

# Algoritmo para mapeamento dos usos do solo e cobertura vegetal a partir do uso do NDVI: um estudo aplicado no nordeste de Mato Grosso Do Sul

Algorithm for mapping land uses and vegetable coverage, from the use of NDVI applied in the northeastern of Mato Grosso Do Sul

Algoritmo para mapeo de usos de suelo y cobertura vegetal, a partir del uso de NDVI aplicado en el noreste de Mato Grosso do Sul

**José Roberto Mantovani**  

Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP, São José Dos Campos (SP), Brasil

[j.mantovani@unesp.br](mailto:j.mantovani@unesp.br)

**Leandro Reginaldo Maximino Lelis**  

Instituto Federal da Paraíba – IFPB, Catolé do Rocha (PB), Brasil

[leandrolelis87@gmail.com](mailto:leandrolelis87@gmail.com)

---

## Resumo

Propõe-se uma metodologia que utiliza conceitos matemáticos da teoria de formação de clusters, o K-means, para obter de forma automática os clusters dos valores do NDVI. As implementações do método K-means existentes em softwares específicos exige a predefinição no número de clusters, sendo a contribuição desta metodologia a determinação do número de clusters automaticamente, sem a necessidade da interferência do tomador de decisões que, pode variar de acordo com o tempo e o espaço de uma imagem para outra, bem como de um sensor para o outro. Foram selecionados diferentes sensores para generalizar essa classificação do índice: *Thematic Mapper* (TM) a bordo do satélite Landsat-5; *Operational Terra Imager* (OLI) a bordo do satélite Landsat-8; *MultiSpectral Instrument* (MSI) a bordo do satélite Sentinel, nível-2A. O mapeamento e a validação do algoritmo são efetuados na região nordeste do estado de Mato Grosso do Sul, a qual, apresenta ao longo de 37 anos (1984- 2021) alterações em sua cobertura vegetal. Os resultados para os três períodos, fornecidos pelo algoritmo distinguiu melhor o comportamento espectral dos *pixels* referentes às classes de água, solo exposto e áreas urbanas; já o JENKS generalizou essas classes, por outro lado, distinguiu com melhor precisão vegetação de baixo porte, vegetação natural e florestas plantadas.



**Palavras-chave:** Cluster. K-Means. Comportamento espectral. Jenks.

---

#### **Abstract**

A methodology is proposed that uses mathematical concepts of the theory of cluster formation, the K-means, to automatically obtain the clusters of the NDVI values, to determine the use and occupation of the land. The implementations of the K-means method existing in specific software require the predefinition of the number of clusters, and the contribution of this methodology is the determination of the number of clusters automatically, without the need for interference of the decision maker, which may vary according to time and space from one image to another, as well as from one sensor to the next. Different sensors were selected to generalize this index classification: Thematic Mapper (TM) on board the Landsat-5 satellite; Operational Terra Imager (OLI) aboard the Landsat-8 satellite; MultiSpectral Instrument (MSI) aboard the Sentinel, level-2A satellite. The mapping and validation of algorithm are carried out in the northeast region of state of Mato Grosso do Sul, which, over 37 years (1984-2021) shows changes in its vegetation cover. The results obtained for the three periods, provided by the algorithm, better distinguished the spectral behavior of pixels referring to water classes, exposed soil and urban areas; on the other hand, the JENKS method generalized these classes, on the other hand, it better distinguished low-sized vegetation, natural vegetation and planted forests.

**Keywords:** Cluster. K-means. Spectral behavior. Jenks.

---

#### **Resumen**

Se propone una metodología que utiliza conceptos matemáticos de la teoría de formación de clusters, K-means, para obtener automáticamente los clusters de los valores del NDVI. Las implementaciones del método K-means existentes em software específico requieren la predefinición del número de clusters, y el aporte de esta metodología es la determinación del número de forma automática, sin necesidad de interferencia del decisor, que puede variar según el tiempo y el espacio de una imagen a outra, así como de um sensor al siguiente. Se seleccionaron diferentes sensores para generalizar esta clasificación del índice: Thematic Mapper (TM) a bordo del satélite Landsat-5; Operational Terra Imager (OLI) a bordo del satélite Landsat-8; MultiSpectral Instrument (MSI) a bordo del satélite Sentinel, nivel-2. El mapeo y la validación del algoritmo se realizan en la región noreste del estado de Mato Grosso do Sul, que, a lo largo de 37 años (1984-2021), muestra cambios en su cobertura vegetal. Los resultados para los tres períodos, proporcionados por el algoritmo, distinguieron mejor el comportamiento espectral de los píxeles referidos a clases de agua, suelo expuesto y áreas urbanas; por otro lado, JENKS generalizó estas clases, pero distinguió con mayor precisión la vegetación de porte bajo, la vegetación natural y los bosques plantados.

**Palabras-clave:** Cluster. K-means. Comportamiento espectral. Jenks.

---

## **Introdução**

Os índices de vegetação são importantes parâmetros de análise utilizados em diferentes áreas da ciência, seja para prever a produtividade de diferentes culturas agrícolas e/ou para o reconhecimento dos vários processos de natureza física, queimadas, alterações na cobertura vegetal, uso do solo e alterações em corpos hídricos, entre outros (E.G. Eastman & Fulk 1993, Anyamba & Eastman 1996; Potter & Brooks 1998; Richard & Pocard 1998). Nos últimos anos, a necessidade de desenvolver

pesquisas e ferramentas computacionais para o controle e planejamento ambiental capazes de atender diferentes critérios simultaneamente experimentou rápido crescimento com os avanços de técnicas analíticas de base de dados, *softwares* e *hardware*. Destacam-se o Sistema de Informação Geográfica (SIG), o geoprocessamento e os produtos do Sensoriamento Remoto, que permitem integrar técnicas a conceitos e postulados matemáticos. Constituem-se em mecanismos importantes de análise, estudos e planejamento do espaço. Através de imagens orbitais de satélite é possível uma análise da estrutura e da dinâmica da paisagem em diferentes escalas espaciais e temporais, variando conforme a configuração do sensor (DE ALBUQUERQUE *et al.*, 2014).

O índice de vegetação de diferença normalizada (NDVI) é um indicador gráfico que pode ser usado para analisar medições de sensoriamento remoto, geralmente a partir de uma plataforma espacial, avaliando se o alvo observado contém ou não vegetação verde viva. Este índice é importante nos estudos referentes ao vigor da vegetação, mudanças na biomassa e estresse hídrico das plantas (CARDOZO *et al.*, 2009).

O índice produz limiares entre -1 e 1, representando o comportamento espectral, onde se tem o retroespalhamento do contraste verde; valores muito baixos e/ou negativos (-1 a 0) correspondem às áreas áridas de rocha, areia ou neve. Valores moderados positivos (0,2 a 0,3) representam arbustos e pastagens, enquanto valores altos (0,6 a 1) indicam florestas temperadas e tropicais (SHIMAKUBURO *et al.*, 1999; PONZONI, 2007, FITZ, 2013; PONZONI, SHIMABUKURO E KUPLICH, 2012).; ESRI, 2019). No entanto, não há uma metodologia direcionada à clusterização (agrupamento), deste índice para delimitar os seus intervalos e determinar os diferentes tipos de cobertura de solo de uma área sob análise, sendo, normalmente, utilizados métodos manuais (supervisionados) ou analógicos que, dependem fundamentalmente da interpretação do tomador de decisão, e que muitas vezes carecem de lógica matemática, podendo ser relativo, tendencioso e sem precisão (MANTOVANI e BUENO, 2020).

Uma outra opção para pra realizar a classificação do NDVI é o K-Means, consiste em uma técnica de clusterização destinada ao agrupamento de n conjunto de dados (instâncias) em k grupos. Ele é matematicamente capaz de lidar com qualquer modelo de dados. As n instâncias são métricas de distância com os k grupos (também conhecidos por protótipos), passando iterativamente a pertencer ao grupo que estiver

mais próximo. Os k grupos também vão sendo reajustados a cada iteração (NOGARE, 2016), porém necessita que o número de clusters seja especificado *a priori* pelo tomador de decisões.

Para superar essas limitações, tendo como referências os artigos de Johann *et al.*, (2013); Zhao *et al.*, (2017) e Pascucci *et al.*, (2018), que utilizam o algoritmo k-means, a partir de imagens orbitais a bordo do satélite Landsat-5, sensor *Thematic Mapper* (TM), para a delimitação dos intervalos do índice NDVI, nessa pesquisa propõe-se uma metodologia que utiliza conceitos matemáticos relativos à variação de funções com base no comportamento de suas derivadas (BOULOS, 1974), integrado ao conceito de K-means. Dessa forma obtém-se automaticamente (de maneira não supervisionada), sem a necessidade da interferência do tomador de decisões, a clusterização do índice NDVI, baseado na análise da frequência do número de *pixels* em função dos dados de NDVI extraídos de imagens orbitais, em diferentes escalas espaciais e temporais.

Para esse experimento, foram utilizados três períodos distintos (1984, 2014 e 2021), para três sensores: TM (*Thematic Mapper*), a bordo do satélite Landsat-5; OLI (*Operational Land Imager*), a bordo do satélite Landsat-8; MSI (*Multispectral Imager*), a bordo do satélite Sentinel, level-2A. A principal contribuição desta metodologia é a determinação do número de clusters relativos ao índice NDVI de maneira automatizada (não supervisionado), para classificar os seus intervalos, obtido a partir da análise e calibração dos dados fornecidos pelas imagens.

A metodologia proposta neste estudo foi desenvolvida e validada adotando uma área situada no nordeste do estado de Mato Grosso do Sul (MS), onde apresenta mudanças da cobertura vegetal e uso do solo ao longo de um período de 37 anos, impulsionada pelo segmento industrial ligado à produção de celulose e de grãos.

## Área de estudo

A área de estudo encontra-se na região nordeste do estado de Mato Grosso do Sul, com uma área de 77.399 km<sup>2</sup>, inserida no domínio morfoclimático e fitogeográfico do bioma Cerrado (2 milhões km<sup>2</sup>), na divisa com os estados de São Paulo (SP), Goiás (GO) e Mato Grosso (MT) (Figura 1). Abrange a Unidade de Gerenciamento dos Recursos hídricos do rio Sucuriú e Verde.



agosto, havendo uma pluviosidade superior aos 100 mm mensais nos meses de outubro a março (SILVA; ASSAD; EVANGELISTA, 2008).

## Material e métodos

Para o desenvolvimento da pesquisa, foram utilizadas imagens do sensor TM a bordo do satélite Landsat-5, imagens do sensor OLI a bordo do satélite Landsat-8, coleção Nível 2, corrigidos atmosféricamente e gerados a partir do LASRC (Código de Refletância de Superfície Landsat-8), que produz refletância de superfície em uma resolução espacial de 30 m (SILVA *et al.*, 2021), adquiridas no site da USGS (<http://earthexplorer.usgs.gov/>). Consistiram-se nas cenas (órbita/ponto): 223/75; 223/74; 223/73; 224/75; 224/74; 224/73; 224/72; 225/74; 225/73; datadas de julho de 1984 e 2014. Também foram usadas imagens do sensor multiespectral MSI a bordo do satélite Sentinel, level-2A, datadas de julho de 2021, com resolução espacial de 10 m, adquiridas no site oficial do programa Copernicus, vinculado à Agência Espacial Europeia (ESA) (<https://scihub.copernicus.eu/dhus/#/home>). Consistiram-se nas cenas: 22KDE; 22KDD; 22KDC; 22KDB; 22KCE; 22KCD; 22KCC; 22KCB; 22KBF; 22KBE; 22KBD; 22KBC; 22KBB; 21KZA; 21KZV; 21KZU; 21KZT; 21KZS; 21KYV; 21KYU; 21KYT; 21KYS. Justifica-se o uso de imagens do mês de julho, por se a estação seca da região (inverno) e não apresentar nuvens.

Para validação a análise do índice NDVI em uma área de estudo teste e a decisão sobre a viabilidade da proposição de uma nova metodologia para sua classificação foram selecionadas 25 amostras, todas na mesma localização espacial para os três períodos e sensores; 5 amostras para cada cluster obtidos, tanto para o algoritmo proposto quanto para o JENKS. Para isso, foi adotado o mesmo número de clusters fornecido pelo algoritmo proposto neste estudo. Para auxiliar na análise das clusterizações, os valores do índice Kappa foram comparados aos limiares estabelecidos por Landis e Koch (1997) disponíveis na literatura para, assim, poder analisar os potenciais e as limitações do novo algoritmo.

