



Mapeamento da Vegetação Nativa do Cerrado na Região de Três Lagoas-MS com o Google Earth Engine

Cerrado Native Vegetation Mapping in a Region of Três Lagoas – MS with Google Earth Engine

*Nayara Vasconcelos Estrabis*¹

*José Marcato Junior*¹

Hemerson Pistori^{2 3}

Recebido em março de 2019.

Aprovado em julho de 2019.

RESUMO

O Cerrado é um dos biomas existentes no Brasil e o segundo mais extenso da América do Sul. Possui grande importância devido a sua biodiversidade, ecossistema e principalmente por servir como um reservatório, ou “esponja”, que distribui água para os demais biomas, além de ser berço de nascentes de algumas das maiores bacias da América do Sul. No entanto, devido às atividades antrópicas praticadas (com destaque para a pecuária e silvicultura) e a redução da vegetação nativa, este bioma está ameaçado. Considerado como *hotspot* em biodiversidade, o Cerrado pode não existir em 2050. Com a necessidade de sua preservação, o objetivo desse trabalho consistiu em investigar o uso de algoritmos de aprendizado de máquina para realizar o mapeamento da vegetação nativa existente na região do município de Três Lagoas, utilizando a plataforma em nuvem Google Earth Engine. O processo foi realizado com uma imagem Landsat-8 OLI, datada de 10 de outubro de 2018, e com os algoritmos *Random Forest* (RF) e *Support Vector Machine* (SVM). Na validação da classificação, o RF e o SVM apresentaram índices kappa iguais a 0,94 e 0,97, respectivamente. O RF, quando comparado ao SVM, apresentou classificação mais ruidosa. Por fim, verificou-se a existência de vegetação nativa de aproximadamente 2556 km² ao adotar o RF e 2873 km² ao adotar SVM.

PALAVRAS-CHAVE: Sensoriamento Remoto. Aprendizado de Máquina. Classificação de Imagens. Preservação Ambiental.

¹Universidade Federal de Mato Grosso do Sul - UFMS, Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia. Programa de Pós-graduação em Tecnologias Ambientais, Brasil. E-mail: {nayara.estrabis@ufms.br; jose.marcato@ufms.br}

²Universidade Federal de Mato Grosso do Sul - UFMS, Faculdade de Computação. Programa de Pós-graduação em Ciência da Computação, Brasil. E-mail: pistori@ucdb.br

³Universidade Católica Dom Bosco - UCDB, Centro de Ciências Agrárias. Programa de Pós-graduação em Desenvolvimento Local e Agroambiental, Brasil. E-mail: pistori@ucdb.br

ABSTRACT

The Cerrado is one of the biomes existent in Brazil and the second largest of the South America. It has a great importance due to its biodiversity, ecosystem and mainly to be like a tank or “sponge”, that distribute water to the others biomes around it. Besides, it is a local of sources for some largest watersheds in South America. However, due to the anthropic activities practiced (with emphasis for the pasture and silviculture) and the reduction of the native vegetation, this biome is threatened. Considered a biodiversity hotspot, the Cerrado could not exist in 2050. With the preservation necessity, this study aimed at investigation of machine learning algorithm for the mapping of Cerrado native vegetation in Três Lagoas city, using the cloud plataform Google Earth Engine. The process was performed using a Landsat-8 OLI image, dated on 10 October 2018, and the Random Forest (RF) and Support Vector Machine (SVM) algorithms. In the classification validation, the kappa index resulted in 0.94 and 0.97, respectively. RF algorithm, when compared to SVM, presented noisy classification. It was verified the existence of native vegetation area about 2,556 km² when was adopted RF and 2,873 km² when SVM was adopted.

KEYWORDS: Remote Sensing. Machine Learning. Image Classification. Environmental Preservation.

* * *

Introdução

O Cerrado é um bioma existente no Brasil, considerado o segundo maior na América do Sul, com extensão que abrange 22% do território brasileiro (2 milhões km²), cobre 12 estados (Maranhão, Piauí, Tocantins, Goiás, Rondônia, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Bahia, Distrito Federal, São Paulo e Paraná) com incidência em outros 3 (Amapá, Roraima e Amazonas) (RATTER *et al.*, 1997; IBGE, 2004; IBGE, 2018; MMA, 2019).

Esta região abriga biodiversidade composta por mais de 4000 espécies endêmicas entre plantas e vertebrados, 11627 espécies catalogadas de vegetação nativa, além de ser local de nascentes das três maiores bacias da América do Sul: bacia Amazônica, bacia do Prata e bacia do Rio São Francisco (MARRIS, 2005; DURIGAN *et al.*, 2007; STRASSBURG *et al.*, 2017, FERREIRA *et al.*, 2017; MMA, 2019). Outra característica observada por Wantzen *et al.* (2003), consiste na ocorrência do efeito filtro natural nos cursos

d'água do Cerrado, que proporciona aspecto cristalino na água. Exemplos deste efeito foram encontrados no “Rio Verde”, “Córrego Brilhante” e “Rio Claro” no estado de Mato Grosso.

A presença da vegetação nativa tem importância significativa para este bioma e para a manutenção dos recursos hídricos do Brasil. No Cerrado existem espécies de plantas nativas que possuem raízes profundas que atuam como uma esponja e canal de transporte de água para o reabastecimento dos aquíferos, diferente das raízes curtas das culturas de plantação e do solo sem cobertura que fazem com que a água evapore antes de alcançar os aquíferos (SCHOLZ et al., 2002; FELLETT, 2017).

Apesar de sua grande importância e rica biodiversidade, o Cerrado é considerado *hotspot* mundial em biodiversidade devido à ameaça de destruição (STRASSBURG et al., 2017; DURINGAN et al., 2007; MARRIS, 2005). Expansões das áreas de pastagem, agricultura, queimadas e desmatamento têm sido alguns dos fatores da redução da cobertura da vegetação nativa do Cerrado, gerando impactos no seu ecossistema e biodiversidade (BEUCHLE et al., 2015; PEREIRA e GAMA, 2010; DURINGAN et al., 2007; WARD et al., 1992). Nos últimos 40 anos, este bioma tem sido o centro da agricultura intensiva e de expansão de pastagem. No entanto, essas alterações afetam o balanço hídrico desta região (SPERA, 2016), gerando consequências negativas para o bioma.

Beuchle et al. (2015), identificaram mudanças na cobertura florestal de dois biomas, o Cerrado e a Caatinga, entre os anos de 1990 e 2010. Os autores verificaram um padrão contínuo de redução na vegetação nativa do Cerrado, com variações médias anuais de 0,6%. Além disso, detectaram que a quantidade de vegetação nativa remanescente é menor que a quantidade de outros tipos de cobertura do solo.

As mudanças na cobertura vegetal no Cerrado têm sido negligenciadas quando comparada a atenção dada a região da Amazônia (BEUCHLE et al., 2015; MARRIS, 2005). Essa situação torna-se ainda mais preocupante, com a

possibilidade do Cerrado não existir em 2050, de acordo com Strassburg et al. (2017).

A legislação federal do Brasil, Lei Federal 12.651 de 2012 (BRASIL, 2012), conhecida como Novo Código Florestal, dispõe sobre a proteção da vegetação nativa, áreas ambientais e sua exploração. De acordo com esta lei, é exigida a preservação da vegetação de florestas na Amazônia Legal em 80% da área da propriedade. No entanto, em áreas de bioma Cerrado, a exigência para preservação da vegetação (reserva legal) é de 20% ou 35%, de acordo com a localização.

