

# DETECÇÃO DE ÁREAS QUEIMADAS UTILIZANDO DADOS TEMPORAIS E MODELAGEM ESTATÍSTICA NÃO SUPERVISIONADA

Rogério G. Negri<sup>1</sup>, Andréa E. O. Luz<sup>1,2</sup>, Alejandro C. O. Frery<sup>3</sup> e Wallace C. O. Casaca<sup>4</sup>

<sup>1</sup>Universidade Estadual Paulista (UNESP), Instituto de Ciência e Tecnologia, São José dos Campos, Brasil

<sup>2</sup>Programa de Pós-Graduação em Desastres Naturais (UNESP/CEMADEN), São José dos Campos, Brasil

<sup>3</sup>Victoria University of Wellington (VUW), School of Mathematics and Statistics, Wellington, New Zealand

<sup>4</sup>Universidade Estadual Paulista (UNESP), Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, Brasil

## RESUMO

As queimadas florestais têm aumentado significativamente nos últimos anos. Eventos desta natureza motivam o desenvolvimento de metodologias automatizadas para o devido mapeamentos e monitoramento. Este trabalho introduz um método capaz de mapear de forma acurada áreas afetadas por fogo utilizando modelagem estatística e series temporais de imagens de sensoriamento remoto. A avaliação desta proposta é realizada por dois estudos de caso envolvendo áreas de floresta no Brasil com frequente histórico de queimadas. São utilizadas imagens obtidas pelos satélites Landsat-8 e Sentinel-2 e apresentadas comparações com um método alternativo.

**Palavras-chave** – Queimadas, índices espectrais, séries temporais, não supervisionado.

## ABSTRACT

*The frequency of forest fires has increased significantly in recent years across the planet. Events of this nature motivate the development of automated methodologies aimed at mapping areas affected by fire. In this context, we propose a method capable of accurately mapping areas affected by fire using time series of remotely sensed multispectral images by statistical modeling. In order to evaluate the introduced proposal, we carry out a case study on a region in Brazil with recurrent history of forest fires. Furthermore, images obtained by the Landsat-8 satellite are used in this case study. Comparisons with an alternative method are included in this analysis.*

**Key words** – Forest fires, spectral index, time series, unsupervised.

## 1. INTRODUÇÃO

As queimadas florestais compreendem um sério problema ambiental e que tem proporcionado diversos danos ecológicos [1]. Quando frequentes, as queimadas são altamente prejudiciais, uma vez que afetam o processo de regeneração florestal e torna o fogo um agente disruptivo das dinâmicas ambientais [2]. Consequentemente, mapear e monitorar áreas afetadas por fogo torna-se essencial para mitigar e prevenir maiores impactos ao meio ambiente. Neste sentido, diferentes técnicas têm sido desenvolvidas para identificar este tipo de evento a partir de imagens de sensoriamento.

Dentre diferentes abordagens encontradas na literatura, o emprego de índices espectrais tem se mostrado uma ferramenta de grande utilidade. Em [3] é proposta uma forma de mensurar a severidade das queimadas com uso dos índices NDVI (*Normalized Difference Vegetation Index*) e NBR (*Normalized Burn Ratio*). De modo similar, [4] e [5] apresentam formas de expressar a severidade das queimadas e estimar a taxa de recuperação florestal nos instantes pós-queimada via índice  $\Delta$ NBR (*Differenced Normalized Burn Ratio*).

Técnicas de Aprendizado de Máquina também figuram como ferramentas poderosas na detecção de áreas queimadas. Por exemplo, [6] utiliza redes neurais como ferramenta primária na identificação de queimadas. De forma simples e eficiente, em [7] é proposto um modelo probabilístico que estima a cobertura florestal e respectiva suscetibilidade à queimadas.

