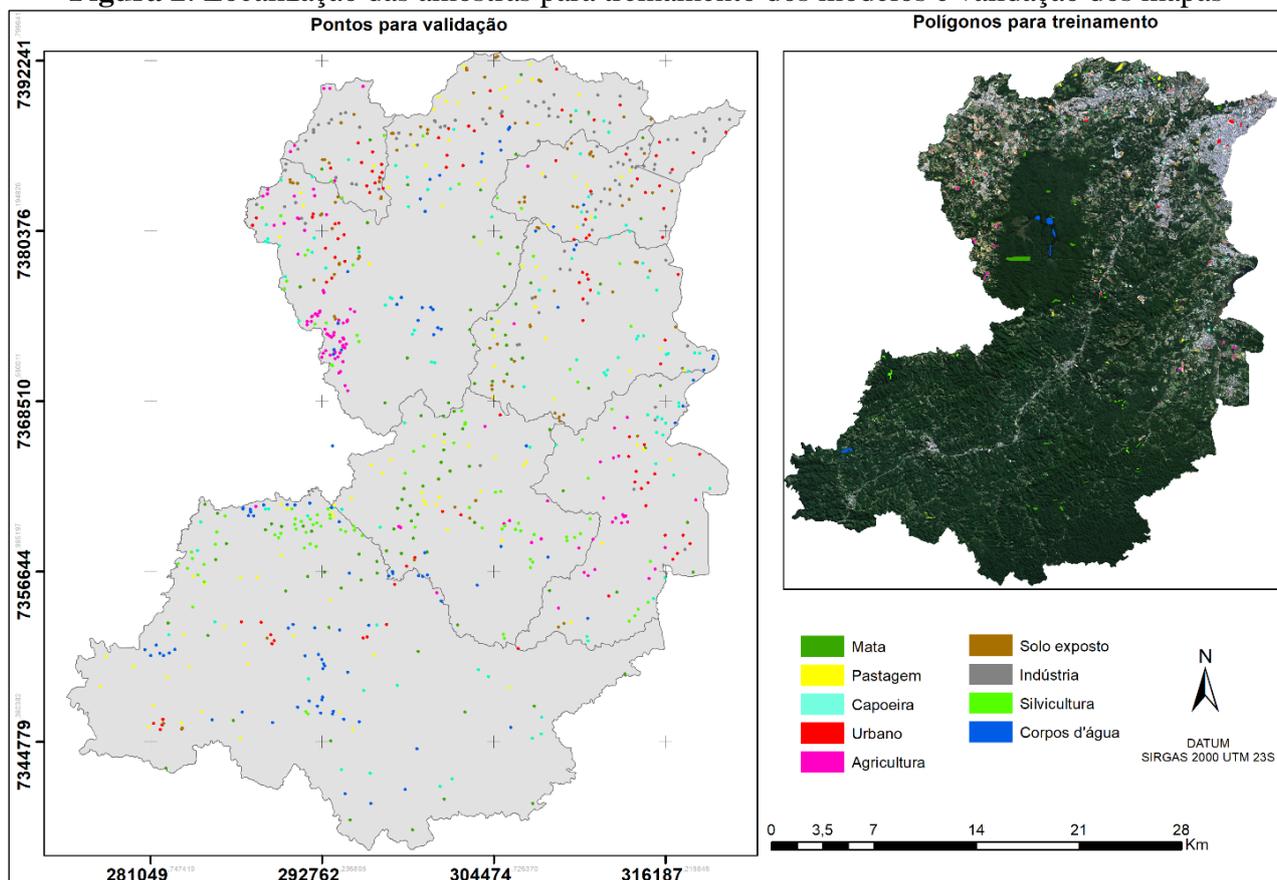
A amostragem para treinamento foi realizada até que os polígonos vetorizados abrangessem pelo menos 10.000 pixels em cada classe de uso. Esta amostragem totalizou 118.005 pixels, sendo 29.733 pertencentes à classe “Mata”, 10.126 à classe “Pastagem”, 13.558 à classe “Capoeira”, 10.612 à classe “Urbano”, 10.882 à classe “Agricultura”, 10.571 à classe “Solo exposto”, 10.093 à classe “Indústria”, 10.065 à classe “Silvicultura” e 12.365 à classe “Corpos d'água”.

A amostragem para validação dos mapas foi feita com a vetorização de 100 pontos em cada classe de uso, distribuídos aleatoriamente por toda a área de estudo, totalizando 900 pontos de validação, como mostra a **figura 2**. A designação da classe de uso foi feita por fotointerpretação de imagens de alta resolução, disponíveis na área de visualização do GEE. Foi utilizado o critério da não

sobreposição dos pontos de validação sobre os polígonos amostrados para o treinamento dos modelos preditivos.

Foram avaliadas estratégias para obter as observações para treinamento em classes balanceadas e desbalanceadas. No treinamento em classes desbalanceadas, foi utilizado o total de observações inicialmente amostradas nos polígonos de cada classe de uso (sem controle da distribuição das amostras). E também uma amostragem com controle da distribuição (frequência) de observações entre as classes. No treinamento em classes desbalanceadas com controle da distribuição das amostras entre as classes, a quantidade de observações em cada classe foi definida avaliando-se visualmente os resultados obtidos na classificação supervisionada, aumentando a frequência das classes subestimadas e diminuindo a frequência das classes superestimadas (10.000 observações nas classes “Mata”, “Pastagem”, “Urbano” e “Solo exposto”, 5.000 observações na classe “Corpos d’água” e “Indústria”, 1.000 observações nas classes “Capoeira”, “Agricultura” e “Silvicultura”). No treinamento em classes balanceadas, foram avaliadas a utilização de 100, 1.000, 5.000 e 10.000 observações em cada classe de uso da terra. Estas estratégias foram aplicadas tanto na abordagem por pixels quanto na abordagem por segmentação. Os segmentos amostrados foram aqueles que se apresentaram sobrepostos aos pixels amostrados inicialmente.

Figura 2: Localização das amostras para treinamento dos modelos e validação dos mapas



Fonte: Elaborado pelos autores (2022).

Na abordagem OBIA, a segmentação da imagem foi realizada utilizando o algoritmo SNIC (Simple Non-Iterative Clustering) (ACHANTA; SUSSTRUNK, 2017), disponível no GEE, sendo considerado um kernel de 9 pixels e o tamanho da vizinhança igual a 18 pixels.

