UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José dos Campos
Instituto de Ciência e Tecnologia

Cemaden
Centro Nacional de Monitoramento e Alertas de Desastres Naturais

UNIDADE DE PESQUISA DO
MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES

PÁTRIA AMADA BRASIL
GOVERNO FEDERAL

ANDRÉA ELIZA DE OLIVEIRA LUZ

MAPEAMENTO DA OCORRÊNCIA E SUSCETIBILIDADE DE FOGO EM ÁREAS VEGETADAS COM USO DE MODELOS DE APRENDIZADO DE MÁQUINA E DADOS DE SENSORIAMENTO REMOTO

2022

ANDRÉA ELIZA DE OLIVEIRA LUZ

MAPEAMENTO DA OCORRÊNCIA E SUSCETIBILIDADE DE FOGO EM ÁREAS VEGETADAS COM USO DE MODELOS DE APRENDIZADO DE MÁQUINA E DADOS DE SENSORIAMENTO REMOTO

Dissertação apresentada ao Instituto de Ciência e Tecnologia, Universidade Estadual Paulista (Unesp), Campus de São José dos Campos; Centro Nacional de Monitoramento e Alertas de Desastres Naturais (Cemaden), como parte dos requisitos para a obtenção do título de MESTRE pelo Programa de Pós-Graduação em DESASTRES NATURAIS.
Área: Desastres naturais. Linha de pesquisa: Instrumentação e análise de dados.

Orientador: Prof. Dr. Rogério Galante Negri

São José dos Campos
2022

Instituto de Ciência e Tecnologia [internet]. Normalização de tese e dissertação [acesso em 2022]. Disponível em http://www.ict.unesp.br/biblioteca/normalizacao

Apresentação gráfica e normalização de acordo com as normas estabelecidas pelo Serviço de Normalização de Documentos da Seção Técnica de Referência e Atendimento ao Usuário e Documentação (STRAUD).

Luz, Andréa Eliza de Oliveira
MAPEAMENTO DA OCORRÊNCIA E SUSCETIBILIDADE DE FOGO EM ÁREAS VEGETADAS COM USO DE MODELOS DE APRENDIZADO DE MÁQUINA E DADOS DE SENSORIAMENTO REMOTO / Andréa Eliza de Oliveira Luz. - São José dos Campos : [s.n.], 2022.
45 f. : il.

Dissertação (Mestrado em Desastres Naturais) - Pós-graduação em Desastres Naturais - Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia; Centro Nacional de Monitoramento e Alertas de Desastres Naturais (Cemaden), São José dos Campos, 2022.
Orientador: Rogério Galante Negri.

1. Incêndios Florestais. 2. Sensoriamento Remoto. 3. Aprendizado de Máquina. I. Negri, Rogério Galante , orient. II. Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia, São José dos Campos. III. Universidade Estadual Paulista 'Júlio de Mesquita Filho' - Unesp. IV. Universidade Estadual Paulista (Unesp). V. Centro Nacional de Monitoramento e Alertas de Desastres Naturais (Cemaden). VI. Título.

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi e Seção Técnica de Informática, ICMC/USP
com adaptações - STATI, STRAUD e DTI do ICT/UNESP.
Renata Aparecida Couto Martins CRB-8/8376

SUMÁRIO

LUZ, O.E.A. **MAPEAMENTO DA OCORRÊNCIA E SUSCETIBILIDADE DE FOGO EM ÁREAS VEGETADAS COM USO DE MODELOS DE APRENDIZADO DE MÁQUINA E DADOS DE SENSORIAMENTO REMOTO**. Dissertação. São José dos Campos: Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia; Centro Nacional de Monitoramento e Alertas de Desastres Naturais (Cemaden), 2022.

RESUMO

Como consequência da combinação de extremos climáticos e das ações humanas, destaca-se nos anos recentes a ocorrência de desastres naturais decorrentes de queimadas em áreas florestais. Tais eventos têm sido observados com mais frequência e de forma generalizada, ocasionando graves danos ambientais e sociais em diversos ecossistemas. Estudos apontam para a necessidade de métodos para o monitoramento de queimadas, os quais podem ser realizados através do mapeamento de áreas afetadas ou suscetíveis a tal evento. Neste contexto, dados extraídos de séries multitemporais de imagens obtidas por sensoriamento remoto e técnicas de aprendizado de máquina são componentes potenciais no desenvolvimento de métodos e ferramentas para esse fim. Nesta pesquisa foram desenvolvidos dois métodos através do emprego de conceitos de modelagem, classificação estatística e detecção de anomalias. O primeiro método é capaz de proporcionar o mapeamento das ocorrências através da precisão da identificação das áreas afetadas por fogo, enquanto o segundo realiza o mapeamento da suscetibilidade ao fogo através da identificação de localizações não anômalas.

Palavras-chave: incêndios florestais; sensoriamento remoto; aprendizado de máquina.

LUZ, O.E.A. ***FIRE OCCURRENCE AND SUSCEPTIBILITY MAPPING IN VEGETATED AREAS WITH THE USE OF MACHINE LEARNING MODELS AND REMOTE SENSING DATA***. *Dissertation. São José dos Campos: São Paulo State University (Unesp), Institute of Science and Technology; National Center for Monitoring and Early Warning of Natural Disasters (Cemaden), 2022.*

ABSTRACT

As a result of the combination of climatic extremes and human actions, the occurrence of natural disasters resulting from fires in forest areas has been highlighted in recent years. Such events have been observed more frequently and in a generalized way, causing serious environmental and social damages in several ecosystems. Studies point to the need for methods for monitoring fires, which can be carried out by mapping areas affected or susceptible to such an event. In this context, data extracted from multitemporal series of images obtained by remote sensing and machine learning techniques are potential components in the development of methods and tools for this purpose. In this research two methods were developed through the use of modeling concepts, statistical classification and anomaly detection. The first method is able to provide the mapping of occurrences through the accuracy of the identification of fire-affected areas, while the second performs the mapping of fire susceptibility through the identification of non-anomalous locations.

Keywords: *forest fires; remote sensing; machine learning.*

1 INTRODUÇÃO

No topo das ameaças globais, as mudanças climáticas são responsáveis por grande parte dos problemas ambientais discutidos atualmente. As consequências dessas variações no clima incluem a extinção de espécies ameaçadas, riscos à saúde alimentar, além de riscos associados à ocorrência de eventos extremos, tais como ondas de calor, chuvas extremas e inundações costeiras (NOBRE; MARENGO; SOARES, 2019). Esses eventos são suficientes para motivar a realização de estudos e concentração de esforços para redução de risco (AALST, 2006).

Em tempo, os eventos extremos refletem diretamente na ocorrência e exposição a desastres naturais. A nível de exemplificação, o aumento das temperaturas que contribuem para a ocorrência, severidade e duração das ondas de calor e secas tem causado aumento no risco de incêndios florestais. Estes incêndios se destacam como as principais causas de degradação florestal nos últimos anos, principalmente pelo registro de grandes incêndios florestais em intervalos muito mais curtos do que no passado (LEWIS et al., 2011; SILVA et al., 2018). De acordo com Malhi et al. (2008), os eventos de seca severa superpostos a atividades de uso do solo aumentam a suscetibilidade das florestas aos incêndios.

Historicamente, o fogo influencia os padrões e processos em diferentes ecossistemas (MORITZ et al., 2014), incluindo a distribuição e estrutura da vegetação. Embora a humanidade e o fogo sempre tenham coexistido, nossa capacidade de controlá-lo permanece imperfeita e tem ainda se tornado mais difícil ao passo que as mudanças climáticas alteram seu regime (BOWMAN et al., 2009). Ademais, nossa capacidade de previsão de regimes futuros de fogo é limitada pela falta de compreensão daquilo que desencadeia e controla os incêndios de alta intensidade (BRANDO et al., 2014).

Vale ressaltar que mudanças induzidas na estrutura da vegetação são responsáveis por tornar ambientes menos densos e úmidos, favorecendo a propagação do fogo e proporcionando um ambiente suscetível a incêndios generalizados (PAN et al., 2011).

No Brasil, por exemplo, as ocorrências de fogo estão fortemente relacionadas ao desmatamento e às queimadas utilizadas para promover a renovação de áreas de pastagem e cultivo (BAYNE et al., 2019; GARCIA et al., 2021). Todavia é importante destacar que a origem do fogo pode se dar de forma natural, acidental ou criminosa (CAÚLA et al., 2015).

Recentemente, no ano de 2020, o Pantanal foi destaque na mídia internacional por apresentar o maior número de incêndios já observados no bioma (MARQUES et al., 2021; INPE, 2021). Tais eventos sem precedentes causaram graves danos sociais e ambientais. Em um passado não distante, maior destaque tem sido dado à região amazônica devido às centenas de milhares de focos de incêndio por ano e devido às frequentes mudanças na sua composição (PRESTES et al., 2020). Com base nessas constatações, destaca-se a necessidade do aprimoramento das metodologias para o mapeamento e monitoramento do fogo como forma de mitigar ou mesmo prevenir a ocorrência de novos incêndios.

Nesse sentido, estudos sobre a avaliação da suscetibilidade a incêndios florestais são

foco em discussões para a redução do risco. Abordagens descritas por Hong et al. (2018) e Pourghasemi et al. (2020), incluem a predição de eventos e a identificação de áreas suscetíveis ao fogo florestal. No entanto, são destacadas fontes de incerteza nos processos de modelagem relacionados à coleta de dados. Além disso, as bases de dados utilizadas dependem de arquivos históricos e levantamentos de campo.

Na mesma perspectiva, o Sensoriamento Remoto tem se destacado em diversas iniciativas dentro da comunidade científica por sua ampla observação da Terra e pelas soluções de modelos automatizados disponíveis para o estudo em questão (YUAN et al., 2020). Ao longo das últimas décadas essas tecnologias foram aprimoradas e atualmente permitem o fornecimento de dados sobre grandes áreas, com alta frequência temporal e baixo custo operacional, favorecendo assim pesquisas a respeito do fogo. Por sua vez, o aumento na quantidade e qualidade de dados fornecidos por avançados sensores imageadores tem possibilitado ainda o aprimoramento de algoritmos para mapeamento, monitoramento e predição de incêndios (MARTÍNEZ-ÁLVAREZ; BUI, 2020). Ademais, soma-se a isto o uso destes dados em combinação as técnicas de Aprendizado de Máquina, as quais tem se mostrado eficientes quanto a predição de eventos climáticos extremos em áreas vegetadas (HART et al., 2019; LEUENBERGER et al., 2018). Em especial, os métodos de classificação de dados aplicados à caracterização do domínio espacial se sobressaem como componentes relevantes, capazes de viabilizar a extração de conhecimento da superfície terrestre de forma automatizada.

Neste contexto, a detecção de anomalias implementada pelos modelos *One-class Support Vector Machine* (OC-SVM) (SCHÖLKOPF et al., 2001) e *Isolation Forest* (IF) (LIU; TING; ZHOU, 2008) surgem como alternativa na identificação e mapeamento de áreas atingidas ou suscetíveis ao fogo. Usualmente, essas técnicas têm sido utilizadas na identificação de fraudes bancárias, verificação de invasores em sistemas de segurança e no apoio a exames médicos (GU et al., 2019), mas recentemente também são apontadas como ferramentas úteis no estudo e monitoramento ambiental (HAVENS; BEZDEK, 2011). Pode-se citar como exemplo, o estudo desenvolvido por Salehi e Rashidi (2018), que demonstra os benefícios das abordagens de detecção de anomalias para prever o risco de incêndio florestal.

Face às discussões apontadas, utilizando-se de técnicas de Aprendizado de Máquina e análise de séries temporais de imagens de Sensoriamento Remoto, foram desenvolvidos dois métodos capazes de mapear a ocorrência e a suscetibilidade ao fogo em áreas vegetadas.

De forma superficial, o primeiro método busca identificar e extrair informações sobre os locais atingidos por fogo, proporcionando o mapeamento das ocorrências de queimadas em um período passado. O segundo método, emprega os modelos OC-SVM e IF a fim de identificar locais com características verossímeis às regiões já atingidas por fogo, proporcionando assim, um mapeamento sobre a suscetibilidade a este tipo de evento.

Este documento encontra-se organizado da seguinte forma: no Capítulo 2 são apresentados os artigos *Anomaly-Driven Approach for Forest Fire Susceptibility Mapping Using Multitemporal Remote Sensing Data* e *Fire Detection with Multitemporal Multispectral Data*

and a Probabilistic Unsupervised Technique; o Capítulo 3 aborda as considerações finais acerca das pesquisas realizadas.

2 ARTIGOS

2.1 *Anomaly-Driven Approach for Forest Fire Susceptibility Mapping Using Multitemporal Remote Sensing Data*

O artigo a seguir foi submetido em 19 de janeiro de 2022 no periódico *Modeling Earth Systems and Environment*.

Anomaly-Driven Approach for Forest Fire Susceptibility Mapping Using Multitemporal Remote Sensing Data

Abstract

The economic and environmental impacts of wildfires have motivated the development of new studies and methodologies to prevent the occurrence of such devastating events. In fact, identifying and mapping fire-susceptible areas arise as critical tasks not only to pave the way for rapid responses to stop the fire spreading, but also to support emergency evacuation plans for families affected by the forest fires. Therefore, in this paper we introduce a new automated, fully unsupervised technique for mapping the risk of fires in forest areas by combining multitemporal Remote Sensing tools and anomaly detection concepts. More specifically, we focus our analysis on two case studies of Brazilian forest areas, assessing the effectiveness of the proposed methodology in two recent real events of forest fire in Brazil. To design our experiments, we take multitemporal images acquired by the Landsat-8 Operational Land Imager and Modis sensors. We experimentally show that the current methodology was capable of predicting the fires in the studied areas at posterior instants, thus demonstrating the general performance and applicably of the proposed method to prevent and mitigate the impact of fires in forest regions.

Keywords: Remote Sensing, multitemporal, anomaly detection, forest fires, spectral indices

1 Introduction

Worldwide, forest fires comprise a phenomenon of great importance due to their severe economic and environmental consequences [1]. It is well-known that the origin of fires may be natural, accidental, or even criminal [2]. However, in recent decades, the climate changes and the intense human activity have substantially contributed to increasing the occurrence and severity of fire events on the terrestrial biomes and ecosystems [3]. Furthermore, since fires may be used to renewal pasture and cultivation areas [4], such a dangerous strategy can also trigger fire events, hence leading to widespread fires as those that have devastated the Brazilian biomes like Amazon and Pantanal [5].

Frequent incidents of fires are hazardous for the forest regeneration processes, compromising the local dynamics and biodiversity [6]. Additionally, according to the Intergovernmental Panel on Climate Change (IPCC), these events are considered the primary source of greenhouse gases emission, thus contributing to global warming [7]. Moreover, as stated by Brando et al. [8] and Pan et al. [9], environmental issues like temperature increasing, precipitation reducing, and the induced changes in the vegetation structure are responsible for the emergence of less dense, humid forest environments, which are more susceptible to fires.

Although fire monitoring comprises a highly relevant task, mapping the risk of fire events occurrence is a challenge [10, 11], since the success in predicting future fires is limited by the lack of understanding of what triggers and controls the dynamic of these events [8]. Among the different approaches devoted to reducing forest fires, the assessment of fire susceptibility is a particularly effective way, which allows both to predict potential events and detect areas of risk [12]. Such a preventive approach is also advocated by Pourghasemi et al. [13], where the authors state the possibility of linking fire-prone areas with an incendiary event, including forest burning.

Concerning the forest fire issue, a branch of scientific community has faced the problem by applying Remote Sensing (RS) tools, such as multitemporal data and computer vision apparatus [14, 15]. Indeed, new RS-based technologies have been proposed over the last years, allowing to acquire reliable data over large monitored areas with high temporal frequency and low operating cost [16]. The increase in quantity and quality of data, as provided by advanced sensors, favors the development of new algorithms for mapping, monitoring and predicting fires [17, 18]. For exemple, the combination of Remote Sensing data with Unsupervised Learning has proven to be consistent for predicting extreme weather events [19, 20]. In fact, the classification methods for spatial domain mapping stand out as a relevant tool capable of enabling the automated extraction of knowledge from the Earth's surface. The works carried out by Dickson et al. [21], Kamalakannan et al. [22], and Hong et al. [12] demonstrate the possibility of developing susceptibility risk models based on Machine Learning techniques and large databases.

In the light of the above-discussed issues, in this paper we introduce a new technique for identifying and mapping fire-susceptible areas which combines the effectiveness of Unsupervised Learning with the accuracy of multitemporal Remote Sensing data. The designed framework is capable of both extracting and classifying fire-related features of different localities affected by fire in a past period. The processed features are then employed to build an anomaly classification model, used to label the locations prone to fire events. We attest to the accuracy performance of the proposed technique by taking in our experimental analysis two real cases of bushfire in Brazil.

This paper is organized as follows: Section 2 provides a brief review of image classification techniques, anomaly detection methods, and spectral indices for burning identification. Section 3 presents the conceptual development of

the proposed technique, including computational implementation details. The experimental results, as well as the description of the study areas, periods and data, are given in Section 4, while Section 5 summarizes our findings and concludes the work.