No contexto das discussões apresentadas, este trabalho introduz um novo método para mapeamento de áreas queimadas utilizando séries temporais de imagens de sensoriamento remoto, índices espectrais e modelagem probabilística baseada em Aprendizado de Máquina. Em linhas gerais, índice NBR é aplicado sobre uma série temporal de imagens com intuito de expressar as variações espectro-temporal dos alvos. Posteriormente, com uso de regressão logística, tais variações são avaliadas e empregadas no mapeamento de áreas afetadas por fogo.

Com objetivo de evidenciar a efetividade da proposta, são apresentados resultados alcançados em dois estudos de caso envolvendo áreas de floresta no Brasil com frequente histórico de queimadas. Imagens obtidas pelos sensores Landsat-8 OLI e Sentinel-2 MSI são utilizadas nestes estudos.

As discussões que seguem estão organizadas da seguinte forma: a Seção 2 exibe conceitos teóricos necessários para a formalização da proposta, apresentada na Seção 3; estudos de caso e respectivos resultados são apresentados na Seção 4, seguido pelas conclusões na Seção 5.

## 2. CONCEITOS ELEMENTARES

Seja  $\mathcal{I}^{(t)}$  uma imagem definida sobre um suporte  $\mathcal{S} \subset \mathbb{N}^2$  e adquirida no instante  $t$ ;  $\mathbf{x}_i^{(t)}$  é um vetor, definido no espaço de atributos  $\mathcal{X}$ , contendo informações registradas pelo sensor sobre na posição  $s_i \in \mathcal{S}$  de  $\mathcal{I}^{(t)}$ .

Índices espectrais, computados a partir das componentes de  $\mathbf{x}_i^{(t)}$  em cada  $s_i$  de  $\mathcal{I}^{(t)}$ , podem auxiliar na identificação e representação de alvos específicos. Dentre diversos índices espectrais encontrados na literatura, alguns exemplos típicos

são NDVI e NDWI. Geralmente, tais índices combinam a informação de bandas espectrais (i.e., componentes de  $x_i^{(t)}$ ) relacionadas às faixas do visível e infravermelho, sendo assim sensível a variações de cor, composição/mistura do solo e clorofila.

O índice NBR [8] é uma alternativa conveniente para identificação de áreas afetadas por fogo. Considerando uma imagem  $\mathcal{I}^{(t)}$  com  $NIR_i^{(t)}$  e  $SWIR_i^{(t)}$  como componentes de  $x_i^{(t)}$ , que representam o comportamento de um alvo na posição  $s_i$  em relação aos comprimentos de onda do infravermelho próximo e de ondas curtas, a imagem de valores de NBR é definida por:

$$\mathcal{I}_{NBR}^{(t)}(s_i) = \frac{NIR_i^{(t)} - SWIR_i^{(t)}}{NIR_i^{(t)} + SWIR_i^{(t)}}, \forall s_i \in \mathcal{S}. \quad (1)$$

Os índices espectrais podem ser usados segundo um único instante  $t$  ou sobre uma série temporal. Quando utilizado em série temporal, a consideração de diferenças entre instante pode auxiliar no mapeamento de eventos específicos. Especificamente, quando o interesse recai no mapeamento de áreas queimadas, as abordagens encontradas na literatura empregam imagem pós-queimada ou combinam dados antes (pré) e após (pós) as queimadas [9]. Neste sentido, [8] propôs o “Delta-NBR” ( $\Delta NBR$ ) com intuito de medir a severidade das queimadas sobre a vegetação. Tal metodologia consiste em caracterizar a severidade da queimada através da diferença  $\Delta NBR = \mathcal{I}_{NBR}^{(pré)} - \mathcal{I}_{NBR}^{(pós)}$ , onde  $\mathcal{I}_{NBR}^{(pré)}$  e  $\mathcal{I}_{NBR}^{(pós)}$  representam o NBR em dois instantes distintos, antes e após o evento de queimada. Segundo [8], evidências de queimadas são caracterizadas por  $\Delta NBR > 0.1$ .