O monitoramento e o controle das atividades nas áreas do bioma Cerrado se fazem cada vez mais necessários e urgentes, para sua preservação e conservação. O Google Earth Engine (GEE), plataforma online gratuita, surge como uma ferramenta que possibilita o mapeamento e monitoramento de grandes áreas, como o Cerrado.

O GEE é uma plataforma em nuvem (TSAI et al., 2018) que desenvolve os processos de forma interativa e rápida para análises de dados espaciais que exigem alta performance computacional (GORELICK et al., 2017), além de permitir edições e criações de algoritmos (KUMAR e MUTANGA, 2018). É considerado um dos maiores avanços no processo de monitoramento da observação da Terra (KUMAR e MUTANGA, 2018). Cada vez mais tem sido explorado, em mapeamentos de uso e cobertura do solo de grandes extensões devido a sua rapidez de processamento e na geração de resultados (DUONG et al., 2018; KUMAR e MUTANGA, 2018; NYLAND et al., 2018; TSAI et al., 2018).

Apesar das vantagens e possibilidades que o GEE proporciona, existem poucos trabalhos utilizando-o para o mapeamento de regiões do Brasil. Alguns exemplos são os estudos por Goldblatt, Ballesteros e Burney (2017) no Sertão Semi-árido, áreas de mineração na Amazônia brasileira por Lobo et al. (2018), análise de temperatura em uma região de São Paulo por Bourscheidt (2017), áreas de pastagem no Brasil por Parente e Ferreira (2018), estudos na região

do Pantanal por Pereira e Gama (2018) e no Cerrado, realizado por Guerra et al. (2017) em Matopiba.

A plataforma online gratuita MapBiomas v.3.0 (Projeto MapBiomas, 2019) foi desenvolvida utilizando recursos do GEE. Esta plataforma produz mapas anuais da cobertura do uso do solo no Brasil desde 1985. A última versão disponibilizada refere-se ao ano de 2017. O MapBiomas utiliza-se de imagens Landsat e classificação com o algoritmo *Random Forest* (RF) realizados no GEE.

O estado de Mato Grosso do Sul possui 61% do seu território composto pelo bioma Cerrado (IBGE, 2018). Três Lagoas é um dos municípios deste estado pertencente à este bioma, e tem obtido destaque na produção de Eucalipto, um tipo de floresta plantada não-nativa do Cerrado, além de outras atividades que alteram constantemente a paisagem natural da região.

Portanto, torna-se necessário o desenvolvimento de estudos voltados para o monitoramento do Cerrado após 2017, data do último mapeamento disponibilizado no MapBiomas. É importante avaliar diferentes métodos de aprendizado de máquina voltados para a classificação de imagens com o GEE.

Assim, este estudo tem o objetivo de mapear a vegetação nativa na região do município de Três Lagoas no ano de 2018, utilizando a plataforma GEE com técnicas de aprendizado de máquina, *Random Forest* (RF) e *Support Vector Machine* (SVM).

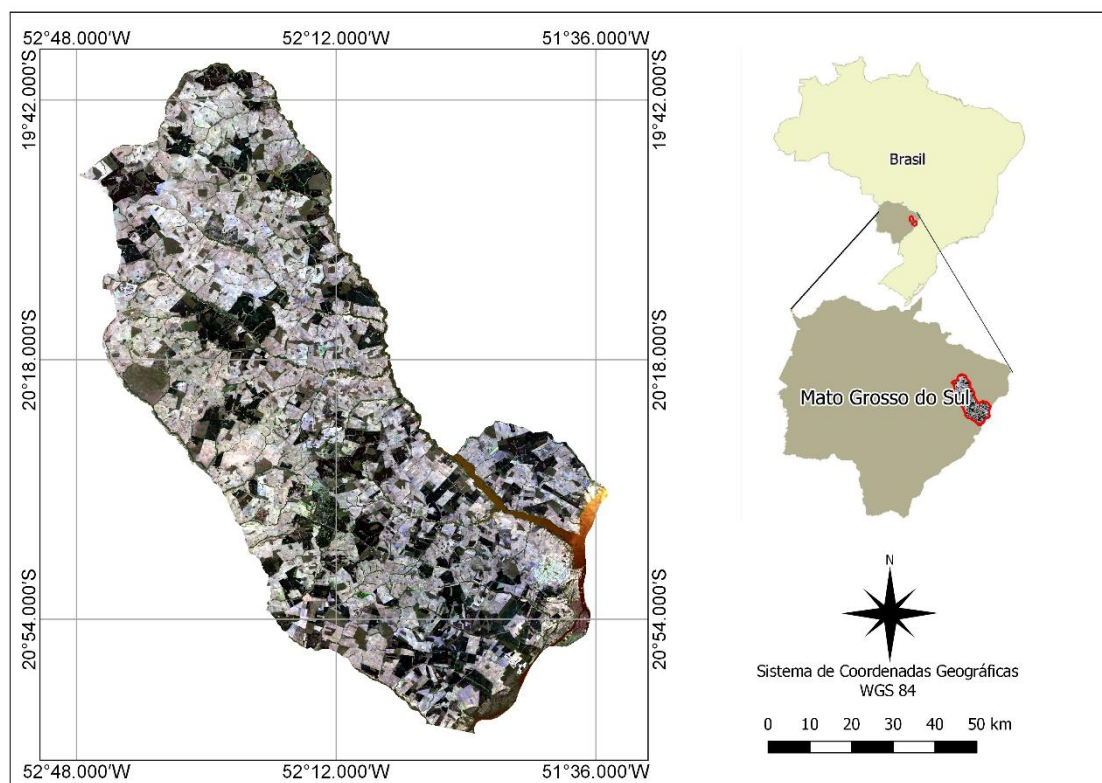
1 Metodologia

1.1 Área de Estudo

O município de Três Lagoas está localizado na porção leste do estado de Mato Grosso do Sul, como pode ser observado na Figura 1. Além das áreas de pastagem existentes na região, a silvicultura é uma atividade econômica em destaque, especialmente o eucalipto. O cultivo e produção do eucalipto fez Três Lagoas ser reconhecido internacionalmente como “Capital da Celulose”

(TRÊS LAGOAS, 2019). A instalação de indústrias de celulose na região contribuiu para o crescimento econômico do município e região, e aumentou a demanda de bens e serviços (TISOTT, SCHMIDT, e WAQUIL, 2017).

Figura 1 – Localização do município de Três Lagoas, da área de estudo e imagem Landsat 8 OLI órbita 223, ponto 074, data de 10 de outubro de 2018



Fonte: Elaborada pelos autores.

1.2 Classificação e Validação

A obtenção da imagem utilizada e a classificação foram realizadas na plataforma GEE. O processo de validação da classificação foi realizado no software QGIS (QGIS Development Team, 2019).

O GEE fornece um repositório de banco de imagens provenientes de diferentes satélites. Desta forma, através deste recurso, foi importada uma imagem Landsat 8 OLI (*Operational Land Imager*) Tier 1, com data de 10 de outubro de 2018. Dentre suas características destacam-se: apresentam alta acurácia posicional (BARROS, MARCATO JUNIOR, e PARANHOS FILHO,

2016); são ortorretificadas e convertidas para reflectância de superfície; possuem bandas multi-espectrais com resolução espacial de 30 metros; banda no infra-vermelho próximo (NIR) e duas bandas no infravermelho de ondas curtas (SWIR) (USGS, 2018). As bandas de 2 a 7 (Tabela 1) foram utilizadas neste estudo.

