MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc m21d/2022/08.09.12.27-TDI

NEAR REAL-TIME FIRE DETECTION AND MONITORING: A MACHINE LEARNING ALGORITHM DEVELOPMENT FOR MATOPIBA REGION, BRAZIL

Mikhaela Aloísia Jéssie Santos Pletsch

Doctorate Thesis of the Graduate Course in Remote Sensing, guided by Drs. Thales Sehn Körting, and Luiz Eduardo Oliveira e Cruz de Aragão, approved in August 15, 2022.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/47DMC92>

INPE
São José dos Campos
2022

PUBLISHED BY:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

BOARD OF PUBLISHING AND PRESERVATION OF INPE INTELLECTUAL PRODUCTION - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):

Chairperson:

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra (CGCT)

Members:

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)

Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

DIGITAL LIBRARY:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

DOCUMENT REVIEW:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

ELECTRONIC EDITING:

Ivone Martins - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc m21d/2022/08.09.12.27-TDI

NEAR REAL-TIME FIRE DETECTION AND MONITORING: A MACHINE LEARNING ALGORITHM DEVELOPMENT FOR MATOPIBA REGION, BRAZIL

Mikhaela Aloísia Jéssie Santos Pletsch

Doctorate Thesis of the Graduate Course in Remote Sensing, guided by Drs. Thales Sehn Körting, and Luiz Eduardo Oliveira e Cruz de Aragão, approved in August 15, 2022.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/47DMC92>

INPE
São José dos Campos
2022

Cataloging in Publication Data

Pletsch, Mikhaela Aloísia Jéssie Santos.

P716n Near real-time fire detection and monitoring: a machine learning algorithm development for matopiba region, Brazil / Mikhaela Aloísia Jéssie Santos Pletsch. – São José dos Campos : INPE, 2022.

xxvi + 113 p. ; (sid.inpe.br/mtc m21d/2022/08.09.12.27-TDI)

Thesis (Doctorate in Remote Sensing) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2022.

Guiding : Drs. Thales Sehn Körting, and Luiz Eduardo Oliveira e Cruz de Aragão.

1. Active fires. 2. Wildfires. 3. Time series analysis. 4. Machine learning. 5. Geotechnologies. I.Title.

CDU 528.8:630*43

Esta obra foi licenciada sob uma Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS
Secretaria de Pós-Graduação - SEPGR

DEFESA FINAL DE TESE DE MIKHAELA ALOÍSIA JÉSSIE SANTOS PLETSCH
REG 136662/2018, BANCA Nº 214/2022

No dia 15 de agosto de 2022, às 13h, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutora em Sensoriamento Remoto.

Novo Título: "NEAR REAL-TIME FIRE DETECTION AND MONITORING: A MACHINE LEARNING ALGORITHM DEVELOPMENT FOR MATOPIBA REGION, BRAZIL"

Membros da banca:

Dra. Liana Oighenstein Anderson – Presidente da banca – INPE
Dr. Thales Sehn Körting – Orientador – INPE
Dr. Luiz Eduardo Oliveira e Cruz de Aragão – Orientador – INPE
Dr. Alberto Waingort Setzer – Membro Interno – INPE
Dr. Michelle Cristina Araujo Piccoli - Membro Externo – UCLouvain, Université catholique de Louvain
Dr. Fabien Hubert Wagner - Membro Externo – UCLA, University of California, Los Angeles

Declaração de aprovação de Dr. Michele Cristina Araujo Piccoli anexa ao processo.

Documento assinado eletronicamente por **Thales Sehn Korting**, **Pesquisador**, em 22/08/2022, às 10:22 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Liana Oighenstein Anderson**, **Pesquisador**, em 22/08/2022, às 10:27 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Fabien Hubert Wagner (E)**, **Usuário Externo**, em 22/08/2022, às 12:26 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Luiz Eduardo Oliveira E Cruz de Aragão**, **Chefe da Divisão de Observação da Terra e Geoinformática**, em 22/08/2022, às 13:09 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Alberto Waingort Setzer**, **Pesquisador**, em 23/08/2022, às 08:33 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.mcti.gov.br/verifica.html, informando o código verificador **10272737** e o código CRC **A2ADF832**.

Referência: Processo nº 01340.006295/2022-53 SEI nº 10272737

"Não conseguimos mudar coisa alguma sem antes aceitá-la. A condenação não libera, oprime".

Carl Gustav Jung
(1875 - 1961)

À Deus, e aos meus pais ***Orides*** *e* ***Roseli****, os quais me deram o dom da vida. E ao meu grande amor e amigo,* ***Felipe****, com quem eu me casarei assim que eu defender essa tese :D*

ACKNOWLEDGEMENTS

There are a myrid of people and entities that I am deeply grateful. Some of them that I am more aware, and some of them (the greater amount maybe?) that I am not. Anyway, this project would not have been possible without the support of many people!