O conjunto de informações e dados utilizados foram organizados em um banco de dados geográficos para tratamento, conversão, processamento e análise dos resultados implementados em um SIG.

## Índice da diferença normalizada de vegetação (NDVI)

O NDVI foi calculado para cada pixel utilizando as refletividades das bandas no vermelho e infravermelho próximo, conforme proposto por Rouse *et al.*, (1973). Para as imagens do sensor TM (1984), foram selecionadas as bandas 3 (0.63 - 0.69  $\mu\text{m}$ ) e 4 (0.76 - 0.90  $\mu\text{m}$ ) do satélite Landsat-5, para o sensor OLI (2014), as bandas 4 (654 nm) e 5 (0.845 - 0.885  $\mu\text{m}$ ) do satélite Landsat-8. Para as imagens MSI (2021), foram utilizadas as bandas 8 (0.83  $\mu\text{m}$ ) e 4 (0.665  $\mu\text{m}$ ) do satélite Sentinel, level-2A. O cálculo no NDVI é dado pela equação 1:

$$NDVI = (V - I_{vp}) / (V + I_{vp}) \quad (1)$$

Onde: **V** = valores de pixels da banda vermelha e **I<sub>vp</sub>** = valores de pixel da banda infravermelha próxima.

Os limiares obtidos pela aplicação do índice foram: -0.48 a 0.73 (TM); 0.99 a 0.99 (OLI); -1 - 0.7 (MSI). Isso evidencia que, apesar do índice ser normalizado, a distribuição espacial e temporal dos respectivos valores não são análogos, variam entre o limiar -1 a 1, e que uma classificação dos mesmos baseada na experiência do pesquisador pode levar a resultados personalizados, sendo está a premissa para desenvolver um algoritmo de classificação que considera o comportamento e a consistência das bases de dados utilizadas.

## Parametrização do índice NDVI

Os resultados obtidos pela aplicação do índice NDVI estão no formato *raster*, 32 bits, pontos flutuantes (*Floating points*). Para parametrização dos respectivos valores de NDVI, visando à obtenção da tabela de atributos, e, desse modo, poder analisar o comportamento espectral dos mesmos, por meio do *software* ArcMap (ESRI, 2015), cada cena foi reamostrada para o formato do tipo inteiro não assinado (*unsigned integer*), 8 bits, por meio da ferramenta “*Copy raster*”, seguido da aplicação de cálculo de estatística, por meio da ferramenta “*Calculate Statistics*”. Ao final desse procedimento, cada imagem NDVI foi transformada em uma escala de 1 a 255 níveis de cinza. Após o procedimento de tratamento e conversão dos dados, as respectivas tabelas

de atributos foram exportadas em formato .dbf, depois convertidas em formato .xlsx, para ser convertido para um arquivo “MAT” para leitura no software MatLab (MATrix LABoratory).

## **Proposta metodológica para o algoritmo de Clusterização do índice NDVI**

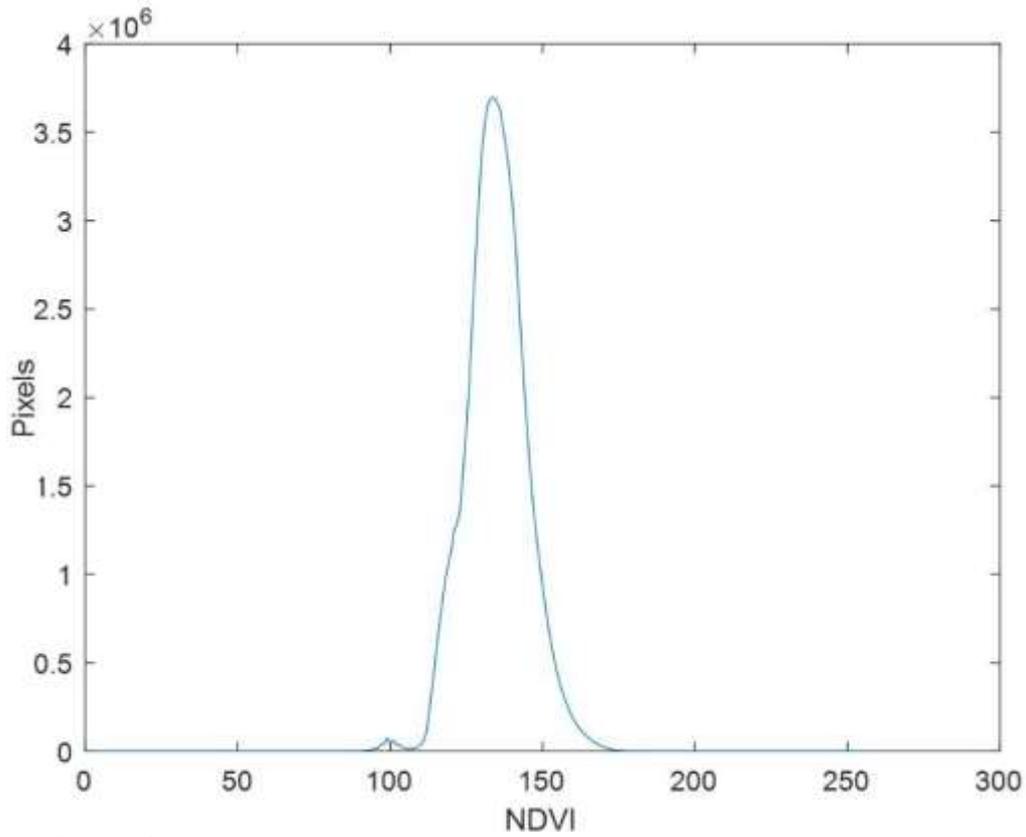
A proposta inicial para classificar os índices NDVI foi o uso do algoritmo de clusterização K-Means, mas o mesmo se mostrou inadequado devido à simetria dos dados práticos relativos ao NDVI vs. Número de *pixels*, conforme ilustrado na Figura 2. O algoritmo K-means original não tem a habilidade de diferenciar no espaço os dados de NDVI com a mesma escala, o qual considera em Clusters diferentes dados que deveriam pertencer a um mesmo Cluster. Para resolver este problema, neste trabalho é proposto um algoritmo K-means modificado que utiliza além da distância média, dados da função da distribuição de probabilidades acumulada.

### **Algoritmo de clusterização K-means modificado**

A análise foi elaborada com base na partição da área de estudos em unidades fisiográficas, considerando a forma e distribuição espacial das características das imagens multiespectrais através da interpretação visual das imagens para os diferentes sensores. Na Figura 2, ilustra-se o gráfico do NDVI vs. Número de *pixels* para o ano 1984.

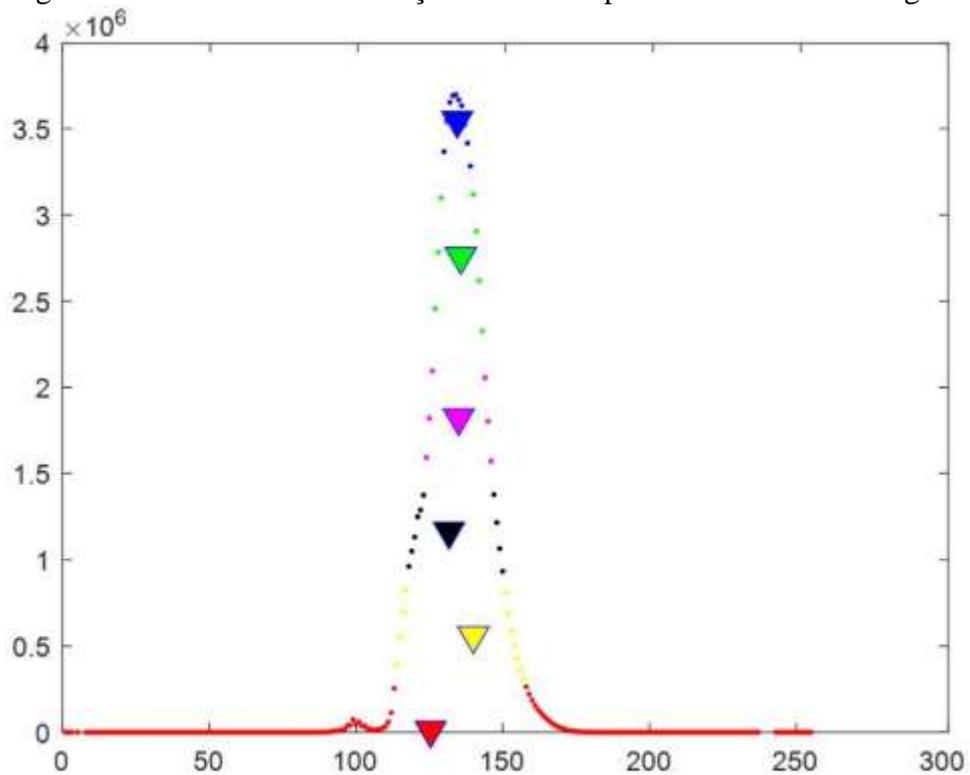
O uso do método de Clusterização K-means, diretamente nos dados originais, gera Clusters praticamente simétricos para ambos os lados da curva da Figura 2, como apresentado na Figura 3.

Figura 2: Gráfico NDVI vs. Número de *pixels* para uma área de estudos



Elaboração: Autores, 2023.

Figura 3: Técnica de Clusterização K-means aplicada aos dados da Figura 2



Elaboração: Autores, 2023.

Para contornar este problema, é necessária uma representação dos dados que elimine a sua simetria e permita o uso do K-means para agrupá-los de forma apropriada. Neste sentido, a função mais adequada para representar estes dados é a função densidade de probabilidade acumulada (*fda*), o que permitiria uma representação dos dados de forma assimétrica. O processo para determinar a *fda* dos dados é descrito a seguir.

### Cálculo da função de probabilidade acumulada (*fda*)

No problema sob análise, considerando que os dados representam a quantidade de *pixels* para cada um dos valores de NDVI, é possível determinar a probabilidade de cada um dos números de *pixels* dados através da equação (2):

$$Prob_p = \frac{NPI_p}{\sum_{j=1}^N (NDVI_j)} \quad p = 1, \dots, N \quad (2)$$

*Sendo:*

$Prob_p$  : Probabilidade de ocorrer o número de *pixels* *b*;

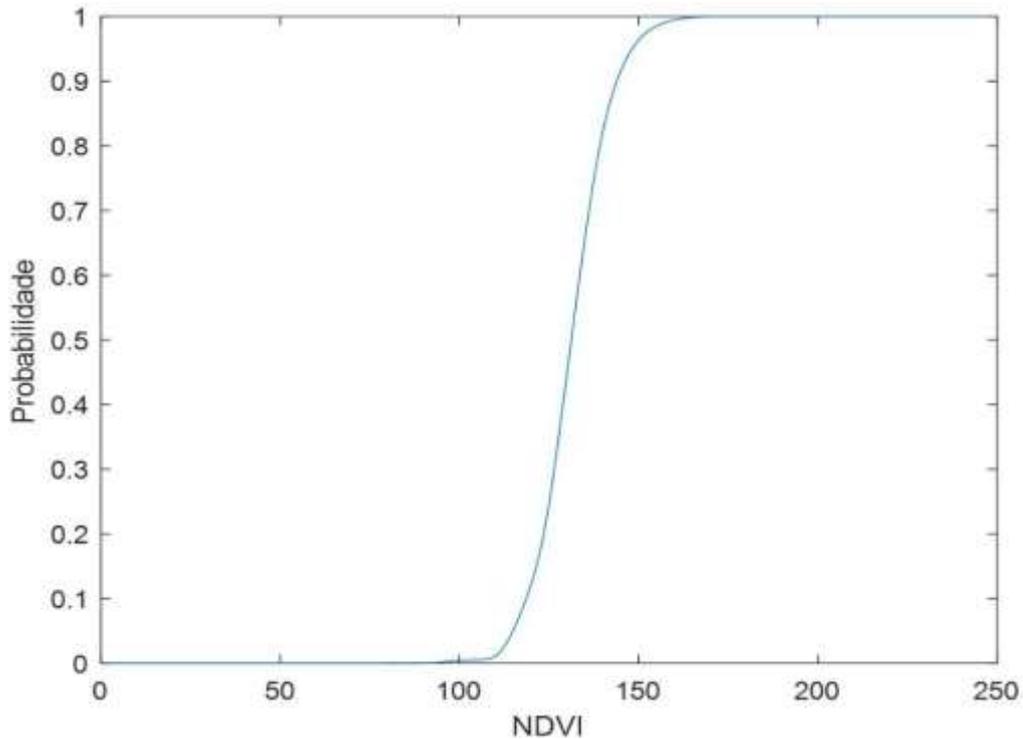
$NPI_p$ : Número de *pixels* do *p*-ésimo índice NDVI;

*p*: *p*-ésimo índice NDVI;

*N*: Número de amostras da base de dados;

Com as probabilidades de cada um dos valores de número de *pixels*,  $Prob_p$ , obtém-se a função densidade de probabilidade acumulada dos dados, que é representada graficamente na Figura 4, para o ano de 1984 do caso exemplo.