O efeito da adição do componente topográfico (declividade), dos índices NDVI e NDWI, das suas respectivas amplitudes anuais e variabilidades espaciais (textura) foi avaliado comparando-se o desempenho dos mapas de uso da terra com e sem a adição dessas informações nos dados de

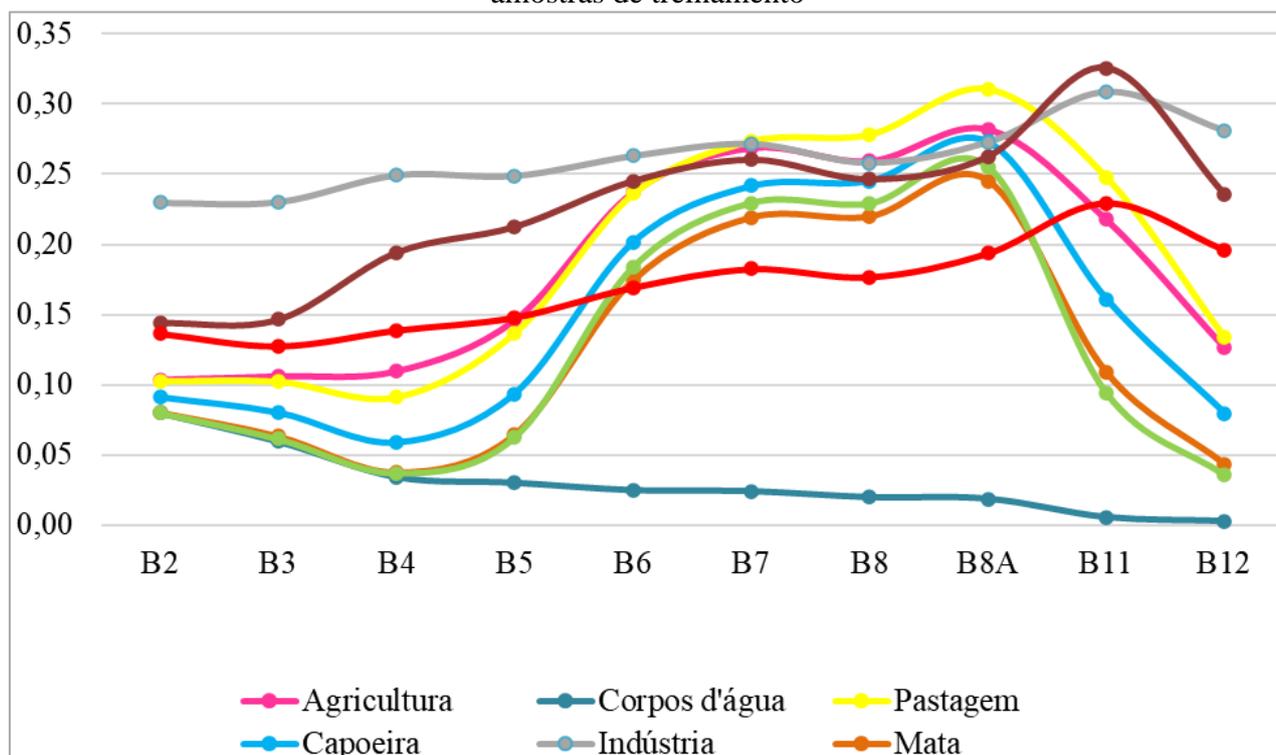
treinamento dos modelos. As métricas utilizadas para medir o desempenho dos mapas foram a acurácia global (WEISS; ZHANG, 2003), o índice Kappa (COHEN, 1960) e as exatidões do usuário e do produtor (BANKO, 1998).

Em todos os casos, o algoritmo de classificação utilizado foi o Random Forest (BREIMAN, 2001), por ser um classificador amplamente aplicado na classificação do uso da terra, que tem superado outros algoritmos de classificação (RODRIGUEZ-GALIANO *et al.*, 2012) e está implementado no GEE.

3. RESULTADOS E DISCUSSÕES

As classes de uso da terra apresentaram diferentes comportamentos espectrais. A assinatura espectral dos alvos pode ser observada no gráfico da média dos valores de reflectância ao longo do espectro eletromagnético do Sentinel-2, conforme mostra a **figura 3**. A classe “Indústria” apresentou alta reflectância e baixa amplitude ao longo do espectro eletromagnético. As áreas urbanas e com solo exposto apresentaram tendência de aumento na reflectância média com o aumento do comprimento de onda. Contudo, a taxa de incremento médio é maior na curva referente às áreas com solo exposto.

Figura 3: Reflectância média nas bandas do Sentinel-2 em cada classe de uso e ocupação nas amostras de treinamento



Fonte: Elaborado pelos autores (2022).

As classes “Mata”, “Silvicultura” e “Capoeira” apresentaram comportamento bastante similar, principalmente entre as duas primeiras classes. Há um declínio na curva da reflectância média com o aumento do comprimento de onda na região do visível, seguido de aumento brusco a partir da banda B5 (Red Edge 1), atingindo o pico de reflectância na banda B8A (Red Edge 4) e caindo drasticamente na região do infravermelho de ondas curtas, conforme mostra a **figura 3**. Este comportamento é típico de áreas cobertas por vegetação. Isso também foi observado nas classes “Pastagem” e “Agricultura”, porém, estas apresentaram maior reflectância média que as demais classes representantes de áreas vegetadas. As curvas das classes “Pastagem” e “Agricultura”

apresentam ligeira diferenciação nas regiões do vermelho, infravermelho próximo e Red Edge 4, B11 (SWIR 1) e B12 (SWIR 2), como mostra a **figura 3**.

As áreas ocupadas por corpos d'água apresentaram os menores valores de reflectância média e uma tendência à diminuição nos valores de reflectância conforme o comprimento de onda aumenta. Nas regiões do infravermelho de ondas curtas, os valores médios de reflectância tendem a zero.

Os dados apresentados na **figura 3** mostram que, apenas com as informações extraídas diretamente dos sensores, pode ser uma tarefa difícil separar de forma satisfatória classes como "Mata" e "Silvicultura" ou "Pastagem" e "Agricultura", por exemplo.