Tabela 1– Característica das Bandas Landsat 8 OLI

Bandas	Nome	Espectro (µm)
2	Azul	0,452 – 0,512
3	Verde	0,533 – 0,590
4	Vermelho	0,636 – 0,673
5	NIR	0,851 – 0,879
6	SWIR-1	1,566 – 1,651
7	SWIR-2	2,107 – 2,294

Fonte: Adaptado de USGS (2018).

Para o processo de treinamento, foram primeiramente estabelecidas as classes que caracterizaram a área (ver Figura 2). Com a consideração da economia local, foram atribuídas cinco classes: (1) vegetação nativa para as vegetações remanescentes ainda existente; (2) silvicultura para eucalipto e pinus, em especial para o cultivo de eucalipto praticado na região; (3) pastagem para as áreas de criação de gado; e (4) água para os cursos d'água, lagos, represas e nascentes.

Na área que abrange o município de Três Lagoas, foram coletadas 25 amostras para cada classe, totalizando-se 100 amostras para o treinamento. As amostras foram delimitadas manualmente em diferentes tamanhos, de acordo com tamanho e disposição dos objetos na imagem.

Vários métodos de aprendizado de máquinas têm sido utilizados para o mapeamento de vegetação nativa e de uso do solo, dentre eles o RF e o SVM (TSAI et al, 2018). Os algoritmos de aprendizado de máquina RF e SVM foram utilizados na classificação. O algoritmo RF é um conjunto de árvores preditoras, onde cada árvore possui um vetor aleatório independente e a seleção de informação ocorre pela escolha da classe mais popular para um determinado dado (BREIMAN, 2001).

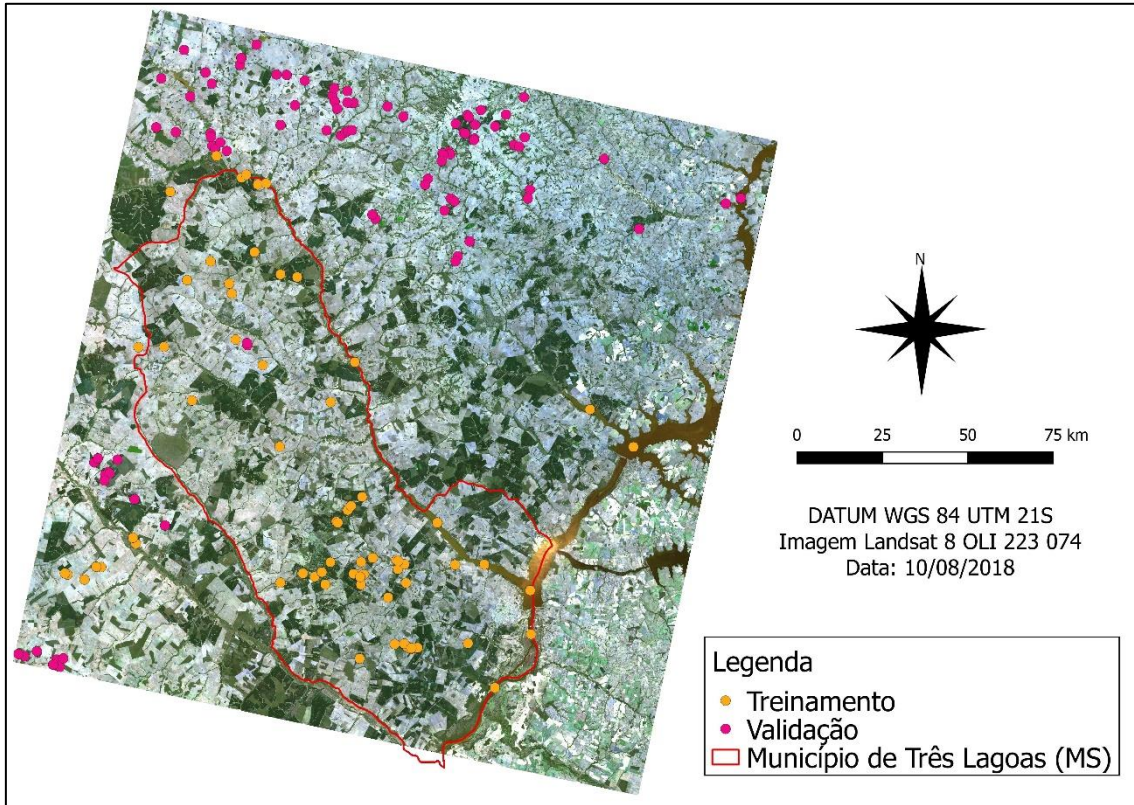
Para este estudo foram realizados testes com RF considerando 10 e 250 árvores de decisão, os quais apresentaram resultados similares. Ao levar em consideração o tempo de processamento, optou-se por utilizar o RF com 10 árvores de decisão.

O algoritmo SVM mapeia os dados para um hiperplano, onde decisões são feitas de forma linear, garantindo alta generalização da máquina de aprendizagem (CORTES e VAPNIK, 1995). Para este algoritmo foram realizados testes com os demais parâmetros existentes no GEE. No entanto, por apresentarem tempo de computação excedido neste estudo, o algoritmo SVM foi configurado para o modo padrão do sistema GEE .

A Figura 2 apresenta os locais das amostras de treinamento (em amarelo) e validação (em rosa). As amostras de treinamento concentraram-se dentro do limite do município de Três Lagoas, com exceção de alguns polígonos. As amostras de validação concentraram-se na região fora do limite do município de Três Lagoas.

A classificação realizada na imagem foi validada no software QGIS utilizando 22 polígonos para cada classe, correspondendo ao todo em 88 polígonos de validação. As amostras de validação (Figura 2) concentraram-se em regiões diferentes da região dos polígonos de treinamento.

Figura 2 – Localização das amostras de treinamento (em amarelo) e validação (em rosa)



Fonte: Elaborada pelos autores.

Visando a validação da classificação, foi estimado o índice kappa, método de avaliação proposto por Cohen (1960) (Equação 1).

$$K = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

onde K é o valor do kappa, P_o é a proporção de concordância e P_e é a proporção de concordância esperada por acaso.

Para fins de validação, estimou-se também a acurácia geral (Equação 2), que apresenta a probabilidade de uma amostra ser corretamente classificada. A acurácia do produtor (Equação 3), relacionada às omissões das amostras em relação a classe certa, apresenta o quão bem uma área pode ser mapeada na Terra. A acurácia do usuário (Equação 4), relacionada aos erros de comissão, apresenta o quão bem o mapa representa o que realmente é no solo, ou seja, a confiabilidade do mapa (STORY e CONGALTON, 1986).

$$AG = \frac{\sum ac}{\sum ar} \quad (2)$$

onde AG é o valor da acurácia geral, ac é o número de amostras corretas e ar é o número total de amostras de referência.

$$AP = \frac{ac}{ar} \quad (3)$$

onde AP é o valor da acurácia do produtor, ac é o número de amostras corretas da classe e ar é o número total de amostras de referência da classe.

$$AU = \frac{ac}{ad} \quad (4)$$

onde AU é o valor da acurácia do usuário, ac é o número de amostras corretas da classe e ad é o número total de amostras da imagem classificadas para a classe.