Figura 4: Função densidade de probabilidade acumulada (fda)

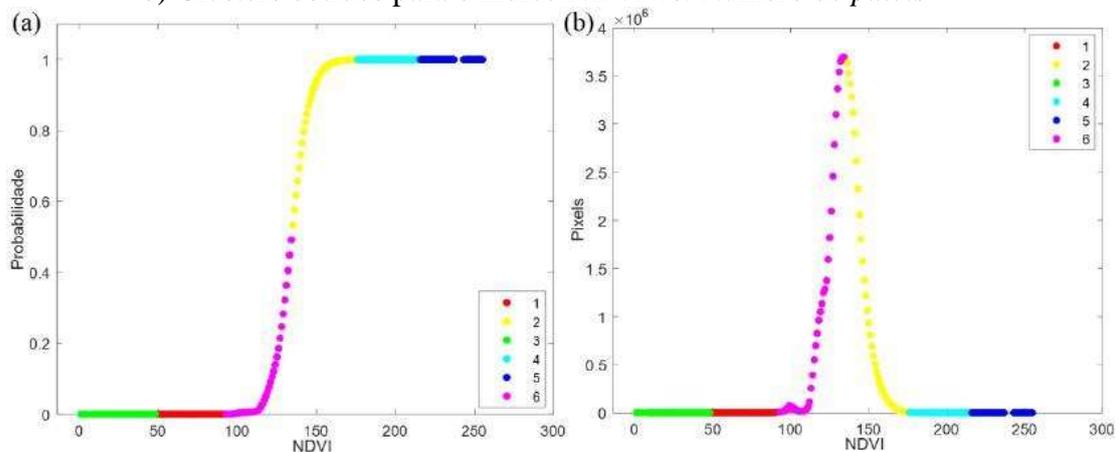


Elaboração: Autores, 2023.

### Classificação de dados a partir da fda

Aplicando o algoritmo k-means, considerando que existem 6 clusters, diretamente na *fda*, obtêm-se os resultados ilustrados na Figura 5. Verifica-se grande acumulação de dados nas fronteiras, à direita e à esquerda, do conjunto que é o domínio do gráfico da Figura 5 (a) (dados que aparecem nos gráficos de forma quase horizontais).

Figura 5: (a) Função dos índices NDVI vs. densidades de probabilidade acumuladas;  
b) Clusters obtidos para o índice NDVI vs. Número de *pixels*



Elaboração: Autores, 2023.

Este comportamento dos índices NDVI faz com que o algoritmo K-means gere clusters desnecessários entre (0-100) e (160-250) (verde, vermelho, cian e azul escuro); além de não fazer uma boa classificação para os dados que aparecem no interior do domínio da função. i.e, entre (100 - 160). Assim, estes dados acumulados nas fronteiras do conjunto do domínio da função (3.a) são classificados *a priori* como pertencentes a dois clusters, sem necessidade de usar a técnica K-means para classificá-los.

### **Classificação dos dados no domínio da função NDVI vs. Número de Pixels**

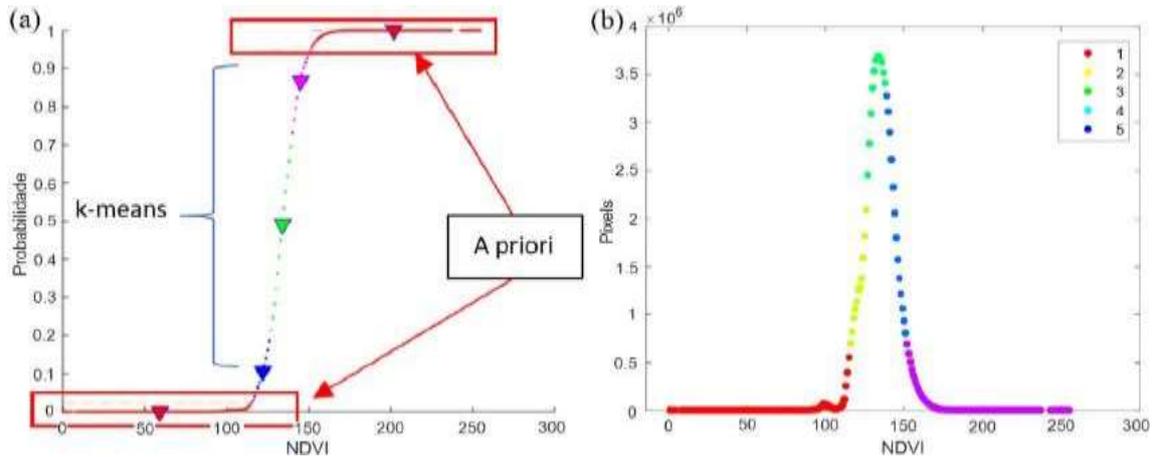
Para determinar quais dados devem pertencer aos clusters nos extremos do conjunto dos índices NDVI, determina-se a derivada dos dados da *fda*. Para os dados na fronteira do conjunto, verifica-se que eles apresentam pouca variação para um determinado intervalo de índices NDVI, porque esses dados apresentam um comportamento quase horizontal, e a derivada da função de probabilidade para este intervalo de valores NDVI é aproximadamente igual a zero. Desta forma, eles são classificados em 2 clusters conforme ressaltado anteriormente.

Analisando a função da esquerda para direita; quando a derivada começa a aumentar indica que a curva está apresentando um comportamento crescente indicando que um cluster pode ser gerado. Esta mesma análise da derivada é efetuada no sentido contrário para classificar os dados no sentido direita-esquerda da curva. Para os dados do ano 1984, é permitido um valor para a derivada inferior a 9,0% para ser considerado dentro da classificação *a priori* proposta, devido o comportamento simétrico e íngreme da *fda*.

### **Classificação dos dados acumulados no interior do domínio da função NDVI vs. Numero de Pixels**

Para os dados no interior do conjunto do domínio da função *fda*, aplica-se k-means para classificá-los. Para o exemplo, nos dados de 1984, aplica-se a técnica k-means para os NDVI entre os valores de 110 e 160, determinando-se inicialmente que existem 3 clusters nos dados sob análise, e o resultado é apresentado na Figura 6.

Figura 6: (a) Ilustra a determinação dos clusters nas fronteiras da função de distribuição de probabilidade; (b) Função NDVI vs. Número de pixels para o caso em estudo



Elaboração: Autores, 2023.

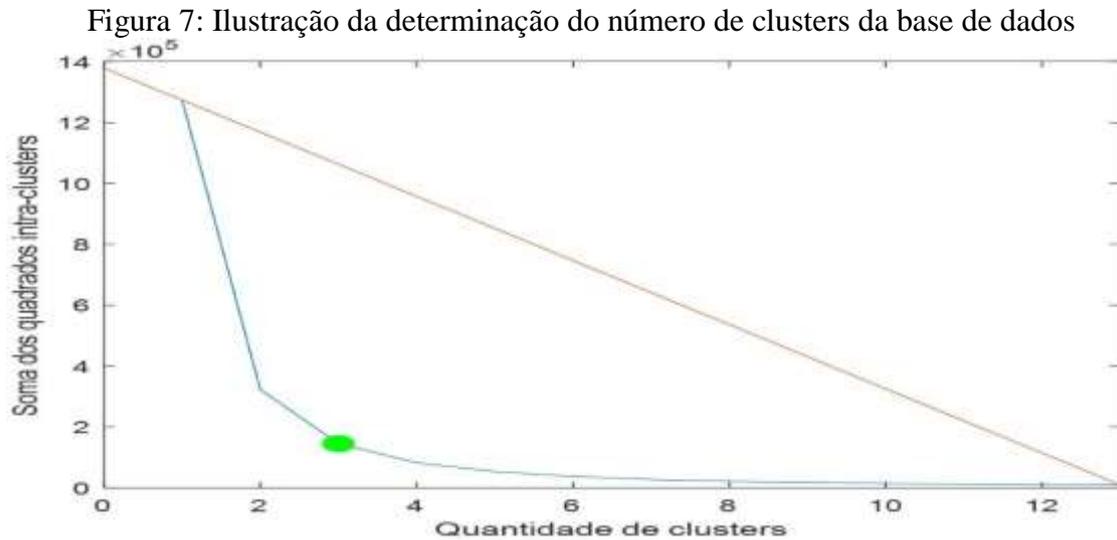
### Determinação do número de clusters

Para determinar um número adequado de clusters para o K-means, utiliza-se um método referenciado na literatura como “método do cotovelo”, o qual é pautado na inspeção visual do comportamento dos dados, baseado no conceito de que quantos mais clusters são determinados na base de dados sob análise, menos homogêneos eles são, porque a distância média entre os dados que fazem parte de cada grupo diminui. Desta forma, deve-se obter o equilíbrio entre a quantidade e qualidade dos clusters (LOPEZ-RUBIO *et al.*, 2018; BHOLOWALIA e KUMAR, 2014; KODINARIYA e MAKWANA, 2013; LIU *et al.*, 2018).

No método K-means, determina-se a distância das observações até o centroide que ela pertence, e o ideal é minimizar esta distância. Desta forma, aplica-se o K-means para vários valores de k (número de clusters) e calcula-se a soma dos quadrados das distâncias intra-cluster, ou seja, calcula-se a soma dos quadrados das distâncias entre as observações e os seus respectivos centroides.

Como exemplo, para os dados do ano de 1984, aplica-se o método considerando que possam ser definidos de 1 até 13 clusters. No gráfico da Figura 7 (curva azul), ilustra-se que quantos mais cluster são considerados, a variação da soma dos quadrados intra-cluster diminui. Isto indica que as observações que estão sendo agrupadas em cada cluster perdem homogeneidade enquanto os centroides estão ficando cada vez mais

homogêneos. Neste sentido, o equilíbrio entre o número de centroides e a função soma dos quadrados intra-cluster está no ponto mais afastado à reta (reta vermelha) que une os extremos do gráfico da função intra-cluster, neste caso, 3 clusters indicados pelo ponto verde no gráfico da Figura 7. Finalmente, para classificar todos os dados do ano 1984, são necessários 5 clusters.



Elaboração: Autores, 2023.

### Algoritmo proposto

- i. Leitura da base de dados;
- ii. Calcular a função  $fda$ ;
- iii. Determinar os clusters nas fronteiras do conjunto considerando o limiar mínimo estabelecido para o valor da primeira derivada;
- iv. Determinação do número de clusters no interior do domínio da função  $fda$ :
  - Especificar um número inicial de cluster;
  - Calcular as distâncias intra-clusters para o número de clusters especificado. Construir a função quantidade de clusters pela soma ao quadrado das distâncias intra-clusters,  $f(x)$ ;
  - Determinar a equação da reta,  $g(x)$ , que une os pontos entre o menor número de clusters e o maior valor do número de clusters da função  $f(x)$ .
  - Determinar a distância máxima entre as funções  $f(x)$  e  $g(x)$ . O ponto onde esta distância é máxima corresponde ao número de clusters  $k$  adequado para o interior do domínio da função  $fda$ , e conseqüentemente para a base

de dados sob análise.