Na abordagem com base em pixels, utilizando apenas os dados de reflectância de imagens do Sentinel-2, o mapa de uso da terra predito pelo Random Forest apresentou acurácia global de 79,33% e índice Kappa de 76,75%. Quando acrescentadas as informações dos índices NDVI e NDWI, houve incremento de 1,23% na acurácia e 1,38% no índice Kappa, como mostra a **tabela 1**.

Quando adicionadas também as amplitudes anuais dos valores do NDVI e NDWI, não houve alteração no desempenho do mapa de uso da terra com base em pixels. Quando acrescentadas as informações sobre a variabilidade espacial (textura) dos índices NDVI e NDWI nos dados de treinamento, a acurácia global do mapa de uso da terra foi de 81,00% e o índice Kappa de 78,63%, representando pequeno aumento em relação ao mapa produzido sem utilizar estas variáveis. O melhor desempenho encontrado neste estudo ocorreu quando adicionada a declividade como variável preditiva na abordagem baseada em pixels, sendo a acurácia global de 82,44% e o índice Kappa de 80,25%, como apresenta a **tabela 1**.

Tabela 1: Desempenho do mapa de uso e ocupação utilizando as informações extraídas dos sensores do Sentinel-2 e o efeito da adição dos índices NDVI e NDWI, suas amplitudes anuais, suas variabilidades espaciais (texturas) e a variável topográfica (declividade)

Variáveis preditivas	PIXELS		OBIA	
	Acurácia	Kappa	Acurácia	Kappa
			%	
Sensor	79,33	76,75	75,58	72,53
Sensor, NDVI e NDWI	80,56	78,13	75,25	72,16
Sensor, NDVI e NDWI, amplitudes	80,56	78,13	76,81	73,91
Sensor, NDVI e NDWI, amplitudes, texturas	81,00	78,63	-	-
Sensor, NDVI e NDWI, amplitudes, texturas, declividade	82,44	80,25	-	-
Sensor, NDVI e NDWI, amplitudes, geometrias	-	-	77,70	74,92
Sensor, NDVI e NDWI, amplitudes, geometrias e declividade	-	-	77,26	74,41
Sensor + amplitudes + geometrias	-	-	78,15	75,42

Fonte: Elaborado pelos autores (2022).

Com a abordagem OBIA, a acurácia global do mapa de uso da terra foi de 75,58% e o índice Kappa de 72,53%, utilizando apenas as informações extraídas dos sensores do Sentinel-2. Quando adicionados os índices NDVI e NDWI, houve uma pequena queda nos valores de acurácia (-0,33%) e no índice Kappa (-0,37%). Adicionados os valores das amplitudes anuais do NDVI e NDWI, a acurácia do mapa de uso da terra foi de 76,81% e o índice Kappa de 73,91%. Quando adicionadas as informações sobre a geometria dos segmentos (área, perímetro, largura e altura) a acurácia do mapa de uso da terra foi de 77,70% e índice Kappa de 74,92%. Quando adicionada a variável topográfica houve uma pequena diminuição no desempenho do classificador, apresentando 77,26% de acurácia e 74,41% de índice Kappa. O melhor resultado para a abordagem OBIA foi obtido retirando-se dos dados de treinamento as variáveis que apresentaram queda no desempenho do mapa de uso da terra. Ou seja, utilizando-se apenas as informações dos sensores do Sentinel-2, as amplitudes anuais dos índices (NDVI e NDWI) e as informações geométricas dos segmentos, alcançando 78,15% de acurácia global e 75,42% de índice Kappa, conforme a **tabela 1**.

Khatami, Mountrakis e Stehman (2016), em uma meta-análise estatística que cobriu 15 anos de pesquisa, observaram que a inclusão da textura aumentou em média 12,1% os valores de acurácia na classificação do uso e ocupação. Contudo, no presente estudo houve apenas um pequeno aumento na acurácia com a inclusão da textura. Melhores vantagens foram alcançadas incluindo os índices (NDVI e NDWI) e a declividade, na abordagem por pixels. Já na abordagem OBIA, as variáveis responsáveis por melhorar o desempenho do classificador nos testes realizados foram a variação anual dos índices (NDVI e NDWI) e as características geométricas dos segmentos.

Utilizando todos os dados amostrados para treinamento, foram obtidos acurácia global de 80,89% e índice Kappa de 78,50% no mapa de uso da terra para a abordagem por pixels e 77,22% de acurácia com 74,38 de índice Kappa na abordagem OBIA. Quando se utilizou o controle subjetivo da quantidade de observações em cada classe para treinamento do classificador, houve uma pequena queda nos valores de acurácia (-2,56%) e no índice Kappa (-2,88%) do mapa de uso da terra para a abordagem por pixels, e um ligeiro aumento na acurácia (0,45%) e no índice Kappa na abordagem OBIA (0,49%), em relação ao treinamento com todos os dados amostrados, como mostra a **tabela 2**.

Tabela 2: Número de observações nas diferentes estratégias de amostragem para treinamento do classificador, acurácias e índices Kappa obtidos nos mapas de uso da terra classificados pelo Random Forest

PIXELS						
	Classes desbalanceadas		Classes balanceadas			
	Sem controle da distribuição	Controle subjetivo da distribuição	100	1.000	5.000	10.000
Qt.	118.005	48.129	900	9.118	45.143	90.006
	----- % -----					
Acurácia	80,89	78,33	75,89	82,44	80,67	81,11
Kappa	78,50	75,63	72,88	80,25	78,25	78,75
OBIA						
	Classes desbalanceadas		Classes balanceadas			
	Sem controle da distribuição	Controle subjetivo da distribuição	100	1.000	5.000	10.000
Qt.	118.005	48.000	900	9.000	45.000	90.000
	----- % -----					
Acurácia	77,03	77,59	73,02	76,59	78,15	77,81
Kappa	74,16	74,79	69,65	73,66	75,42	75,04

Fonte: Elaborado pelos autores (2022).

Estes resultados sugerem que podem ser alcançados resultados com qualidade muito próxima, se a distribuição das amostras entre as classes for definida arbitrariamente por um especialista com conhecimento da área a ser mapeada, utilizando uma quantidade menor de amostras para treinamento. Contudo, os resultados do treinamento em classes balanceadas, com quantidade suficiente de amostras por classe, superaram as demais abordagens.

Além disso, em relação à quantidade de observações para treinamento, os resultados indicam que não é necessária uma amostragem exaustiva. A partir de 1.000 observações por classe, já encontramos ajustes com os melhores resultados observados entre os testes.