I would like to thank my supervisors Dr. Thales Sehn Körting and Dr. Luiz Eduardo Oliveira e Cruz de Aragão, who helped me in their own ways with: guidance throughout the project, financial and emotional support, advice and encouragement words, and the numerous revisions that helped make some sense of the confusion. Particularly, I also thank Dr. Thales Sehn Körting that accepted to be my advisor in my Master degree in Remote Sensing in 2016. Without that opportunity, this thesis would not be possible, and I would not have gotten this far in the scientific carrier. Thales believed me when I first said that I would like to learn how to code and use it for environmental management. I do not know if he is aware of that, but he changed my life by believing in me, and I am deeply grateful for that!

I let here a special thanks to Dr. Victor Velázquez Fernandez, my advisor of the Scientific Initiation at College and my scientific father. Velázquez is an example of teacher, mentor and father. By letting me in in the scientific world, and accepting to be my first advisor, he allowed me to see my own wings and my strength to start flying. He showed me that dreams can come true, that with effort stars could be reached. I'm proud of, and grateful for, my time working with him.

Many thanks also to Dr. Celso H. L. Silva-Junior, a great and inspiring friend that supported me during our time together in the postgraduate degree. Felipe Carraro Morita, a great data scientist and analyst, my fiance and future husband, that always believed in me and support this thesis in several ways.

I also thank the National Institute for Space Research (INPE) and all their amazing members/employees, as well as INPE/Fire Monitoring Program members for the partnership, clarification and sharing of existing gaps and opportunities for action. Thanks to the National Council for Scientific and Technological Development (CNPq) for the scholarship that provided me resources to complete this project.

Immense gratitude to Rosangela Galindo, my friend and therapist, for her patience and support.

Also thanks to my committee members, Dra. Liana Oighenstein Anderson, Dra.

Michelle Cristina Araujo Picoli, Dr. Alberto Waingort Setzer, Dr. Fabien Hubert Wagner, whose expertise supports this project.

I would like also to extend my sincere thanks to AdeAgro for the opportunity to work with such a nice team, and for remembering me my own value.

And finally, thanks to my family, parents, and numerous friends who endured this long process with me, always offering support and love.

If you are reading this, many thanks! No information is relevant if it does not reach people like you and me.

ABSTRACT

MATOPIBA is an agricultural frontier, located mainly in Cerrado. MATOPIBA is an acronym for its location over: Maranhão, Tocantins, Piauí and Bahia states. In this region, fires are essential for its biodiversity maintenance. However, the increase in its recurrence and intensity, as well as accidental fires can lead to socioeconomic and environmental losses. Due to this dual relationship with fire, near real-time fire management is required throughout the region. In this context, we developed, to the best of our knowledge, the first Machine Learning algorithm based on time series obtained from GOES-16 ABI, an optical sensor able to detect and monitor Active Fires in near real-time in MATOPIBA. ABI is onboard GOES-16, and it presents 16 spectral bands, with a nominal spatial resolution of 2 km at nadir, and a full disk image every 10 minutes over Americas. To do so, we analyzed the best combination of three Machine Learning algorithms and how long it takes to consider a historical time series able to support accurate active fire predictions. We used the most accurate combination for the final model development. The results show that the final model ensures an overall accuracy rate of approximately 80%. The final model potential is remarkable not only for single detections but also for a consecutive sequence of positive predictions. Roughly, the final model achieves an accuracy rate peak after around 20 h of consecutive active fire detections, but there is an important trade-off between the accuracy and the time required to assemble more fire indications, which can be decisive for firefighters in real life.

Keywords: Active Fires. Wildfires. Time Series Analysis. Machine Learning. Geotechnologies.

DETECÇÃO E MONITORAMENTO DE FOGOS ATIVOS QUASE EM TEMPO REAL: DESENVOLVIMENTO DE UM ALGORITMO DE MACHINE LEARNING PARA A REGIÃO DO MATOPIBA, BRASIL

RESUMO

MATOPIBA é uma fronteira agrícola, localizada principalmente no Cerrado. A área se chama MATOPIBA por ser um acrônimo da sua localização ao longo dos estados do Maranhão, Tocantins, Piauí e Bahia. Nessa região, as queimadas são essenciais para a manutenção da biodiversidade. No entanto, o aumento de sua recorrência e intensidade, bem como incêndios acidentais podem acarretar prejuízos socioeconômicos e ambientais. Devido à essa relação dupla com o fogo, o manejo do fogo quase em tempo real é necessário em toda a região. Neste contexto, desenvolvemos, até onde sabemos, o primeiro algoritmo de Machine Learning baseado em séries temporais obtidas a partir do GOES-16 ABI, um sensor óptico capaz de detectar e monitorar Fogos Ativos quase em tempo real no MATOPIBA. ABI está a bordo do GOES-16, e apresenta 16 bandas espectrais, com resolução espacial nominal de 2 km no nadir, e uma imagem completa a cada 10 minutos sobre as Américas. Para isso, analisamos a melhor combinação de três algoritmos de Machine Learning e quanto tempo leva para considerar uma série temporal histórica capaz de suportar previsões precisas de fogos ativos. Usamos a combinação mais precisa para o desenvolvimento do modelo final. Os resultados mostram que o modelo final garante uma taxa de acerto geral de aproximadamente 80%. O potencial do modelo final é notável não apenas para detecções simples, mas também para uma sequência consecutiva de previsões positivas. A grosso modo, o modelo final atinge um pico de taxa de precisão após cerca de 20 horas de detecções consecutivas de fogos ativos, mas há um importante trade-off entre a precisão e o tempo necessário para reunir mais indicações de fogos, o que pode ser decisivo para os bombeiros na vida real.