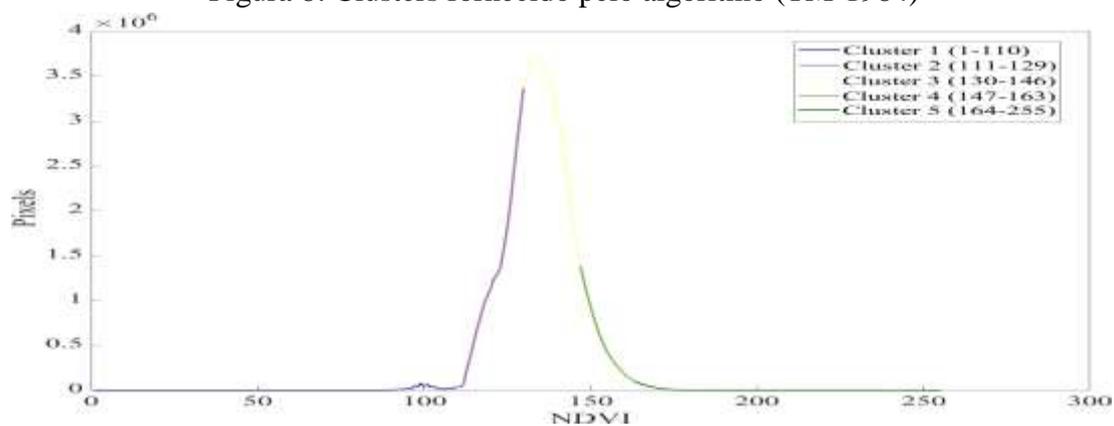
- v. Usar o método K-means nos dados no interior do domínio da função *fda*, considerando os *k* clusters determinados no passo *iv*.

O algoritmo K-means modificado e implementado no software MatLab foi aplicado para análise e classificação dos índices NDVI considerando o horizonte de tempo predeterminado para a área de estudos.

## Resultados

Baseado nos resultados obtidos e apresentados no gráfico da Figura 8, para o sensor TM (1984), verifica-se que a metodologia de clusterização proposta delimitou o índice NDVI em 5 clusters, compreendendo os seguintes intervalos: cluster 1: 1-110; cluster 2: 111-129; cluster 3: 130-146; cluster 4: 147-163; cluster 5: 164-255.

Figura 8: Clusters fornecido pelo algoritmo (TM-1984)



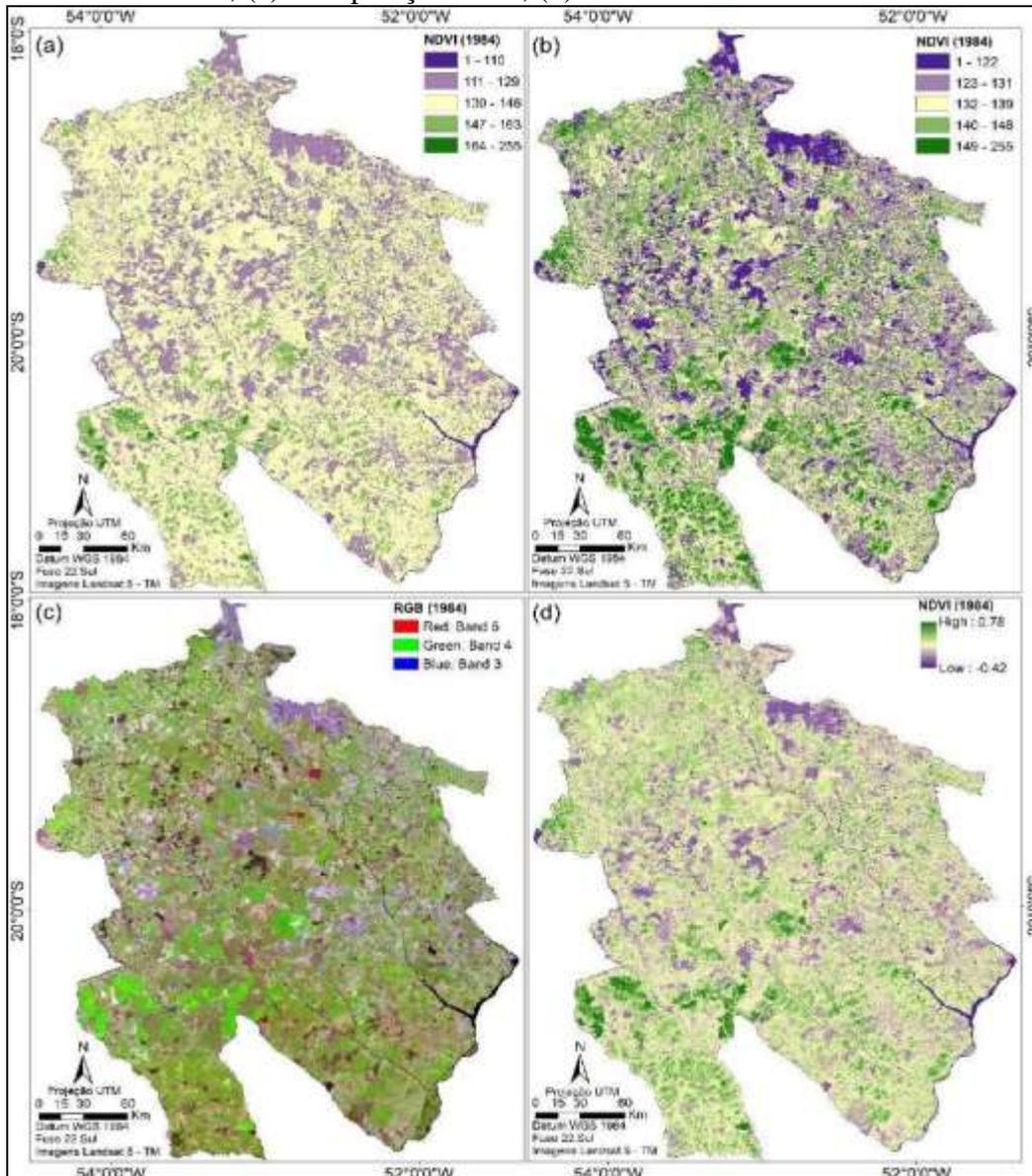
Elaboração: Autores, 2023.

A Figura 9 (a) e (b) apresentam resultados fornecidos tanto para o algoritmo proposto quanto para o método de JENKS, a Figura 9 (c) a composição falsa cor e a Figura 9 (d) o índice NDVI normalizado (sem a parametrização). Verificou-se que, no cluster 1 gerado pelo algoritmo (valores mais baixos de NDVI), foram mapeadas superfícies líquidas (rio Sucuriu e rio Paraná), já no JENKS, o cluster 1 generalizou e agrupou superfícies líquidas ao solo exposto, áreas queimadas e urbanizadas. No algoritmo proposto, essas áreas consideradas de solo exposto, áreas queimadas e

urbanizadas foram agrupadas no cluster 2, ao contrário do JENKS, onde o cluster 2 não diferenciou essas coberturas.

Ressalta-se assim uma vantagem no algoritmo proposto em relação ao JENKS, uma vez que consegue distinguir superfícies líquidas de outras formações, principalmente nos clusters 1 e 2, sendo o cluster 3 o mais generalizado do algoritmo proposto. Por outro lado, no JENKS mapeia com melhor precisão o cluster 4, formado por florestas de baixo porte, e o cluster 5, formadas por extratos vegetais mais densos (floresta plantada) inferidas como áreas de silvicultura.

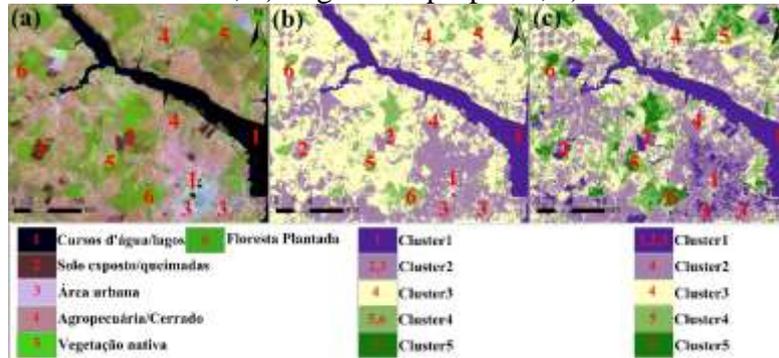
Figura 9: (a) Clusters fornecidos pelo algoritmo proposto; (b) Clusters fornecidos pelo JENKS; (c) Composição RGB; (d) Índice NDVI normalizado



Elaboração: Autores, 2023.

Para o algoritmo proposto, essas áreas estariam inseridas no cluster 3, predominantemente formadas por formação Campestre, Savanas ou Campo Cerrado e pastagens. Já para o algoritmo proposto, os clusters 4 e 5 são praticamente os mesmos, formados por florestas de baixo porte, não havendo essa distinção claramente entre floresta e áreas plantadas (Figura 10).

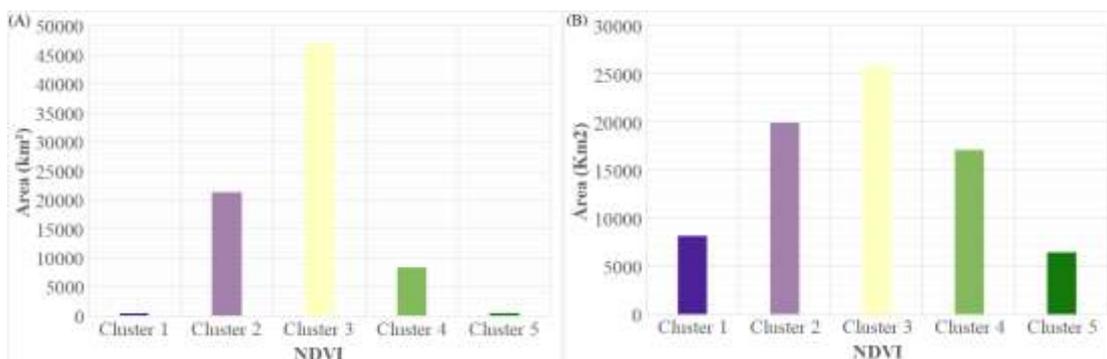
Figura 10: Chave de interpretação dos alvos mapeados; a) Imagem RGB – TM 1984; b) Algoritmo proposto; c) JENKS



Elaboração: Autores, 2023.

Assim, verifica-se, uma vantagem do algoritmo proposto para os clusters 1 (superfícies líquidas apenas) e 2 (áreas urbanas, queimadas e solo exposto) e 3 (formação Campestre, Savanas ou Campo Cerrado e pastagens, sobretudo por esta região nesse período ser mais ligada ao setor agropecuário fomentado pela criação de gado. Conforme o gráfico da Figura 11, a proporção dos clusters mostra exatamente a distribuição (km<sup>2</sup>) dos respectivos clusters, ressaltando, cabe salientar que o JENKS leva vantagem na classificação dos clusters 4 e 5, pois estava em ascensão visado pela plantação de árvores (setor da silvicultura) situadas a sul do oeste da área de estudos.