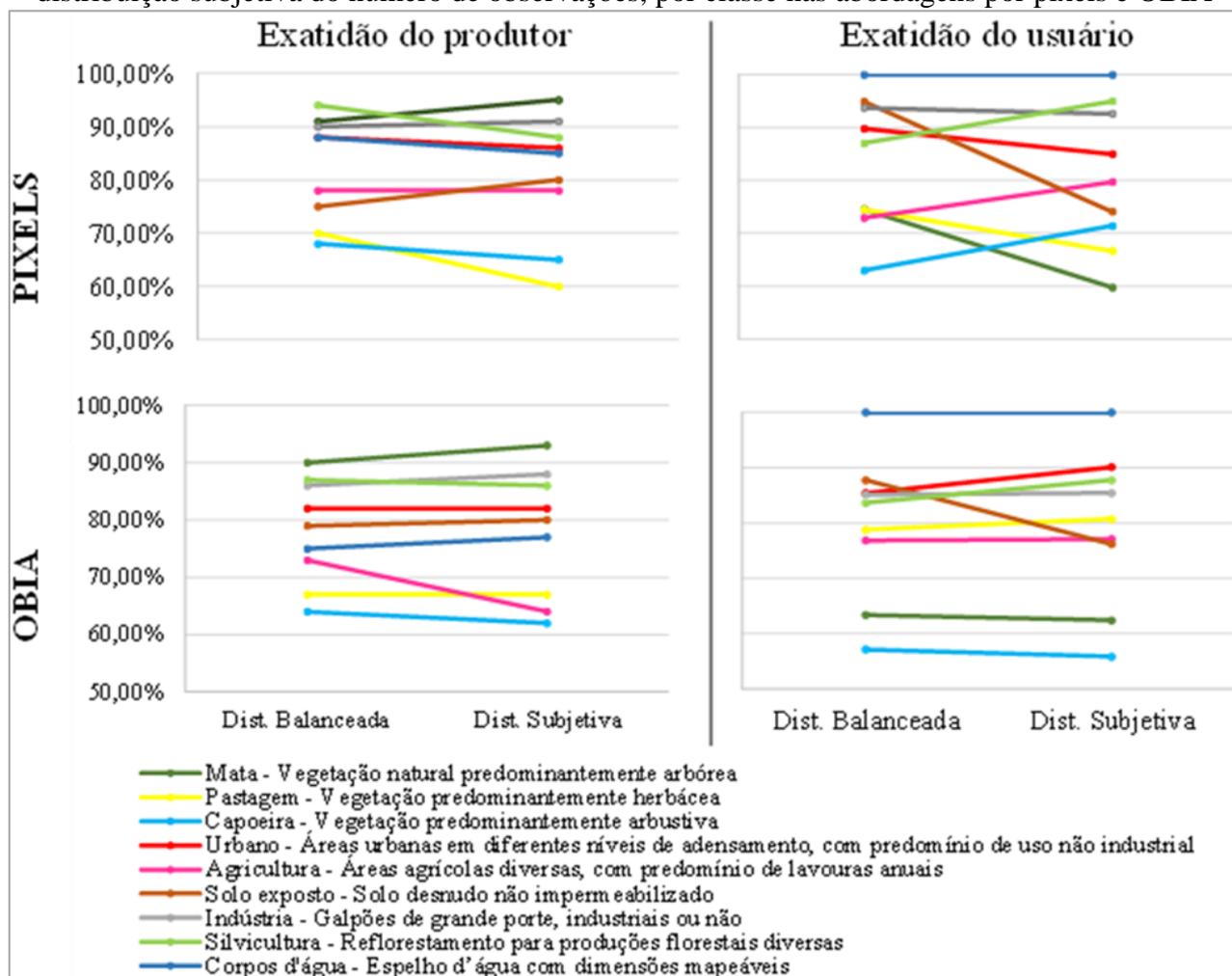
Avaliando-se o efeito da quantidade de observações no treinamento em classes balanceadas, a maior diferença nos valores de acurácia e no índice Kappa ocorreu quando se aumentou de 100 para 1.000 observações em cada classe, tanto na abordagem por pixels quanto por OBIA.

Li *et al.* (2014), avaliando diferentes algoritmos de classificação para mapeamento da cobertura de áreas urbanas, observaram que, a partir de 200 observações por classe, os classificadores começavam a alcançar o seu melhor desempenho. No presente estudo, entre 1.000 e 10.000

observações de treinamento por classe, a variação nos valores de acurácia e índice Kappa foram pequenos, sendo alcançado o melhor desempenho com 1.000 observações por classe na abordagem por pixels (acurácia de 82,44% e Kappa de 80,25%) e com 5.000 observações por classe na abordagem OBIA (acurácia de 78,15% e Kappa de 75,42%), como mostra a **tabela 2**.

Na abordagem por pixels, quando utilizada distribuição subjetiva da quantidade de observações por classe para treinamento, a chance de uma observação de campo ser corretamente predita no mapa (exatidão do produtor) diminuiu em 10% na classe “Pastagem”, 6% na classe “Silvicultura”, 3% nas classes “Capoeira” e “Corpos d’água”, e 2% na classe “Urbano”, no treinamento com 1.000 observações por classe. No mesmo comparativo, a exatidão do produtor aumenta em 4% na classe “Mata”, 5% na classe “Solo exposto” e 1% na classe “Indústria”. A exatidão do produtor permanece igual para a classe “Agricultura” neste comparativo, como mostra a **figura 4**.

Figura 4: Exatidões do usuário e do produtor no treinamento em classes balanceadas e com distribuição subjetiva do número de observações, por classe nas abordagens por pixels e OBIA



Fonte: Elaborado pelos autores (2022).

Na abordagem por pixels, quando utilizada distribuição subjetiva da quantidade de observações por classe para treinamento, a chance de a classe predita no mapa corresponder corretamente a uma observação de campo (exatidão do usuário) diminuiu em 21% na classe “Solo exposto”, 15% na classe “Mata”, 8% na classe “Pastagem”, 4% na classe “Urbano” e 1% na classe “Indústria”. Foi observado incremento na exatidão do usuário nas classes “Capoeira” (8%), “Silvicultura” (8%) e “Agricultura” (7%). A exatidão do usuário para a classe “Corpos d’água” foi a mesma em todos os testes realizados neste estudo, com valor igual a 100%, conforme a **figura 4**.