2 Theoretical Aspects and Background

2.1 Anomaly Detection as Classification

In an elementary point-of-view, a classifier is a function $F : \mathcal{X} \to \mathcal{Y}$ which assigns an element $\mathbf{x}$ from the attribute space $\mathcal{X}$ to a specific class in $\Omega = \{\omega_1, \omega_2, \ldots, \omega_c\}$, by setting a class indicator in a subset of natural numbers $\mathcal{Y} = \{1, 2, \ldots, c\}$. Under these conditions, if $\mathbf{x} \in \mathcal{X}$ and $y \in \mathcal{Y}$, $y = F(\mathbf{x})$ indicates that $\mathbf{x}$ belongs to the class ω_y.

The image classification task comprises the application of F on each pixel of image $\mathcal{I}$. More specifically, the image $\mathcal{I}$ is defined on a support $\mathcal{S} \subset \mathbb{N}^2$, where each pixel $s \in \mathcal{S}$ is assigned to a vector $\mathbf{x}$. Reciprocally, $\mathcal{I}(s) = \mathbf{x}$ may be used to denote the assignment between s and the vector $\mathbf{x}$, giving rise to the formula $\mathcal{C}(s) = \omega_y$ to express that $F(\mathbf{x}) = y$.

Basically, most image classification methods in the literature stand for different ways of defining $F : \mathcal{X} \to \mathcal{Y}$ and applying it on $\mathcal{I}$. Supervised and unsupervised are the most usual learning paradigms adopted to model F. In the supervised approach, the formulation of F demands taking a set of training samples whose class indicator is known in advance. On the other hand, in the unsupervised case, the classifier does not depend on pre-labeled training data so that the method accounts for automatically identifying patterns, clusters, and particular relationships over the data.

The anomaly detection task can be understood as a specific application of unsupervised classification which aims at detecting events of rare occurrences or incidents that may conflict with a set of observations [23]. One-Class Support Vector Machine [24] and Isolation Forest [25] are representative examples of anomaly detection techniques studied in the literature. Such approaches have been successfully applied to detect bank frauds, verify intruders into security systems, and support medical examinations [26]. In addition, anomaly detection techniques have been used as very useful tools to cope with environmental monitoring issues [27, 28].

2.1.1 One-Class Support Vector Machine

Support Vector Machine (SVM) is a very popular classification method in Remote Sensing applications. Its solid mathematical formulation, simple algorithmic architecture and generalization capability are a few of attractive features of this method [29, 30].

From the original conception of SVM method, different variants have been proposed, for example, the Laplacian [31], transductive [32], context sensitive

[29, 33] and "one-class" SVMs [24]. Particularly, the one-class SVM (OC-SVM) relies on the problem of quantile estimation for anomaly detection.

Conceptually, starting from a set of observations $\mathcal{I}$, the OC-SVM method provides an unsupervised model that distinguishes the attribute vectors $\mathbf{x}$ as part of a set of non-anomalous elements according to a probability ν of false positive occurrence. Formally, we can write $F: \mathcal{I} \subset \mathcal{X} \rightarrow \{+1, -1\}$, where the output $+1$ implies that the data input is in $\mathcal{I}$, and -1 otherwise. The definitive classifier, F, is given as follows:

$$F(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{m} \alpha_i K(\mathbf{x}, \mathbf{x}_i) - b\right), \tag{1}$$

where $b = \sum_{j=1}^{m} \alpha_j K(\mathbf{x}_i, \mathbf{x}_j)$ to some $\mathbf{x}_i \in \mathcal{I}$, and $K(\cdot, \cdot)$ is a kernel function.

The coefficients α_i, $i = 1, \ldots, m$, are obtained by solving the following Optimization Problem:

$$\begin{array}{l} \min\limits_{\alpha_1,\ldots,\alpha_m} \sum_{i,j=1}^{m} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \\ s.t. \begin{cases} \alpha_i \in [0, \frac{1}{\nu m}] \\ \sum_{i=1}^{m} \alpha_i = 1 \end{cases} \end{array} \tag{2}$$

It is worth noting that the OC-SVM is parameterized by $\nu \in [0, 1]$ as well as by other parameters related to the kernel function taken. Further details on kernel functions are discussed in [34].

2.1.2 Isolation Forest

Isolation Forest (IF) is one of the most recent methods applied to detect anomalies. Unlike other anomaly detection methods, the IF lies in the fact that it does not depend on distance measurements or density-based models to select anomalous features in the image [35]. More precisely, the IF focus on looking for isolate anomalies rather than profiling regular patterns (i.e., non-anomalies). The anomaly characterization is performed through a set of binary decision trees, denominated "isolation trees" (IT), which are components of ensemble-type classifiers , the so-called "isolation forest".

Two main properties are the basis of the IF classification method [23]: (i) anomalous instances are minority present in a dataset; and (ii) the data gather attribute values very different from those observed in regular instances. Furthermore, the IF method can be tuned by varying three major parameters: the number n_{it} of IT in the ensemble; the sub-sampling size m_{sub} of elements randomly drawn from the analyzed dataset; and the maximum depth d allowed for the IT growth.

Formally speaking, let $\mathcal{I}$ be a dataset whose elements $\mathbf{x}$ are defined on an n-dimensional space. We denote $\widetilde{\mathcal{I}}$ as the set resulting from a sub-sampling performed on $\mathcal{I}$ with m_{sub} elements. From $\widetilde{\mathcal{I}}$, a binary decision tree is constructed where each node T, and its associated data, can be subdivided into

the descendants T_l and T_r. Such division is accomplished by randomly choosing an attribute $q \in \{1, \ldots, n\}$ and a value p, which may assume m_{sub} possible values. The division is then performed when the dataset associated with the node is greater than one; there are different values w.r.t. the selected attribute and the binary tree has not yet reached the maximum depth d.

We now assume that the process leading to the construction of an IT is replicated n_{it} times under the condition of different $\widetilde{\mathcal{I}}$, with distinct q and p selected for its construction. Therefore, as mentioned, the set of IT originates the definitive IF model.

Once the IF is properly built, a given vector $\mathbf{x}$ is classified as anomaly according to a root-to-leaf path length $h(\mathbf{x})$ taken over each IT in the IF. Based on this concept, the following function is used to measure the anomaly level of $\mathbf{x}$:

$$R(\mathbf{x}, m_{sub}) = 2^{-\frac{E(h(\mathbf{x}))}{c(n)}}, \tag{3}$$

where $c(n) = 2\ln(n-1) + \zeta - \frac{2(n-1)}{n}$; with $\zeta = 0.577215664$ representing an approximation for the Euler's constant.

Therefore, whereas $R(\mathbf{x}, m_{sub}) \to 1$, $\mathbf{x}$ is characterized as an anomaly; otherwise, if $R(\mathbf{x}, m_{sub}) < 0.5$, the pattern $\mathbf{x}$ is considered regular; finally, if $R(\mathbf{x}, m_{sub}) \approx 0.5$, the whole set is classified as regular.

2.2 Remote Sensing, Spectral Indexes and Burn Detection

Remote Sensing plays a key role in the analysis of environmental data. Thanks to the advances promoted by this technology, several space programs have been launched, allowing the acquisition of large amounts of images from the Earth's surface with high temporal frequency and low operating cost [16, 36].

Concerning the development of innovative fire monitoring apparatus, advances have been made in terms of identify as well as quantifying the extent of burned areas [37, 38]. However, Ban et al. [36] point out the absence of robust and automated methods for this purpose.

In this sense, the use of spectral indices is a well-established, robust approach to get discriminative features and useful information for a desirable set of targets [39]. These include as targets the wildfire areas captured by a remotely sensed scene, so that the indices allow to assess the burn severity of these areas while still making use of low computational resources [40, 41] Some spectral indices traditionally used by the Remote Sensing community are the following: the Normalized Difference Vegetation Index (NDVI) [42], the Normalized Difference Water Index (NDWI) [43], and the Normalized Burn Ratio (NBR) [44].

Generally, spectral indices like NDVI and NBR combine information from visible to shortwave-infrared spectral bands (whose are sensitive to variations in color, soil composition, moisture, vegetation chlorophyll, etc.) in order to capture different characteristics for soil and vegetation areas affected by fire

[45]. Moreover, it is worth mentioning that due to its low-reflectance behavior, the NBR index may be unable to distinguish water bodies and burned areas. In order to circumvent this issue, the NDWI becomes an useful index to quantify areas not associated with fires.

In more mathematical terms, let $\mathcal{I}(s) = \mathbf{x}$ be an attribute vector positioned at pixel $s \in \mathcal{S}$, where $\mathbf{x}$ is formed by the reflectance intensities at green, red, near-infrared and shortwave-infrared wavelengths bands, x_{Green}, x_{Red}, x_{NIR} and x_{SWIR}, respectively. $\mathcal{I}_{\text{NDVI}}$, $\mathcal{I}_{\text{NDWI}}$ and $\mathcal{I}_{\text{NBR}}$ establish the representation of image $\mathcal{I}$ in terms of NDVI, NDWI and NBR spectral indices:

$$\mathcal{I}_{\text{NDVI}} = \frac{x_{\text{NIR}} - x_{\text{Red}}}{x_{\text{NIR}} + x_{\text{Red}}}, \quad \mathcal{I}_{\text{NDWI}} = \frac{x_{\text{Green}} - x_{\text{NIR}}}{x_{\text{Green}} + x_{\text{NIR}}}, \quad \mathcal{I}_{\text{NBR}} = \frac{x_{\text{NIR}} - x_{\text{SWIR}}}{x_{\text{NIR}} + x_{\text{SWIR}}}.$$

Notice that the above-described indices can be used as time-series images, as the difference between two images taken at different instants may allow identifying and mapping specific events. As stated by Sobrino et al. [46], most fire mapping methods that make use of Remote Sensing data have been proposed by assuming post-fire or pre-fire images. For example, the "Delta-NBR" methodology (ΔNBR), introduced by Key and Benson [44], was designed to numerically assess the fire severity on vegetation areas and it comprises the characterization of fire events as a function of the values obtained by:

$$\Delta\text{NBR} = \text{NBR}^{(\text{pre})} - \text{NBR}^{(\text{pos})}, \tag{4}$$

where $\text{NBR}^{(\text{pre})}$ and $\text{NBR}^{(\text{pos})}$ account for the NBR index computed at two distinct instants, before and after the fire event. In our approach, burning signs are determined by values above 0.1 [44]. A high value of ΔNBR implies an elevated level of severity for burning.

3 Fire Susceptibility Mapping

3.1 Computational Methodology

Now, we describe the main steps of the proposed methodology for mapping fire-prone areas by combining multitemporal Remote Sensing resources and unsupervised anomaly detection techniques. Figure 1 shows an overview of the main stages of our approach, i.e., (i) analysis configuration and data acquisition, (ii) design of the fire susceptibility mapping model, and (iii) generation of output results for analysis and validation purposes.

We start by delimiting the study areas and describing the sensor used in our analysis. More precisely, we take a sequence of images as multitemporal data, i.e., a set of remotely sensed images that were captured during three different periods.

Pre-fire period: comprises an image series taken before the fire occurred. The goal here is to compute the central tendency in each position of the study area, and posteriorly use such an information to generate the NBR index as a

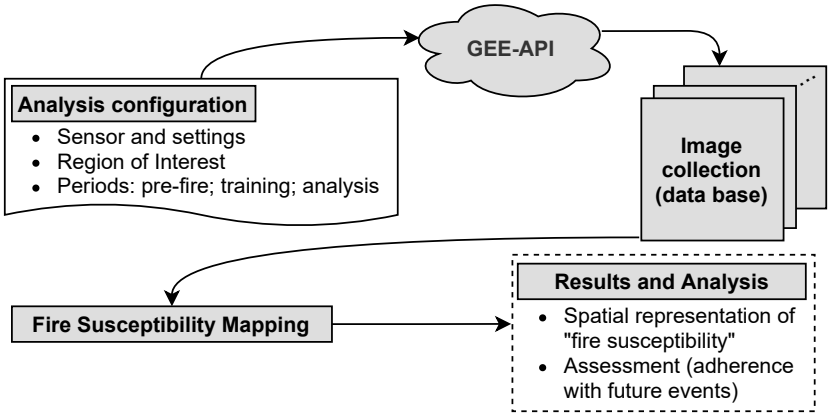

Fig. 1 Overview of the proposed methodology for mapping fire-susceptible areas.

benchmark before the presence of fire (i.e., $\text{NBR}^{(pre)}$), according to the ΔNBR method.

Training period: covers a time interval whose the data instances are exploited to identify fire events and, subsequently, building an anomaly detection model which learns the behavior of the fires immediately before they spread.

Analysis period: consists of the test period, which our trained anomaly detection model is applied to classify the fire-susceptible regions.

Once the collection of input images are collected, they are used to automatically build a training dataset so that a unsupervised fire identification model is generated (see Figure 2). First, we compute discriminative spectral maps from the pre-fire images. These include NDVI, NDWI, and NBR, where this latter is used to calculate the derivative spectral indices ΔNBR, when constructing our training dataset. Here, the number of instants is regulated by the parameter "time-lapse" (ℓ), expressed in days. Such parameter is introduced to arrange the data series in a regularly-spaced instant fashion. In our approach, if there is more than one image within the same time-lapse interval, we priority the most recent image. Moreover, areas affected by cloud/shadow occurrence are removed and recursively filled with the corresponding areas (i.e., not affected by cloud/shadow) taken from the immediately preceding instants.

The computation of ΔNBR is performed for the median of NBR as calculated in the "pre-fire" period. In this process, the observed NBR values in the training period are taken as "post-fire" (i.e., $\text{NBR}^{(\text{pos})}$). From the ΔNBR computed at each instant, if it exceeds a pre-fixed threshold $\tau \geq 0.1$ [1], then it is labeled as an event associated with the occurrence of fire. Aiming at minimizing false positives, the identified fire-affected areas is passed through the Modis Burned Area product [47] (with confidence above 95%) at respective instant. In this case, the NDVI, NDWI and NBR spectral indices are recorded at each fire-affected position, but with regard to the immediately previous instant. We then construct a fire-prone database in a fully unsupervised way so that the potential fire-related events are discriminated in terms of spectral indices. Finally, the

[1] Preliminary tests show that $\tau \geq 0.1$ provides a suitable mapping of fire-affected locations.

generated database is used to create an anomaly detection model by using OC-SVM or IF methods.

Concerning the application of the anomaly detection model, it is carried out on each image of the "analysis period", immediately after the computation of the spectral indices. Notice that, in our classification model, the identification of non-anomalous locations is interpreted as *areas that demonstrate fire susceptibility*. Finally, the percentage of anomaly occurrence w.r.t. the number of instants in the "analysis period" at each position into the region of interest generates the definitive fire susceptibility map.

When the percentage values tend to 100% in a given position, it means the such a location is similar to other fire-affected areas in all instants of the "analysis period". It is important to emphasize that, once the model is built from fire-affected areas, a regular behavior (i.e., non-anomaly) is assigned to elements with areas similar to the observed fire-affected regions. Reversely, if the percentage values tend to zero, the assigned location rarely behaves like a fire-affected target, inferring a small susceptibility to fire.

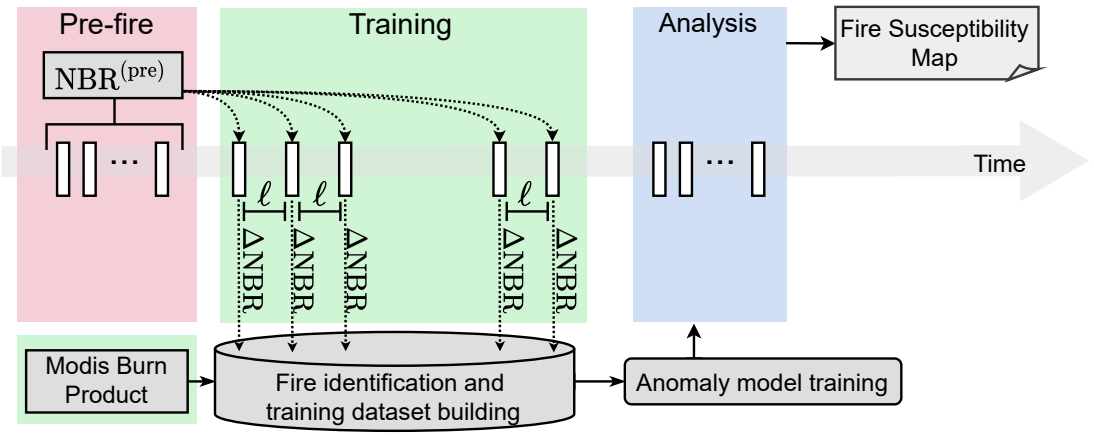

Fig. 2 Step diagram of the proposed method for susceptibility mapping.

3.2 Data Resources, Tools and Parameter Tuning

In this section, we focus on covering a few implementation aspects of our methodology as well as the data and tools used during our research.

To implement our computing prototypes, we use the Python 3.8 programming language [48]. Concerning data organization and processing, we take the Numpy [49] and Pandas [50] libraries. Finally, to train and build the OC-SMV and IF anomaly detection models, we run the Scikit-Learn [51] library.