2 Resultado e Discussão

O treinamento realizado pelos algoritmos apresentou concordâncias superiores a 95%, em quase todas as classes (Figura 3). Na própria plataforma, foi possível verificar os valores de acurácia geral do treinamento obtidos pelo RF e SVM, com valores de 0,99998 e 0,99995, respectivamente.

Os resultados do processo de validação com o RF e SVM podem ser verificados nas Tabelas 2 e 3, respectivamente. Os valores na diagonal, em destaque, representam a quantidade de alvos que apresentaram concordância e os demais são as confusões que ocorreram com as outras classes.

Ao analisar a Tabela 2 é possível verificar concordâncias acima de 90% na maioria das classes para a classificação do RF. As classes pastagem e água apresentaram os maiores resultados com concordâncias acima de 99% e a classe silvicultura apresentou o menor resultado com valores acima de 86% de concordância.

Tabela 2– Matriz de Confusão – Validação RF (%)

Classe	Referência			
	Vegetação Nativa	Silvicultura	Pastagem	Água
Vegetação Nativa	91,51	13,78	0	0,09
Silvicultura	8,47	86,22	0	0
Pastagem	0	0	99,95	0,09
Água	0	0	0,05	99,81

Fonte: Elaborada pelos autores.

A Tabela 3 apresentou valores de concordâncias para a classificação do SVM acima de 95% na maioria das classes. As classes vegetação nativa, pastagem e água apresentaram resultados similares e melhores concordâncias. Já a classe silvicultura apresentou o menor resultado com 92,44%.

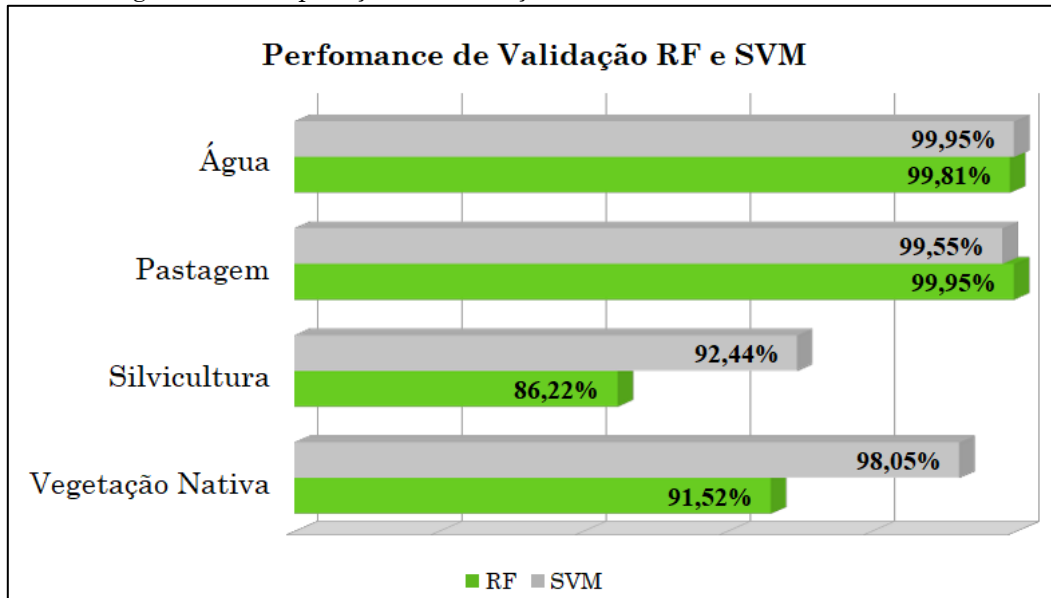
Tabela 3– Matriz de Confusão – Validação SVM (%)

Classe	Referência			
	Vegetação Nativa	Silvicultura	Pastagem	Água
Vegetação Nativa	98,05	7,56	0,39	0,03
Silvicultura	1,95	92,44	0	0
Pastagem	0	0	99,55	0,02
Água	0	0	0,05	99,95

Fonte: Elaborada pelos autores.

A comparação entre os resultados da validação obtidos para cada algoritmo é apresentada na Figura 3. As classes água e pastagem apresentaram semelhanças na validação para cada algoritmo, nas demais classes houve diferença entre 6,22% à 6,53% entre os resultados. A classe de vegetação nativa, apresentou a maior diferença (6,53%) entre os resultados de validação para os algoritmos.

Figura 3 – Comparação da validação do RF em verde e SVM em cinza



Fonte: Elaborada pelos autores.

Na Tabela 4 são apresentados os valores do índice kappa, da acurácia geral, acurácia do usuário e do produtor, estimados no processo de validação através do software QGIS. A acurácia do produtor e do usuário estão relacionados aos erros de omissão e comissão, e o índice kappa refere-se as concordâncias obtidas para cada classe e para cada classificação.

Tabela 4– Resultado da validação: acurácia geral (%) e índice kappa

Classes	Acurácia do Produtor		Acurácia do Usuário		Kappa por classe	
	RF	SVM	RF	SVM	RF	SVM
Vegetação Nativa	91,52%	98,05%	90,71%	94,31%	0,88	0,93
Silvicultura	86,22%	92,44%	87,32%	96,98%	0,85	0,96
Pastagem	99,95%	99,55%	99,97%	99,99%	1,00	1,00
Água	99,81%	99,95%	99,79%	99,80%	1,00	1,00
	Acurácia Geral				Kappa Geral	
RF	95,69%				0,94	
SVM	98,09%				0,97	

Fonte: Elaborada pelos autores.

De forma geral as classificações apresentaram validação com índices kappa superiores à 0,90. Os índices kappa obtidos foram de 0,94 para o RF e 0,97 para o SVM, considerados respectivamente, como concordâncias “quase perfeitas”, de acordo com Landis e Koch (1977).

A validação apresentou acurácia do produtor, relacionada ao erro de comissão ou inclusão, acima de 80% para o RF e acima de 90% para o SVM, conforme Tabela 4. A acurácia do usuário, relacionada ao erro de omissão, obteve valores acima de 85%, indicando erros de omissão inferiores à 15%.

A classe vegetação nativa apresentou 91,51% (Tabela 2) e 98,05% (Tabela 3) de concordância entre os pontos de validação e a classificação realizada, respectivamente, pelo RF e SVM. Não houve correspondência em 8,48% das amostras de validação com a classificação para o RF e 1,95% para o SVM. As confusões encontradas durante o processo de validação para esta classe (vegetação nativa) ocorreram com a classe silvicultura, principalmente em decorrência das semelhanças espectrais existentes entre ambas as classes.