Palavras-chave: Fogo Ativo. Incêndios. Análise de Séries Temporais. Machine Learning. Geotecnologias.

LIST OF FIGURES

Page

LIST OF TABLES

LIST OF ABBREVIATIONS

AF	–	Active fire
AFI	–	Active Fire Index
AHI	–	Advanced Himawari Imager
AI	–	Artificial Intelligence
API	–	Application Programming Interface
AVHRR	–	Advanced Very High Resolution Radiometer
BA	–	Burned area
BAI	–	Burned area index
BFAST	–	Breaks For Additive Seasonal and Trend
CBERS	–	China Brazil Earth Resources Satellite
CNN	–	Convolutional Neural Networks
CRISP-DM	–	Cross Industry Standard Process for Data Mining
DC	–	Data cube
DL	–	Deep learning
DM	–	Data mining
FAF	–	Far active fire
FDCA	–	Fire Detection and Characterization Algorithm
FIRMS	–	Fire Information for Resource Management System
FM	–	Final model
FN	–	False negative
FP	–	False positive
FRP	–	Fire Radiative Power
FT-RGB	–	Fire Temperature RGB
GOES	–	Geostationary Operational Environmental Satellites
Gr	–	Grasslands
GRU	–	Gated Recurrent Units
HRR	–	Heat release rate
INPE	–	Instituto Nacional de Pesquisas Espaciais
KDD	–	Knowledge Discovery from Databases
LR	–	Logistic regression
LULC	–	Land use and land cover
MATOPIBA	–	States of Maranhão, Tocantins, Piauí, Bahia
ML	–	Machine learning
MODIS	–	Moderate-Resolution Imaging Spectroradiometer
MIRBI	–	Mid-infrared burn index
NAF	–	Near active fire
NetCDF	–	Network Common Data Form
NF	–	Natural formation
NASA	–	National Aeronautics and Space Administration
NBR	–	Normalized burn ratio
NBR+	–	Normalized burn ratio plus

NBRT	–	Normalized burn ratio thermal
NDVI	–	Normalized Difference Vegetation Index
NF	–	Natural Formation
NOAA	–	National Oceanic and Atmospheric Administration
NRT	–	Near real-time
NIR	–	Near infrared
OLS-MOSUM	–	ordinary least squares residual moving sum
PNMC	–	National Policy on Climate Change
PPCerrado	–	Prevention and Control of Deforestation and Forest Fires in the Cerrado
RF	–	Random Forest
RGB	–	Red, Green and Blue
RNN	–	Recursive Neural Network
RS	–	Remote Sensing
SF	–	Savanna formation
S-NPP	–	Suomi National Polar-orbiting Partnership
SI	–	Spectral index
SVM	–	Support Vector Machine
SWIR	–	Short-wave infrared
TIF	–	Tagged Image File Format
TM	–	Thematic Mapper
TN	–	True Negative
TP	–	True positive
TRMM	–	Tropical rainfall measuring mission
TS	–	Time series
US	–	United States
VCI	–	Vegetation Condition Index
VIIRS	–	Visible Infrared Imaging Radiometer Suite
WF_ABBA	–	Wild Fire Automated Biomass Burning Algorithm
XGBoost	–	Extreme Gradient Boosting

LIST OF SYMBOLS

km^2	–	square kilometers
W	–	Watts
%	–	Percentage
d	–	days
kW	–	Kilowatt
MW	–	Megawatt
$km^2\ d^{-1}$	–	Square kilometers per day
μm	–	Micrometre

CONTENTS

Page

1 INTRODUCTION

Fire incidence has the potential to consume and modify large areas of vegetation (BOND et al., 2005), to decrease the surface-to-atmosphere water transfer, to increase surface warming (IVO et al., 2020), and to release aerosol and gases that contributes to global climate change (LASHOF, 1991; WERF et al., 2010). In Brazil, fires are also an important source of air pollution with harmful health consequences (SILVA et al., 2016). It has become a burden for the public health system (URIARTE et al., 2009; ARAGÃO et al., 2020) owing to an increase in respiratory diseases during the fire season (MACHADO-SILVA et al., 2020). The fire season is a period of the year when the fires are most likely to occur (PIVELLO et al., 2021), which takes place mainly during the transition from dry to the wet season, July-October (RISSI et al., 2017).