In order to access the multitemporal Remote Sensing images, we adopt the Google Earth Engine Application Programming Interface (API) [52], which is compatible with Python language. This API favors access to high-performance computing resources for processing geospatial datasets and it allows the automation of the image search process for a given period and region of interest. The images used to define the pre-fire, training and analysis periods are taken from

the Operational Land Imager (OLI) sensor, on-board the Landsat-8 satellite, with 30 m of spatial resolution, 16 days of temporal resolution and surface reflectance information from ultra-blue to shortwave-infrared wavelengths. Concerning the ancillary Modis Burned Area data, used as support to identify the fire events and build the training dataset, it comprises a 500 m resolution product containing the monthly burned area per pixel according to a confidence value [47].

Moreover, the images selected to define the pre-fire, training and analysis periods are limited to 50% of cloud/shadow occurrence with respect to the area comprised by the region of interest (defined in the "analysis configuration"). The adopted percentage level was established after preliminary tests. The cloud/shadow detection is determined by using the so-called image band "pixel_qa". More details about this procedure are found in [53]. With respect the parameter "time lapse" (ℓ), once the current implementations focus on Landsat-8 images, we adopt $\ell = 15$ so as to consider only one image per each instant in the time series.

According to Section 2.1, the OC-SVM and IF methods require the tuning of parameters. Face with the high freedom degrees associated with the process of selecting appropriate parameters for the methods, we apply the well-established Grid Search [54] procedure to calibrate the definitive models. Basically, such a procedure consists of exhaustive tests, over a defined parameter space-search, the best set of parameters which ensures higher performance. For OC-SVM, the search space that determines the tested settings is given from the $\nu \in \{10^{-1}, 10^{-2}, \ldots, 10^{-7}\}$ value and a RBF kernel with $\gamma \in \{10^{-1}, 10^{-2}, \ldots, 10^{-7}\}$. Regarding the IF parameters, we take $n_{it} \in \{50, 100, 200\}$, $m_{sub} \in \{\sqrt{dim(\mathcal{X})}, 100\%, 75\%, 50\%\}$, and $d \in \{1, 2, \ldots, 30\}$. In addition, this procedure is replicated in the decision rules of the data set for each possible configuration, according to a 10-fold cross-validation process.

Finally, we freely provide the codes and data used to run our experiments at https://github.com/anonimous/FSM/.

4 Experiments and Results

4.1 Experiment Overview

Section 3 introduced the conceptualization and implementation of a fire susceptibility mapping method. We now assess the proposed method for different scenarios concerning two distinct study areas and epochs. As mentioned in Section 3.2, we take images acquired by the Operational Land Imager (OLI) sensor, on-board the Landsat-8 satellite to conduct the experiments. The obtained fire susceptibility maps are then assessed using the burn events identified by the Modis Burned Area product as reference (i.e., ground-truth) data. Moreover, for sake of representation and comparison, the fire susceptibility values are grouped into four classes: very low $[0, 0.25[$, mid-low $[0.25, 0.5[$, mid-high $[0.5, 0.75[$ and high $[0.75, 1]$.

Details about the study areas and periods are shown in Section 4.2. The obtained results and discussions are presented in Section 4.3. Furthermore, the experiments are performed using both OC-SVM and IF methods as anomaly detection models. The respective output mappings are compared in order to conclude about the most appropriate model in the context of the proposed method.

4.2 Study Areas, Data and Evaluated Periods

We apply the proposed approach in two distinct study areas and epochs. Figure 3 depicts the spatial location of such regions. The first study area (Area 1) comprises an area of Brazilian Amazon, i.e., a portion of São Félix do Xingu city, State of Pará, Brazil. The second area (Area 2) includes a portion of Cáceres city, State of Mato-Grosso, Brazil, which is another area of legal Amazon. Such areas have been place of recent burn events with distinct intensities.

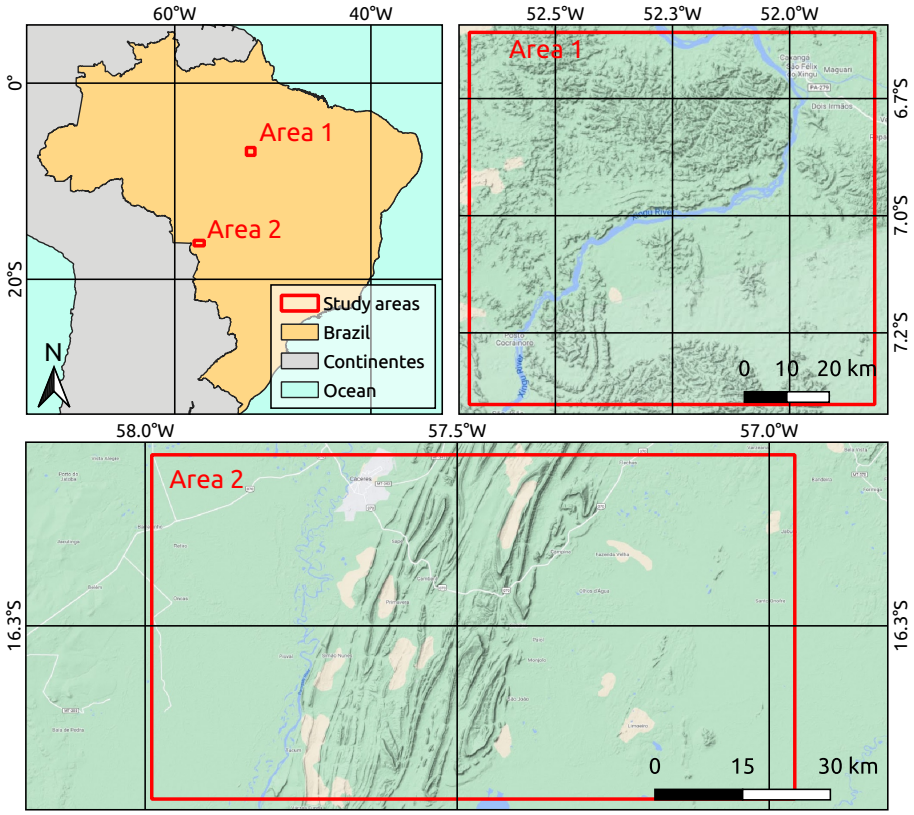

Fig. 3 Study areas location.

Multitemporal image series registered by the OLI sensor are used as input data to the proposed method. For sake of data exhibition, Figure 4 shows the median image computed from January 1st 2017 to December 31st 2019 regarding each study area.

The experiments comprise three distinct epochs: 2018, 2019, and 2020. Table 1 summarizes the epochs (I, II, and III) and the respective pre-fire, training, and analysis periods expressed as a function of the reference year Y. In

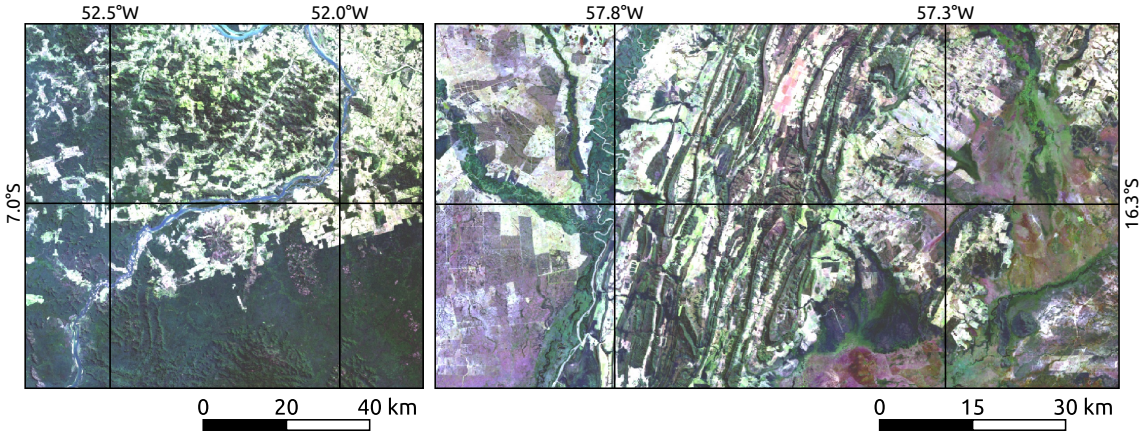

Fig. 4 The median images, based on the period January 1st 2017 to December 31st 2019, regarding Areas 1 (left) and 2 (right). Representations in natural color composition.

addition, the "assessment period" shown in Table 1 stands for the burned areas registered by the Modis Burned Area product over the period September 1st to December 31st Y, adopted to assess the fire susceptibility mappings. Figures 5 and 6 presents such reference dataset for Areas 1 and 2 at distinct epochs.

Table 1 Summary about the epochs and respective periods expressed as a function of the reference year Y.

Reference	**Training**	**Analysis**	**Assessment**
Jan. 1st (Y-3) to Dec. 31st (Y-1)	Jun. 1st (Y-1) to Mar. 31st Y	Jul. 1st Y to Aug. 31st Y	Sep. 1st Y to Dec. 31st Y
Epochs	I	II	III
Reference year (Y)	2018	2019	2020

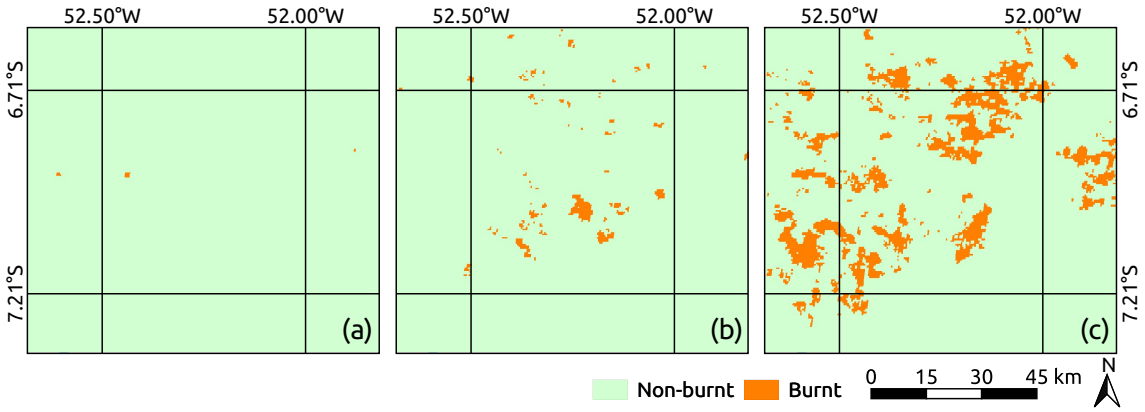

Fig. 5 Fire-affected areas according to Modis Burn Area product for Area 1. The assessment periods regarding the epochs I, II and III are represented by the maps (a), (b) and (c), respectively.

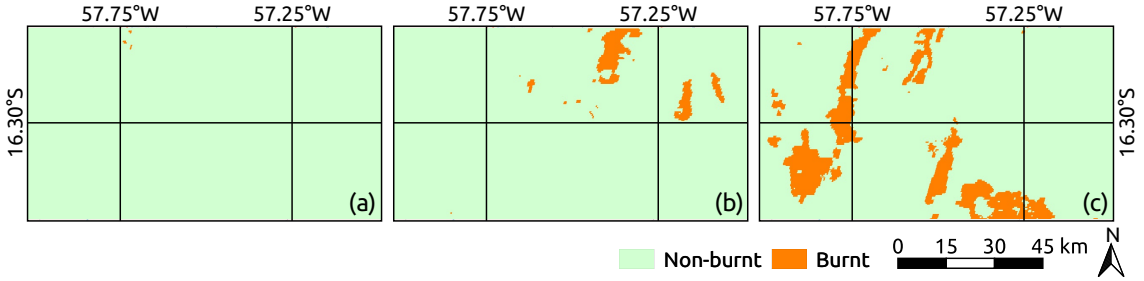

Fig. 6 Fire-affected areas according to Modis Burn Area product for Area 3. The assessment periods regarding the epochs I, II and III are represented by the maps (a), (b) and (c), respectively.

4.3 Results and Discussions

In consonance with the experiment overview presented in Section 4.1, from Figures 7 to 12, we present the computed fire susceptibility maps. More precisely, Figures 7 and 8 comprise the obtained maps for Areas 1 and 2 at epochs I, II and III, in which the proposed method was equipped with IF as the anomaly detection model. Similarly, Figures 9 to 10 embraces the results by employing the OC-SVM as the anomaly detection model.

In a qualitative point of view, one can notice that the IF method provides more regularized mapping (i.e., less "isolate" pixels) when compared against OC-SVM method. Figures 9(b) and 10(b) endorse the previous results when applying the OC-SVM classifier. Targets with no historic of burn events are allocated in the low susceptibility class (i.e., $[0, 0.25[$). Usually, such targets are related to water bodies or low biomass portions. This pointed out behavior reveals consistency in the proposed method.

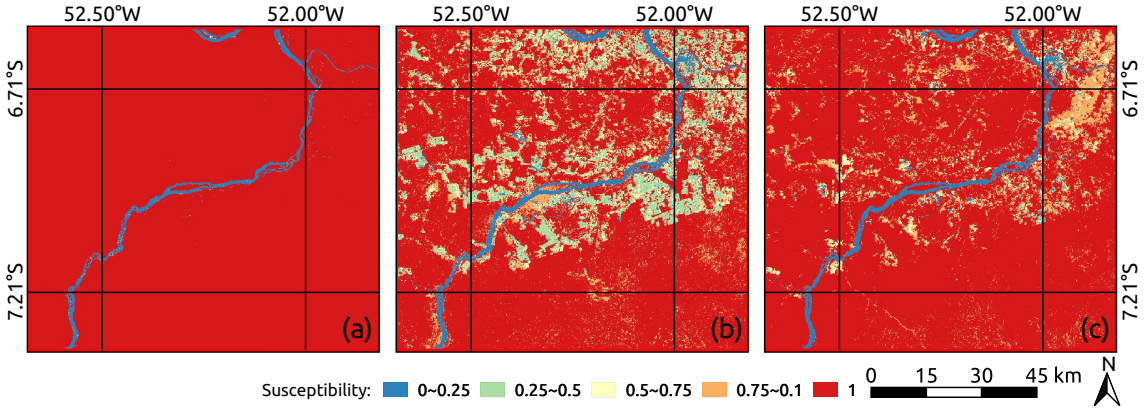

Fig. 7 Fire susceptibility maps for Area 1 using the IF method as anomaly detection model. Sub-figures (a), (b) and (c) refers to mappings based on the "analysis period" of epochs I, II and III, respectively.

Posterior, the adherence of presented fire susceptibility maps are analyzed in comparison with post-fire events registered by the Modis Burned Area product (Figures 5 and 6). As previously mentioned, only the burned locations are

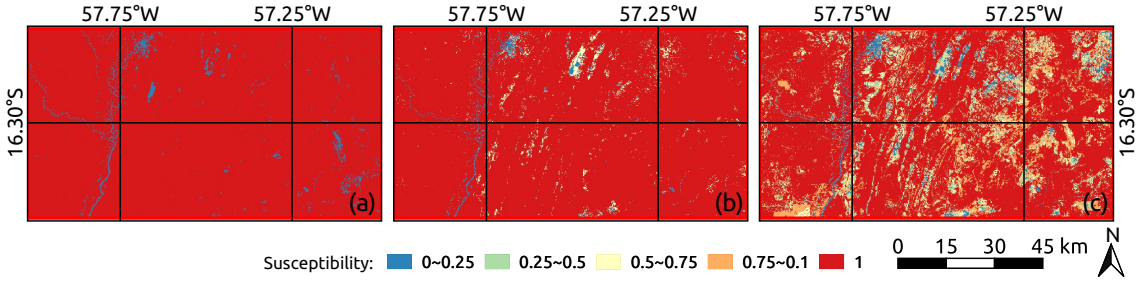

Fig. 8 Fire susceptibility maps for Area 3 using the IF method as anomaly detection model. Sub-figures (a), (b) and (c) refers to mappings based on the "analysis period" of epochs I, II and III, respectively.

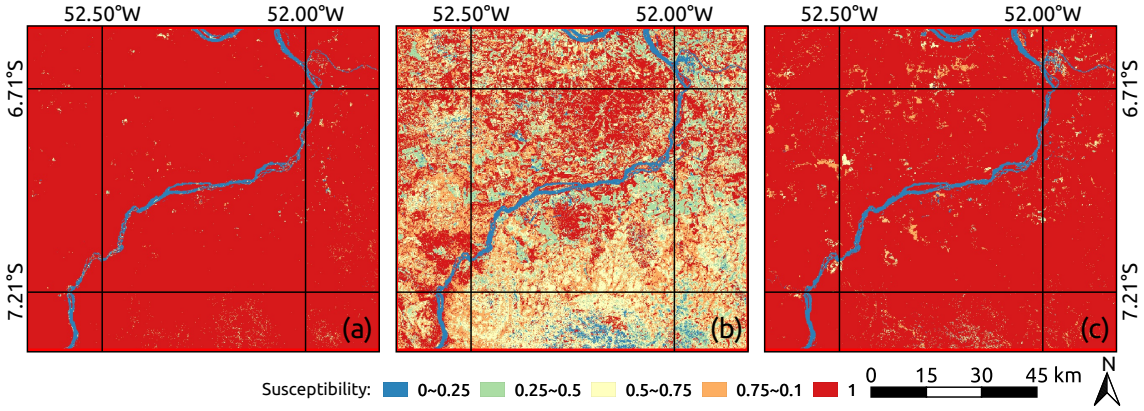

Fig. 9 Fire susceptibility maps for Area 1 using the OC-SVM method as anomaly detection model. Sub-figures (a), (b) and (c) refers to mappings based on the "analysis period" of epochs I, II and III, respectively.