Na abordagem OBIA, quando utilizada a distribuição subjetiva da quantidade de observações por classe para treinamento, a exatidão do produtor diminui em 9% na classe “Agricultura”, 2% na classe “Capoeira” e 1% na classe “Silvicultura”, em relação ao treinamento com 5.000 observações por classe. No mesmo comparativo, a exatidão do produtor aumenta em 3% na classe “Mata”, 2% nas classes “Indústria” e “Cursos d’água” e 1% na classe “Solo exposto”. A exatidão do produtor permanece a mesma nas classes “Pastagem” e “Urbano”. A exatidão do usuário diminui em 12% na classe “Solo exposto” e 1% nas classes “Mata” e “Capoeira”. O incremento na exatidão do usuário é observado nas classes “Urbano” (4%), “Silvicultura” (4%) e “Pastagem” (2%). A exatidão do usuário permanece a mesma nas classes “Agricultura”, “Indústria” e “Cursos d’água”, como mostra a **figura 4**.

Comparando visualmente os mapas de uso da terra obtidos com controle subjetivo da distribuição de observações por classe e aqueles obtidos utilizando 1.000 observações por classe na abordagem por pixels e 5.000 observações por classe na abordagem OBIA, o resultado é aparentemente mais satisfatório nos mapas obtidos com controle subjetivo da distribuição de observações. É perceptível que o classificador superestima algumas classes (Agricultura, Capoeira e Silvicultura), enquanto subestima outras classes (Mata e Urbano), quando o treinamento é feito em classes balanceadas, como mostra a **figura 5**.

Apesar de os resultados apresentados na **tabela 1** e na **tabela 2** indicarem objetivamente que a qualidade dos ajustes obtidos pela classificação com base em pixels foi superior à classificação com base em objetos, a **figura 5** mostra que, visualmente, os resultados apresentados pela classificação com base em pixels ainda apresentam efeito “sal e pimenta”, mesmo com a utilização das texturas e da declividade no modelo preditivo para atenuar este efeito.

Uma avaliação subjetiva de um especialista com conhecimento da área, frente a resultados objetivos tão semelhantes encontrados utilizando as duas abordagens de classificação, é importante, pois, a validação está limitada aos pontos de observação sorteados na etapa de testes. Observa-se neste estudo que o método por pixels foi mais preciso, porém, os limites e os contatos entre as classes ficaram mais bem definidos quando utilizada a classificação por objetos, conforme mostra a **figura 5**. O mapa de uso e ocupação produzido na classificação com base em pixels para toda a área de estudo é apresentado no apêndice A.

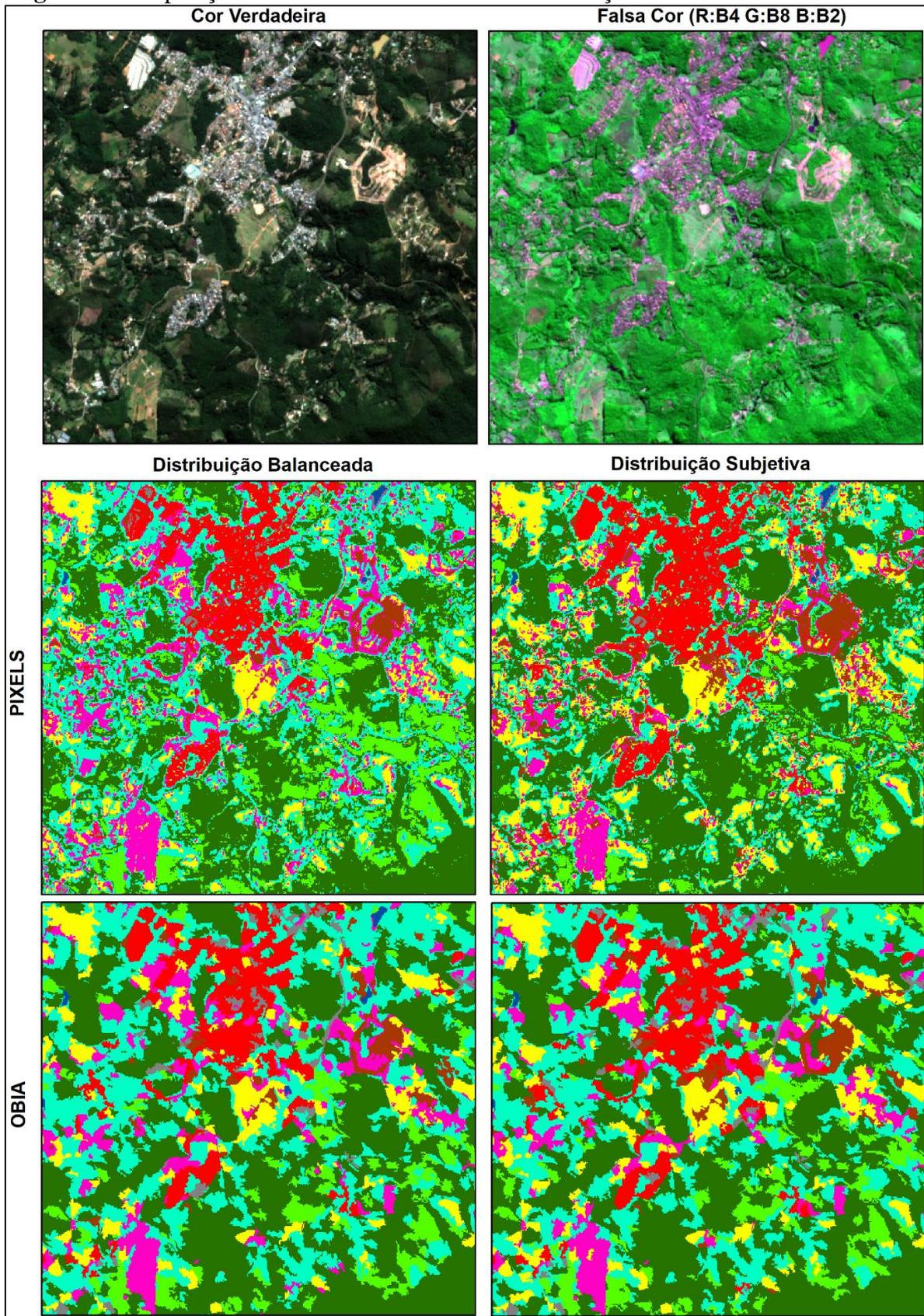
A área predita para a classe “Mata” foi de 73.405 ha no treinamento com 1.000 observações por classe e 89.016 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 85.250 ha no treinamento em classes balanceadas e de 86.991 ha no treinamento com distribuição subjetiva do número de observações por classe, como observado na **figura 6**.