Apesar do  $\Delta NBR$  oferecer uma boa perspectiva em relação aos efeitos da severidade das queimadas, é importante destacar que interferências atmosféricas ou até mesmo variações nos limites das bandas NIR e SWIR demandam atenção durante a interpretação dos resultados. Assim, podemos evitar esta limitação construindo um mapa binário que incorpora as informações radiométricas derivada de um modelo de regressão logística.

### 3. PROPOSTA

A Figura 1 ilustra uma visão geral do método proposto.

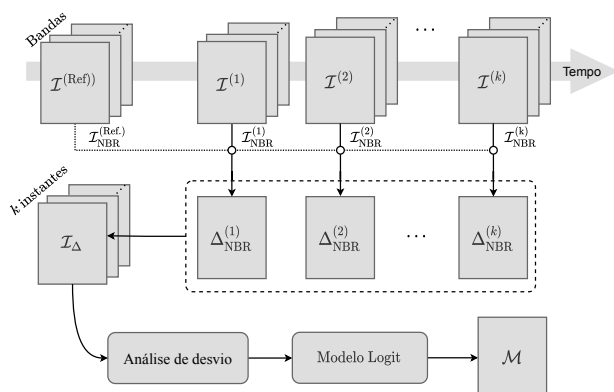


Figura 1: Etapas do método proposto.

O bloco superior exibe o cálculo do NBR. Inicialmente, é assumido como referência a imagem  $\mathcal{I}^{(Ref)}$ , que expressa o comportamento usual dos alvos no domínio espacial  $\mathcal{S}$ . Ainda, seja  $\mathcal{I}^{(t)}$ , com  $t = 1, 2, \dots, k$ , uma série temporal de imagens também definidas sobre  $\mathcal{S}$ . Dado  $s_i \in \mathcal{S}$ , denotamos por  $x_i^{(t)} = [x_{i1}^{(t)}, \dots, x_{in}^{(t)}] \in \mathcal{X}$ , para  $t \in \{Ref, 1, 2, \dots, k\}$ , um vetor de atributos com medidas registradas em  $n$  bandas espectrais. Utilizando componentes apropriadas de  $x_i^{(t)}$ , é computado o NBR em cada posição de  $\mathcal{S}$  a fim de determinar  $\mathcal{I}_{NBR}^{(t)}$ .

Baseado nas discussões anteriores sobre o índice  $\Delta NBR$ , admitindo  $\mathcal{I}_{NBR}^{(Ref)}$  como uma imagem “pré-fogo”, podemos definir  $\Delta_{NBR}^{(t)} = \mathcal{I}_{NBR}^{(Ref)} - \mathcal{I}_{NBR}^{(t)}$ , para  $t = 1, 2, \dots, k$ . Denotamos  $\Delta_{NBR}^{(t)}(s_i) = \delta_i^{(t)} \in \mathbb{R}$ . Admitindo ordem cronológica sobre  $\Delta NBR^{(t)}$  é determinada a imagem  $\mathcal{I}_{\Delta}$ , onde em cada posição  $s_i$  temos  $\mathcal{I}_{\Delta}(s_i) = \delta_i = [\delta_i^{(1)}, \dots, \delta_i^{(k)}]$ .

Vale destacar que  $\mathcal{I}_{\Delta}$  é uma representação espaço-temporal dos valores de  $\Delta NBR$ . Locais que não sofrem alterações ao longo do tempo tendem a exibir valores similares de  $\delta_i^{(t)}$  independentemente de  $t$ . Por outro lado, a ausência de correlação entre duas ou mais componentes de  $\delta_i$  surge como evidência de mudança temporal.

Neste contexto, a medida  $d_i$  pode ser empregada para quantificar a correlação entre as componentes de  $\delta_i$ :

$$d_i = \left\| \delta_i - \frac{\langle \delta_i, \mathbf{w} \rangle \mathbf{w}}{\|\mathbf{w}\|^2} \right\|, \quad (2)$$

onde  $\mathbf{w} = [1, 1, \dots, 1] \in \mathbb{R}^k$  é o vetor diretor da “linha de identidade” usada como referência (i.e., onde  $\delta_i^{(t)} = \delta_i^{(\ell)}, \forall t, \ell$ );  $\|\cdot\|$  e  $\langle \cdot, \cdot \rangle$  correspondem à norma e ao produto interno em um espaço Euclidiano. A submissão de  $\delta_i$  associado a cada  $s_i \in \mathcal{S}$  em (2) proporciona como resultado a imagem  $\mathcal{I}_D$ .