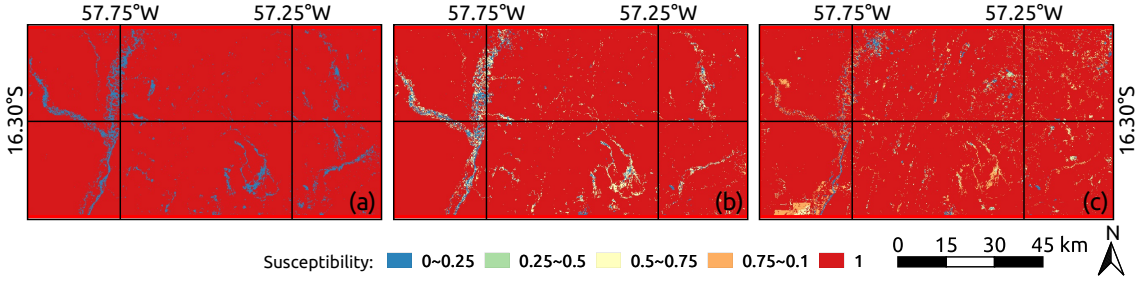

Fig. 10 Fire susceptibility maps for Area 3 using the OC-SVM method as anomaly detection model. Sub-figures (a), (b) and (c) refers to mappings based on the "analysis period" of epochs I, II and III, respectively.

considered in our adherence analysis, where the expected outputs should stand for high susceptibility values.

The plots of Figures 11 and 12 present the observed frequencies for each susceptibility class when using the IF- and OC-SVM-based models. Both set of results are quite similar to each other. As expected, the fire-affected areas (according to the adopted reference dataset – Figs. 5-6) are usually assigned to high susceptibility levels (i.e., the class $[0.75, 1.0]$).

In order to demonstrate the significance level of the results (i.e., mid-high and high – $[0.5, 1.0]$) in comparison to the low susceptibilities frequencies (i.e., low

and mid-low $[0, 0.5[$), a single-tailed (unilateral) hypothesis test for population proportion [55] was computed. Under a significance of 1%, the statistical tests reveal that the higher susceptibilities occurs in proportion above 75%.

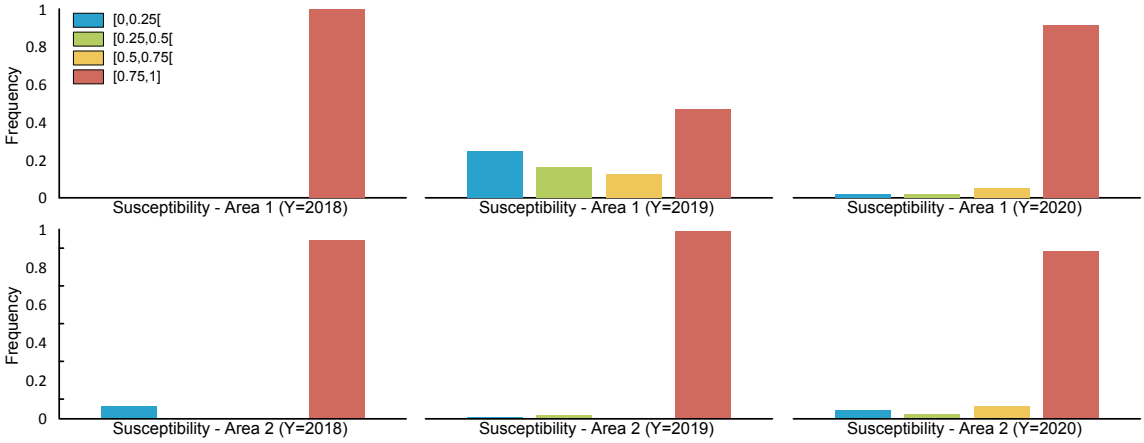

Fig. 11 Observed susceptibility values, regarding the fire-affected areas in the reference dataset for each study area and epoch, using the IF method as anomaly detection model.

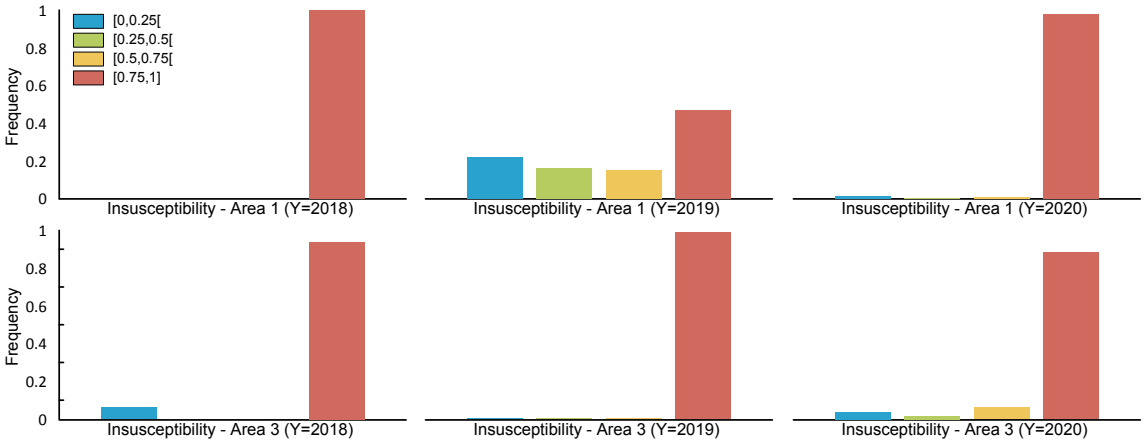

Fig. 12 Observed susceptibility values, regarding the fire-affected areas in the reference dataset for each study area and epoch, using the OC-SVM method as anomaly detection model.

5 Conclusions

In this paper, we developed a new methodology for mapping the fire susceptibility in forest areas using multitemporal Remote Sensing images and unsupervised anomaly detection methods. We demonstrate the effectiveness and accuracy of the proposed approach by quantitatively as well as qualitatively analyzing two case studies covering areas heavily affected by fire in the legal Amazon forest.

From the obtained results, we showed that the current methodology was capable of assigning a high fire-susceptibility level to the real fire-damaged areas, by comparing the resulting maps with reference maps representing the

fire occurrences in period posterior to those adopted to infer the mentioned susceptibility maps. Concerning the IF and OCSVM as unsupervised anomaly detection methods, we found that IF delivered more regularized, accurate maps than OCSVM.

Future perspectives for this study include: (i) test other anomaly detection models; (ii) analyze the use of other spectral indices in the anomaly detection model building; (iii) apply the proposed method in other study areas; (iv) include a confidence value for the output maps; (v) extend the implementation to other multispectral sensors such as Sentinel-2 MSI and Terra/Aqua Modis; (vi) evaluate the proposed method through adaptation to other environmental issues, such as deforestation, floods, oil spills, melting glaciers, etc.

References

[1] Artés, T., Oom, D., De Rigo, D., Durrant, T.H., Maianti, P., Libertà, G., San-Miguel-Ayanz, J.: A global wildfire dataset for the analysis of fire regimes and fire behaviour. Scientific data **6**(1), 1–11 (2019)

[2] Caúla, R., Oliveira-Júnior, J.F., Lyra, G.B., Delgado, R., Heilbron Filho, P.: Overview of fire foci causes and locations in brazil based on meteorological satellite data from 1998 to 2011. Environmental Earth Sciences **74**(2), 1497–1508 (2015)

[3] Cochrane, M.A., Barber, C.P.: Climate change, human land use and future fires in the amazon. Global Change Biology **15**(3), 601–612 (2009)

[4] Garcia, L.C., Szabo, J.K., de Oliveira Roque, F., Pereira, A.d.M.M., da Cunha, C.N., Damasceno-Júnior, G.A., Morato, R.G., Tomas, W.M., Libonati, R., Ribeiro, D.B.: Record-breaking wildfires in the world's largest continuous tropical wetland: Integrative fire management is urgently needed for both biodiversity and humans. Journal of environmental management **293**, 112870 (2021)

[5] INPE: Instituto Nacional de Pesquisas Espaciais- Banco de Dados de queimadas. Accessed: 2021-03-29 (2021). http://www.inpe.br/queimadas/bdqueimadas

[6] Prestes, N.C.C.d.S., Massi, K.G., Silva, E.A., Nogueira, D.S., de Oliveira, E.A., Freitag, R., Marimon, B.S., Marimon-Junior, B.H., Keller, M., Feldpausch, T.R.: Fire effects on understory forest regeneration in southern amazonia. Frontiers in Forests and Global Change **3**, 10 (2020)

[7] Field, C.B., Barros, V., Stocker, T.F., Dahe, Q.: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, ??? (2012)

[8] Brando, P.M., Balch, J.K., Nepstad, D.C., Morton, D.C., Putz, F.E., Coe, M.T., Silvério, D., Macedo, M.N., Davidson, E.A., Nóbrega, C.C., *et al.*: Abrupt increases in amazonian tree mortality due to drought–fire interactions. Proceedings of the National Academy of Sciences **111**(17), 6347–6352 (2014)

[9] Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., *et al.*: A large and persistent carbon sink in the world's forests. Science **333**(6045), 988–993 (2011)

[10] Birch, D.S., Morgan, P., Kolden, C.A., Abatzoglou, J.T., Dillon, G.K., Hudak, A.T., Smith, A.M.: Vegetation, topography and daily weather influenced burn severity in central idaho and western montana forests. Ecosphere **6**(1), 1–23 (2015)

[11] Mann, M.L., Batllori, E., Moritz, M.A., Waller, E.K., Berck, P., Flint, A.L., Flint, L.E., Dolfi, E.: Incorporating anthropogenic influences into fire probability models: Effects of human activity and climate change on fire activity in california. PLoS One **11**(4), 0153589 (2016)

[12] Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A.-X., Xu, C.: Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. the case of dayu county, china. Science of the total environment **630**, 1044–1056 (2018)

[13] Pourghasemi, H.R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., Cerda, A.: Assessing and mapping multi-hazard risk susceptibility using a machine learning technique **10**(1), 1–11

[14] Sausen, T.M., Lacruz, M.S.P.: Sensoriamento Remoto Para Desastres. Oficina de Textos, ??? (2015)

[15] de Oliveira, L.V., Negri, R.G., Santos, L.B.L.: Análise de técnicas de detecção de mudança para mapeamento de desastres com uso de dados de sensoriamento remoto. Revista Brasileira de Cartografia **72**(1), 177–189 (2020)

[16] Pereira, A.A., Pereira, J., Libonati, R., Oom, D., Setzer, A.W., Morelli, F., Machado-Silva, F., De Carvalho, L.M.T.: Burned area mapping in the brazilian savanna using a one-class support vector machine trained by active fires. Remote Sensing **9**(11), 1161 (2017)

[17] Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., *et al.*: Deep learning in environmental remote sensing: Achievements and challenges. Remote Sensing of Environment **241**, 111716

(2020)

[18] Martínez-Álvarez, F., Tien Bui, D.: Advanced Machine Learning and Big Data Analytics in Remote Sensing for Natural Hazards Management. Multidisciplinary Digital Publishing Institute (2020)

[19] Hart, E., Sim, K., Kamimura, K., Meredieu, C., Guyon, D., Gardiner, B.: Use of machine learning techniques to model wind damage to forests. Agricultural and forest meteorology **265**, 16–29 (2019)

[20] Leuenberger, M., Parente, J., Tonini, M., Pereira, M.G., Kanevski, M.: Wildfire susceptibility mapping: Deterministic vs. stochastic approaches. Environmental Modelling & Software **101**, 194–203 (2018)

[21] Dickson, B.G., Prather, J.W., Xu, Y., Hampton, H.M., Aumack, E.N., Sisk, T.D.: Mapping the probability of large fire occurrence in northern arizona, usa **21**(5), 747–761

[22] Kamalakannan, J., Chakrabortty, A., Bothra, G., Pare, P., Kumar, C.P.: Forest fire prediction to prevent environmental hazards using data mining approach. In: Proceedings of the 2nd International Conference on Data Engineering and Communication Technology, pp. 615–622. Springer

[23] Akoglu, L., Tong, H., Koutra, D.: Graph based anomaly detection and description: a survey. Data mining and knowledge discovery **29**(3), 626–688 (2015)

[24] Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. Neural computation **13**(7), 1443–1471 (2001)

[25] Liu, F.T., Ting, K.M., Zhou, Z.: Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422 (2008). https://doi.org/10.1109/ICDM.2008.17

[26] Gu, S., Hu, Q., Cheng, Y., Bai, L., Liu, Z., Xiao, W., Gong, Z., Wu, Y., Feng, K., Deng, Y., *et al.*: Application of organic fertilizer improves microbial community diversity and alters microbial network structure in tea (camellia sinensis) plantation soils. Soil and Tillage Research **195**, 104356 (2019)

[27] Dereszynski, T.G. Ethan W e Dietterich: Spatiotemporal models for data-anomaly detection in dynamic environmental monitoring campaigns. ACM Transactions on Sensor Networks (TOSN) **8**(1), 1–36 (2011)

[28] Havens, T.C., Bezdek, J.C.: An efficient formulation of the improved visual assessment of cluster tendency (ivat) algorithm. IEEE Transactions on

Knowledge and Data Engineering **24**(5), 813–822 (2011)

[29] Bruzzone, L., Persello, C.: A novel context-sensitive semisupervised svm classifier robust to mislabeled training samples. IEEE Transactions on Geoscience and Remote Sensing **47**(7), 2142–2154 (2009)

[30] Mountrakis, G., Im, J., Ogole, C.: Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing **66**(3), 247–259 (2011)

[31] Gu, Y., Feng, K.: Optimized laplacian svm with distance metric learning for hyperspectral image classification. IEEE journal of selected topics in applied earth observations and remote sensing **6**(3), 1109–1117 (2013)

[32] Li, Y., Wang, Y., Bi, C., Jiang, X.: Revisiting transductive support vector machines with margin distribution embedding. Knowledge-Based Systems **152**, 200–214 (2018)

[33] Negri, R.G., Dutra, L.V., Sant'Anna, S.J.S.: An innovative support vector machine based method for contextual image classification. ISPRS journal of photogrammetry and remote sensing **87**, 241–248 (2014)

[34] Shawe-Taylor, J., Cristianini, N., *et al.*: Kernel Methods for Pattern Analysis. Cambridge university press, ??? (2004)

[35] Liu, F.T., Ting, K.M., Zhou, Z.-H.: Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data (TKDD) **6**(1), 1–39 (2012)

[36] Ban, Y., Zhang, P., Nascetti, A., Bevington, A.R., Wulder, M.A.: Near real-time wildfire progression monitoring with sentinel-1 sar time series and deep learning. Scientific Reports **10**(1), 1–15 (2020)

[37] Alonso-Canas, I., Chuvieco, E.: Global burned area mapping from envisat-meris and modis active fire data. Remote Sensing of Environment **163**, 140–152 (2015)

[38] Giglio, L., Loboda, T., Roy, D.P., Quayle, B., Justice, C.O.: An active-fire based burned area mapping algorithm for the modis sensor. Remote sensing of environment **113**(2), 408–420 (2009)

[39] Kumar, A., Bhandari, A.K., Padhy, P.: Improved normalised difference vegetation index method based on discrete cosine transform and singular value decomposition for satellite image processing. IET Signal Processing **6**(7), 617–625 (2012)

[40] Schepers, L., Haest, B., Veraverbeke, S., Spanhove, T., Vanden Borre,

J., Goossens, R.: Burned area detection and burn severity assessment of a heathland fire in belgium using airborne imaging spectroscopy (apex). Remote Sensing **6**(3), 1803–1826 (2014)

[41] Veraverbeke, S., Harris, S., Hook, S.: Evaluating spectral indices for burned area discrimination using modis/aster (master) airborne simulator data. Remote Sensing of Environment **115**(10), 2702–2709 (2011)

[42] Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., *et al.*: Monitoring vegetation systems in the great plains with erts. NASA special publication **351**(1974), 309 (1974)

[43] Gao, B.-C.: Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment **58**(3), 257–266 (1996)

[44] Key, C.H., Benson, N.C.: Landscape assessment (la). In: Lutes, Duncan C.; Keane, Robert E.; Caratti, John F.; Key, Carl H.; Benson, Nathan C.; Sutherland, Steve; Gangi, Larry J. 2006. FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. LA-1-55 **164** (2006)

[45] Tran, B.N., Tanase, M.A., Bennett, L.T., Aponte, C.: Evaluation of spectral indices for assessing fire severity in australian temperate forests. Remote Sensing **10**(11), 1680 (2018)

[46] Sobrino, J.A., Llorens, R., Fernández, C., Fernández-Alonso, J.M., Vega, J.A.: Relationship between soil burn severity in forest fires measured in situ and through spectral indices of remote detection. Forests **10**(5), 457 (2019)

[47] USGS: MODIS/Terra+Aqua Burned Area Monthly L3 Global 500 m SIN Grid. Accessed: 2021-03-29 (2021). https://lpdaac.usgs.gov/products/mcd64a1v006/

[48] Van Rossum, G., Drake, F.L.: The Python Language Reference Manual. Network Theory Ltd., ??? (2011)

[49] Van Der Walt, S., Colbert, S.C., Varoquaux, G.: The numpy array: a structure for efficient numerical computation. Computing in science & engineering **13**(2), 22–30 (2011)

[50] McKinney, W., *et al.*: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56 (2010). Austin, TX

[51] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., *et al.*: Scikit-learn: Machine learning in python. the Journal of machine Learning research **12**, 2825–2830 (2011)

[52] GEE-API: Google Earth Engine API. Accessed: 2021-11-22 (2021). https://github.com/google/earthengine-api

[53] 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 **202**, 18–27 (2017)

[54] LaValle, S.M., Branicky, M.S., Lindemann, S.R.: On the relationship between classical grid search and probabilistic roadmaps. The International Journal of Robotics Research **23**(7-8), 673–692 (2004)

[55] Devore, J.L.: Probability and Statistics for Engineering and the Sciences. Cengage Learning, ??? (2011)

unesp

Rogerio Galante Negri <rogerio.negri@unesp.br>

MESE-D-22-00065 - Submission Confirmation

1 mensagem

Modeling Earth Systems and Environment <em@editorialmanager.com> 19 de janeiro de 2022 15:06
Responder a: Modeling Earth Systems and Environment <jebamalar.jayapal@springer.com>
Para: Rogério Galante Negri <rogerio.negri@unesp.br>

Dear Prof. Dr. Negri,

Thank you for submitting your manuscript, "Anomaly-Driven Approach for Forest Fire Susceptibility Mapping Using Multitemporal Remote Sensing Data", to Modeling Earth Systems and Environment

The submission id is: MESE-D-22-00065

Please refer to this number in any future correspondence.