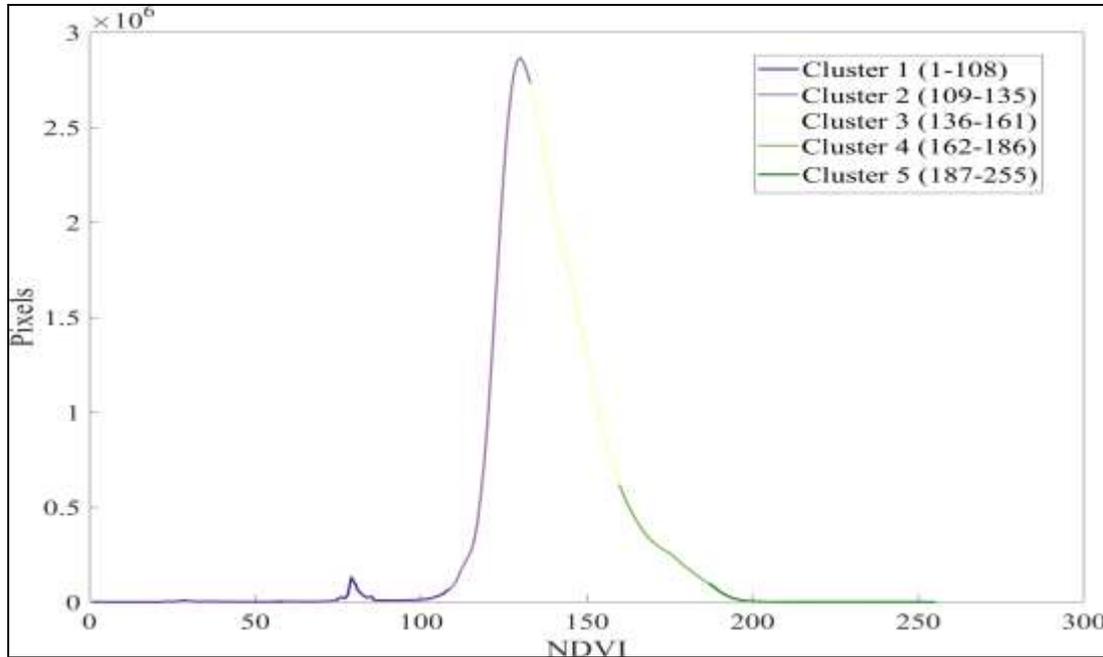
Figura 11: Distribuição dos clusters; A) Algoritmo proposto; B) JENKS



Elaboração: Autores, 2023.

No gráfico da Figura 12, apresentam-se os resultados obtidos pela clusterização, para o sensor OLI (2014). O algoritmo proposto delimitou o índice NDVI em 5 Clusters, compreendendo os seguintes intervalos: Cluster 1: 1-108; Cluster 2: 109-135; Cluster 3: 136-161; Cluster 4: 162-186, Cluster 5: 187-255.

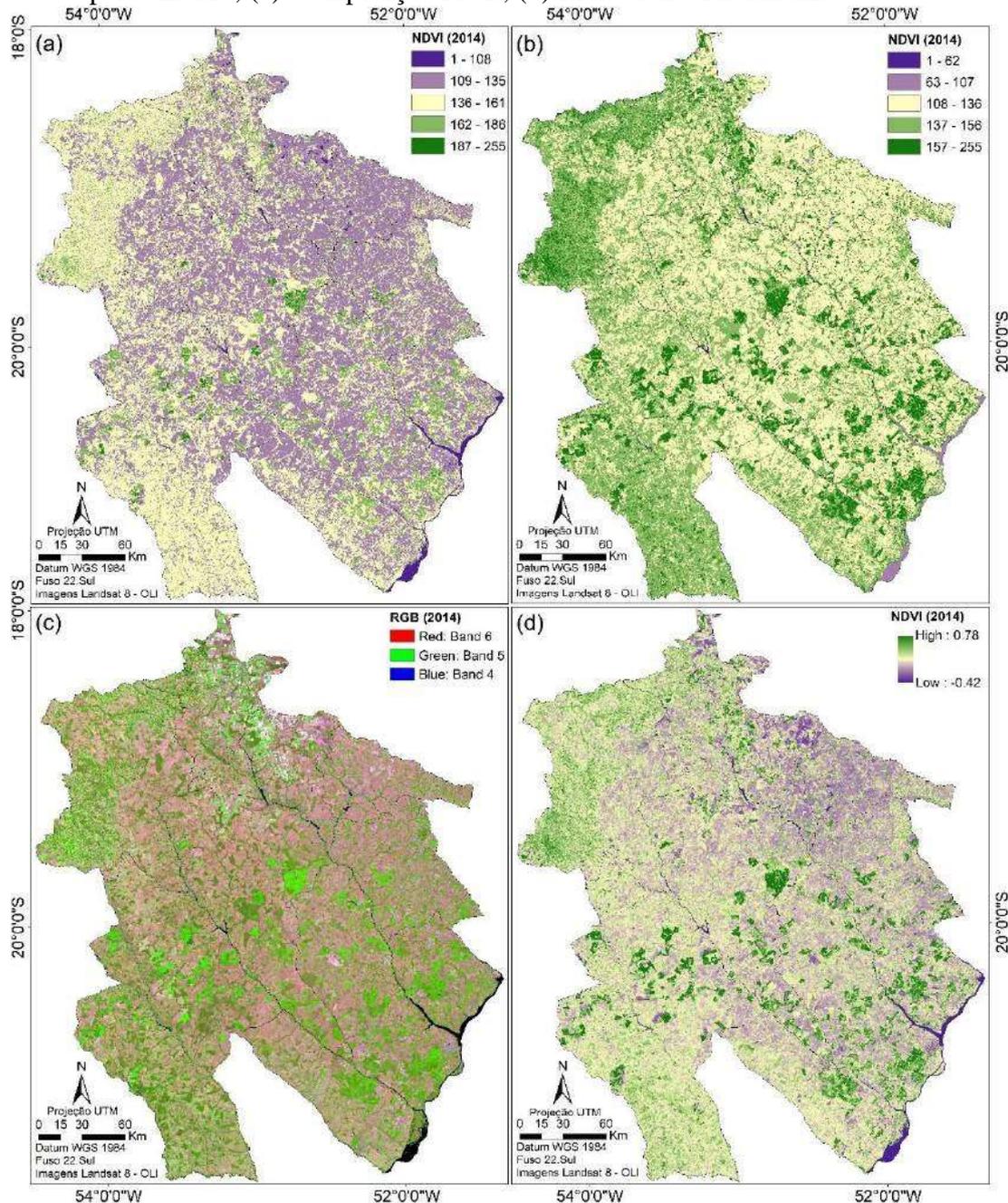
Figura 12: Clusters fornecido pelo algoritmo (OLI-2014)



Elaboração: Autores, 2023.

Para o algoritmo proposto, no cluster 1 foram mapeadas áreas com superfície líquida, além do rio Sucuriu e rio Paraná. Nesse período, já haviam construído a Usina Hidrelétrica de Porto Primavera-SP (inaugurada em 1999), mostrando áreas alagadas pelo reservatório, próximo ao município de Brasilândia-MS, situadas no sul da área de estudos. Por outro lado, o JENKS não classificou áreas no cluster 1, as superfícies líquidas estão postas no cluster 2. Porém, ao contrário do mapeamento de 1984, no cluster 2, o algoritmo proposto generalizou áreas urbanas, queimadas e solo exposto com as formações Campestres, Savanas ou Campo Cerrado e pastagens. No cluster 3, o algoritmo proposto delimitou a área oeste da área de estudos, misturando as classes do cluster 2 com áreas de vegetação de baixo porte, ao contrário do cluster 3 do JENKS, que mapeia generalizadamente áreas urbanas, queimadas e solo exposto com as formações Campestres, Savanas ou Campo Cerrado e pastagens (Figura 13).

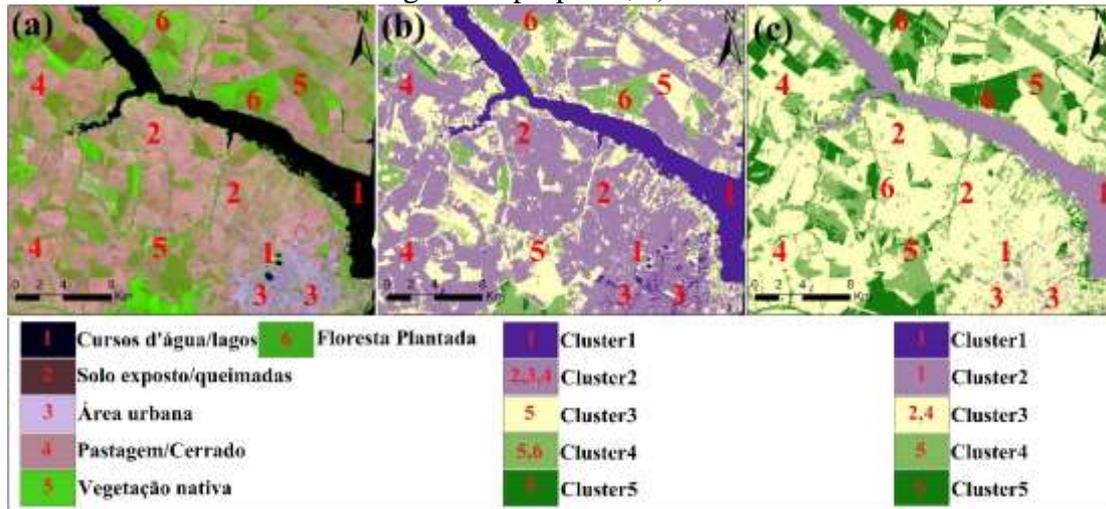
Figura 13: (a) Clusters fornecidos pelo algoritmo proposto; (b) Clusters fornecidos pelo JENKS; (c) Composição RGB; (d) Índice NDVI normalizado



Elaboração: Autores, 2023.

Para o cluster 4 do algoritmo proposto, ocorre o mesmo que no período anterior, pois o mapeamento de áreas de floresta de baixa densidade são confundidas com áreas de floresta densa (floresta plantada), prevalecendo as áreas de floresta baixa, ao contrário do JENKS, que mapeia separadamente essas áreas, principalmente as áreas de floresta plantada pela expansão da silvicultura (Figura 14).

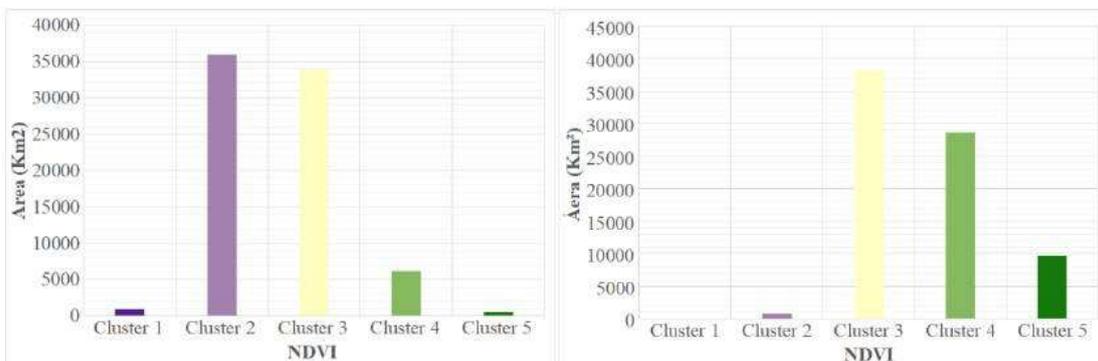
Figura 14: Chave de interpretação dos alvos mapeados; a) Imagem RGB – OLI 2014; b) Algoritmo proposto; c) JENKS



Elaboração: Autores, 2023.

Em síntese, os resultados são similares aos do período anterior, tendo em conta que o algoritmo proposto tem melhor vantagem nos cluster 1, 2 e 3, porém, o JENKS se sobressai nos clusters 4 e 5.