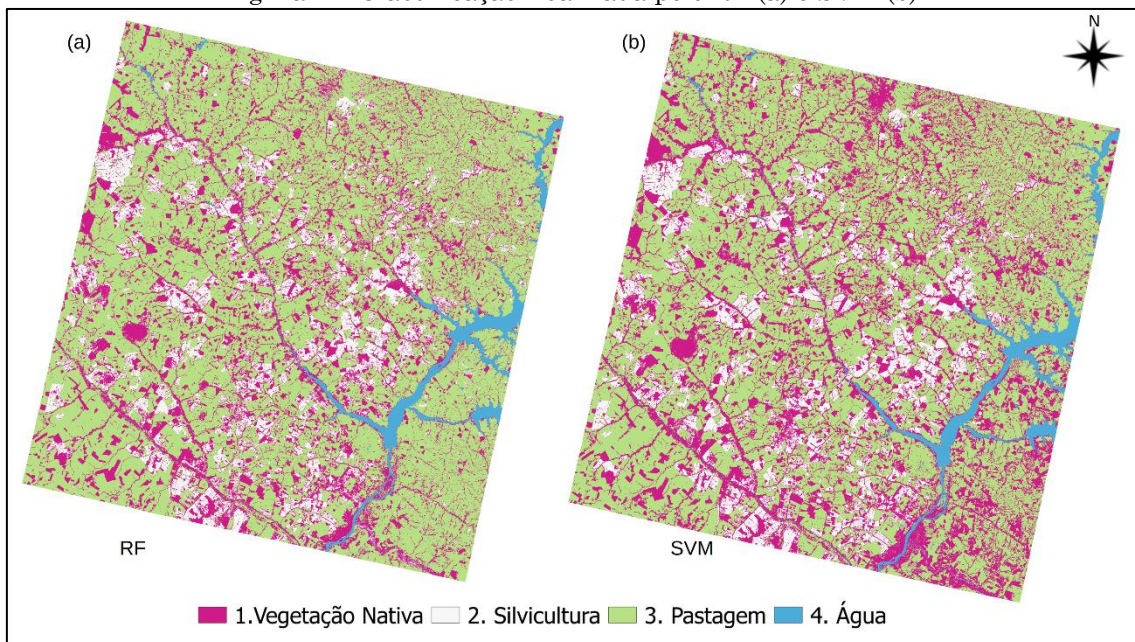
A silvicultura, classe composta por vegetação de eucalipto, apresentou respectivamente 86,22% e 92,44% (Tabelas 2 e 3) de concordância dos polígonos com a classificação do RF e do SVM na validação. Não houve correspondência em 7,56% das amostras para o SVM e 13,78% para o RF. O SVM obteve resultado superior a 6% em relação ao obtido na classificação realizada pelo RF.

A classe pastagem apresentou resultado de validação de 99,95% para o RF e 99,55% para o SVM (Figura 3). As confusões inferiores à 1% ocorreram com a classe água em ambos os algoritmos. Essas confusões podem ser explicadas pela possível presença de vegetação em águas com características espectrais similares com as vegetações existentes nas pastagens. Na classificação gerada pelo SVM, houve também confusões com a vegetação nativa. Algumas árvores existentes em áreas de pastagem podem ter influenciado durante a classificação do algoritmo e gerado confusão entre as classes.

A água apresentou quase 100% de polígonos concordantes na validação, sendo 99,81% ao considerar a classificação do RF e 99,95% para a classificação com o SVM, apresentando menos que 0,2% de amostras não concordantes. Apesar de possuir características espectrais diferentes das demais classes, esta classe apresentou confusões com a classe vegetação nativa e pastagem. Esse resultado pode ser explicado pela existência de vegetação em áreas úmidas e em margens de cursos d'água, lagos, nascentes e olhos d'água.

Na Figura 4 é possível verificar a classificação da imagem realizada pelos algoritmos RF (a) e pelo SVM (b), respectivamente. Comparando a classificação realizada por cada algoritmo na Figura 4.a e 4.b, foi possível verificar, para a mesma área de uma determinada classe, a presença de alguns pixels ou detecções pertencentes a outras classes e que não fazem correspondência à classe de referência.

Figura 4 – Classificação realizada pelo RF (a) e SVM (b)

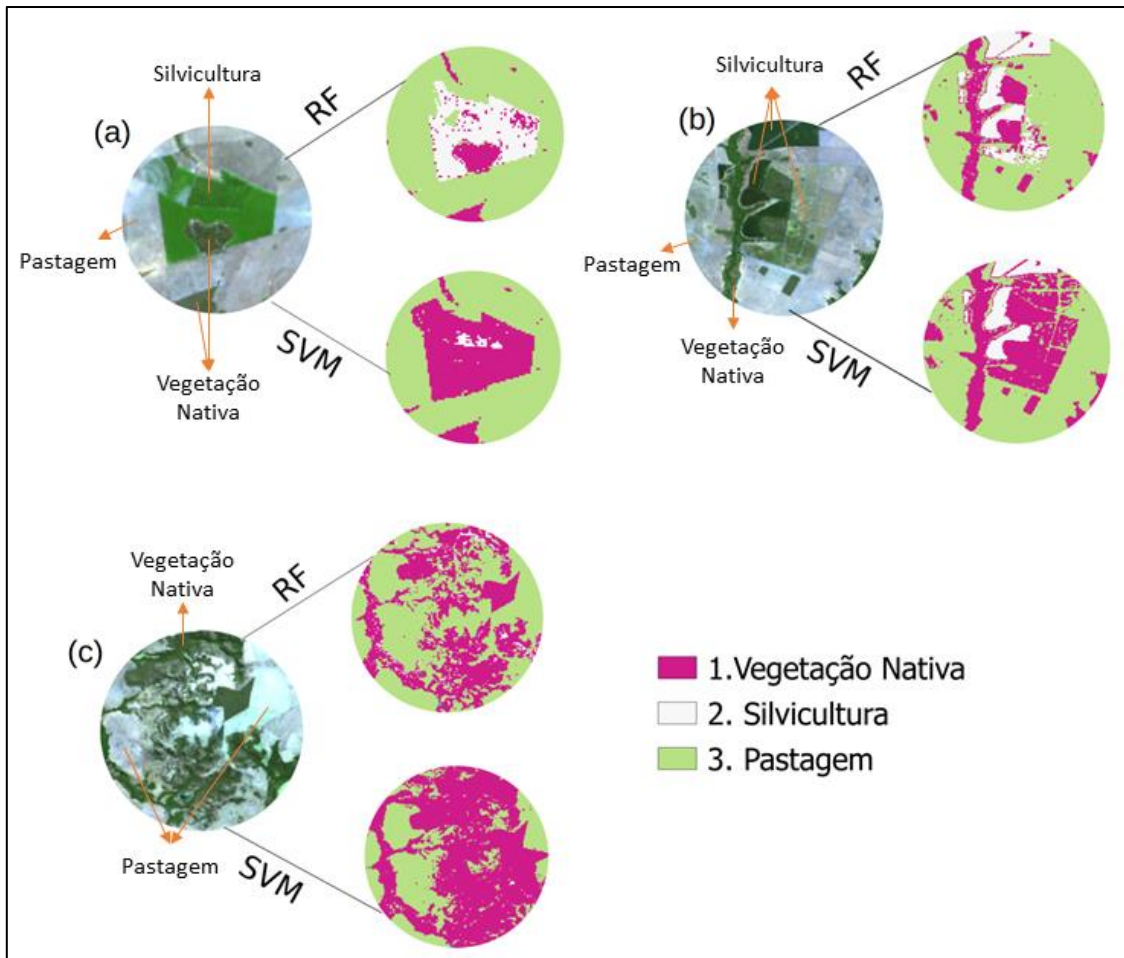


Fonte: Elaborada pelos autores.