A staggering number of fires in 2019 and 2020 in different Brazilian biomes revealed the national criminal fire management unpreparedness (PIVELLO et al., 2021), especially in August, 2019, when fires reached a turning point that was widely covered by the media around the world (ALENCAR et al., 2020; BENCHERIF et al., 2020). Different from the Brazilian Amazon and Pantanal biomes, fires in the Brazilian tropical savanna Cerrado can be associated with both human land-use activities and natural drivers (RAMOS-NETO; PIVELLO, 2000; FIDELIS, 2020). It is the easiest and cheapest way to boost the fresh grass growth for cattle ranching, as well as to open new agricultural areas (MIRANDA et al., 2002; KLINK; MACHADO, 2005; PIVELLO, 2011). At the same time, natural fire ignitions can be caused by lightning (RAMOS-NETO; PIVELLO, 2000), making Cerrado species adapted for fires. Without this phenomenon, the region would be dominated by grasses, and over time, forest encroachment could cause a biodiversity loss (ABREU et al., 2017). While human-induced fires are frequent and intense, impacting the biodiversity and aboveground biomass, natural fires are usually rapid, have low intensity and do not spread over large areas. In addition, natural fires occur every 3-6 years, maintaining the regional biodiversity and ecological processes (PIVELLO et al., 2021). Thus, Cerrado presents a singular dual relationship with fire: its incidence is necessary for biodiversity preservation (FIDELIS et al., 2018; FIDELIS, 2020), but the increase in its recurrence and intensity, as well as accidental fires, has the potential to cause negative social, economic and environmental impacts (MIRANDA et al., 2010; PIVELLO, 2011). Nonetheless, different from what was speculated, meteorological conditions did not play a role in the increasing number of fires in 2019. Instead of that, changes in Land Use and Land Cover (LULC) are likely to be the main drivers (KELLEY et al., 2021).

Cerrado is a global biodiversity hotspot (DAMASCO et al., 2018). Even though, nearly half of its original vegetation has already disappeared, mainly due to advancing agricultural frontiers (SCHMIDT; ELOY, 2020). Brazil's most recent agricultural frontier is located in northern Cerrado, in a region known as MATOPIBA (an acronym for its location over: Maranhão, Tocantins, Piauí and Bahia states) (MIRANDA et al., 2014). MATOPIBA contains 38% of the Cerrado biome, and it has experienced almost half of Cerrado's deforestation (INPE, 2022). Besides, MATOPIBA presents, at the same time, the largest undisturbed remnants of Cerrado vegetation and a quarter of the Cerrado's soybean area (MIRANDA et al., 2014; SOTERRONI et al., 2019), both important for environmental and economic issues, respectively. Nonetheless, over the last decade, the combination of climate change and land-use change has severely increased drought conditions in the region, which contribute to a higher fire risk, mainly during the fire season (MARENGO et al., 2022), and jeopardize biodiversity and food security.

Given that fire has a dual-character in Cerrado and that it is considered a highly dynamic phenomenon, the use of near real-time (NRT) remote sensing datasets available from geostationary satellites has provided promising results for fire monitoring in this biome (PLETSCH et al., 2019). Even though such datasets usually present a trade-off between spatial and temporal resolutions (WOOSTER et al., 2021), the Advanced Baseline Imager (ABI) onboard the new generation of Geostationary Operational Environmental Satellite-R (GOES-R) Series was designed to overcome it by improving spatial, temporal and radiometric characteristics of the previous GOES Imager (SCHMIT et al., 2005; SCHMIT et al., 2017). GOES-16 ABI can be considered an ultrahigh temporal resolution sensor, with temporal resolution of 10 minutes, and spatial resolution from 0.5 to 2 km (NOAA AND NASA, 2022b). Although ABI as well as others geostationary satellite sensors are available, they are yet underused (ABDI et al., 2022).

Due to the data deluge, high velocity production and Earth surface target diversity (LANEY et al., 2001), the ABI dataset can also be considered big data. While it presents a vast amount of unexplored information, its access, process, and comprehension become impossible by means of traditional methods that rely on handmade procedures. To overcome this challenge, scientists (from NASA/FIRMS portal (NASA, 2021) and INPE/Fire Monitoring Program (INPE, 2021)) have been developing algorithms and Machine Learning (ML) models using different remote sensing data for fire detection and monitoring, and releasing the results to support fire management. Although some developments were already done for ABI, they were

developed based on the previous sensors. An example is the Fire Detection and Characterization Algorithm (FDCA), which was created based on the Wildfire Automated Biomass-Burning Algorithm (WF-ABBA) (SCHMIDT et al., 2010). Currently, FDCA has been undergoing validation and refinement (SCHMIDT, 2019; SCHMIDT, 2020). However, some authors have indicated that the reliable AF detection challenge remains (HALL et al., 2019; LI et al., 2020), once it is necessary to adapt the algorithms to the current sensor characteristics (CHUVIECO et al., 2020; HALL et al., 2019). Besides, the lack of ground truth data also is a hindrance for the validation process (SCHMIDT, 2019).

The Moderate Resolution Imaging Spectroradiometer (MODIS) (GIGLIO et al., 2016) and the Visible Infrared Imaging Radiometer Suite (VIIRS) (SCHROEDER et al., 2014) present well-designed and already established fire products. Nonetheless, MODIS' temporal resolution, when considering both Aqua and Terra satellites, is four times a day, whereas VIIRS is only twice a day. Even when MODIS and VIIRS data are used together, it does not provide an NRT dataset.