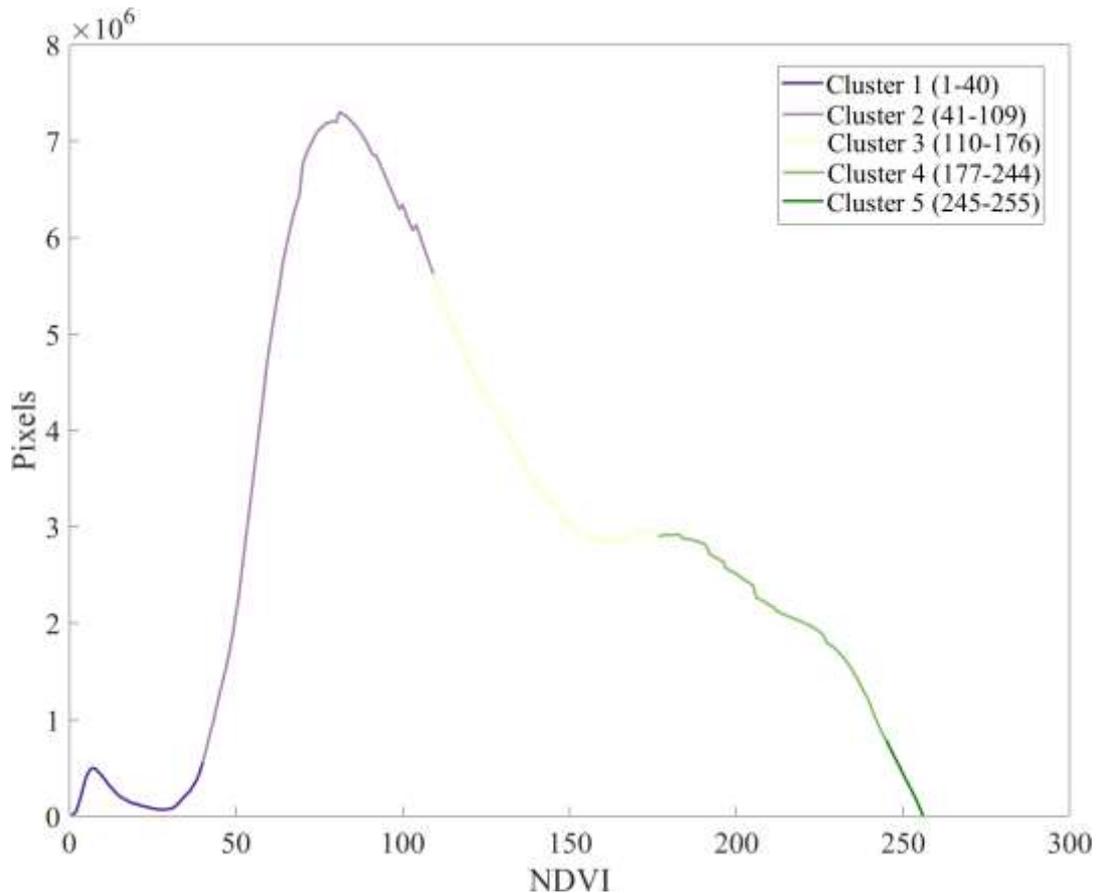
Figura 15: Distribuição dos clusters; A) Algoritmo proposto; B) JENKS



Elaboração: Autores, 2023.

No gráfico da Figura 16, apresentam-se os resultados obtidos pela clusterização, para o sensor MSI (2021). A metodologia delimitou o índice NDVI em 5 Clusters, compreendendo os seguintes intervalos: Cluster 1: 1-154; Cluster 2: 155-181; Cluster 3: 182-206; Cluster 4: 207-232, Cluster 5: 233-255.

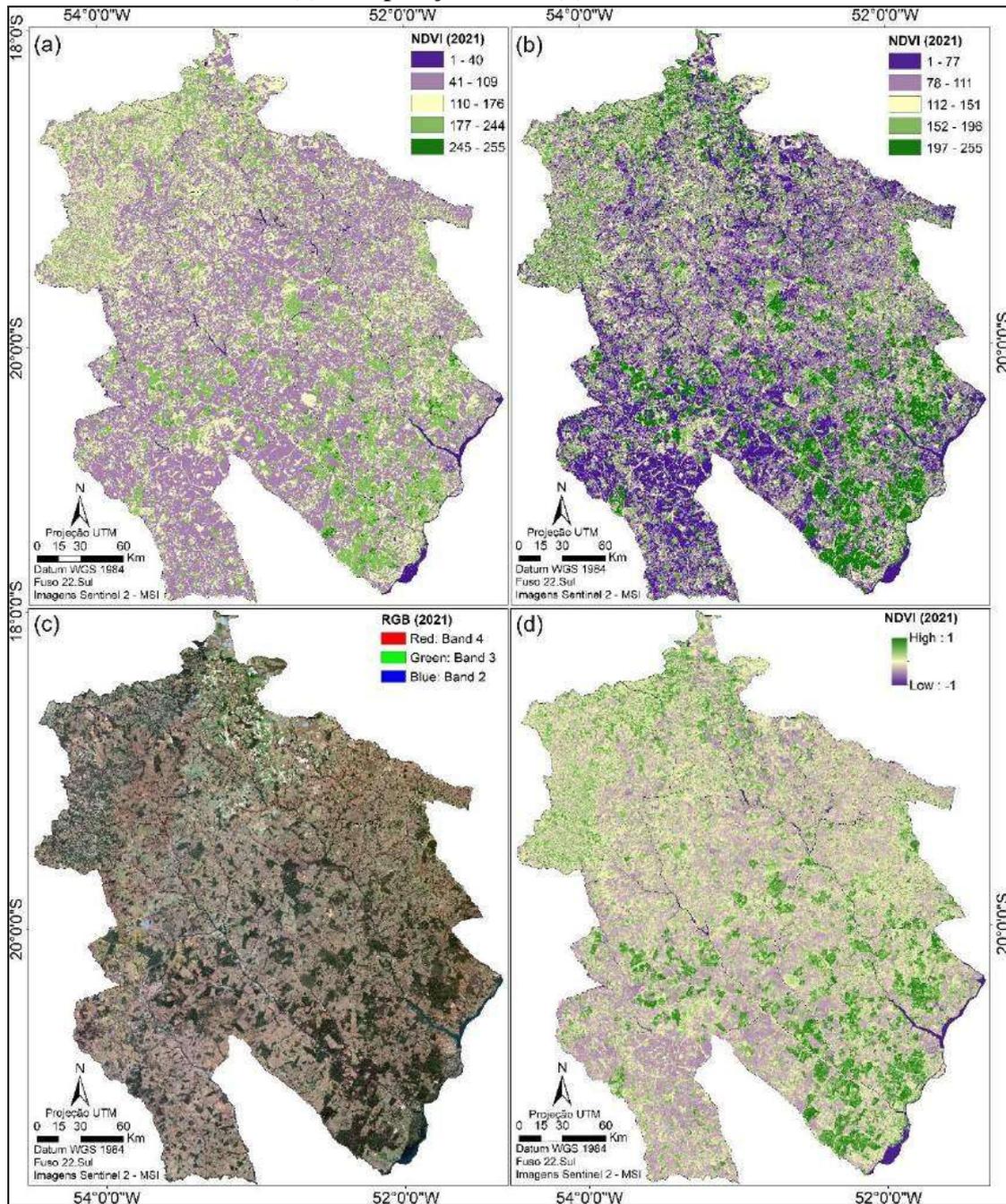
Figura 16: Clusters fornecido pelo algoritmo (MSI-2021)



Elaboração: Autores, 2023.

Para o último período aplicado, os clusters obtidos pelo algoritmo proposto mantiveram o mesmo aspecto mapeado; Cluster 1, presença de superfícies líquidas, no cluster 2 houve a mesma generalização de áreas urbanas, queimadas, solo exposto, porém agregou-se com as formações Campestres, Savanas ou Campo Cerrado e pastagens. Já no JENKS, houve uma generalização que não havia ocorrido nos períodos anteriores. O cluster 1 generalizou superfícies líquidas com áreas urbanas, queimadas, solo exposto com as formações Campestres, Savanas ou Campo Cerrado com vegetação de baixo porte, o qual para os períodos anteriores estava alocada sempre no cluster 4 (Figura 17).

Figura 17: (a) Clusters fornecidos pelo algoritmo proposto; (b) Clusters fornecidos pelo Jenks; (c) Composição RGB; (d) Índice NDVI normalizado

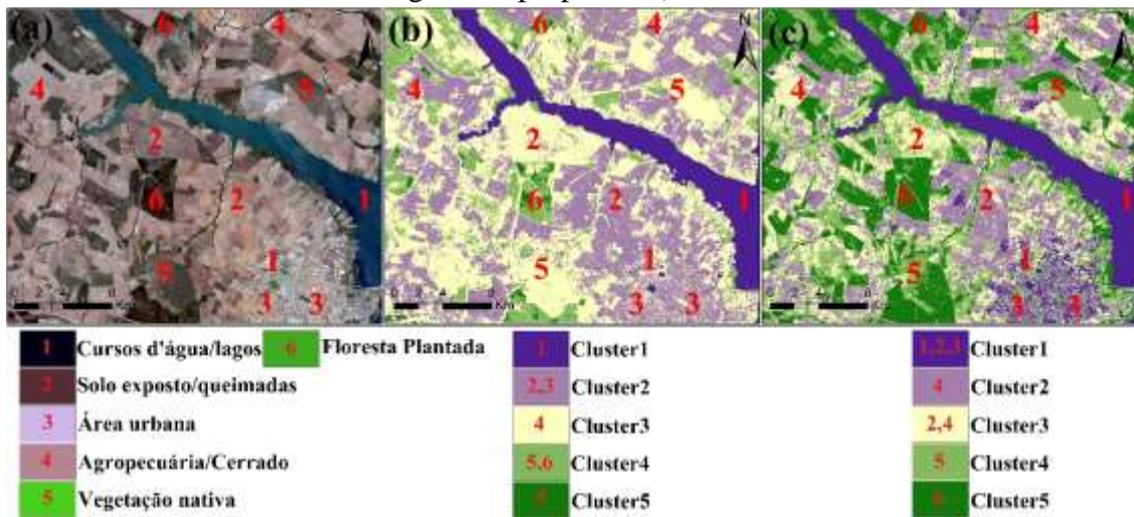


Elaboração: Autores, 2023.

Quanto ao cluster 2, nota-se que as pastagens foram predominantes. No cluster 3, o algoritmo proposto mapeou áreas proeminentes as formações de cerrado e pastagem, comparando com o índice bruto (sem parametrização) e a imagem falsa cor (RGB), especialmente correspondentes. Por outro lado, não se pode fazer tal afirmação

com relação ao cluster 3 gerado pelo JENKS, onde a distribuição está aleatoriamente distribuída por toda área de estudos. No cluster 4 do algoritmo proposto, há uma generalização das florestas de baixo porte e vegetação densa, não havendo distinção com o cluster 5. Para o JENKS, mostra-se um comportamento contrário, seguindo o mesmo padrão de classificação dos períodos anteriores, sendo delineadas com melhor precisão essa diferença entre o cluster 4 e o cluster 5, nos quais são expressivas as áreas com vegetação de alta densidade, ainda mais intensificadas pela expansão da silvicultura nessa região do estado de Mato Grosso do Sul (Figura 18).

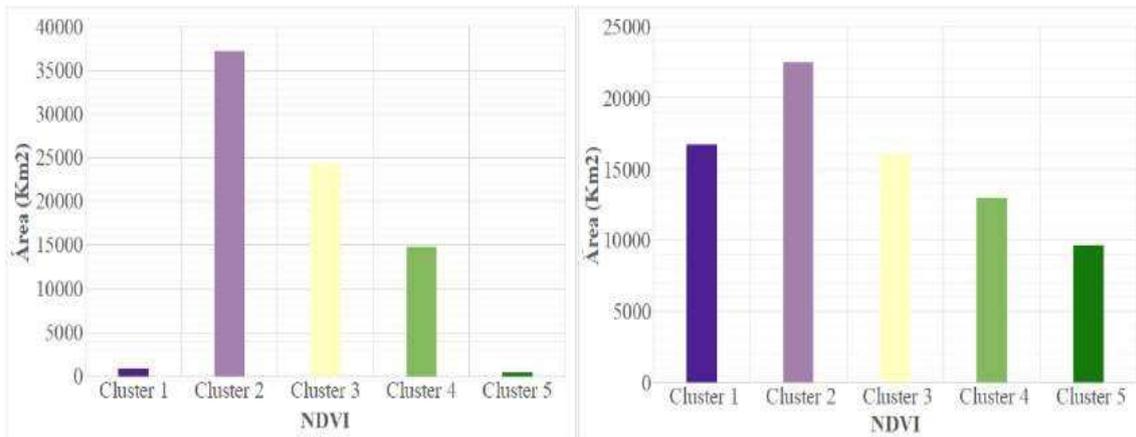
Figura 18: Chave de interpretação dos alvos mapeados; a) Imagem RGB – OLI 2021; b) Algoritmo proposto; c) JENKS



Elaboração: Autores, 2023.

Para o JENKS, mostra-se um comportamento contrário, seguindo o mesmo padrão de classificação dos períodos anteriores, sendo delineadas com melhor precisão essa diferença entre o cluster 4 e o cluster 5, nos quais são expressivas as áreas com vegetação de alta densidade, ainda mais intensificadas pela expansão da silvicultura nessa região do estado de Mato Grosso do Sul (Figura 19).