During the review process, you can keep track of the status of your manuscript by accessing the Editorial Manager Website.

Your username is: rogeriogalante
If you forgot your password, you can click the 'Send Login Details' link on the EM Login page at https://www.editorialmanager.com/mese/

With kind regards
Springer Journals Editorial Office
Modeling Earth Systems and Environment

This letter contains confidential information, is for your own use, and should not be forwarded to third parties.

Recipients of this email are registered users within the Editorial Manager database for this journal. We will keep your information on file to use in the process of submitting, evaluating and publishing a manuscript. For more information on how we use your personal details please see our privacy policy at https://www.springernature.com/production-privacy-policy. If you no longer wish to receive messages from this journal or you have questions regarding database management, please contact the Publication Office at the link below.

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: https://www.editorialmanager.com/mese/login.asp?a=r). Please contact the publication office if you have any questions.

2.2 *Fire Detection with Multitemporal Multispectral Data and a Probabilistic Unsupervised Technique*

O artigo a seguir foi submetido em 2 de fevereiro de 2022 no periódico *IEEE Transactions on Geoscience and Remote Sensing*.

Fire Detection with Multitemporal Multispectral Data and a Probabilistic Unsupervised Technique

Andréa E. O. Luz, Rogério G. Negri, Alejandro C. Frery, *IEEE Senior Member*, Maurício. A. Dias

***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 and classification. The spatial resolution of the mapping provided is equivalent to the multispectral data used. In order to evaluate the introduced proposal, we carry out case studies on regions located in Brazil and Bolivia that present a recurrent history of forest fires. Furthermore, images obtained by the Landsat-8 and Sentinel-2 satellites are used in these case studies and comparisons with an alternative method are included in the analyses. The results obtained show the high adherence between the mappings generated by the introduced method and the reference data.**

***Index Terms*—Remote Sensing, forest fires, spectral index, multitemporal, unsupervised mapping.**

I. Introduction

CLIMATE CHANGE is one of the most significant environmental challenges facing humanity. Occasionally, these changes in climate increase the probability of extreme events that alter the rainfall and drought patterns [23, 35] in several regions of the planet, besides triggering the process of natural disasters arising mainly from floods, landslides, droughts, and forest fires [36].

Extreme droughts are among the worst climatic disasters since they produce large economic and social losses [1, 30]. Moreover, these phenomena also cause forest fires [8], an environmental issue that has been observed over the years, leading to severe ecological damages [28]. Frequent fires are hazardous because they affect forest regeneration, making fire a disruptive agent of this environment's dynamics [29]. Consequently, fire mapping and monitoring are essential to identify, mitigate and even prevent harmful impacts.

In this perspective, different techniques have been developed and applied to Remote Sensing images to extract data from the Earth's surface and identify fire-affected areas [2, 15]. According to Birch et al. [6], the use of time series of images obtained by Remote Sensing is essential for studies about monitoring, management, and recovering of burned areas.

Among the different approaches found in the literature, spectral indices have helped identify fire-affected areas. Escuin et al. [9] measure the severity of fire events using the Normalized Difference Vegetation Index (NDVI) and the Normalized Burn Ratio (NBR). Similarly, Liu [22] and Zhao et al. [42] express the burning severity and to estimate the vegetation recovery in the post-fire instants with the Differenced Normalized Burn Ratio (dNBR).

Machine Learning methods provide a powerful alternative to detect and map fire-affected areas [21]. Ban et al. [3] used neural networks as a primary tool for the automatic identification of burned areas. In a simple but efficient way, Matricardi et al. [24] propose the use of probabilistic models to estimate canopy cover and fire potential in forest areas.

Face with the presented discussions, this study introduces a novel method for fire-affected areas mapping using multitemporal Remote Sensing images and probabilistic models.

Overall, we use multitemporal Remote Sensing images to identify targets variations related to the Normalized Burn Ratio index. Such variations are turned into a fire-affected map by an unsupervised processing which uses logistic regression and spatial smoothing with Markov Random Fields modeling.

In order to demonstrate the effectiveness of the proposed method, case studies were carried out in regions of Brazil and Bolivia using Landsat-8 and Sentinel-2 images. The selected study areas have been placed frequent fires events in recent years.

Contributions. In summary, the main contributions of this paper are the following:

- An accurate unsupervised method designed to map fire affected areas in multitemporal image series;
- The proposed method can be applied to data acquired by sensors with distinct spatial radiometric resolutions;
- The proposed method is modular and, therefore, flexible regarding the use of other models, in addition to those presented in the following formalization;
- A conceptual formalization that, after convenient changes and adaptations, can be applied to other remote sensing studies beyond the mapping of fire-affected areas.

This article is organized as follows: Section II provides a brief review of fundamental concepts and techniques required for the proposed method, which is formalized in Section III; Section IV presents the application of the proposed method

A. E. 0. Luz is with the Graduate Program in Natural Disasters, Sciences and Technology Institute, São Paulo State University (UNESP) and Center for Monitoring and Early Warning of Natural Disasters (CEMADEN), São José dos Campos, São Paulo, Brazil. e-mail: andrea.luz@unesp.br

R. G. Negri is with the Department of Environmental Engineering, Sciences and Technology Institute, São Paulo State University (UNESP), São José dos Campos, São Paulo, Brazil. e-mail: rogerio.negri@unesp.br

A. C. Frery is with the School of Mathematics and Statistics, Victoria University of Wellington (VUW), Wellington, New Zealand. e-mail: alejandro.frery@vuw.ac.nz

M. A. Dias is with Department of Mathematics and Computer Science, School of Sciences and Technology, São Paulo State University (UNESP), Presidente Prudente, São Paulo, Brazil. e-mail: ma.dias@unesp.br

Manuscript received XX YY, 20ZZ; revised WW U.U., 20VV.

to map fire affected areas in two distinct areas using different sensors; Finally, section V concludes this article.

II. Background

A. Definitions and notation

Let $\mathcal{I}^{(t)}$ be an image defined on the support $\mathcal{S} \subset \mathbb{N}^2$ acquired at an instant t, $\mathcal{X}$ be the set of possible values at each position, and $\boldsymbol{x}_i^{(t)} = \mathcal{I}^{(t)}(s_i)$ the observation $\boldsymbol{x}_i^{(t)} \in \mathcal{X}$ at position $s_i \in \mathcal{S}$ of $\mathcal{I}^{(t)}$.

In the remote sensing context, the components of $\boldsymbol{x}_i^{(t)}$ are the values measured by the sensor, or derived features, over a specific Earth surface position.

B. Spectral indexes for burn detection

Spectral indices allow the extraction and analysis of remotely sensed data. In the face of a feature of interest, a spectral index can assist in its identification. This approach is also essential due to the impossibility of modifying orbiting imaging sensors and the difficulty of obtaining field data with the same spatial and temporal resolution [39]. Generally speaking, spectral indices are derived from algebraic operations on the attributes of $\mathcal{X}$ that characterize the behavior of $\boldsymbol{x}_i^{(t)}$ at every pixel s_i of $\mathcal{I}^{(t)}$.

Among dozens of spectral indices in the literature [41], typical examples are the Normalized Difference Vegetation Index (NDVI) [32], the Normalized Difference Water Index (NDWI) [12], and the Soil Adjusted Vegetation Index (SAVI) [18].

Usually, such indices combine the information from visible, near-infrared, and mid-infrared spectral bands, being sensitive to variations in color, composition and soil moisture, and vegetation chlorophyll. The Normalized Burn Ratio (NBR) [19] is a convenient index for identifying areas affected by fire.

Consider an image $\mathcal{I}^{(t)}$ with $\text{NIR}_i^{(t)}$ and $\text{SWIR}_i^{(t)}$ as components of the attribute vector $\boldsymbol{x}_i^{(t)}$, which represents the target behavior at position s_i, with respect the near and short-wave infrared wavelengths. The image of NBR values is defined as:

$$\mathcal{I}_{\text{NBR}}^{(t)}(s_i) = \frac{\text{NIR}_i^{(t)} - \text{SWIR}_i^{(t)}}{\text{NIR}_i^{(t)} + \text{SWIR}_i^{(t)}}, \text{ for every } s_i \in \mathcal{S}. \quad (1)$$

Spectral indices may be used at a single time t or as a time series. In the latter case, the use of differences between instants may help identifying and mapping specific events. When the interest is mapping burned areas through remote sensing images, the approaches found in the literature use only post-fire images or combine pre- and post-fire data [33]. In particular, Key and Benson [19] propose the "Delta-NBR" feature (ΔNBR), aimed at measuring the severity of fire events on the vegetation. Such methodology comprises a characterization of fire severity through the following difference:

$$\Delta\text{NBR} = \mathcal{I}_{\text{NBR}}^{(\text{pre})} - \mathcal{I}_{\text{NBR}}^{(\text{post})}, \quad (2)$$

where $\mathcal{I}_{\text{NBR}}^{(\text{pre})}$ and $\mathcal{I}_{\text{NBR}}^{(\text{post})}$ are the NBR index computed at two distinct instants, before and after the fire event. Burning signs are characterized by values above 0.1. Table I displays a fire severity categorization also proposed in Ref. [19].

TABLE I
Burning severity categorization

Severity	**ΔNBR** ($\times 10^3$)
Enhanced Regrowth, High	-500 to -251
Enhanced Regrowth, Low	-250 to -101
Unburned	-100 to 99
Low Severity	100 to 269
Moderate-Low Severity	270 to 439
Moderate-High Severity	440 to 659
High Severity	660 to 1300

Although the DNBR offers good perspectives regarding the fire severity effects on the landscapes, it is important to highlight that atmospheric interference, or even variations in the limits of NIR and SWIR spectral bands, may require adjustments in the burn severity categories (Table I). Another limitation stems from the intrinsic features of each ecosystem, that may impose significant mismatches between the severity classes and those intervals [38].

We avoid the aforementioned limitations by building a binary map that incorporates both radiometric-derived and contextual information. The former enters through a logistic regression (Section II-C); the latter by using the ICM algorithm (Section II-D).

C. Logistic model

The Logistic Regression has been widely used to estimate the probability of membership in "element-classes" in cases of binary association [17]. It acts as a class indicator provided a set of observations $\mathcal{D} = \{(\boldsymbol{x}_i, y_i) \in \mathcal{X} \times \mathcal{Y} : i = 1, \ldots, m\}$, where $\mathcal{Y} = \{0, 1\}$.

This regression belongs to the class of Generalized Linear Models [26], in which the response y_i follows a Bernoulli law with probability of success

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \frac{1}{1 + \exp\{-[\mathbf{1}, \boldsymbol{x}] \cdot \boldsymbol{\theta}^T\}}, \quad (3)$$

where $\boldsymbol{\theta} = [\theta_0, \theta_1, \ldots, \theta_n]$ is the parameter vector with one dimension more than the input vector $\boldsymbol{x}$, and $\mathbf{1}$ is a column of ones. The behavior of $f(\boldsymbol{x}; \boldsymbol{\theta})$, due to its sigmoid structure, has an image defined in the range $[0, 1]$.

Beyond regression purposes, the model $f(\boldsymbol{x}; \boldsymbol{\theta})$ may also be applied for binary classification tasks. Eq. (4) represents the classification result $\mathcal{M}$ resulting from applying f on each position s_i of $\mathcal{I}$:

$$\mathcal{M}(s_i) = \begin{cases} 0 & \text{if } f(\boldsymbol{x}_i; \boldsymbol{\theta}) < 0.5, \\ 1 & \text{if } f(\boldsymbol{x}_i; \boldsymbol{\theta}) \geq 0.5. \end{cases} \quad (4)$$

where the values 0 and 1 are class indicators.

The parameter $\boldsymbol{\theta}$ that characterizes f can be estimated by minimizing the following objective function [4]:

$$J(\boldsymbol{\theta}) = -\frac{1}{m} \sum_{i=1}^{m} \big[y_i \log f(\boldsymbol{x}_i; \boldsymbol{\theta}) + (1 - y_i) \log \big(1 - f(\boldsymbol{x}_i; \boldsymbol{\theta})\big) \big]. \quad (5)$$

Although the derivative of (5) with respect to $\boldsymbol{\theta}$ does not have explicit form, its convexity grants that optimization algorithms such as the Descending Gradient converge to a global minimum [20].

D. Iterated Conditional Modes

The ICM algorithm is a classification technique that incorporates both the site evidence and the local information [5]. It is based on modeling the former with a suitable marginal distribution, and the latter with a Markov random field for the underlying (unobserved) classes. It consists of solving the following maximization problem [11]:

$$\boldsymbol{x}_i \text{ is assigned to } \omega_j \Leftrightarrow$$
$$j = \underset{k=1,\ldots,c}{\arg\max} \left\{ P(\boldsymbol{x}_i \mid \omega_k) + \beta \sum_{\eta \in W} \sum_{t \in \mathcal{V}(i)} \big(1 - \delta(y_i, y_t)\big) \right\}, \tag{6}$$

where ω_k denotes $y_i = k$, $\beta \geq 0$ weights the influence of neighboring pixels (the context), η is a clique[1] that contains i, W is the set of cliques, $\mathcal{V}(i)$ is the coordinate set of neighboring pixels of s_i, and

$$\delta(y_a, y_b) = \begin{cases} 1 & \text{if } y_a = y_b, \\ 0 & \text{otherwise} \end{cases}$$

is the Kronecker delta function. Eq. (6) is solved iteratively until a convergence criterion is met. The parameter β is estimated, at each iteration, by maximum pseudolikelihood using the previous classification as input; see details in Ref. [11]. In the limit $\beta = 0$, the solution of (6) is the maximum likelihood classification map, while when $\beta \to \infty$ the observations are discarded in favor of the most frequent class in the neighborhood.

III. Method Proposal

A. Conceptual formalization

Figure 1 depicts a high-level overview of the proposed method for fire-affected area detection.

1) Multitemporal NBR computing: The red block outlines how the NBR is computed. Initially, assume there is a reference image $\mathscr{I}^{(\text{Ref})}$ expressing the usual behavior of targets regarding a spatial domain $\mathscr{S}$. Also, let $\mathscr{I}^{(t)}$, for $t = 1, 2, \ldots, k$, be a multitemporal image series also regarding $\mathscr{S}$. Given $s_i \in \mathscr{S}$, denote $\boldsymbol{x}_i^{(t)} = [x_{i1}^{(t)}, \ldots, x_{in}^{(t)}] \in \mathscr{X}$, for $t \in \{\text{Ref}, 1, 2, \ldots, k\}$ as an attribute vector with measures registered at n spectral wavelengths. Using the appropriate components of $\boldsymbol{x}_i^{(t)}$, it is possible to compute the NBR index at each position of $\mathscr{S}$, and then determine $\mathscr{I}_{\text{NBR}}^{(t)}$.

[1] A clique of an undirected graph $\mathscr{G} = (\mathscr{A}, \mathscr{B})$ is a set of vertices $\bar{\mathscr{A}} \subseteq \mathscr{A}$ such for all $i, j \in \bar{\mathscr{A}}$, there is an edge between i and j.