Therefore, to ensure Brazilian biodiversity by means of fire management, we developed, to the best of our knowledge, the first ML algorithm (hereafter, Final Model-FM) able to detect and monitor Active Fires (AF) in NRT in MATOPIBA. FM uses GOES-16 ABI imagery and is focused on the LULC of Natural Formation, which is composed of Natural Forest, Savanna Formation, and Grasslands. For the FM development, we first analyzed the performance of three ML algorithms and established how much historical data (expressed in days, hereafter called lag) from before a fire event the FM is required to make accurate AF classifications. Then, the most accurate algorithm and lag were selected for the FM development. In this process, we used MODIS and VIIRS AF products for comparison as reference satellites and filtering purposes and manually mapped BA on Sentinel-2 imagery. This procedure was required to have access to the most reliable fire data possible. In addition, its finer spatial resolution data (10-20 m) are commonly used for validation processes from coarser-resolution satellites (LI et al., 2020).

The novelty of this study is mainly regarding the development of a Machine Learning model and the use of an ultrahigh temporal resolution dataset of 10 minutes. In addition, it also complements the literature by comparing ML models for AF detection and monitoring and the comprehension of how much historical data are required to train an ML model to accurately classify AF. Finally, we also present in this thesis, to the best of our knowledge, the first analysis of the nature of fire

based on the Aristotle's four causes (material, formal, efficient, and final), which is essential to have a complete comprehension about fires.

1.1 Objectives

Considering the importance of ensuring the Brazilian biodiversity by means of fire management and the gaps presented, the central goal of this thesis is to develop, to the best of our knowledge, the first ML algorithm (hereafter Final Model - FM) based on GOES-16 ABI sensor able to detect and monitor Active Fires (AF) in NRT in MATOPIBA's Natural Formations.

Furthermore, this thesis aims at answering the following research questions:

- What is the overall performance of the FM? Does Land Used and Land Cover (LULC) play an important role in FM accuracy?
- Does the size of the burned area (BA) influence the FM accuracy? Is the FM influenced by BA found in the surroundings of a central ABI pixel grid?
- What is the potential of FM considering a sequence of positive fire predictions? What is its agreement with MODIS and VIIRS dataset?
- Assuming that we have a certain number of consecutive AF detections by FM, what is the fire reality in the remaining data along MATOPIBA?
- How does FM perform in detail considering single detections, the reference satellites and the manually mapped BA along the time?

To answer these questions, we organized the thesis in seven specific objectives:

- Chronologically structure ABI data in the known Data Cube;
- Develop algorithms able to apply and validate the ML models in large datasets in automated processes;
- Identify among three ML algorithms, Random Forest, Logistic Regression and Extreme Gradient Boosting, which is the most suitable for fire detection and monitoring;
- Analyse how many days (lag) are indicated to consider in a historical time series able to support accurate fire predictions;

- Manually develop a burned area mapping based on Sentinel-2 able to support the FM assessment;
- Develop a FM (based on the most indicated ML algorithm and lag) able to support the monitoring of fires in the daily activities of fire managers;
- Comprehend FM potential and variables that influence its performance.

1.2 Thesis outline

This thesis is divided in 5 chapters. Firstly, we present a general background about the main issues already presented (Section 2). After that, we present the methodology (Section 3), followed by results (Section 4), and discussions (Section 5). Finally, we present in Section 6 the conclusions and recommendations for future work.

2 LITERATURE REVIEW

In the first subsection of this chapter, we present a literature review about Cerrado, its fire studies and initiatives for fire management in the region. In the second subsection, we explain, as far as we know, for the first time the fire causality based on the Aristotle's four (be)causes, answering the paramount question: *Why does fire comes into existence?* In the same section, we also explain the fire regimes, fire behaviour and wildfires.