Figura 19: Distribuição dos clusters; A) Algoritmo proposto; B) JENKS

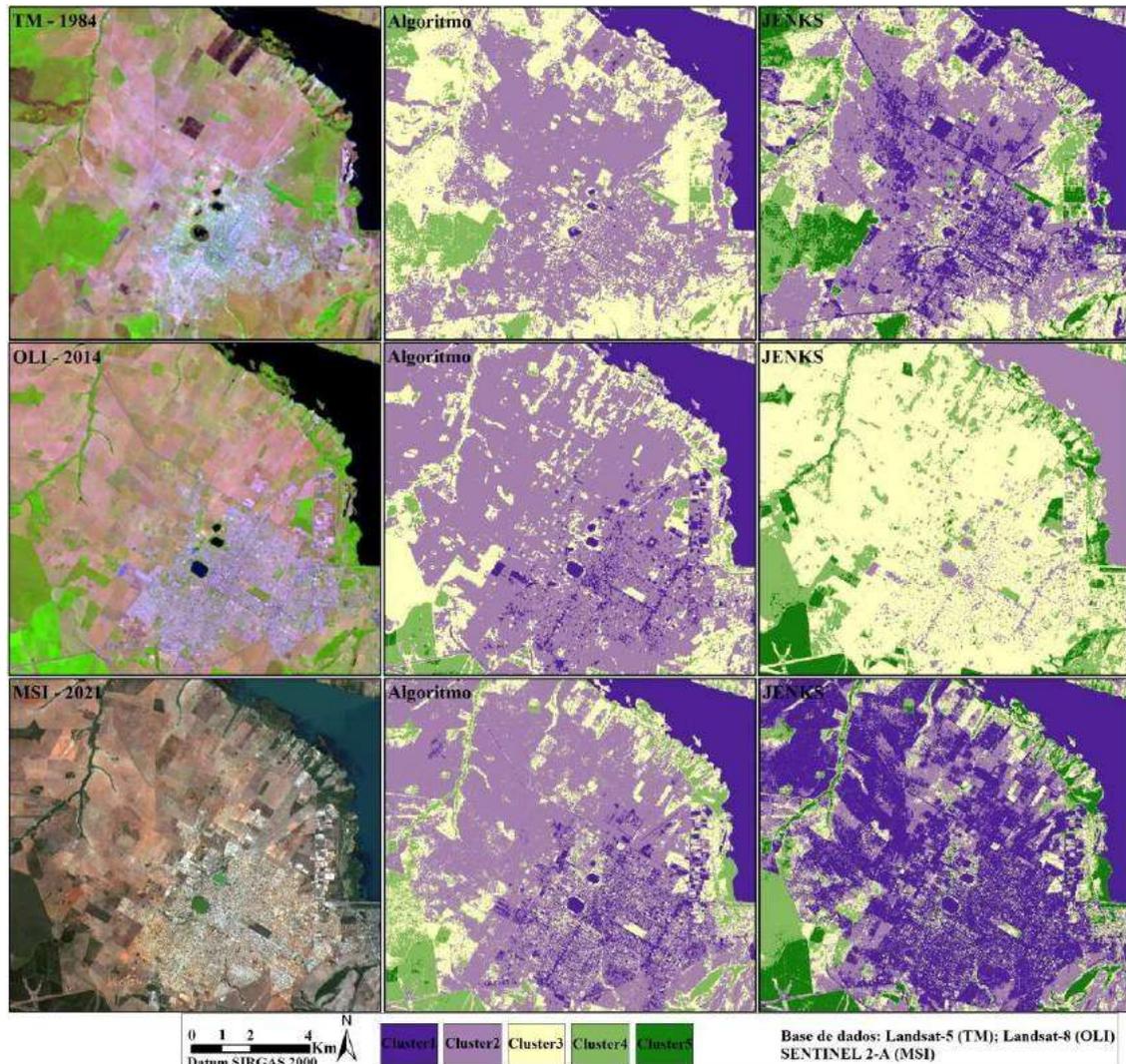


Elaboração: Autores, 2023.

Na Figura 20, apresenta-se a área amostral, situada no sul da área de estudos, principalmente onde há a maior concentração de áreas de silvicultura, determinada para verificação da correspondência entre a clusterização obtida para o algoritmo proposto e para o JENKS. Baseado nas amostras adquiridas para a validação estatística, na tabela 1 apresenta-se os resultados do índice Kappa.

Verifica-se que o algoritmo proposto distingue com melhor concordância espacial para os três períodos o comportamento espectral dos pixels que seriam referentes às classes de água do solo exposto e áreas urbanas, agrupadas no cluster 1, sendo nítida essa questão no que concerne a distinção da área urbana do município de Três Lagoas; já o JENKS generaliza essas classes; por outro lado, o JENKS distingue com melhor precisão as áreas de vegetação de baixo porte, vegetação natural e florestas plantadas (silvicultura), contudo, o algoritmo proposto generaliza essas classes especificamente no cluster 4, considerado aqui vegetação de baixo porte.

Figura 20: Comparação do índice NDVI; algoritmo proposto e Jenks



Elaboração: Autores, 2023.

A acurácia das classificações provenientes das imagens NDVI de diferentes datas e sensores distribuídas ao longo do ano e das imagens NDVI foi identificada através da matriz de confusão por meio do índice de Kappa (Figura 21). Com base na matriz de confusão, a avaliação dos resultados se consistiram: sensor TM/1984, o algoritmo proposto apresentou concordância “muito boa” (0.6), para o Jenks concordância “razoável” (0.40); sensor OLI/2014, para o algoritmo proposto concordância “razoável” (0.40), para o Jenks concordância “razoável” (0.39); sensor MSI/2021, algoritmo proposto concordância “boa” (0.5), Jenks concordância “boa” (0,5).

Figura 21: Matriz de confusão; a) algoritmo (1984/TM); b) jenks (1984/TM); c) algoritmo (2014/OLI); d) jenks (2014/OLI); e) algoritmo (2021/MSI); f) jenks (2021/MSI)

<b>(a)</b>									<b>(b)</b>								
ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa	ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	5.00	0.00	0.00	0.00	0.00	5.00	1.00	0.00	C_1	2.00	0.00	3.00	0.00	0.00	5.00	0.40	0.00
C_2	0.00	2.00	2.00	1.00	0.00	5.00	0.40	0.00	C_2	0.00	0.00	4.00	1.00	0.00	5.00	0.00	0.00
C_3	0.00	0.00	5.00	0.00	0.00	5.00	1.00	0.00	C_3	0.00	0.00	5.00	0.00	0.00	5.00	1.00	0.00
C_4	0.00	0.00	1.00	0.00	4.00	5.00	0.00	0.00	C_4	0.00	0.00	1.00	1.00	3.00	5.00	0.20	0.00
C_5	0.00	0.00	0.00	0.00	5.00	5.00	1.00	0.00	C_5	0.00	0.00	0.00	0.00	5.00	5.00	1.00	0.00
Total	5.00	2.00	8.00	1.00	9.00	25.00	0.00	0.00	Total	2.00	0.00	13.00	2.00	8.00	25.00	0.00	0.00
P_Accuracy	1.00	1.00	0.63	0.00	0.56	0.00	0.68	0.00	P_Accuracy	1.00	0.00	0.38	0.50	0.63	0.00	0.52	0.00
Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40

<b>(c)</b>									<b>(d)</b>								
ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa	ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	5.00	0.00	0.00	0.00	0.00	5.00	1.00	0.00	C_1	1.00	1.00	1.00	0.00	2.00	5.00	0.20	0.00
C_2	1.00	2.00	2.00	0.00	0.00	5.00	0.40	0.00	C_2	1.00	0.00	4.00	0.00	0.00	5.00	0.00	0.00
C_3	0.00	0.00	4.00	0.00	1.00	5.00	0.80	0.00	C_3	0.00	0.00	3.00	0.00	2.00	5.00	0.60	0.00
C_4	0.00	0.00	4.00	0.00	1.00	5.00	0.00	0.00	C_4	0.00	0.00	0.00	5.00	1.00	6.00	0.83	0.00
C_5	0.00	0.00	3.00	0.00	2.00	5.00	0.40	0.00	C_5	0.00	0.00	1.00	0.00	5.00	6.00	0.83	0.00
Total	6.00	2.00	13.00	0.00	4.00	25.00	0.00	0.00	Total	2.00	1.00	9.00	5.00	10.00	27.00	0.00	0.00
P_Accuracy	0.83	1.00	0.31	0.00	0.50	0.00	0.52	0.00	P_Accuracy	0.50	0.00	0.33	1.00	0.50	0.00	0.52	0.00
Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39

<b>(e)</b>									<b>(f)</b>								
ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa	ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	5.00	0.00	0.00	0.00	0.00	5.00	1.00	0.00	C_1	5.00	0.00	0.00	0.00	0.00	5.00	1.00	0.00
C_2	0.00	5.00	0.00	0.00	0.00	5.00	1.00	0.00	C_2	0.00	5.00	0.00	0.00	0.00	5.00	1.00	0.00
C_3	0.00	0.00	0.00	0.00	5.00	5.00	0.00	0.00	C_3	0.00	0.00	0.00	0.00	5.00	5.00	0.00	0.00
C_4	0.00	0.00	0.00	0.00	5.00	5.00	0.00	0.00	C_4	0.00	0.00	0.00	0.00	5.00	5.00	0.00	0.00
C_5	0.00	0.00	0.00	0.00	5.00	5.00	1.00	0.00	C_5	0.00	0.00	0.00	0.00	5.00	5.00	1.00	0.00
Total	5.00	5.00	0.00	0.00	15.00	25.00	0.00	0.00	Total	5.00	5.00	0.00	0.00	15.00	25.00	0.00	0.00
P_Accuracy	1.00	1.00	0.00	0.00	0.33	0.00	0.60	0.00	P_Accuracy	1.00	1.00	0.00	0.00	0.33	0.00	0.60	0.00
Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50

Elaboração: Autores, 2023.

## Discussão

Em síntese, nota-se, através das análises dos mapas, gráficos e figuras, bem como o resultado do índice Kappa, que os clusters obtidos pela metodologia proposta comparada com o JENKS, correlacionadas com as classes de uso e cobertura da terra, que as maiores mudanças ocorridas no período 1985 a 2017 foram a redução da vegetação campestre e pastagem e o aumento da área agrícola. No período 2017 a 2021, as mudanças ocorreram em toda a área analisada. Esse aumento ocorreu, sobretudo, em municípios localizados mais ao norte da área estudada, como Chapadão do Sul e Costa Rica, que registraram, nas últimas décadas a expansão de monoculturas voltadas ao mercado externo, precipuamente a soja e o milho.

Em 2019, de acordo com dados da Produção Agrícola Municipal (PAM), levantamento realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), Chapadão do Sul possuía 90 mil hectares ocupados por soja e 34 mil hectares por milho,

enquanto Costa Rica contava com 75 mil hectares ocupados por soja e 43 mil hectares por milho. Conseqüentemente, houve uma redução da importância econômica da pecuária, atividade tradicional na região (KUDLAVICZ, 2011). Entre os anos de 2005 e 2019, por exemplo, segundo dados da Pesquisa da Pecuária Municipal (PPM), do IBGE, o efetivo bovino total dos municípios situados na área analisada apresentou decréscimo de 26,4%, passando de 5.491.113 cabeças, em 2005, para 4.039.253 cabeças, em 2019.

O cluster 4 e 5, por sua vez, evidenciou o aumento de áreas verdes em decorrência da expansão do monocultivo de eucalipto, principal matéria-prima utilizada para a produção de celulose em Mato Grosso do Sul, que está entre os principais estados produtores de eucalipto e de celulose em razão da instalação de duas plantas industriais fabricantes de celulose, sendo uma de propriedade da Suzano e a outra da Eldorado Brasil, as quais, juntas, possuem capacidade para produzir quase 5 milhões de toneladas de celulose por ano (SUZANO, 2020; ELDORADO BRASIL, 2020).