Na classe “Pastagem”, a área predita foi de 5.918 ha no treinamento com 1.000 observações por classe e 10.645 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 5.512 ha no treinamento em classes balanceadas e de 6.054 ha no treinamento com distribuição subjetiva do número de observações por classe, conforme a **figura 6**.

A área predita para a classe “Capoeira” foi de 30.116 ha no treinamento com 1.000 observações por classe e 21.838 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 30.806 ha no treinamento em classes balanceadas e de 28.353 ha no treinamento com distribuição subjetiva do número de observações por classe, como mostra a **figura 6**.

Na classe “Urbano”, a área predita foi de 9.124 ha no treinamento com 1.000 observações por classe e 10.745 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 7.475 ha no treinamento em classes balanceadas e de 7.879 ha no treinamento com distribuição subjetiva do número de observações por classe, como observado na **figura 6**.

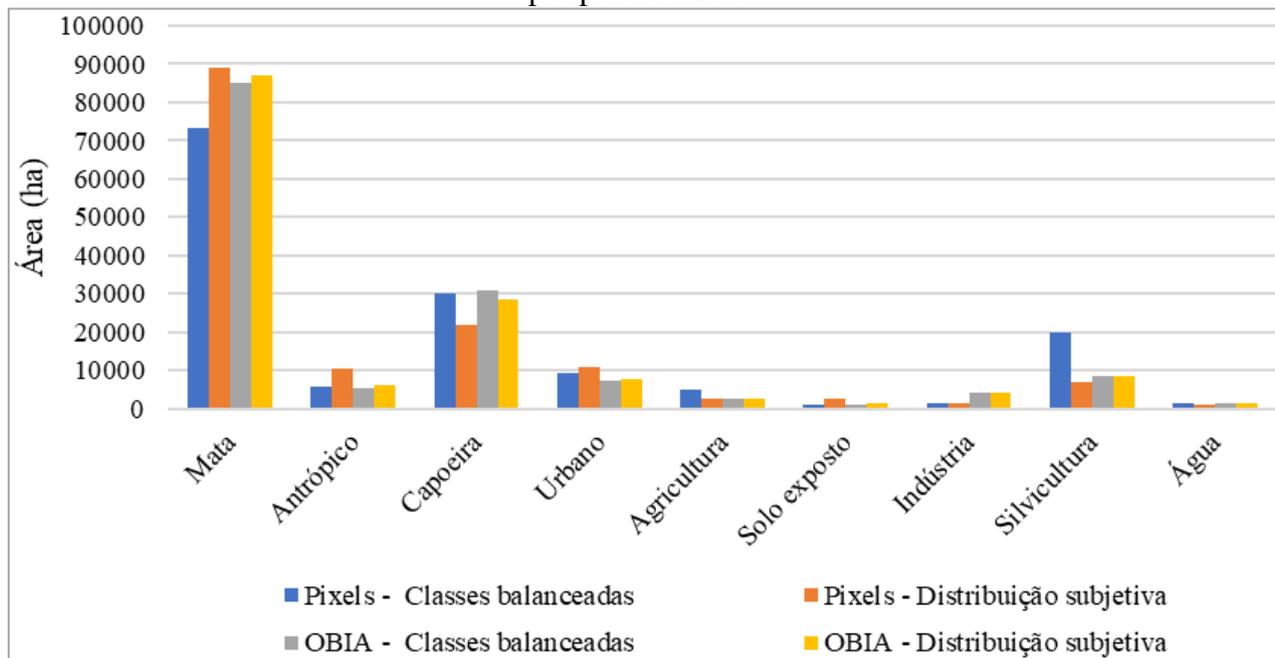
Figura 5: Composições coloridas e resultados da classificação em trecho da área de estudo



Legenda: (Verde escuro: Mata; Amarelo: Pastagem; Ciano: Capoeira; Vermelho: Urbano; Rosa: Agricultura; Marrom: Solo exposto; Cinza: Indústria; Verde claro: Silvicultura; Azul: Cursos d'água).

Fonte: Elaborado pelos autores (2022).

Figura 6: Área predita pelo Classificador Random Forest em cada classe de uso da terra no treinamento em classes balanceçadas e com distribuição subjetiva das observações, nas abordagens por pixels e OBIA



Fonte: Elaborado pelos autores (2022).

A área predita para a classe “Agricultura” foi de 4.872 ha no treinamento com 1.000 observações por classe e 2.626 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 2.834 ha no treinamento em classes balanceçadas e de 2.513 ha no treinamento com distribuição subjetiva do número de observações por classe, como exposto na **figura 6**. Na classe “Solo exposto”, a área predita foi de 1.101 ha no treinamento com 1.000 observações por classe e 2.775 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 1.182 ha no treinamento em classes balanceçadas e de 1.649 ha no treinamento com distribuição subjetiva do número de observações por classe, conforme a **figura 6**.

A área predita para a classe “Indústria” foi de 1.382 ha no treinamento com 1.000 observações por classe e 1.287 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 4.379 ha no treinamento em classes balanceçadas e de 4.231 ha no treinamento com distribuição subjetiva do número de observações por classe, como exposto na **figura 6**.

Na classe “Silvicultura”, a área predita foi de 20.002 ha no treinamento com 1.000 observações por classe e 7.028 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 8.599 ha no treinamento em classes balanceçadas e de 8.340 ha no treinamento com distribuição subjetiva do número de observações por classe, como mostra a **figura 6**.

A área predita para a classe “Cursos d’água” foi de 1.298 ha no treinamento com 1.000 observações por classe e 1.258 ha no treinamento com distribuição subjetiva da quantidade de observações por classe, na abordagem por pixels. Na abordagem OBIA, a área predita foi de 1.331 ha no treinamento em classes balanceçadas e de 1.358 ha no treinamento com distribuição subjetiva do número de observações por classe, conforme a **figura 6**.