Essas confusões podem ser observadas com destaque na Figura 5. Para realizar uma análise qualitativa, detalhes de algumas regiões das classificações foram destacadas. Na Figura 5 apresentam-se 3 áreas (5.a,

5.b e 5.c), onde são indicadas a área visualizada na imagem Landsat e suas respectivas classificações obtidas pelo RF e pelo SVM.

Figura 5 – Detalhamento de amostras de vegetação nativa e silvicultura classificadas pelo RF e SVM



Fonte: Elaborada pelos autores.

Ao observar a Figura 5 é possível verificar detecções diferentes realizadas pelos algoritmos. Na Figura 5.a a região de pastagem foi classificada igual por ambos os algoritmos, no entanto as áreas de vegetação localizada ao centro foi classificada diferente. A vegetação em verde mais claro na imagem Landsat, compreende uma área de silvicultura ao lado de uma área menor de vegetação nativa, em verde mais escuro. A classificação realizada pelo RF conseguiu delimitar essa área melhor do que o SVM. Foram observadas alguns pixel ou pontos da classe de vegetação nativa na área de

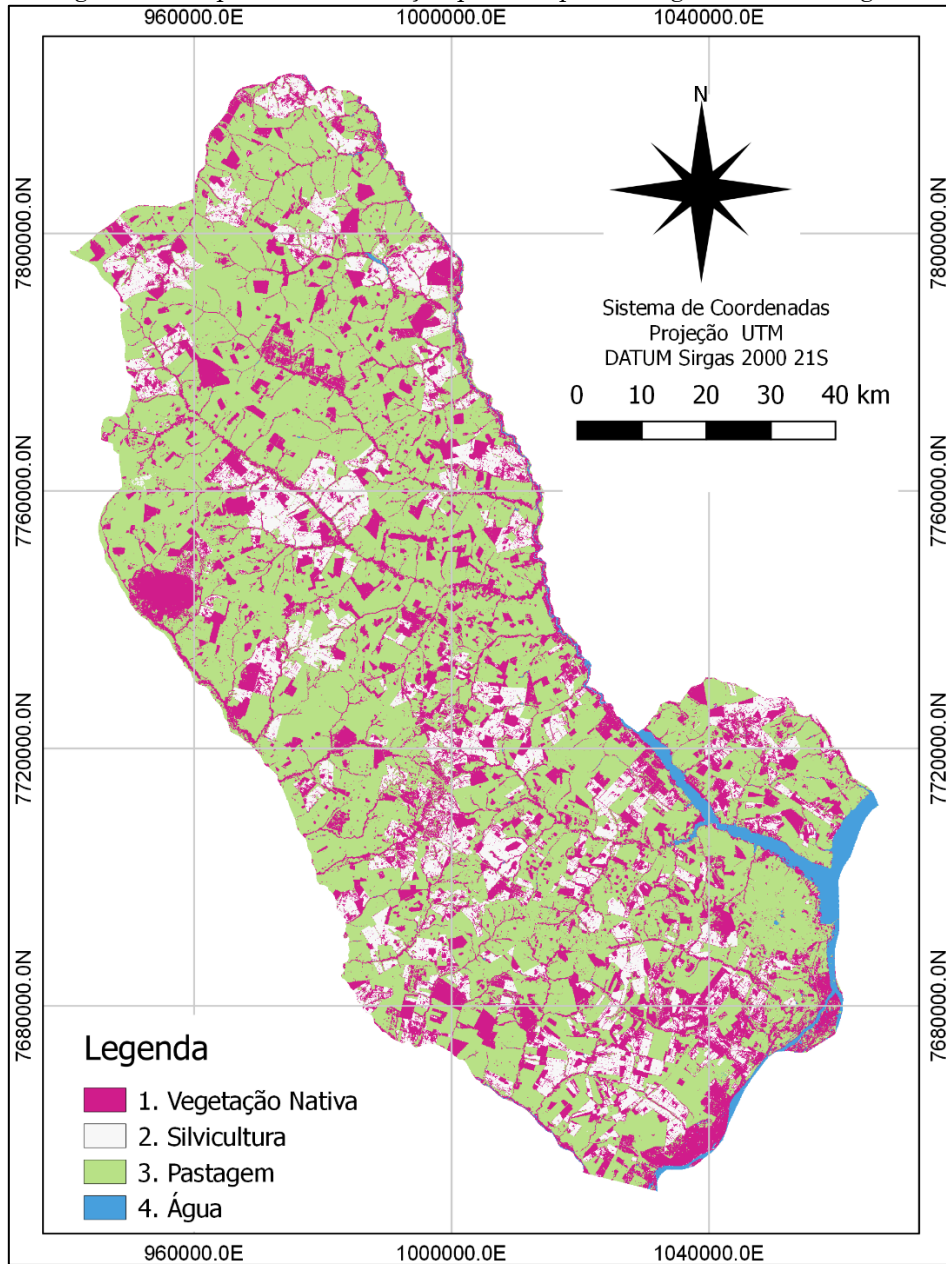
silvicultura. O algoritmo SVM confundiu a maior parte da área de silvicultura (cor branca) com a área de vegetação nativa. Essas confusões podem ser explicadas pela existência de semelhanças espectrais da vegetação com a silvicultura. E também, pela existência de omissões e inclusões ocorridos no processo de classificação, cuja apresentação qualitativa obteve esse aspecto com pixels ou pontos de uma classe inseridos em outra classe.

Ao observar a Figura 5.b é possível notar que o algoritmo SVM conseguiu perceber a existência de vegetação melhor que o RF. No entanto essa área foi identificada como vegetação nativa pelo SVM e como pastagem pelo RF, quando na verdade é uma área de silvicultura com pouco eucalipto. Por se tratar de classes com semelhanças espectrais, como já mencionado anteriormente, podem ter gerado maior influência e confusão para a performance na identificação e classificação pelos algoritmos.

Na Figura 5.c é notada uma área de pastagem com áreas de vegetação nativa. O algoritmo RF delimitou a vegetação nativa mais densa na área e nas áreas com vegetação menos densa classificou como pastagem. O algoritmo SVM detectou áreas com vegetação menos densa e classificou como vegetação nativa. As semelhanças espectrais e os erros de comissão e omissão podem ter influenciado na confusão de classes no processo de classificação pelos algoritmos. Contudo, os algoritmos desempenharam boa performance na classificação da imagem.