No âmbito desta proposta, assumindo que pequenos valores de  $d_i$  representam mudanças irrelevantes em  $s_i$  no tempo, tal posição deve estar associada a uma área não queimada. Baseado nesta premissa, é definido inicialmente um limite superior  $\Xi = \max \left\{ d_i : \max_{j=1, \dots, k} \{ \delta_i^{(j)} \} \leq \xi \right\}$  de acordo com os valores de  $d_i$  relacionados ao vetor  $\delta_i$  cujas componentes são limitadas por uma tolerância  $\xi$ . Por conveniência, denotamos  $D(\xi) = \{ d_i : \max_{j=1, \dots, k} \{ \delta_i^{(j)} \} \leq \xi \}$ .

De forma similar, é possível definir inicialmente um limite inferior  $\Lambda = \Xi + \alpha \sigma_{\Xi}$  para os valores de  $d_i$  relacionado às posições que apresentam uma mudança relevante em consequência de um evento de queimada. Nesta expressão,  $\sigma_{\Xi}$  é o desvio padrão observado em  $D(\xi)$  e  $\alpha \in \mathbb{R}_+$  é um parâmetro que controla a transição entre mudanças relevantes e irrelevantes.

Perante as discussões apresentadas, a informação derivada de uma série temporal de imagens multiespectrais podem discriminar localizações afetadas (i.e., queimadas) e não afetadas (i.e., não-queimada) por fogo. Os limites  $\Xi$  e  $\Lambda$  são elementos críticos na construção de um modelo capaz de expressar a probabilidade das mudanças temporais como consequência de um evento de queimadas. Com intuito de construir tal modelo, é definido o conjunto de referência

$\mathcal{D} = \{(d_i, y_i) \in \mathbb{R} \times \{0, 1\} : i = 1, \dots, m\}$ , onde  $y_i = 0$  ou 1 quando  $d_i \geq \Lambda$  ou  $d_i \leq \Xi$ , respectivamente. Vale observar que, enquanto  $y_i = 1$  representa um evento de não-queimada em  $s_i$  no período analisado (i.e., mudança irrelevante),  $y_i = 0$  representa uma mudança relevante decorrente de queimada. Consequentemente, podemos observar que  $\Xi$  e  $\Lambda$  definem uma margem entre as mudanças relevantes e irrelevantes. A associação intuitiva “mudança-é-queimada” é uma consequência dos estágios iniciais onde os valores de NBR e  $\Delta\text{NBR}$  dão lugar a  $\mathcal{I}_\Delta$  e, por sua vez, a  $\mathcal{D}$ .

Por fim, com base nas informações em  $\mathcal{D}$ , um Modelo Logit (i.e., de regressão logística) é construído e aplicado em cada posição de  $s_i \in \mathcal{S}$ , produzindo assim o mapa  $\mathcal{M}$  composto por mudanças relevantes e/ou irrelevantes.

Mas discussão que segue, o método proposto é denominado por UFD (*Unsupervised Fire Detection*).

#### 4. EXPERIMENTOS E RESULTADOS

Com objetivo de avaliar o método UFD, foram realizados experimentos envolvendo áreas de estudo localizadas em São Félix do Xingu (Área 1) e em uma região na fronteira Bolívia-Brasil (Área 2). Tais experimentos utilizaram imagens obtidas no período de 1 de agosto a 31 de outubro de 2020 pelos sensores Landsat-8 OLI (ref. Área 1) e Sentinel-2 MSI (ref. Área 2). Para cada uma das áreas,  $\mathcal{I}^{(\text{Ref})}$  é definida como a imagem mediana de todas as imagens disponíveis no período entre 1 de junho de 2019 a 31 de junho de 2020 com cobertura de nuvem inferior a 50%.