A aparente melhora no mapa predito, quando se controla subjetivamente a quantidade de observações em cada classe para o treinamento do classificador, pode ser explicada pelo incremento na exatidão do usuário para as classes “Capoeira”, “Silvicultura” e “Agricultura”.

Trabalhando com treinamento em classes balanceadas, apesar de alcançar melhor acurácia e índice Kappa, é perceptível uma grande confusão na predição entre as classes “Mata” e “Silvicultura”, principalmente. Também é possível detectar falhas na predição das classes “Agricultura”, “Capoeira” e “Pastagem”.

O efeito da distribuição subjetiva da quantidade de observações em cada classe para treinamento é pequeno quando utilizada a abordagem OBIA. Isto pode ser observado nos resultados de acurácia e índice Kappa, na avaliação visual dos mapas preditos e analisando-se a pequena diferença entre o tamanho da área predita em cada classe, como apresentado na **figura 6**.

Na abordagem por pixels, o efeito da distribuição subjetiva da quantidade de observações em cada classe para treinamento é relevante, principalmente em termos de área predita. O resultado com o treinamento em classes balanceadas subestima a classe “Mata”, ao mesmo tempo em que superestima as classes “Capoeira”, “Agricultura” e “Silvicultura”. Estas são justamente as classes nas quais a exatidão do usuário aumenta cerca de 8% com a distribuição subjetiva da quantidade de observações por classe.

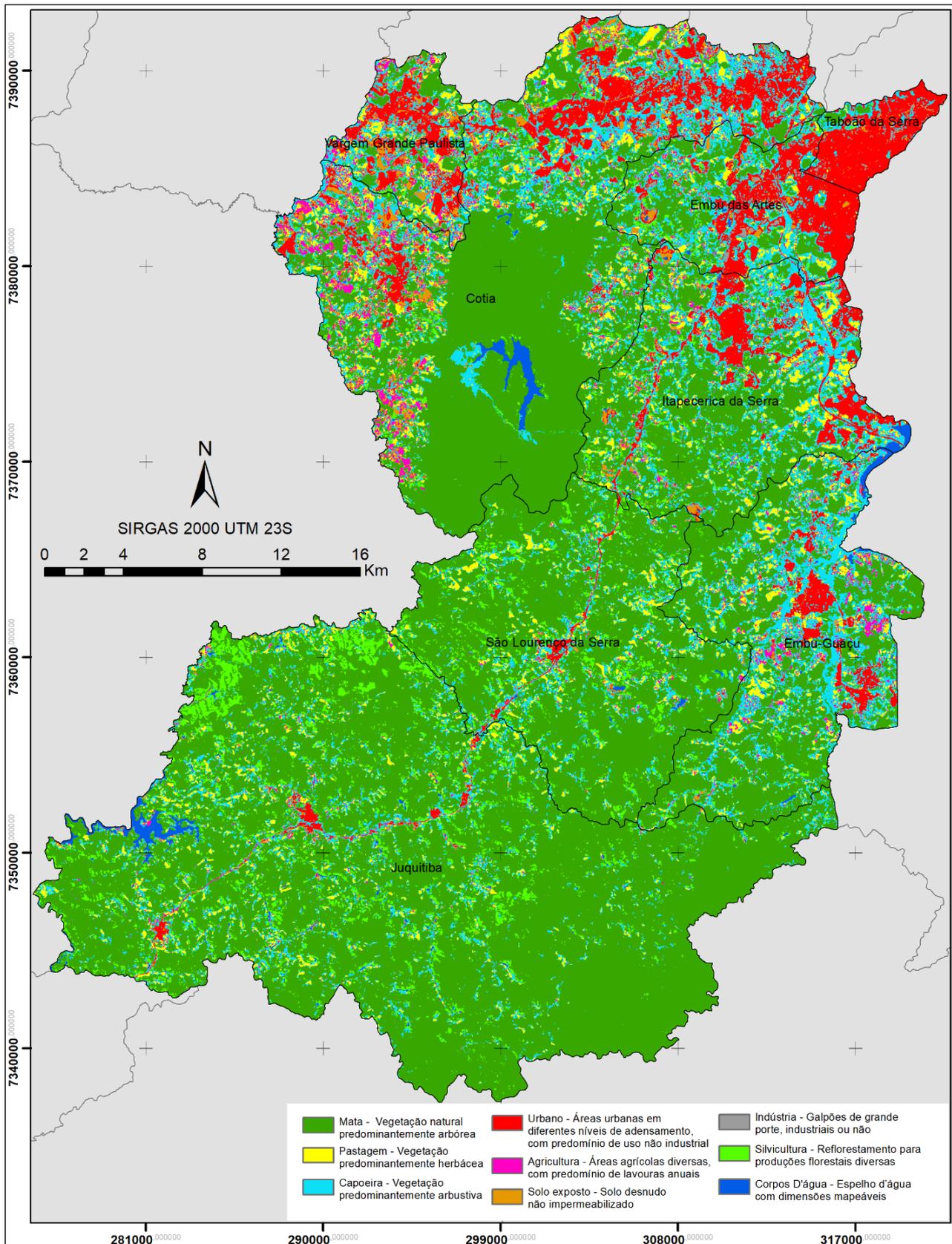
4. CONCLUSÃO

A utilização de informações sobre a amplitude anual dos índices NDVI e NDWI foi importante para melhorar o desempenho dos mapas de uso da terra, tanto na abordagem por pixels quanto na OBIA. Na abordagem por pixels, a adição da variabilidade espacial e da variável topográfica também resultou na melhora do mapa de uso e ocupação.

A utilização de aproximadamente 50.000 observações foi suficiente para gerar mapas de uso da terra satisfatórios no treinamento em classes desbalanceadas. Os melhores resultados, avaliando a acurácia global e o índice Kappa, foram obtidos no treinamento em classes balanceadas utilizando 1.000 e 5.000 observações nas abordagens por pixels e OBIA, respectivamente. Isso indica que a última exige maior número de observações para treinamento que a primeira. A necessidade de maior número de observações para treinamento na abordagem OBIA pode ser devido à diminuição da variabilidade das variáveis preditivas quando os dados são reamostrados para cada objeto.

A distribuição subjetiva do número de amostras para treinamento diminuiu os erros de omissão da classe “Mata” e os erros de comissão das classes “Capoeira” e “Agricultura” na abordagem por pixels. Isso resultou em um aspecto mais coerente do mapa de uso da terra, comparado ao resultado com melhor ajuste global. É, portanto, uma estratégia a ser considerada, sobretudo quando se conhece a distribuição geral das classes de uso na área de estudo.