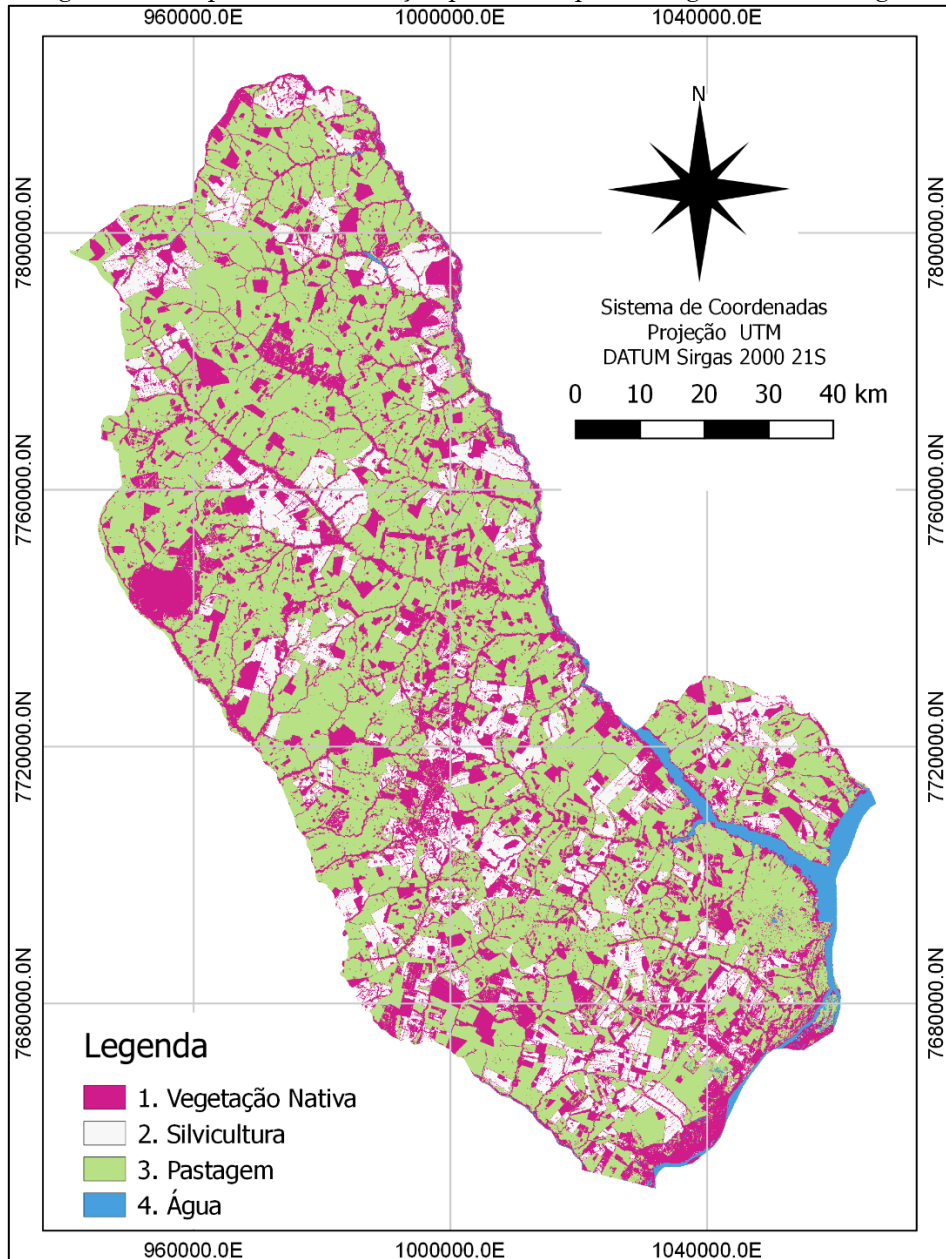
Nas Figuras 6 e 7 são apresentados os mapeamentos resultantes do processo de classificação pelos algoritmos RF e SVM para a região de Três Lagoas. O mapeamento com RF demonstrou que a área de vegetação nativa está distribuída em 25% da área total. Ao considerar o mapeamento realizado pelo SVM, a região obteve 28% de vegetação nativa.

Figura 6 –Mapa de Classificação pelo RF para a região de Três Lagoas



Fonte: Elaborada pelos autores.

Figura 7 – Mapa de Classificação pelo SVM para a região de Três Lagoas



Fonte: Elaborada pelos autores.

A região do município, em maior parte é composta por áreas de pastagem, seguida pela área de vegetação nativa e silvicultura. As áreas de vegetação nativa existem em menor porção, comparado às áreas de pastagem. Ao comparar com as áreas de silvicultura, possui valores próximos as área de vegetação nativa. De acordo com o mapeamento foram identificados

aproximadamente 2556 km² de vegetação nativa pelo RF e 2873 km² pelo SVM na região de Três Lagoas.

Estes resultados obtidos, referentes ao ano de 2018, apresentaram-se maiores que o valor identificado pelo MapBiomias em 2017 para vegetação nativa. Nessa plataforma foram encontrados 2121 km² de vegetação nativa, sendo que 1445 km² de vegetação nativa do tipo formação florestal, 476 km² de formação savânica e 200 km² do tipo formação campestre.

3 Conclusão

O processamento rápido e o ambiente interativo da plataforma GEE demonstraram-se vantajosos para a realização da classificação. O repositório de imagens de diferentes satélites, presente na plataforma, facilita o processo de escolha e obtenção da imagem desejada.

Os algoritmos RF e SVM apresentaram boa performance na classificação e validação. O algoritmo SVM demonstrou, de forma geral, melhor acurácia e índice kappa do que o algoritmo RF.

Através do mapeamento, verificou-se que a área de vegetação nativa do Cerrado está presente em 25% da área total da região ao considerar o RF e 28% ao considerar SVM. As atividades antrópicas ocupam a maior parcela da região, o que contribui com a supressão de vegetação nativa do Cerrado. Desta forma, o monitoramento se faz cada vez mais urgente e necessário para a preservação da vegetação nativa e do bioma Cerrado.

Sugere-se para trabalhos futuros a ampliação desse estudo para todo o Cerrado brasileiro e utilização de outros parâmetros dos algoritmos que possibilite a redução de erros de omissão e comissão no processo de classificação de sua área.

Agradecimentos

Os autores gostariam de agradecer o apoio da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil) através da bolsa de mestrado (Código de Financiamento 001), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e FUNDECT (Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul), através do financiamento de projetos de pesquisa (p. 456149/2014-7, p. 433783/2018-4 e p. 59/300.066/2015). O presente trabalho também foi realizado com financiamento de projeto de pesquisa pela CAPES através do programa Print-CAPES.

Contribuição dos autores

Nayara Vasconcelos Estrabis foi responsável pela confecção, coleta de dados, realização dos experimentos e discussão. Hemerson Pistori foi responsável pela supervisão e revisão. José Marcato Junior foi responsável pela orientação, supervisão e revisão.

Referências

- BARROS, W. M.; MARCATO JUNIOR, J.; PARANHOS FILHO, A. C.. Rodovias como elemento de controle da acurácia posicional de imagens OLI/Landsat-8: aplicação na região do Pantanal. **Revista Brasileira de Cartografia**, v.68, n.5, 2016, pp. 921-928.
- BEUCHLE, R.; GRECCHI, R. C.; SHIMABUKURO, Y. E.; SELIGER, R.; EVA, H. D.; 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, 2015, pp. 116–127.
- BOURSCHEIDT, V. Análise de tendência da temperatura de superfície a partir de imagens Landsat 5: contribuições da plataforma Google Earth Engine. **Anais**

do XVIII Simpósio Brasileiro de Sensoriamento Remoto SBSR. Santos, 2017. pp.7401-7407.

BRASIL. Lei 12.651, de 12 de maio de 2012. Dispõe sobre a proteção da vegetação nativa; altera as Leis nos 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as Leis nos 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória no 2.166-67, de 24 de agosto de 2001; e dá outras providências. **Diário Oficial da União**, 28 de maio de 2012.

BREIMAN, L. Random Forests. **Machine Learning**, vol. 45, 2001. pp. 5-32.

COHEN, J. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, 1960, n. 1, pp. 37-46.

CORTES, C; VAPNIK, V. Support-Vector Networks. **Machine Learning**, vol. 20, 1995. pp. 273-297.