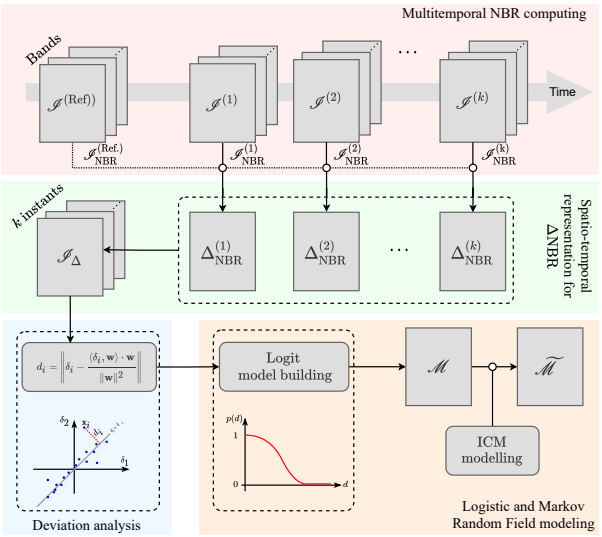

Fig. 1. Proposed method workflow.

2) Spatio-temporal representation for ΔNBR: The ΔNBR comprises a derived spectral index, obtained by the difference of NBR at "pre" and "post" instants, in order to distinguish burn-affected or recovered burnt vegetation from other targets. In this sense, fixing $\mathscr{I}_{\text{NBR}}^{(\text{Ref})}$ as a "pre-fire event" image, we define $\Delta_{\text{NBR}}^{(t)} = \mathscr{I}_{\text{NBR}}^{(\text{Ref})} - \mathscr{I}_{\text{NBR}}^{(\text{t})}$, for $t = 1, 2, \ldots, k$. We denote $\Delta_{\text{NBR}}^{(t)}(s_i) = \delta_i^{(t)} \in \mathbb{R}$. Following the temporal order $t = 1, 2, \ldots, k$ for the computed $\Delta\text{NBR}^{(t)}$ are combined to determine $\mathscr{I}_\Delta$. In such representation, for a given s_i, we have $\mathscr{I}_\Delta(s_i) = \boldsymbol{\delta}_i = [\delta_i^{(1)}, \ldots, \delta_i^{(k)}]$. These operations are represented in the "Spatio-temporal representation for ΔNBR" block (Fig. 1, green box).

3) Deviation analysis: It is worth observing that $\mathscr{I}_\Delta$ is a spatio-temporal representation for the ΔNBR values. Locations s_i without changes due to burns or recovering tends to show similar $\delta_i^{(t)}$ values independently of t. Reversely, the absence of a significant correlation between two or more components of $\boldsymbol{\delta}_i$ is evidence of a temporal change.

The "point-to-line" distance is way of quantifying the correlation between the components of $\boldsymbol{\delta}_i$, where $\boldsymbol{\delta}_i$ is the "point", and the "line" is the geometric place where $\delta_i^{(t)} = \delta_i^{(\ell)}$ with $t, \ell = 1, \ldots, k$. This process is summarized in the "Deviation analysis" block (Fig. 1, blue box)

Using Linear Algebra concepts, for a generic vector $\boldsymbol{\delta}_i$, the mentioned distance, herein denoted as d_i, is computed by:

$$d_i = \left\| \boldsymbol{\delta}_i - \frac{\langle \boldsymbol{\delta}_i, \mathbf{w} \rangle \, \mathbf{w}}{\|\mathbf{w}\|^2} \right\|, \tag{7}$$

where $\mathbf{w} = [1, 1, \ldots, 1] \subset \mathbb{R}^k$ is the vector director for the line used as reference (i.e., where $\delta_i^{(t)} = \delta_i^{(\ell)}$); $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ are the norm and the inner product in an Euclidean vector space. The submission of $\boldsymbol{\delta}_i$ assigned to each $s_i \in \mathscr{S}$ into (7) gives place to a "deviation image" $\mathscr{I}_D$.

In the context of this proposal, assuming that small distances for d_i represent no/irrelevant changes at s_i over time, such position should represent non-burned or non-recently-burned areas. However, as consequence of the Hughes' phenomenon [13],

the dimension (k) of $\boldsymbol{\delta}_i$ plays a direct influence on the values of d_i since it express a distance measure.

Based on this premise, an initial upper-bound value Ξ is defined according to d_i values related to $\boldsymbol{\delta}_i$ vectors whose components are limited by a tolerance ξ. Such a margin acts as a parameter in the proposed method (and has a physical meaning according to the NBR spectral index, as discussed in Ref. [19]):

$$\Xi = \max\Big\{d_i : \max_{j=1,\ldots,k}\big\{\delta_i^{(j)}\big\} \leq \xi\Big\}. \tag{8}$$

For convenience, denote $D(\xi) = \big\{d_i : \max_{j=1,\ldots,k}\big\{\delta_i^{(j)}\big\} \leq \xi\big\}$.

Similarly, it is possible to define an initial lower-bound Λ for the d_i values related to positions that may represent a relevant change as consequence of a burn event. A convenient definition for Λ is:

$$\Lambda = \Xi + \alpha\sigma_{\Xi}, \tag{9}$$

where σ_{Ξ} is the observed standard deviation of $D(\xi)$ and $\alpha \in \mathbb{R}_+^*$ is an parameter that controls the separation/transition from irrelevant-to-relevant changes.

According to the above-presented discussions, the derived information from the original multitemporal image series allows discriminating between affected (burnt) and unaffected locations over the analyzed period. The upper (Ξ) and lower (Λ) bound values are the critical elements to determine for building a model able to express the probability of temporal changes as a consequence of a burning event.

Denote $\mathscr{D} = \{(d_i, y_i) \in \mathbb{R} \times \{0,1\} : i = 1,\ldots,m\}$ the reference set, where $y_i = 0$ or 1 when $d_i \geq \Lambda$ or $d_i \leq \Xi$, respectively. It is worth observing that, while $y_i = 1$ represents a non-burning event at s_i in the analyzed period (i.e., irrelevant changes), $y_i = 0$ stands for relevant changes due to burning. Consequently, we may observe that Ξ and Λ define a margin between irrelevant and relevant changes. The intuitive assignment "change-is-burn" is a consequence of the initial stages where the NBR and ΔNBR values give place to $\mathscr{I}_{\Delta}$ and then $\mathscr{D}$.

4) Logistic regression and MRF modeling: With basis on $\mathscr{D}$, a logistic regression model (Section II-C) $f(d;\boldsymbol{\theta})$ is built. In the light of (4), such model classifies each position $s_i \in \mathscr{S}$ and results in a map $\mathscr{M}$ of either irrelevant or relevant changes. In the sequence, we incorporate the spatial context by using the map $\mathscr{M}$ and the probabilities interpreted from $f(d;\boldsymbol{\theta})$ as input for the ICM algorithm (Section II-D). With this, we generate the regularized map $\widetilde{\mathscr{M}}$. All the presented discussions, from the definition of the $\mathscr{D}$ dataset, to the logistic regression modeling and ICM application, are encompassed by the "Logistic and Markov Random Field Modeling" block (Fig. 1, yellow box).

In the following experiments and discussions, the proposed method is referenced MRF+UFD since it combines MRF concepts into an unsupervised fire detection (UFD) approach.

B. Implementation details

The previous formalization is implemented in two stages, namely: (i) data acquisition and pre-processing; and (ii) model application and processing. The code of the proposed framework is freely available at https://github.com/anonymous/project.

In the first stage, we used Python 3.8 [37] and the Application Programming Interface for Google Earth Engine (API-GEE) [16] to access and obtain the multitemporal image series. Additionally, we relied on functions provided in Pandas [27] and GDAL [40] libraries for data manipulation. More specifically, the remote sensing images $\mathscr{I}^{(t)}$, $t = 1,\ldots,k$ (Landsat-8 OLI and Sentinel-2, see Section IV-B) were obtained after defining a given period and region of interest. Only images with less than $50\,\%$ of cloud/shadow cover are considered. Also, the cloud/shadow-affected regions are removed by a simple masking operation and replaced with the median information computed from an immediate past period. Preliminary tests indicated that six months comprise a sufficiently long history.

The reference image $\mathscr{I}^{(\text{Ref})}$ is the median image computed in a second past period regarding the same region of interest. We then compute the NBR spectral index for each image (i.e., $\mathscr{I}_{\text{NBR}}^{(t)}$, $t \in \{\text{Ref}, 1,\ldots,k\}$) followed by the calculus of Delta-NBR index at instant t as $\Delta_{\text{NBR}}^{(t)} = \mathscr{I}_{\text{NBR}}^{(\text{Ref})} - \mathscr{I}_{\text{NBR}}^{(t)}$, for $t = 1,\ldots,k$.

Regarding the second stage, we implemented the processes comprised of Eqs. (7) to (9), including the Logit model and ICM algorithm, in the Interactive Data Language (IDL) v. 8.8 [10]. We verified the overall performance with values of α (Eq. (9)) in the grid $\alpha \in \{0.01, 0.05, 0.1, 0.25, 0.5, 1.0, 1.5, 2.0\}$ for different datasets. Although the results did not change dramatically, $\alpha = 0.5$ produced the best outcomes.

IV. Experiments

A. Experiment overview

The following subsections present applications on the mapping of fire-affected areas based on our proposal. Such applications include different areas and sensors. Further details are presented in Section IV-B.

For comparison purposes, the proposed method will be compared with the classical approach called ΔNBR (Section II-B) extended to multitemporal data series. This multitemporal (MT) extension of ΔNBR method is herein denoted as MTDNBR. In this case, locations whose value of $\mathscr{I}_{\Delta\text{NBR}}^{(t)}$ surpasses ξ for some instant $t = 1,\ldots,k$ are taken as "occurrence of fire". Formally, let $\mathscr{M}^{\star}$ be the mapping derived from this approach, the occurrence or not of fire at position s_i is identified by

$$\mathscr{M}^{\star}(s_i) = \begin{cases} 1 & \text{if } \mathscr{I}_{\Delta\text{NBR}}^{(t)} < \Xi;\ t = 1,\ldots,k \\ 0 & \text{otherwise.} \end{cases}$$

In the following applications, motivated by the minimum threshold for fire occurrence (i.e., the "low severity" category) presented in Table I, both MRF+UFD and MTDNBR assume $\xi = 0.1$ as a parameter to map the fire-affected areas.

The results are evaluated based on adherence with estimates of burnt areas derived from the Moderate-Resolution Imaging Spectroradiometer (MODIS) sensor. We employ the overall accuracy [34], F1-Score [31], kappa coefficient, and the

variance of kappa [7] as accuracy measures, as well as True/False Positive (T.P.; F.P.) and True/False Negative (T.N.; F.N.) rates to analyze and compare the performance of the methods. The Matthews correlation coefficient (MCC) [25] is also employed as a complementary measure to compare the obtained results with a reference map of burn events.

We used a computer with an Intel Intel i-7 processor (8 core, 3.5 GHz), and 16 GB of RAM running the Ubuntu Linux version 20.04 operating system.

B. Study areas and data description

In order to evaluate the proposed MRF+UFD method and compare it to the MTDNBR approach to map fire-affected areas, applications involving two study areas and different sensors were carried out. Figure 2 illustrates the study areas location. The first study area (Area 1) comprises a region of São Félix do Xingu city, state of Pará, Brazil. The other study area (Area 2) includes a region on the border between Brazil and Bolivia. Both regions have a recurrent history of forest fires between August and October.

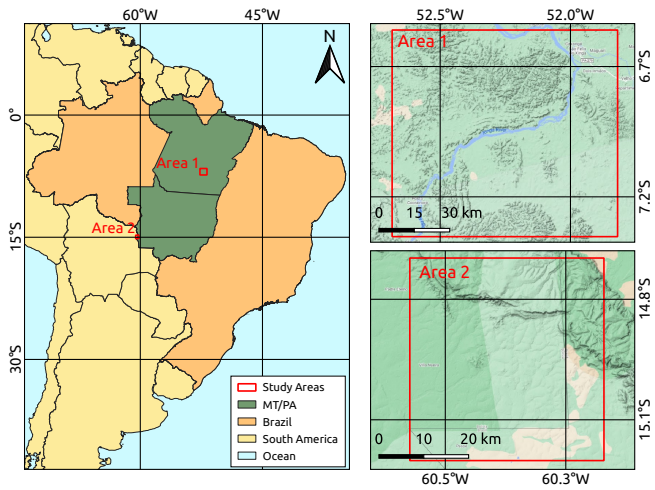

Fig. 2. Study areas location.

Were used images obtained by the Operating Land Imager (OLI – 30 m spatial resolution of size 883 × 809 pixels), and Multispectral Instrument (MSI – 10 m, 660 × 689 pixels) sensors onboard the Landsat-8 and Sentinel-2 satellites, respectively. Once known the tendency of fire incidence in these regions, all images available from August 1st to October 31st 2020, but limited to 50 % of cloud/shadow coverage over the study area, were considered in the analyses. Table II presents the dates of the images that meet these criteria.

TABLE II
DATES OF THE IMAGES CONSIDERED IN THE EXPERIMENTS.

		August	September	October
Area 1	Days	13; 29	14	
Area 2		10; 25	9; 24	9

Regarding the reference image (i.e., $\mathcal{J}^{(\mathrm{Ref})}$, Section III) required by the UFD+MRF method, we adopted the median image from all images acquired by the respective sensors between November 1st 2019 and June 30th 2020, a period characterized by the low occurrence of fires in the regions. Figures 3 and 4 illustrate images of the study areas according to the reference image and the latter considered instant.

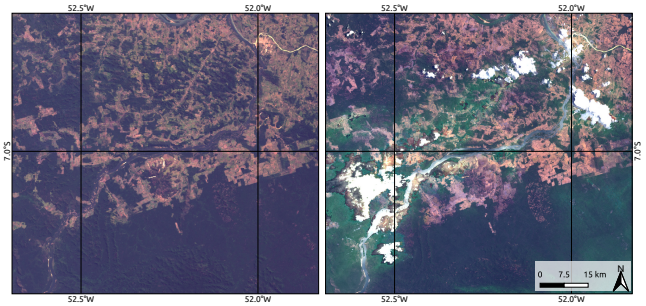

Fig. 3. Reference image (left) and the latter instant (right – September 14th, 2020) for Area 1. Representations in natural color composition.

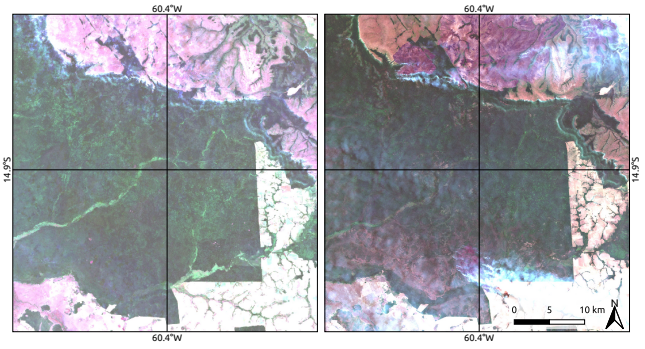

Fig. 4. Reference image (left) and the latter instant (right – October 9th, 2020) for Area 2. Representations in natural color composition.

In order to assess the adherence between the results provided by the analyzed methods and the expected outputs, estimates of burned areas from the product "MCD64A1 MODIS/Terra+Aqua Burned Area Monthly" [14] (500 m of spatial resolution) are employed as reference data. Based on this product, positions/pixels in the study area where the indicative of fire shows confidence superior to 90 % are taken as a "burn" event; otherwise, they are considered "non-burn".

Despite the broad difference of spatial resolution between the MCD64A1 product and the OLI and MSI images, it should be stressed that such a product allows verifying the agreement between the mappings generated by the analyzed methods and the expected results. Consequently, the assessments are limited to verifying if the frequencies and spatial distribution of mapped fire-affected areas are compatible with the expectations. Moreover, the use of MCD64A1 product as reference dataset comes from the absence of multitemporal data collected in situ.

C. Results

Prior to applying the MTDNBR and UFD+MRF methods, for data inspection purposes, figures 5 and 6 show the behavior

of the NBR index concerning the reference images at selected instants in each study area.

With respect to Area 1 (Figure 5), it is possible to notice that, in comparison to the reference image, changes in NBR values occur progressively from August to October. For simplicity of identification, the analysis of this study area and period will be named "Case I".

In Area 2, the most significant changes are only highlighted at the last instant (i.e., October, 9th 2020). This behavior motivates the evaluation of the methods in this area of study according to the August-September and August-October periods, thus defining the "Cases II and III", respectively.

The mappings resulting from the application of the UFD+MRF and MTDNBR methods in each of the established cases (I, II and III) are shown in figures 7 and 8.

The mappings performed on Area 1 verify the consistency of the proposed method (Fg. 7(b)) in relation to the expected result (Fg. 7(a)). On the other hand, the mapping generated by the MTDNBR method (Fg. 7(a)) demonstrates less robustness regarding the identification of areas not affected by the fire.