APÊNDICE A: Mapa de uso e ocupação da sub-região sudoeste da Região Metropolitana de São Paulo



REFERÊNCIAS

ACHANTA, Radhakrishna; SUSSTRUNK, Sabine. Superpixels and Polygons Using Simple Non-iterative Clustering. **2017 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr)**, [s.l.], p. 4651-4660, jul. 2017. IEEE. <http://dx.doi.org/10.1109/cvpr.2017.520>

BANKO, Gebhard. **A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory**. International Institute for Applied Systems Analysis: Laxenburg, Austria, 42 p. 1998. Available at: <http://pure.iiasa.ac.at/5570/>.

BELGIU, Mariana; CSILLIK, Ovidiu. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. **Remote Sensing of Environment**, [s.l.], v. 204, p. 509-523, jan. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2017.10.005>.

BLASCHKE, Thomas *et al.*; Geographic Object-Based Image Analysis – Towards a new paradigm. **Isprs Journal of Photogrammetry and Remote Sensing**, [s.l.], v. 87, p. 180-191, jan. 2014. Elsevier BV. <http://dx.doi.org/10.1016/j.isprsjprs.2013.09.014>.

BREIMAN, Leo. Random Forests. **Machine Learning**, [s.l.], v. 45, n. 1, p. 5-32, 2001. Springer Science and Business Media LLC. <http://dx.doi.org/10.1023/a:1010933404324>.

COHEN, Jacob. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, [S.L.], v. 20, n. 1, p. 37-46, abr. 1960. SAGE Publications. <http://dx.doi.org/10.1177/001316446002000104>.

DRUSCH, Matthias *et al.*; Sentinel-2: esa's optical high-resolution mission for gmes operational services. **Remote Sensing of Environment**, [s.l.], v. 120, p. 25-36, maio 2012. Elsevier BV. <https://doi.org/10.1016/j.rse.2011.11.026>.

FARR, Tom G. *et al.* . The Shuttle Radar Topography Mission. **Reviews of Geophysics**, [s.l.], v. 45, n. 2, p. 1-33, 19 maio 2007. American Geophysical Union (AGU). <http://dx.doi.org/10.1029/2005rg000183>.

GAO, Bo-Cai. NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, [S.L.], v. 58, n. 3, p. 257-266, dez. 1996. Elsevier BV. [http://dx.doi.org/10.1016/s0034-4257\(96\)00067-3](http://dx.doi.org/10.1016/s0034-4257(96)00067-3).

GÓMEZ, Cristina; WHITE, Joanne C.; WULDER, Michael A. Optical remotely sensed time series data for land cover classification: a review. **Isprs Journal of Photogrammetry and Remote Sensing**, [s.l.], v. 116, p. 55-72, jun. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.isprsjprs.2016.03.008>.

GORELICK, Noel *et al.* Google Earth Engine: planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, [s.l.], v. 202, p. 18-27, dez. 2017. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2017.06.031>.

HAY, Geoffrey J.; CASTILLA, Guilherme. Geographic Object-Based Image Analysis (GEOBIA): a new name for a new discipline. **Lecture Notes in Geoinformation and Cartography**, [s.l.], p. 75-89, 2008. Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-77058-9_4.

KHATAMI, Reza; MOUNTRAKIS, Giorgos; STEHMAN, Stephen V. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: general guidelines for practitioners and future research. **Remote Sensing of Environment**, [s.l.], v. 177, p. 89-100, maio 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2016.02.028>.

LI, Congcong *et al.* Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. **Remote Sensing**, [s.l.], v. 6, n. 2, p. 964-983, 24 jan. 2014. MDPI AG. <http://dx.doi.org/10.3390/rs6020964>.

LIU, Desheng; XIA, Fan. Assessing object-based classification: advantages and limitations. **Remote Sensing Letters**, [s.l.], v. 1, n. 4, p. 187-194, dez. 2010. Informa UK Limited. <http://dx.doi.org/10.1080/01431161003743173>.

MA, Lei *et al.* A review of supervised object-based land-cover image classification. **Isprs Journal of Photogrammetry and Remote Sensing**, [s.l.], v. 130, p. 277-293, ago. 2017. Elsevier BV. <http://dx.doi.org/10.1016/j.isprsjprs.2017.06.001>.

MORE. **Mecanismo online para referências, versão 2.0**. Florianópolis: UFSC: Rexlab, 2013. Disponível em: <http://www.more.ufsc.br/>. Acesso em: 19 de jun. de 2023.

RODRIGUEZ-GALIANO, Victor F. *et al.* An assessment of the effectiveness of a random forest classifier for land-cover classification. **Isprs Journal of Photogrammetry and Remote Sensing**, [s.l.], v. 67, p. 93-104, jan. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002>.

ROSA, Marcos Reis. Classificação do Padrão de Ocupação Urbana de São Paulo Utilizando Aprendizagem de Máquina e Sentinel 2. **Geography Department University of Sao Paulo**, [s.l.], p. 15-21, 9 set. 2018. Universidade de São Paulo, Agência USP de Gestão da Informação Acadêmica (AGUIA). <http://dx.doi.org/10.11606/rdg.v0ispe.145784>.

SHELESTOV, Andrii *et al.*; Exploring Google Earth Engine Platform for Big Data Processing: classification of multi-temporal satellite imagery for crop mapping. **Frontiers in Earth Science**, [s.l.], v. 5, p. 1-10, 24 fev. 2017. Frontiers Media SA. <http://dx.doi.org/10.3389/feart.2017.00017>.

WEISS, Sholom M.; ZHANG, Tong. Performance analysis and evaluation. In: The handbook of Data Mining. **Lawrence Erlbaum Associates Publishers**, Mahwah, NJ, vol. 14, p. 425 – 440, 2003. <https://doi.org/10.1201/b12469>.



Informações sobre a Licença

Este é um artigo de acesso aberto distribuído nos termos da Licença de Atribuição Creative Commons, que permite o uso irrestrito, distribuição e reprodução em qualquer meio, desde que o trabalho original seja devidamente citado.

License Information

This is an open access article distributed under the terms of the Creative Commons Attribution License, which allows for unrestricted use, distribution and reproduction in any medium, as long as the original work is properly cited.