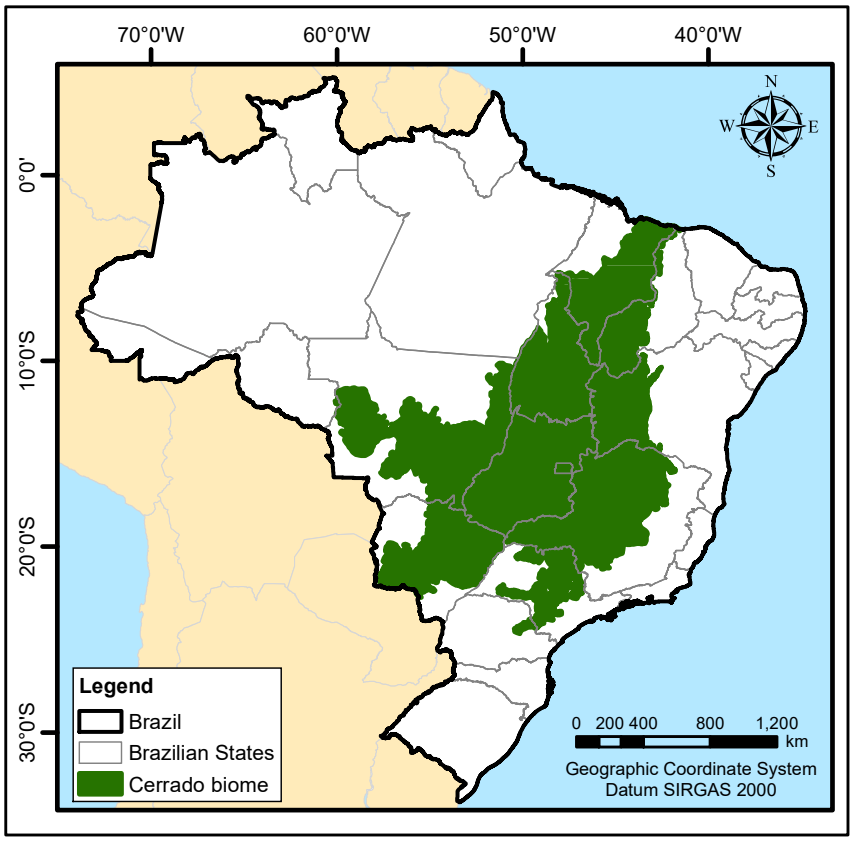
After that, we present the whole conundrum related to fire monitoring, remote sensing, time series, data cube, data mining and machine learning. In the last section, we bring a literature review about the related works, such as the use of ABI and ML for fire management in and out of Cerrado. It is important to highlight that in order to avoid citations redundancy, the complementary bibliography is presented in the different sections of the chapter Results.

2.1 Cerrado biome

The Cerrado is the second largest biome in South America, with more than 2 million km^2, and about 24% of the Brazilian territory (RIBEIRO; WALTER, 2008; MMA, 2009) (Figure 2.1). This biome extends over the Federal District and the states of Goiás, Tocantins, Maranhão, Mato Grosso do Sul, Minas Gerais, Mato Grosso, Piauí, São Paulo, Bahia, Paraná and Rondônia.

Comprehending 22.4° in latitude, Cerrado varies considerably, from the sea level to 1,800 m of elevation (SANO et al., 2010; MMA, 2014). According to the climate Köppen-Geiger classification, Cerrado predominantly presents dry winter (Aw), April-September, and hot summer (Cwa), October-March (PEEL et al., 2007; RIBEIRO; WALTER, 2008; ALVARES et al., 2013). Considered a savanna, this heterogeneous environment presents average annual precipitation ranging from 1,300 to 1,600 mm, and temperature of 20,1 °C (RIBEIRO; WALTER, 2008), yet both vary over the years (FERREIRA et al., 2018).

Figure 2.1 - Map of Brazil and the Brazilian States with highlight in the green of the Cerrado biome.

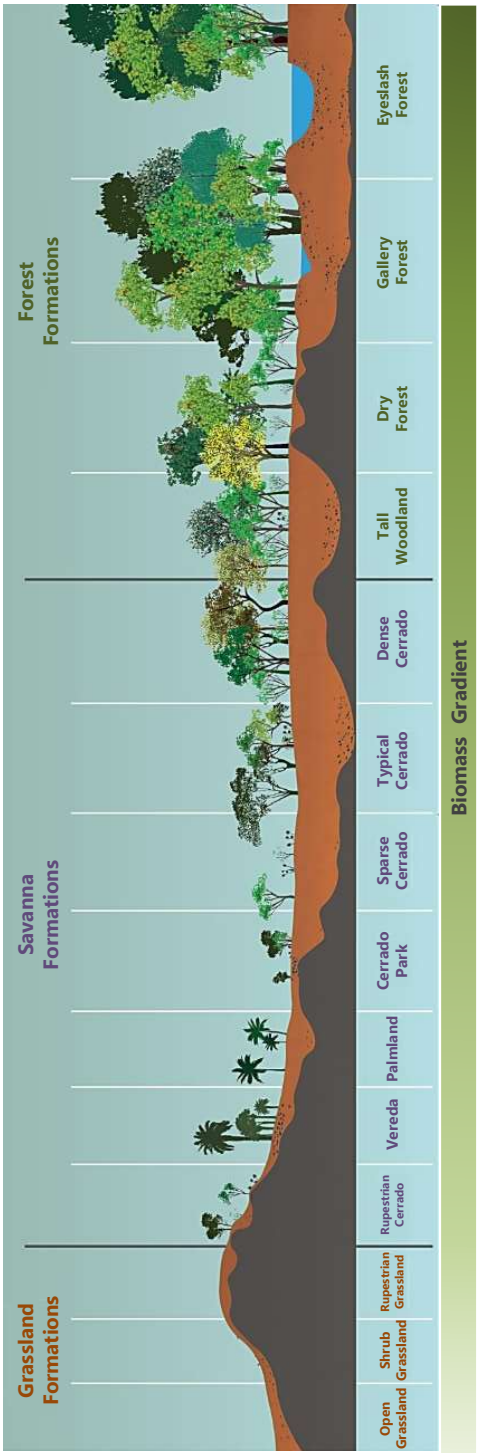


SOURCE: Author's own elaboration.