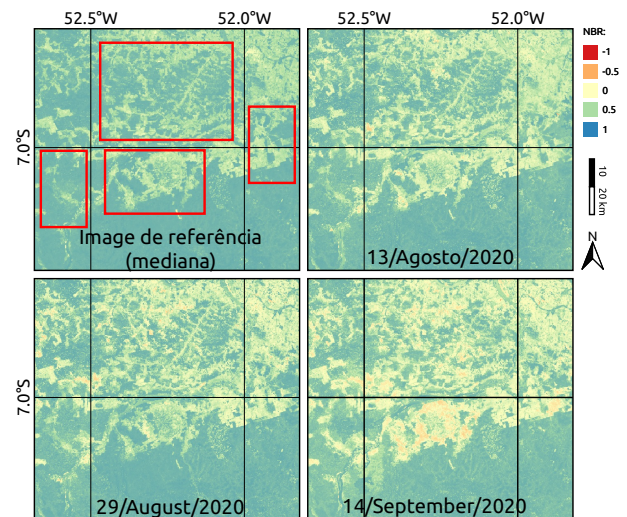
A Figura 2 apresenta alguns instantes incluídos no período analisado, em termos do índice NBR. As regiões retangulares destacam porções com intensa variação de NBR ao longo do tempo, indicando assim evidências de queimada.

Para fins de comparação, a abordagem clássica  $\Delta\text{NBR}$  estendida para dados multitemporais (denominada MTDNBR – *Multitemporal  $\Delta\text{NBR}$* ) será usada como base de referência. Neste caso, as localizações cujo valor de  $\mathcal{I}_{\Delta\text{NBR}}^{(t)}$  supera  $\xi$  em algum instante  $t = 1, \dots, k$  é tomado como “queimada”. Formalmente, seja  $\mathcal{M}^*$  o mapeamento derivado desta abordagem, a ocorrência ou não de queimada na posição  $s_i$  é denotado por:

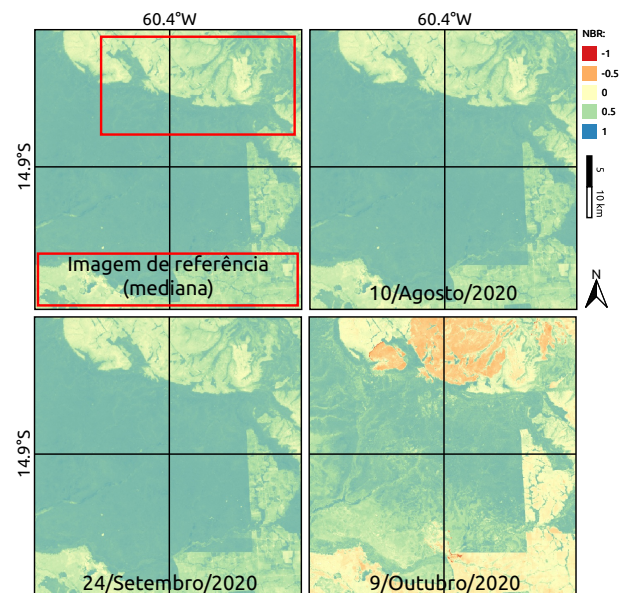
$$\mathcal{M}^*(s_i) = \begin{cases} 1 & \text{se } \mathcal{I}_{\Delta\text{NBR}}^{(t)} < \Xi; t = 1, \dots, k \\ 0 & \text{caso contrário.} \end{cases}$$

Motivado pelo limiar mínimo para ocorrência de queimada, conforme mencionado na Seção 2, ambos abordagens UFD e MTDNBR assumem  $\Xi = 0.1$  como parâmetro para mapear áreas afetadas por fogo. Ainda, a fim de avaliar a consistência dos mapeamentos gerados pelos métodos UFD e MTDNBR, estimativas geradas pelo produto “MCD64A1 Modis Burned Area Monthly” [10] (500 m de resolução espacial) foram adotadas como referência. Segundo este produto, posições onde há indicativo de queimada com confiança acima de 90 % são tomadas como áreas “queimadas”; caso contrário, “não-queimada”. O Acerto Global (OA) e o coeficiente kappa [11] foram adotados como medidas de acurácia.