Regarding the mappings achieved in Area 2, the MTDNR method (Figs. 8(e) and (f) – Cases II and III) fails again when trying to separate the fire occurrences from the reference outputs (Figs. 8(a) and (b)).

Differently from the MTDNBR method, UFD+MRF shows consistency in cases where there are practically no fire outbreaks (i.e., Case II) as well as localized occurrences (i.e., Case III), as illustrated in Figures 8(c) and (d), respectively.

Figure 9 shows the amounts of True/False Positive/Negative accounted for in each of the study areas and cases. As already noted, MTDNBR is prone to errors errors, i.e., FP. The proportions of errors (FP+FN) committed by the proposed method are significantly lower than the competitor MTDNBR.

Table III confirms the superiority of the proposed method.

TABLE III
ACCURACY VALUES SUMMARY.

Case	Method	OA	F1-Score	MCC	kappa	var.kappa
I	UFD+MRF	0.871	0.564	0.492	0.489	1.5×10^{-5}
	MTDNBR	0.575	0.390	0.271	0.190	2.1×10^{-6}
II	UFD+MRF	0.999	0.006	0.007	0	2.8×10^{-6}
	MTDNBR	0.390	0.001	0.007	0	1.9×10^{-9}
III	UFD+MRF	0.889	0.615	0.596	0.557	2.6×10^{-5}
	MTDNBR	0.507	0.409	0.261	0.165	2.1×10^{-6}

It is worth observing that the low values of the F1-Score, MCC and kappa measures in Case II are mainly due to the imbalance between the number of burn samples concerning "non-burn" samples. However, in this situation, the OA measure justifies the accuracy of the UFD+MRF method.

Regarding Cases I and III, the superiority of the proposed method is proven once again. The significance between the results achieved by the UFD+MRF method proved by applying the hypothesis test derived from the kappa coefficient [7], whose resulting p-values are inferior to 10^{-4} in all comparisons. The adherence between the estimated and expected results is also measured through the MCC coefficient, which shows that the MRF+UFD results agree more with the reference than those provided by TDNBR.

Finally, it is essential to point out that the proposed method requires a higher computational cost when compared to the MTDNBR approach. If not including the computational time spent in the data series acquisition phase, which is common in both methods, the UFD+MRF method spends $316.8\,\mathrm{s}$, $211.9\,\mathrm{s}$ and $199.6\,\mathrm{s}$ for completing Cases I, II and III, respectively. The estimation of the parameters of the Logit model, obtained by the Descending Gradient strategy, and the application of the ICM algorithm are responsible for the additional computational time. MTDNBR requires less than $0.0017\,\mathrm{s}$ in all cases, since this procedure involves simple thresholding according to a defined ξ value.

V. CONCLUSIONS

We proposed an unsupervised approach for identifying fire-affected areas with multitemporal remote sensing data. Spectral index, logistic regression, and Markov Random Field modeling are the main concepts that support this proposal. The central idea lies in measuring the progressive deviations over time from a fixed reference image and then modeling these deviations as a probability function able to map locations affected by fire in some instant of the time series.

Two study cases, considering distinct regions and remote sensors, were carried out to assess the proposed method. Comparisons with a conventional approach extended to multi-temporal data were also included in the experiments.

The results showed a high adherence between the proposed method and a coarse resolution product derived from the Modis sensor for fire mapping purposes. Reversely, the concurrent method was not able to deliver consistent results.

The proposed method's robustness comes from using a deviation measure computed over time, which is posteriorly submitted to a probabilistic treatment to include the contextual information. On the other hand, the concurrent method (MTDNBR), as an alternative to mapping fire-affected areas in temporal series with basis on the ΔNBR idea, is highly subject to errors due to the use of fixed thresholds as well as by the targets' spectral variation over time.

Although the proposed method demonstrates higher computational run-time when compared to the simple threshold-based competitor, it is worth highlighting that it still comprises a light-cost algorithm.

As future work, we plan to: (i) adapt and investigate the use of the proposed framework on other environmental issues, like deforestation, flooding, oil spills, glaciers melting, etc. (ii) analyze other probability modeling approaches; (iii) include a confidence measure concerning the resulting mappings; (iv) extend the proposed method beyond binary outputs.

ACKNOWLEDGMENT

The authors acknowledge the support from FAPESP (Grants 2021/01305-6).

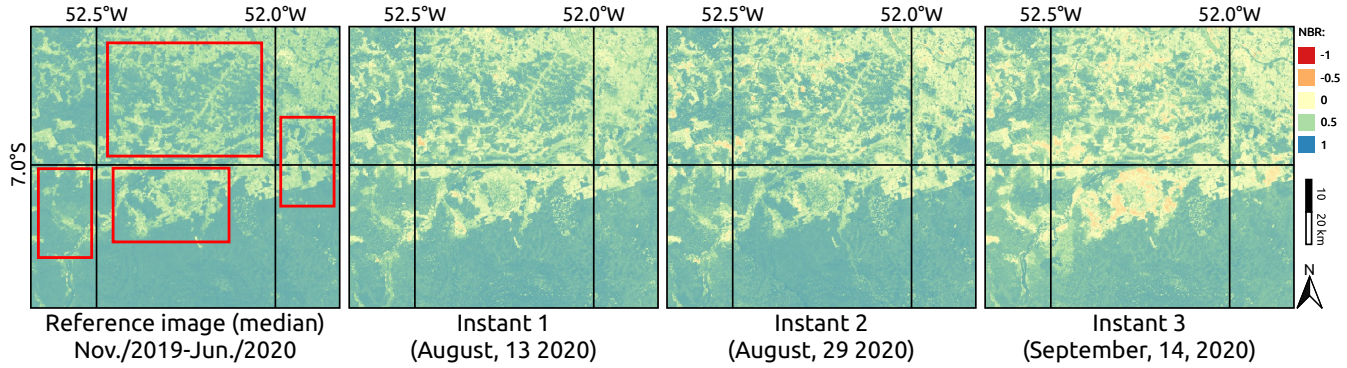

Fig. 5. The NBR behavior for the reference image and individual instants regarding the Area 1. Regions delimited by red rectangles highlight places with potential changes over time.

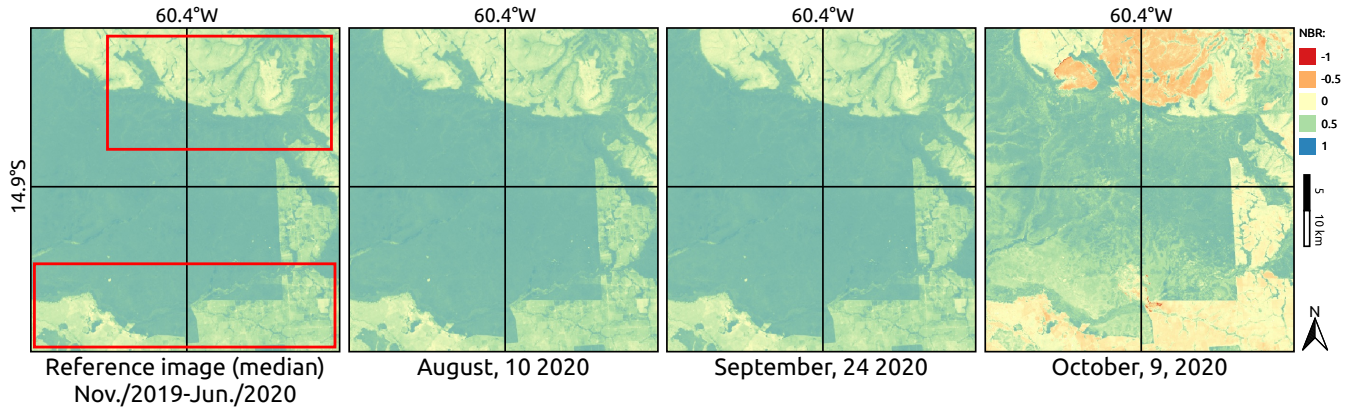

Fig. 6. The NBR behavior for the reference image and individual instants regarding the Area 2. Regions delimited by red rectangles highlight places with potential changes over time.

REFERENCES

[1] A. AghaKouchak, "A multivariate approach for persistence-based drought prediction: Application to the 2010–2011 East Africa drought," *Journal of Hydrology*, vol. 526, pp. 127–135, 2015.

[2] I. Alonso-Canas and E. Chuvieco, "Global burned area mapping from ENVISAT-MERIS and MODIS active fire data," *Remote Sensing of Environment*, vol. 163, pp. 140–152, 2015.

[3] Y. Ban, P. Zhang, A. Nascetti, A. R. Bevington, and M. A. Wulder, "Near real-time wildfire progression monitoring with Sentinel-1 SAR time series and deep learning," *Scientific Reports*, vol. 10, no. 1, pp. 1–15, 2020.

[4] R. A. Berk, *Statistical learning from a regression perspective*, 2nd ed., ser. Springer texts in statistics. Switzerland: Springer, 2016.

[5] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. B-48, pp. 259–302, 1986.

[6] D. S. Birch, P. Morgan, C. A. Kolden, J. T. Abatzoglou, G. K. Dillon, A. T. Hudak, and A. M. Smith, "Vegetation, topography and daily weather influenced burn severity in central Idaho and western Montana forests," *Ecosphere*, vol. 6, no. 1, pp. 1–23, 2015.

[7] R. G. Congalton and K. Green, *Assessing the accuracy of remotely sensed data: principles and practices*, 3rd ed. CRC press, 2019.

[8] S. S. da Silva, P. M. Fearnside, P. M. L. de Alencastro Graça, I. F. Brown, A. Alencar, and A. W. F. de Melo, "Dynamics of forest fires in the southwestern Amazon," *Forest Ecology and Management*, vol. 424, pp. 312–322, 2018.

[9] 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*, vol. 29, no. 4, pp. 1053–1073, 2008.

[10] Exelis, "Idl – interactive data language, version 8.8," Exelis Visual Information Solutions, Boulder, Colorado, 2021. [Online]. Available: https://www.l3harrisgeospatial.com/docs/using_idl_home.html

[11] A. C. Frery, A. H. Correia, and C. C. Freitas, "Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3098–3109, 2007.

[12] B. Gao, "NDWI – a normalized difference water index for remote sensing of vegetation liquid water from space," *Remote sensing of environment*, vol. 58, no. 3, pp. 257–266, 1996.

[13] A. Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2017.

[14] L. Giglio, L. Boschetti, D. Roy, A. A. Hoffmann, M. Humber, and J. V. U. o. M. Hall, *Collection 6 MODIS Burned Area Product User's Guide (Version 1.3)*, National Aeronautics and Space Administration, 2020. [Online]. Available: https://lpdaac.usgs.gov/documents/875/MCD64_User_Guide_V6.pdf

[15] L. Giglio, T. Loboda, D. P. Roy, B. Quayle, and C. O. Justice, "An active-fire based burned area mapping algorithm for the MODIS sensor," *Remote sensing of environment*, vol. 113, no. 2, pp. 408–420, 2009.

[16] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote sensing of Environment*, vol. 202,

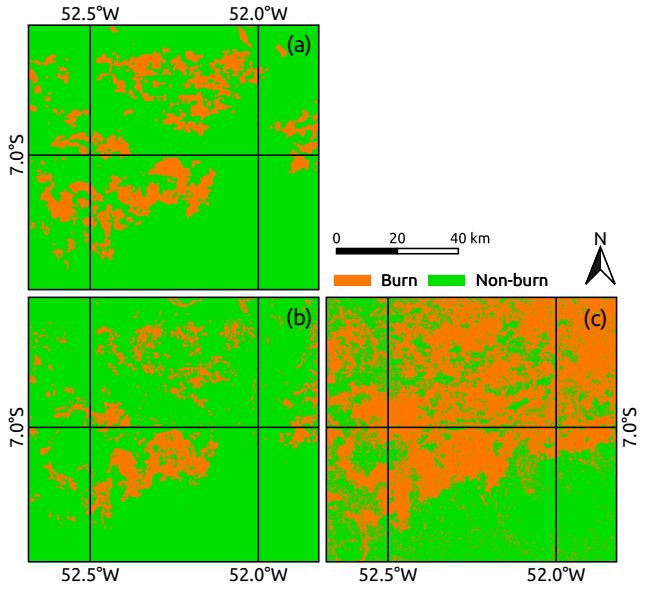

Fig. 7. The mapping of fire affected locations provided by the (b) UFD+MRF and (c) MTDNBR methods for Area 1/Case I. The reference mapping is shown in subfigure (a).

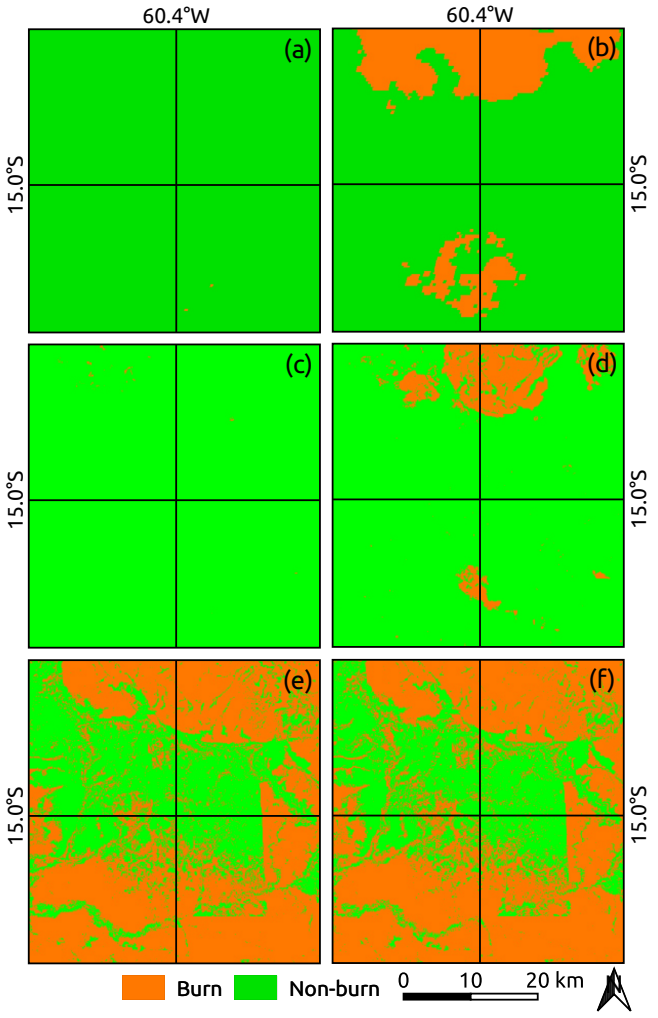

Fig. 8. The mapping of fire affected locations for Area 2. The reference mapping, UFD+MRF and MTDNBR results regarding Case II are show in subfigures (a), (c) and (e), respectively. The reference mapping, UFD+MRF and MTDNBR results regarding Case III are show in subfigures (b), (d) and (f), respectively.

pp. 18–27, 2017.

[17] D. Hosmer, S. Lemeshow, and R. Sturdivant, *Applied Logistic Regression*, ser. Wiley Series in Probability and Statistics. Wiley, 2013.

[18] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote sensing of environment*, vol. 25, no. 3, pp. 295–309, 1988.

[19] C. Key and N. Benson, "Landscape assessment: Ground measure of severity, the composite burn index; and remote sensing of severity, the normalized burn ratio," USDA Forest Service, Rocky Mountain Research Station, USGS Publications Warehouse, Tech. Rep. RMRS-GTR-164-CD: LA 1-51, 2006. [Online]. Available: http://pubs.er.usgs.gov/publication/2002085

[20] M. Kochenderfer and T. Wheeler, *Algorithms for Optimization*. MIT Press, 2019.

[21] M. Leuenberger, J. Parente, M. Tonini, M. G. Pereira, and M. Kanevski, "Wildfire susceptibility mapping: Deterministic vs. stochastic approaches," *Environmental Modelling & Software*, vol. 101, pp. 194–203, 2018.

[22] Z. Liu, "Effects of climate and fire on short-term vegetation recovery in the boreal larch forests of northeastern China," *Scientific reports*, vol. 6, no. 1, pp. 1–14, 2016.

[23] J. Luterbacher, L. Paterson, R. von Borries, K. Solazzo, R. Devillier, and S. Castonguay, "United in science 2021: A multi-organization high-level compilation of the latest climate science information," World Meteorological Organization (WMO), Global Carbon Project (GCP), Intergovernmental Panel on Climate Change (IPCC), United Nations Environment Programme (UNEP), World Health Organization (WHO), the Met Office (United Kingdom, UK), Tech. Rep., 2021. [Online]. Available: https://library.wmo.int/doc_num.php?explnum_id=10794

[24] E. A. Matricardi, D. L. Skole, M. A. Pedlowski, and W. Chomentowski, "Assessment of forest disturbances by selective logging and forest fires in the Brazilian Amazon using Landsat data," *International Journal of Remote Sensing*, vol. 34, no. 4, pp. 1057–1086, 2013.

[25] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta – Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[26] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed. New York: Chapmann and Hall, 1989.

[27] W. McKinney *et al.*, "Data structures for statistical computing in Python," in *Proceedings of the 9th Python in Science Conference*, vol. 445. Austin, TX, 2010, pp. 51–56.