DUONG, P.; TRUNG, T; NASAHARA, K.; TADONO, T. JAXA High-Resolution Land Use/Land Cover Map for Central Vietnam in 2007 and 2017. **Remote Sensing**, v. 10, 2018, n. 9, pp. 1406.

DURIGAN, G.; SIQUEIRA, M.F de; FRANCO, G.A.D.C. Threats to the Cerrado remnants of the state of São Paulo, Brazil. **Scientia Agricola**, v. 64, 2007, n. 4, pp. 355-363.

FELLET, J. Como as raízes do Cerrado levam água a torneiras de todas as regiões do Brasil. BBC Brasil. 2017. Site <<https://www.bbc.com/portuguese/brasil-39391161>>, acessado em janeiro de 2019.

FERREIRA, F.G.; MACHADO, E. L. M.; SILVA NETO, C. DE M. E ; SILVA JÚNIOR, M. C.; MEDEIROS, M. M.; GONZAGA, A. P. D.; SOLÓRZANO, A.; VENTUROLI, F.; FAGG, J. M. F. Diversity and indicator species in the cerrado biome, Brazil. **Australian Journal of Crop Science AJCS**, v. 11, 2017, n. 08, pp. 1042-1050.

GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, 2017, pp. 18-27.

GOLDBLATT, R.; RIVERA BALLESTEROS, A.; BURNEY, J. High Spatial Resolution Visual Band Imagery Outperforms Medium Resolution Spectral

- Imagery for Ecosystem Assessment in the Semi-Arid Brazilian Sertão. **Remote Sensing**. 2017, vol. 9, n. 12, p. 1336.
- GUERRA, J. B.; SCHULTZ, B.; SANCHES, I. D. Mapeamento automático da expansão da agricultura anual no Matopiba entre 2002 e 2015 utilizando a plataforma Google Earth Engine. **Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto SBSR**. Santos, 2017. pp.6850-6857.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Brasil em síntese: território. 2018, Site < <https://brasilemsintese.ibge.gov.br/territorio.html> >. Acessado em dezembro 2018.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Mapa de Biomas e de Vegetação, 2004. Site <<https://ww2.ibge.gov.br/home/presidencia/noticias/21052004biomashtml.shtm>>. Acessado em dezembro 2018.
- KUMAR, L.; MUTANGA, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. **Remote Sensing**, vol, 10, 2018. pp. 1509.
- LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, vol. 33, 1977. pp. 159-174.
- LOBO, F. de L.; SOUZA-FILHO, P. W. M.; NOVO, E. M. L. de M.; CARLOS, F. M.; BARBOSA, C. C. F. B. Mapping Mining Areas in the Brazilian Amazon Using MSI/Sentinel-2 Imagery (2017). **Remote Sensing**, vol.10, 2018, p. 1178.
- MARRIS, E. The forgotten ecosystem. **Nature**, vol. 437, 2005, n 7061. pp. 944–945.
- MMA, Ministério do Meio Ambiente. O Bioma Cerrado. 2019. Site < <http://www.mma.gov.br/biomas/cerrado> >. acessado em janeiro 2019.
- NYLAND, K.E.; GUNN, G. E.; SHIKLOMANOV, N. I.; ENGSTROM, R.N.; STRELETSKIY, A. Land Cover Change in the Lower Yenisei River Using Dense Stacking of Landsat Imagery in Google Earth Engine. **Remote Sens.**, vol. 10, 2018, n. 8. p. 1226.
- PARENTE, L.; FERREIRA, L. Assessing the Spatial and Occupation Dynamics of the Brazilian Pasturelands Based on the Automated Classification of MODIS Images from 2000 to 2016. **Remote Sensing**, vol. 10, 2018, n.4. p. 606.
- PEREIRA, A.C.; GAMA, G V.F. Anthropization on the Cerrado biome in the Brazilian Uruçuí-Una Ecological Station estimated from orbital images. **Brazilian Journal of Biology**. vol. 70, 2010, n. 4. pp. 969-976.

- PEREIRA, G. H. de A. et al. Uso da plataforma Google Earth Engine para análise multitemporal de imagens SAR para detecção de variações de áreas inundadas no Pantanal. **Anais 7º Simpósio de Geotecnologias no Pantanal**, Embrapa Informática Agropecuária/INPE, Jardim, 2018. pp. 274-282.
- Projeto MapBiomas – Coleção 3.0 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil, acessado em 2019 através do link: <http://mapbiomas.org/map#coverage>
- QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project., 2019.
- RATTER, J. The Brazilian Cerrado Vegetation and Threats to its Biodiversity. **Annals of Botany**. vol. 80, 1997, n. 3, pp. 223–230.
- SCHOLZ, F. G. Hydraulic redistribution of soil water by Neotropical savanna trees. **Tree Physiol**. vol.22, 2002. pp. 603-612.
- SPERA, S. A.; GALFORD, G. L.; COE, M. T., MACEDO, M. N.; MUSTARD, J. F. Land-use change affects water recycling in Brazil's last agricultural frontier. **Global Change Biology**. vol.22, 2016, pp. 3405–3413.
- STORY, M.; CONGALTON, R. G. Accuracy assessment: A user's perspective. **Photogrammetric Engineering and Remote Sensing**. vol.52, 1986, pp. 397–399.
- STRASSBURG, B. B. N.; BROOKS, T.; FELTRAN-BARBIERI, R.; IRIBARREM, A.; CROUZEILLES, R.; LOYOLA, R.; LATAWIEC, A.E.; OLIVEIRA FILHO, F.J.B.; SCARAMUZZA, C. A. DE M.; SCARANO, F. R.; SOARES-FILHO, B.; BALMFORD, A. Moment of truth for the Cerrado hotspot. **Nature Ecology & Evolution**. vol.1, n. 0099, 2017, pp.1-3.
- TISOTT, S. T.; SCHMIDT, V.; WAQUIL, P. D. Atividade Florestal e o Desenvolvimento Socioeconômico em Três Lagoas e Região: Uma Análise Baseada na Abordagem de Cluster. **Desenvolvimento em Questão**. vol.15, 2017. pp. 228-260.
- TRÊS LAGOAS. A Cidade. Site <<http://www.treslagoas.ms.gov.br/cidade/>>. acessado em janeiro 2019.
- TSAI, Y.H.; STOW, D.; CHEN, H.L.; LEWISON, R.; AN, L.; SHI, L. Mapping Vegetation and Land Use Types in Fanjingshan National Nature Reserve Using Google Earth Engine. **Remote Sensing**, vol. 10, 2018, pp. 927.

USGS, United State Geological Survey. Section 2 – Observatory Overview. 2018. Site <<https://landsat.usgs.gov/landsat-8-l8-data-users-handbook-section-2>>.

acessado em dezembro 2018.

WANTZEN, K. M. Cerrado streams – characteristics of a threatned freshwater ecosystem type on the Tertiary Shields of Central South America. **Amazoniana**. vol.XVII, 2003. pp. 481-502.

WARD, D. E.; SUSOTT, R. A; KAUFFMAN, J. B ; BABBITT, R. E.; CUMMINGS, D. L.; DIAS, B.; HOEBEN, N.; KAUFMAN, Y. J.; RASMUSSEN, R. A.; SETZER, A. W. Smoke and fire characteristics for cerrado and deforestation burns in Brazil: BASE-B Experiment. **Journal of Geophysical Research**. vol. 97, 1992. pp. 14.601-14.619.