As Figura 3 e 4 apresentam os resultados obtidos, juntamente com os respectivos valores de acurácia; onde podemos verificar a consistência entre o método proposto



(a) Área 1 – Landsat-8 OLI.



(b) Área 2 – Sentinel-2 MSI.

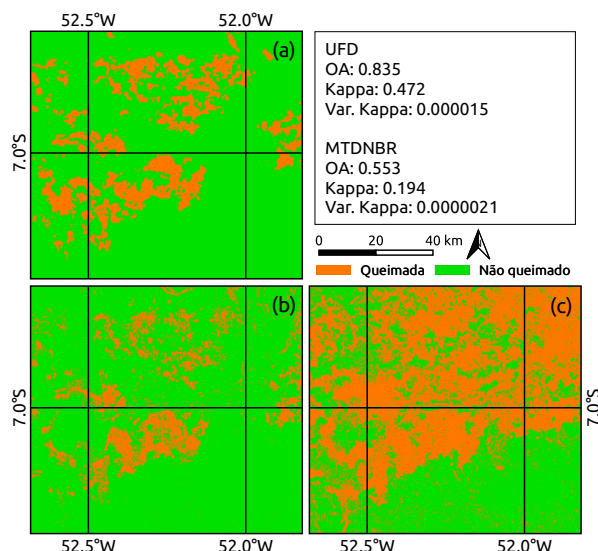
**Figura 2: Áreas de estudo expressas em termos de NBR em diferentes instantes, incluindo a imagem de referência.**

(Fig. 3(b) e 4(b)) em relação aos resultados esperados (Fig. 3(a) e 4(a)). Por outro lado, os mapeamentos gerados pelo método MTDNBR (Fig. 3(c) e 4(c)) exibem falta de robustez em relação às áreas não afetadas por fogo. Os valores computados demonstra superioridade da proposta UFD, cuja significância estatística ( $p$ -valores  $< 10^{-4}$ ) é verificada segundo teste de hipótese baseado em coeficiente kappa [11].

#### 5. CONCLUSÕES

Este trabalho propôs uma nova abordagem não supervisionada para identificação de áreas queimadas utilizando séries temporais de imagens. Índices espectrais e modelagem estatística são os principais conceitos que suportam tal proposta. Foram realizados estudos de caso envolvendo imagens Landsat-8 OLI e Sentinel-2 MSI, incluindo comparações com método alternativo.





**Figura 3:** Mapeamentos obtidos para Área 1 segundo os métodos UFD (b) e MTDNBR (c); (a) produto “Modis Burned Area”.

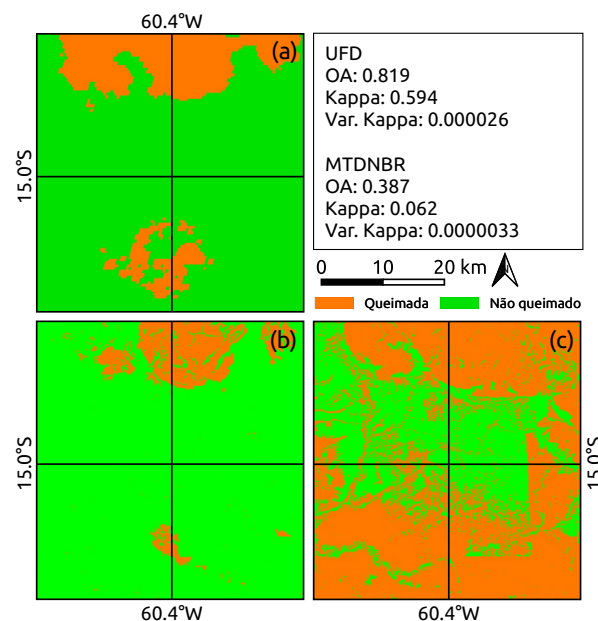
Os resultados exibiram alta aderência entre o método proposto e os resultados esperados. Por outro lado, o método concorrente não foi capaz alcançar resultados consistente.