A expansão dos eucaliptais aconteceu, principalmente, em Três Lagoas, onde se localizam as duas plantas industriais fabricantes de celulose, e em Ribas do Rio Pardo, Brasilândia, Água Clara, Selvíria e Inocência, municípios próximos. Conforme dados da Produção da Extração Vegetal e da Silvicultura (PEVS), do IBGE, a área estudada possuía, em 2019, 885.039 hectares com plantio de eucalipto. Desse total, 98,9% estava concentrado nos municípios citados.

### **Considerações finais**

Neste trabalho, foi proposto um algoritmo para determinar e classificar os usos do solo e a cobertura vegetal, a partir do uso do índice NDVI, usando imagens multiespectrais para diferentes períodos e sensores, sem a necessidade da interferência do tomador de decisões na determinação do número de clusters.

As imagens temporais dos diferentes sensores multiespectrais mapeados pelo algoritmo proposto foram correlacionadas com o algoritmo JENKS. Os clusters mapeados apresentaram boa correlação com as imagens utilizadas, indicando que o algoritmo proposto pode ser uma alternativa para um mapeamento e/ou classificação de imagens, para mapeamentos de uso do solo e cobertura da terra, de maneira rápida e

eficaz, uma vez que não prescinde da experiência do tomador de decisões (não supervisionado), ao contrário de métodos interpretativos como no processo de classificação de imagens (supervisionado). Outro aspecto a ser ressaltado diz respeito ao uso de diferentes sensores multiespectrais para diferentes períodos, mostrando que não há divergência e limitação do algoritmo proposto.

Ao se comparar espacialmente os clusters com as imagens falsa cor (RGB), houve uma correspondência entre ambas, como aumento de florestas plantadas pela expansão da silvicultura. Quanto às divergências encontradas entre os mapeamentos obtidos pelo método proposto de clusterização e os mapas classificados pelo algoritmo JENKS, podem ser atribuídas fundamentalmente à incompatibilidade entre os parâmetros de clusterização dos algoritmos.

Juntamente com estas dificuldades, há uma grande diversidade de bases cartográficas, com diferentes escalas e realizadas segundo diferentes critérios de mapeamento, que não coincidem cartograficamente entre elas, o que inviabiliza a comparação com outras bases cartográficas existentes na literatura, como, por exemplo, mapas do MapBiomas. Isso dificulta ainda mais para o pesquisador desenvolver com precisão metodologias para análise e planejamento tanto ao nível científico como para orientar políticas públicas de uso e ocupação do solo. Outro aspecto a ressaltar se refere ao número de clusters fornecido pelo algoritmo para simular o método de JENKS, pois como o mesmo se comportaria se tivesse que atribuir um número maior ou menor de clusters para simular o JENKS, como seriam os resultados fornecidos por ele? Ressalta-se a necessidade da realização um trabalho de campo para averiguação dessas inconsistências entre o algoritmo proposto e o JENKS.

Por fim, o algoritmo aplicado neste estudo será aplicado em outras áreas para experimentos, visando sua generalização em qualquer área geográfica, utilizando qualquer sensor multiespectral e período para, dessa forma, verificar sua estabilidade em outras áreas, não descartando acrescentar outros métodos no aperfeiçoamento do algoritmo proposto, como, por exemplo, lógica difusa ou lógica fuzzy.

## Referências

- ANYAMBA; A.; EASTMAN, J. R. Interannual variability of NDVI over Africa and its relation to El Niño/Southern Oscillation. **International Journal of Remote Sensing, London**, [S./l.], v.13, p.2533-2548, 1996.
- BHOLOWALIA, P. KUMAR, A. Ebk-means: A clustering technique based on elbow method and k-means in wsn. **International Journal of Computer Applications**, [S./l.], v. 105, n. 9, p. 17-24, 2014.
- BOULOS, P. **Introdução ao Cálculo – Cálculo diferencial**. São Paulo: Editora Edgard Blücher Ltda, 1974.
- CARDOZO, F. S.; PEREIRA, G.; SILVA; G. B. S.; SILVA, F. B.; SHIMABUKURO, Y. E.; MORAES, E. C. Discriminação de áreas alagadas no Pantanal sul-matogrossense a partir de imagens orbitais. In: Anais 2º Simpósio de Geotecnologias no Pantanal, 2009, Corumbá. **Anais [...]**, Corumbá: Embrapa Informática Agropecuária/INPE, 2009.
- DE ALBUQUERQUE, E. M. et al. Análise do comportamento do NDVI e NDWI sob diferentes intensidades pluviométricas no Município de Sousa-PB. **Revista Estudos Geoambientais**, [...], v. 1, n. 1, p. 1-11, 2014.
- EASTMAN, J.R.; FULK, M. Long sequence time series evaluation using standardized principal components. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.8, p.1307-1312, 1993.
- ELDORADO BRASIL. **Plano de Manejo Florestal**. 9. ed. 2020. Disponível em: <https://eldoradobrasil.com.br/img/plano-de-manejo-florestal-2020.pdf>. Acesso em: 03 jun. 2021.
- ESRI – Environmental Systems Research Institute Inc. **ArcGis versão 10.8**. EUA: Environmental Systems Research Institute, 2019.
- FITZ, P. R. **Geoprocessamento sem complicação**. 3. ed. revisada e ampliada. São Paulo, 2013. 26p
- IBGE. Instituto Brasileiro de Geografia e Estatística. **Pesquisa da Pecuária Municipal**. Disponível em: <https://sidra.ibge.gov.br/pesquisa/ppm/quadros/brasil/2019>. Acesso em: 02 jun. 2021.
- IBGE. Instituto Brasileiro de Geografia e Estatística. **Produção Agrícola Municipal**. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pam/tabelas>. Acesso em: 02 jun. 2021.
- IBGE. Instituto Brasileiro de Geografia e Estatística. **Produção da Extração Vegetal e da Silvicultura**. Disponível em: <https://sidra.ibge.gov.br/pesquisa/pevs/quadros/brasil/2019>. Acesso em: 03 jun. 2021.

- JOHANN, Jerry A. et al. Data mining techniques for identification of spectrally homogeneous areas using NDVI temporal profiles of soybean crop. **Eng. Agríc.**, Jaboticabal, v. 33, n. 3, p. 511-524, June 2013.
- Köppen, W. Versuch Einer Klassifikation Der Klimate, Vorzugsweise Nach Ihren Beziehungen Zur Pflanzenwelt. **Geographische Zeitschrift**, [S./l.], v. 6, no. 11, pp. 593–611, 1900.
- KODINARIYA, T. M.; MAKWANA, P. R. Review on determining number of Cluster in K-Means Clustering. **International Journal of Advance Research in Computer Science and Management Studies**, [S./l.], v. 1, n. 6, p. 90-95, 2013.
- KUDLAVICZ, Mieczslau. **Dinâmica agrária e a territorialização do complexo Celulose/papel na microrregião de Três Lagoas/MS**. 2011. 176 f. Dissertação (Mestrado em Geografia) - Universidade Federal de Mato Grosso do Sul, Três Lagoas, 2011.
- LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. **Biometrics**, [S./l.], v.33, n. 1. p. 159-174, 1977.
- LIU, L., Peng, Z., Wu, H., Jiao, H., Yu, Y., and Zhao, J. Fast identification of urban sprawl based on k-means clustering with population density and local spatial entropy. **Sustainability**, [S./l.], v. 10, n. 8, p. 2-16, 2018.
- MANTOVANI, J. R.; BUENO, G. T. Uma proposta metodológica para mapear a dissecação do relevo e aplicá-la no Parque Nacional da Serra da Canastra-MG. **GEOUSP Espaço e Tempo (Online)**, [S. l.], v. 25, n. 1, p. e-170745, 2021.
- MATLAB. Disponível em: <<https://www.mathworks.com/help/stats/k-means-clustering.html>>. Acesso em 19/12/2021.
- NOGARE, D. **Engenharia do Conhecimento e Sistemas Especialistas**. 2016. Disponível em: < <http://www.diegonogare.net/2015/08/entendendo-como-funciona-o-algoritmode-cluster-k-means/>>. Acesso em: 10 maio. 2023.
- OLIVEIRA JUNIOR, E. **Caracterização geoambiental e análise do uso do solo na microbacia do Córrego da Moeda Três Lagoas-MS**. 2005. Monografia (Graduação em Geografia)- Universidade Federal do Mato Grosso do Sul, Três Lagoas, 2005.
- PASCUCCI, S.; CARFORA, M.F.; PALOMBO, A.; PIGNATTI, S.; Casa, R.; Pepe, M.; Castaldi, F. A. Comparison between Standard and Functional Clustering Methodologies: Application to Agricultural Fields for Yield Pattern Assessment. **Remote Sens.**, v. 10, n.4, p. 1-22, 2018.
- POTTER, C.S.; BROOKS, V. Global analysis of empirical relations between annual climate and seasonality of NDVI. **International Journal of Remote Sensing**, London, v.19, n.15, p. 2921-2948, 1998.

PONZONI, F.J.; SHIMABUKURO, Y. E. **Sensoriamento Remoto no Estudo da Vegetação**. 1. ed. São José dos Campos: Editora Parêntese, 2007.

PONZONI, F. J.; SHIMABUKURO, Y.W.; KUPLICH, T. M. **Sensoriamento remoto da vegetação**. 2. ed. São Paulo: Oficina de Textos. 2012.

RICHARD, Y.; POCCARD, I. A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa. **International Journal of Remote Sensing**, London, v. 19, n. 15, p. 2907-2920, 1998.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. **In: EARTH RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM, 3.**, 1973, Washington. Proceedings... v. 1, sec. A, p. 309-317.

SHIMABUKURO, Y. et al. **Classificação e monitoramento da cobertura vegetal do estado do Mato Grosso através de imagens NOAA-AVHRR**. São José dos Campos, SP. INPE, 1999.

SILVA, F. A. M. da; ASSAD, E. D.; EVANGELISTA, B. A. Caracterização climática do bioma Cerrado. In: SANO, S. M. (ed.); ALMEIDA, S. P. de. (ed.); RIBEIRO, J. F. (ed.). **Cerrado: ecologia e flora**. Brasília: Embrapa Informação Tecnológica, 2008. p. 69-88.

ZHAO, C.; LIU, B.; PIAO, S.; WANG, X.; LOBELL, D. B.; HUANG, Y.; HUANG, M.; YAO, Y.; BASSU, S.; CIAIS, P.; DURAND, J.-L.; ELLIOTT, J.; EWERT, F.; JANSSENS, I. A.; LI, T.; LIN, E.; LIU, Q.; MARTRE, P.; MÜLLER, C.; PENG, S.; PEÑUELAS, J.; RUANE, A. C.; WALLACH, D.; WANG, T.; WU, D.; LIU, Z.; ZHU, Y.; ZHU, Z.; ASSENG, S. Temperature increase reduces global yields of major crops in four independent estimates. **Proc. Natl. Acad. Sci.**, [S./l.], v. 114, n. 35. P. 1-6, 2017:

---

#### Autores

---

**José Roberto Mantovani** – É Graduado em Geografia pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP); Mestre em Geografia pela Universidade Federal do Mato Grosso do Sul (UFMS) e Doutor em Geografia pela Universidade Federal de Goiás (UFG). Atualmente atua como pesquisador de Pós-Doutorado, vinculado ao programa de Desastres Naturais, junto a Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/CEMADEN), no Instituto de Ciência e Tecnologia - Câmpus de São José dos Campos.

**Endereço:** Av. Eng. Francisco José Longo, 777 - Jardim Sao Dimas, São José dos Campos, SP, Brasil, CEP 12.245-000.

**Leandro Reginaldo Maximino Lelis** – É Graduado em Geografia pela Universidade Estadual Paulista (UNESP/Campus de Presidente Prudente); Mestre em Geografia pela Universidade Federal de Mato Grosso do Sul (UFMS/Campus de Três Lagoas) e Doutor em Geografia pela Universidade Federal do Rio Grande do Norte (UFRN). Atualmente é Professor no Instituto Federal da Paraíba (IFPB/Campus Catolé do Rocha).

**Endereço:** R. Projetada I, 1 - Catolé do Rocha, PB, Brasil, CEP 58.884-000.

---

**Artigo recebido em: 14 de abril de 2022.**

**Artigo aceito em: 10 de junho de 2023.**

**Artigo publicado em: 01 de julho de 2023.**