Cerrado holds the richest biodiversity among the tropical savannas (SANO et al., 2010), especially regarding the flora. That is mainly because of the high range of edaphic-climatic factors, which results in a diversity of plant-available moisture regime, latitude, chemistry of the soil, geomorphology, topography, and frequency of fire processes (CIANCIARUSO et al., 2005; RIBEIRO; WALTER, 2008). As a result, Cerrado presents 11 different types of physiognomies, varying from open cerrado grasslands (e.g. Open Grassland - *Campos Limpos*) to dense cerrado woodland (e.g. Tall Woodland - *Cerradão*) (Figure 2.2) (RIBEIRO; WALTER, 2008). These vegetation types vary both spatially and temporally. Spatially due to the Cerrado extension, and temporally by means of the season influences. Moreover, about 70% of the alive biomass is underground, representing thus an important storage of carbon (CASTRO; KAUFFMAN, 1998; KLINK; MOREIRA, 2002; MMA, 2014).

Nonetheless, the role of Cerrado is not limited to biodiversity, it also includes food security (KLINK; MOREIRA, 2002). Currently, Cerrado is one of the top grain and beef-producing regions in the world (PEREIRA et al., 2012). Even though in a first moment the Cerrado was considered unsuitable for cultivation, with the technological advances in the recent decades it has become an agricultural frontier, especially in the in the northern Cerrado, in a region known as MATOPIBA (an acronym of its location along: Maranhão, Tocantins, Piauí and Bahia states) (MIRANDA et al., 2014).

Figure 2.2 - Vegetation physiognomies in the Cerrado biome. In a biomass gradient, from the smallest - grasslands and savannas, to the largest forest formations, from the left to the right, respectively.



SOURCE: Adapted from Ribeiro and Walter (2008) and MMA (2017).

Although MATOPIBA is mainly located in Cerrado, it also embraces a small portion of Amazonia and Caatinga biomes, covering an area of approximately 730.000 km^2, which is two times larger than Germany. MATOPIBA presents, at the same time, the largest undisturbed remnants of the Cerrado vegetation and a quarter of the Cerrado's soybean area (MIRANDA et al., 2014; SOTERRONI et al., 2019), equally important for environmental and economic issues.

From this perspective, biodiversity and agriculture aspects unravel Cerrado's importance, yet only 6% of its native vegetation is located in integral protection areas[1] (FRANÇOSO et al., 2015). The severe changes in Land Use and Land Cover (LULC) in the last few decades have been threatened Cerrado (FEARNSIDE, 2001; BEUCHLE et al., 2015). Currently, it is estimated that only about 52% of Cerrado's vegetation remains (INPE, 2018a), and that the main driver of Cerrado conversion is the agricultural expansion (GIBBS et al., 2015).

Along the region, the indiscriminate use of fire to boost fresh grass growth and to open new agricultural areas is an important source challenge to overcome (COUTINHO, 1990; KLINK; MACHADO, 2005; MMA, 2014). This is especially true during the dry season, since there is a precipitation deficit and extreme vegetation conditions (MATAVELI et al., 2018). To support the mapping of fire, the Brazilian INPE/ *Programa Queimadas* [2] is the main initiative that uses RS data to continuously detect, monitor, and store the AF data. Coordinated by the Brazil's National Institute for Space Research (INPE), the Program uses RS data provided by 19 different satellites, and its database also provides the detection mapping since 1998 (INPE, 2018b). According to the dataset, in August, 2019, a total of 560,000 AF were identified in Cerrado (INPE, 2018b). However, fire detection in the region is complex because of its heterogeneity. Due to such an amount of AF, we selected August, 2019, to work in this study.

2.1.1 Fires in Cerrado

The interest in the effects of fire in Cerrado biome date back to the 1970s, and the idea of an integrated and cooperative project with such purpose was born in 1985, yet just in 1988, effective and extensive studies were conducted under the *Fire Project* (DIAS; MIRANDA, 2010). Since then, even though some issues are controversial and gaps remain, a lot of knowledge has been gathered. For instance, due

[1] Government states that this sum is of 9% instead of 6%, but the 3% remaining difference corresponds to deforested areas according to Françoso et al. (2015).

[2] https://queimadas.dgi.inpe.br/queimadas/portal

to the lack of a deep research on the problem, firstly, it was thought that the fire incidence was strictly human induced (COUTINHO, 1990; NASCIMENTO, 2001; DIAS; MIRANDA, 2010). However, fire signs date before the human establishment in the region (FERRAZ-VICENTINI, 1999; NASCIMENTO, 2001; MIRANDA et al., 2009), indicating the presence of natural fires. In this context, natural and human induced fire processes coexisted for more than 10,000 years in the region (MIRANDA et al., 2010; PIVELLO, 2011).