Para evolução desta proposta, investigações futuras podem incluir: (i) inclusão de informação contextual no cálculo das estimativas/mapeamentos efetuados pelo método UFD; (ii) adaptação e aplicação em outras questões ambientais; (iii) extensão da proposta para além de mapeamentos binários.

**Agradecimento:** Os autores agradecem à FAPESP (Procs.: 2021/01305-6, 2021/03328-3) e ao CNPq (Proc. 316228/2021-4) pelo auxílio financeiro.

## 6. REFERÊNCIAS

- [1] Alireza Ghomi Motazeh, Ehsan Farahi Ashtiani, Roghaye Baniyasi, and Fereidun Masoumpoor Choobar. Rating and mapping fire hazard in the hardwood Hyrcanian forests using GIS and expert choice software. *Forestry Ideas*, 19(2) (46):141–150, 2013.
- [2] Nayane Cristina Candida Santos Prestes, Klécia Gili Massi, Estevão Alves Silva, Denis Silva Nogueira, Edmar Almeida de Oliveira, Renata Freitag, Beatriz Schwantes Marimon, Ben Hur Marimon-Junior, Michael Keller, and Ted R Feldpausch. Fire effects on understory forest regeneration in southern Amazonia. *Frontiers in Forests and Global Change*, 3:10, 2020.
- [3] S Escuin, R Navarro, and P Fernandez. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing*, 29(4):1053–1073, 2008.
- [4] Zhihua Liu. Effects of climate and fire on short-term vegetation recovery in the boreal larch forests of northeastern China. *Scientific reports*, 6(1):1–14, 2016.



**Figura 4:** Mapeamentos obtidos para Área 2 segundo os métodos UFD (b) e MTDNBR (c); (a) produto “Modis Burned Area”.

- [5] Bailu Zhao, Qianlai Zhuang, Narasinha Shurpali, Kajar Köster, Frank Berninger, and Jukka Pumpanen. North American boreal forests are a large carbon source due to wildfires from 1986 to 2016. *Scientific reports*, 11(1):1–14, 2021.
- [6] Yifang Ban, Puzhao Zhang, Andrea Nascetti, Alexandre R Bevington, and Michael A Wulder. Near real-time wildfire progression monitoring with Sentinel-1 SAR time series and deep learning. *Scientific Reports*, 10(1):1–15, 2020.
- [7] Eraldo AT Matricardi, David L Skole, Marcos A Pedlowski, and Walter Chomentowski. Assessment of forest disturbances by selective logging and forest fires in the Brazilian Amazon using Landsat data. *International Journal of Remote Sensing*, 34(4):1057–1086, 2013.
- [8] C.H. Key and N.C. Benson. Landscape assessment: Ground measure of severity, the composite burn index; and remote sensing of severity, the normalized burn ratio. Technical Report RMRS-GTR-164-CD: LA 1-51, USDA Forest Service, Rocky Mountain Research Station, USGS Publications Warehouse, 2006.
- [9] Jose Antonio Sobrino, Rafael Llorens, Cristina Fernández, José M Fernández-Alonso, and José Antonio Vega. Relationship between soil burn severity in forest fires measured in situ and through spectral indices of remote detection. *Forests*, 10(5):457, 2019.
- [10] L. Giglio, L. Boschetti, D. Roy, A. A. Hoffmann, M. Humber, and J. V. University of Maryland Hall. *Collection 6 MODIS Burned Area Product User’s Guide (Version 1.3)*. National Aeronautics and Space Administration, 2020.
- [11] Russell G Congalton and Kass Green. *Assessing the accuracy of remotely sensed data: principles and practices*. CRC press, 3 edition, 2019.