[28] A. G. Motazeh, E. F. Ashtiani, R. Baniasadi, and F. M. Choobar, "Rating and mapping fire hazard in the hardwood Hyrcanian forests using GIS and expert choice software," *Forestry Ideas*, vol. 19, no. 2 (46), pp. 141–150, 2013. [Online]. Available: https://forestry-ideas.info/files/issue/Forestry_Ideas_BG_2013_19_2_3.pdf

[29] N. C. C. S. Prestes, K. G. Massi, E. A. Silva, D. S. Nogueira, E. A. de Oliveira, R. Freitag, B. S. Marimon, B. H. Marimon-Junior, M. Keller, and T. R. Feldpausch, "Fire effects on understory forest regeneration in southern Amazonia," *Frontiers in Forests and Global Change*, vol. 3, p. 10, 2020.

[30] D. Rajsekhar, V. P. Singh, and A. K. Mishra, "Multivariate drought index: An information theory based approach for integrated drought assessment," *Journal of Hydrology*, vol. 526, pp. 164–182, 2015.

[31] C. J. V. Rijsbergen, *Information Retrieval*, 2nd ed. Butterworth-

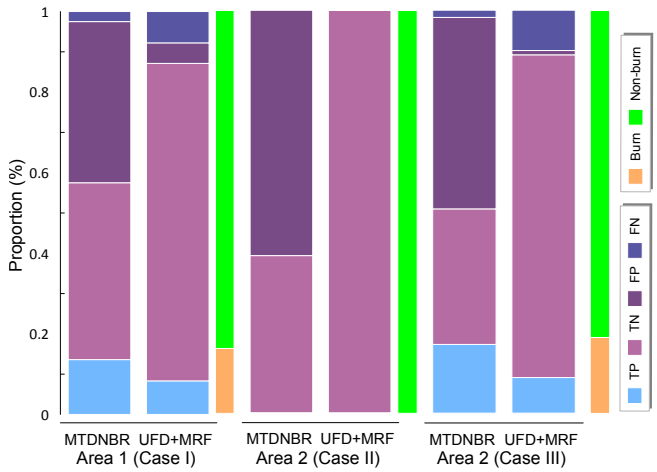

Fig. 9. True/False Positive/Negative detection proportions.

Heinemann, 1979.
[32] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," NASA, NASA special publication 351, 1974, third Earth Resources Technology Satellite-1 Symposium. Volume 1: Technical Presentations, section A.
[33] J. A. Sobrino, R. Llorens, C. Fernández, J. M. Fernández-Alonso, and J. A. Vega, "Relationship between soil burn severity in forest fires measured in situ and through spectral indices of remote detection," *Forests*, vol. 10, no. 5, p. 457, 2019.
[34] M. Story and R. G. Congalton, "Accuracy assessment: a user's perspective," *Photogrammetric Engineering and Remote Sensing*, vol. 52, pp. 397–399, 1986.
[35] P. A. Stott, N. Christidis, F. E. Otto, Y. Sun, J.-P. Vanderlinden, G. J. van Oldenborgh, R. Vautard, H. von Storch, P. Walton, P. Yiou *et al.*, "Attribution of extreme weather and climate-related events," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 7, no. 1, pp. 23–41, 2016.
[36] M. K. Van Aalst, "The impacts of climate change on the risk of natural disasters," *Disasters*, vol. 30, no. 1, pp. 5–18, 2006.
[37] G. van Rossum and F. L. Drake, *The Python Language Reference Manual*. Network Theory Ltd., 2011.
[38] S. Veraverbeke, P. Dennison, I. Gitas, G. Hulley, O. Kalashnikova, T. Katagis, L. Kuai, R. Meng, D. Roberts, and N. Stavros, "Hyperspectral remote sensing of fire: State-of-the-art and future perspectives," *Remote Sensing of Environment*, vol. 216, pp. 105–121, 2018.
[39] M. M. Verstraete and B. Pinty, "Designing optimal spectral indexes for remote sensing applications," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1254–1265, 1996.
[40] E. Westra, *Python geospatial development*. Packt Publishing Birmingham, 2010.
[41] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *Journal of Sensors*, vol. 2017, 2017.
[42] B. Zhao, Q. Zhuang, N. Shurpali, K. Köster, F. Berninger, and J. Pumpanen, "North American boreal forests are a large carbon source due to wildfires from 1986 to 2016," *Scientific reports*, vol. 11, no. 1, pp. 1–14, 2021.

Andréa E. O. Luz received graduate degree in Information Systems (2010) from Faculdade de Pindamonhangaba (FAPI) and M.Sc. in Natural Disasters (2022) at Universidade Estadual Paulista (UNESP). She has experience with algorithms and applications for Remote Sensing and Geoprocessing. She is a fellow at the UNESP Innovation Agency (AUIN).

Rogério G. Negri received the graduated degree in Mathematics (2006) at Universidade Estadual Paulista (UNESP), and the M.Sc. (2009) and Ph.D. (2013) in Applied Computation at Intituto Nacional de Pesquisas Espaciais (INPE). His main research interests include pattern recognition and image processing. He is currently Professor at UNESP, São José dos Campos, São Paulo, Brazil.

Alejandro C. Frery (S'92–SM'03) received the B.Sc. degree in Electronic and Electrical Engineering from the Universidad de Mendoza, Mendoza, Argentina (1983). His M.Sc. degree was in Applied Mathematics (Statistics) from the Instituto de Matemática Pura e Aplicada (IMPA, Rio de Janeiro, 1990) and his Ph.D. degree was in Applied Computing from the Instituto Nacional de Pesquisas Espaciais (INPE, São José dos Campos, Brazil, 1993). He was the founder of LaCCAN – *Laboratório de Computação Científica e Análise Numérica*, Universidade Federal de Alagoas, Maceió, Brazil. He was the Editor-in-Chief of IEEE Geoscience and Remote Sensing Letters (2014–2018). In 2018 he received the IEEE GRSS Regional Leader Award. His research interests are statistical computing and stochastic modeling. He is currently Professor of Statistics and Data Science with the Victoria University of Wellington, New Zealand

Maurício A. Dias received the M.Eng and the Ph.D. degrees from Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil, in 2002 and 2007, respectively. He was a Postdoctoral Researcher (Visiting Researcher in Sabbatical Break) twice: (1) with the Electronics Department of the Polytechnic School of the University of Alcalá (UAH), Alcalá de Henares, Spain, in 2008; (2) with the Centre for Vision, Speech and Signal Processing (CVSSP) of the Faculty of Electronics and Physical Sciences of the University of Surrey, Guildford, England, in 2018. He is currently an Assistant Professor with the State University of São Paulo (UNESP), Presidente Prudente, Brazil. His research interests include digital image processing, digital image analysis, and mathematical morphology.

unesp

Rogerio Galante Negri <rogerio.negri@unesp.br>

IEEE TGRS-2022-00367- Manuscript Received

1 mensagem

Transactions on Geoscience and Remote Sensing <onbehalfof@manuscriptcentral.com> 2 de fevereiro de 2022 08:02
Responder a: john.wright@ieee.org
Para: andrea.luz@unesp.br, rogerio.negri@unesp.br, acfrery@gmail.com, alejandro.frery@vuw.ac.nz, ma.dias@unesp.br

Dear Prof. Negri:

The Editor of the IEEE Transactions on Geoscience and Remote Sensing acknowledges receipt of the following manuscript:

TGRS-2022-00367
Fire Detection with Multitemporal Multispectral Data and a Probabilistic Unsupervised Technique

Your paper has been submitted as a(n) Regular paper.

It is understood that this manuscript is entirely original, has not been copyrighted, published, submitted, or accepted for publication elsewhere, and all necessary clearances and releases have been obtained. If the material in this paper has been published before in any form, it is imperative that you inform me immediately.

Please refer to the paper number in any communications regarding your manuscript. You may check the review status anytime via the IEEE ScholarOne Manuscripts website. When the review of your manuscript has been completed, you will be notified of its disposition by e-mail and at that time reviewer comments will also be made available to you.

*Please allow at least 90 days for the review process before making inquiries as to the status of your manuscript**

TGRS has instituted the following policy regarding payment of page charges of manuscripts accepted for publication by the Editorial Board:

(1) Payment of page charges for the first six printed pages is voluntary. The rate is $110 per page. If these optional sustaining page charges are paid, one hundred free reprints of your article will be provided.

(2) A mandatory excessive page length charge of $230 per page (beginning with page 7 and beyond) is required for pages in excess of six printed pages.

While your manuscript is undergoing review, you may also submit it in its current form to TechRxiv, IEEE's preprint server. By uploading your manuscript to TechRxiv you can quickly disseminate your work to a wide audience and gain community feedback. Any comments you receive on the TechRxiv submission have no bearing on the peer review process. Submit your manuscript at https://www.techrxiv.org/.

Sincerely,

Dr. Simon Yueh
Editor, IEEE Trans. on Geoscience and Remote Sensing

S7

3 CONSIDERAÇÕES FINAIS

Os desastres naturais decorrentes de queimadas e incêndios florestais têm uma clara relação com as mudanças climáticas. Consequentemente, grandes prejuízos são previstos com o aumento das variações no clima e a ocorrência de eventos extremos como secas e estiagens prolongadas.

Diante da necessidade de prevenção ou mesmo mitigação de impactos causados pelo fogo em diferentes ecossistemas terrestres, é fundamental o desenvolvimento de metodologias para o mapeamento de tais ocorrências. Face a este cenário, foram desenvolvidos dois métodos com diferentes finalidades: (i) mapeamento de áreas afetadas pelo fogo em um período passado; (ii) mapeamento de áreas suscetíveis ao fogo em um período futuro; Tais métodos são baseados em conceitos de detecção de anomalias, índices espectrais, processamento de séries multitemporais de imagens obtidas por Sensoriamento Remoto e modelos de aprendizado de máquina.

A formulação e experimentação destas novas propostas foram apresentadas através de dois artigos científicos distintos.

Em aporte ao primeiro método, o artigo *Fire Detection with Multitemporal Multiespectral Datas and a Probabilistic Unsupervised Technique* demonstra a utilização de séries temporais de imagens multiespectrais para o mapeamento de áreas atingidas por incêndios. Este estudo se baseia no emprego de conceitos de modelagem e classificação estatística em uma estrutura probabilística. Em sua validação, são realizados dois casos de estudo, considerando regiões e sensores distintos. Os resultados obtidos apontaram precisão do método na identificação das ocorrências, comprovada pela aplicação do teste de hipótese derivado do coeficiente *kappa*. Ademais, a aderência entre os mapeamentos gerados em relação ao produto de validação foi considerada alta.

Diferente do anterior, e em aporte ao segundo método, o artigo *Anomaly-Driven Approach for Forest Fire Susceptibility Mapping Using Multitemporal Remote Sensing Data* aborda o mapeamento de áreas suscetíveis à ocorrência de incêndios, ou seja, identifica (também de forma totalmente automática) as áreas com alta suscetibilidade ao fogo. Este estudo se baseia no emprego de conceitos de detecção de anomalias através dos modelos OC-SVM e IF. Os resultados obtidos mostraram que a proposta foi capaz de identificar as áreas com alta suscetibilidade ao fogo, o que foi verificado na comparação dos mapas resultantes com os mapas de referência adotados para validá-los. Em relação aos dois modelos de detecção de anomalias, o IF entregou mapas mais aderentes do que o OC-SVM.

Ambos os métodos propostos demonstram contribuições aos estudos relacionados ao mapeamento do fogo, tanto para as análises que dependem de uma base histórica de ocorrências, quanto no suporte e previsão de novos incêndios, através do estudo da suscetibilidade. Ainda, acredita-se que a extensão destas propostas para outros problemas e análises ambientais pode ser alcançada através de ajustes menores sobre os algoritmos desenvolvidos.

Por fim, como perspectivas futuras, são citadas: (i) adaptar e aplicar os métodos desenvolvidos para outras questões ambientais, como desmatamento, inundações, derramamentos de óleo, derretimento de geleiras, etc; (ii) avaliar o comportamento de outros índices espectrais de vegetação no processo de detecção e quantificação da suscetibilidade ao fogo; (iii) incluir uma medida de confiança relacionada aos mapeamentos resultantes; e (iv) desenvolver estratégias para redução do custo computacional, especialmente em relação ao método proposto para mapeamento de suscetibilidade.

REFERÊNCIAS

AALST, M. K. V. The impacts of climate change on the risk of natural disasters. **Disasters**, **Wiley Online Library**, v. 30, n. 1, p. 5–18, 2006.

BAYNE, K. M. et al. Fire as a land management tool: Rural sector perceptions of burn-off practice in new zealand. **Rangeland Ecology & Management**, Elsevier, v. 72, n. 3, p. 523–532, 2019.

BOWMAN, D. M. et al. Fire in the earth system. **Science**, American Association for the Advancement of Science, v. 324, n. 5926, p. 481–484, 2009.

BRANDO, P. M. et al. Abrupt increases in amazonian tree mortality due to drought–fire interactions. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 111, n. 17, p. 6347–6352, 2014.

CAÚLA, R. et al. Overview of fire foci causes and locations in brazil based on meteorological satellite data from 1998 to 2011. **Environmental Earth Sciences**, Springer, v. 74, n. 2, p. 1497–1508, 2015.

GARCIA, L. C. et al. Record-breaking wildfires in the world's largest continuous tropical wetland: Integrative fire management is urgently needed for both biodiversity and humans. **Journal of environmental management**, Elsevier, v. 293, p. 112870, 2021.

GU, S. et al. Application of organic fertilizer improves microbial community diversity and alters microbial network structure in tea (camellia sinensis) plantation soils. **Soil and Tillage Research**, Elsevier, v. 195, p. 104356, 2019.

HART, E. et al. Use of machine learning techniques to model wind damage to forests. **Agricultural and forest meteorology**, Elsevier, v. 265, p. 16–29, 2019.

HAVENS, T. C.; BEZDEK, J. C. An efficient formulation of the improved visual assessment of cluster tendency (ivat) algorithm. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 24, n. 5, p. 813–822, 2011.

HONG, H. et al. Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. the case of dayu county, china. **Science of the total environment**, Elsevier, v. 630, p. 1044–1056, 2018.

INPE. **Instituto Nacional de Pesquisas Espaciais- Banco de Dados de queimadas.** 2021. Accessed: 2021-03-29. Disponível em: <http://www.inpe.br/queimadas/bdqueimadas>.

LEUENBERGER, M. et al. Wildfire susceptibility mapping: Deterministic vs. stochastic approaches. **Environmental Modelling & Software**, Elsevier, v. 101, p. 194–203, 2018.

LEWIS, S. L. et al. The 2010 amazon drought. **Science**, American Association for the Advancement of Science, v. 331, n. 6017, p. 554–554, 2011.

LIU, F. T.; TING, K. M.; ZHOU, Z. Isolation forest. In: **2008 Eighth IEEE International Conference on Data Mining**. [S.l.: s.n.], 2008. p. 413–422.

MALHI, Y. et al. Climate change, deforestation, and the fate of the amazon. **Science**, American Association for the Advancement of Science, v. 319, n. 5860, p. 169–172, 2008.

MARQUES, J. F. et al. Fires dynamics in the pantanal: Impacts of anthropogenic activities and climate change. **Journal of Environmental Management**, Elsevier, v. 299, p. 113586, 2021.

MARTÍNEZ-ÁLVAREZ, F.; BUI, D. T. **Advanced Machine Learning and Big Data Analytics in Remote Sensing for Natural Hazards Management**. [S.l.]: Multidisciplinary Digital Publishing Institute, 2020.

MORITZ, M. A. et al. Learning to coexist with wildfire. **Nature**, Nature Publishing Group, v. 515, n. 7525, p. 58–66, 2014.

NOBRE, C. A.; MARENGO, J. A.; SOARES, W. R. **Climate change risks in Brazil**. [S.l.]: Springer, 2019.

PAN, Y. et al. A large and persistent carbon sink in the world's forests. **Science**, American Association for the Advancement of Science, v. 333, n. 6045, p. 988–993, 2011.

POURGHASEMI, H. R. et al. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. **Scientific reports**, Nature Publishing Group, v. 10, n. 1, p. 1–11, 2020.

PRESTES, N. C. C. d. S. et al. Fire effects on understory forest regeneration in southern amazonia. **Frontiers in Forests and Global Change**, Frontiers, v. 3, p. 10, 2020.

SALEHI, M.; RASHIDI, L. A survey on anomaly detection in evolving data: [with application to forest fire risk prediction]. **ACM SIGKDD Explorations Newsletter**, ACM New York, NY, USA, v. 20, n. 1, p. 13–23, 2018.

SCHÖLKOPF, B. et al. Estimating the support of a high-dimensional distribution. **Neural computation**, MIT Press One Rogers Street, v. 13, n. 7, p. 1443–1471, 2001.

SILVA, S. S. da et al. Dynamics of forest fires in the southwestern amazon. **Forest Ecology and Management**, Elsevier, v. 424, p. 312–322, 2018.

YUAN, Q. et al. Deep learning in environmental remote sensing: Achievements and challenges. **Remote Sensing of Environment**, Elsevier, v. 241, p. 111716, 2020.