In Cerrado, natural fires are caused by showers with thunderstorms and lightning (RAMOS-NETO; PIVELLO, 2000). Besides, Brazil is one of the countries with the highest incidence of lightning, and the numbers have been increasing probably due to climatic events (PIVELLO, 2011). Such amount of lightning events also has the potential to influence the fire incidences in Cerrado. Rarely consuming taller woody plants, the fires in Cerrado occur normally on the surface level (MISTRY, 1998; MIRANDA et al., 2009), and in a high speed, presenting a velocity between 0.2 ms^{-1} and 0.5 ms^{-1}, with the potential to be faster according to the environment conditions (COUTINHO, 1990; MIRANDA et al., 2010). Without the incidence of fire in this environment, grasses may dominate the region, making it inappropriate for the rich fauna in Cerrado. For instance, fire enables the regrowth of several herbaceous stratum species, producing a suitable habitat for species such as the pampas deer (*veado-campeiro*) (ROSS, 1996). Moreover, in areas with fully forest encroachment, the loss reaches 27% in plant species, and 35% in ants (ABREU et al., 2017).

Cerrado presents an adapted environment to fire processes with morphological and physiological vegetation adaptations (PIVELLO, 2011; DURIGAN; RATTER, 2016). Nevertheless, human induced fires are more intense and frequent, impacting this environment on different levels, including biodiversity, nutrient cycling, and human health (COUTINHO, 1990; MISTRY, 1998; NASCIMENTO, 2001; MMA, 2009; MMA, 2014). Yet it is also used for hunting by indigenous (ANDERSON; POSEY, 1987; LEEUWENBERG; SALIMON, 1999; MELO, 2004). Currently, fire is the easiest and the cheapest way of land management. It enables the boosting of fresh grass growth for cattle ranching, and can also be used to open new agricultural areas (ROSS, 1996; MISTRY, 1998; RAMOS-NETO; PIVELLO, 2000; NASCIMENTO, 2001; MIRANDA et al., 2002; KLINK; MACHADO, 2005; PIVELLO, 2011).

Between 2003 and 2012, all the physiognomies presented an average superior to 7,500 AF per year (MMA, 2014). As identified by Coutinho (1990), fires in the Cerrado biome usually begin in May and reaches a peak in August. Almost 30 years

after Coutinho (1990) publication, it was also reported that the incidence of fire is concentrated in the dry season presenting variations within the biome, but mainly in September and October, since there is a precipitation deficit and extreme vegetation conditions (MATAVELI et al., 2017; MATAVELI et al., 2018). From both points of view, the interval between August and October presents the potential to affect huge areas, due to the low relative air humidity during the hottest hours of the day. In this manner, due to human interferences, fire processes in the region have changed.

The frequency of natural fire incidence in Cerrado is estimated to occur from one to nine years or more, while the interval of burn occurrence induced by humans drops from one to four years (COUTINHO, 1990). In such a way, since the earliest times, the main cause of fire in the Cerrado is human induced, as well as worldwide savannas' trends (ARCHIBALD et al., 2012).

2.1.2 Fire management initiatives in Cerrado

As a response to the lack of conservation approaches until the beginning of the century (NASCIMENTO, 2001; SILVA et al., 2006), different efforts have been recently made in order to manage the threats faced by Cerrado. However, efforts to monitor and manage the changes in LULC in Cerrado have been yet overlooked when compared with Brazilian Amazon (KLINK; MACHADO, 2005; BEUCHLE et al., 2015; STRASSBURG et al., 2017).

Due to the dual character of fire, which is necessary for the environmental management and at the same time destructive if more frequent and intense, *Zero-Fire* public policies were the source of intense discussions, once it aimed the total avoidance of the incidence of fire in Cerrado. On the one hand, fire processes control and support the development of a range of plants. Besides, longer periods without fire enable the increase of dead biomass, invasive grasses, and consequently the gathering of fuel load, which may lead to wildfires[3] mainly towards more open physiognomies once it is more fire prone than forest patches. In Cerrado, regardless of fire season, the percentage of dead fuel (mainly dead biomass of graminoids), is one of the most important variables related to fire intensity and flame height (RISSI et al., 2017). That means that the frequency of fires is inversely proportional to their intensity, and Zero-fire policies are not indicated (AGEE, 1998; RAMOS-NETO; PIVELLO, 2000; MORGAN et al., 2001; KLINK; MACHADO, 2005; FRIZZO et al., 2011; FIDELIS et al., 2018).

[3]Fire that is out of control (WHELAN, 1995).

Some governmental efforts implemented early dry season prescribed burning in protected areas with the objective of reducing the area and severity of late dry season fires (FRANKE et al., 2018). However, the main factors influencing fire intensity in areas with fire exclusion for two years are the dead fuel and fuel load, not the fire season (RISSI et al., 2017). In this manner, it would be suitable to control cool fires every four to six years in order to reduce fuel (PIVELLO, 2011), yet there is not a consensus about the most ideal practice. Since fire presents different perspectives, studies that focus on comprehending and gathering knowledge about fire characteristics are essential in order to support solid public policies and a suitable use of fire by local managers (MISTRY, 1998; RAMOS-NETO; PIVELLO, 2000; DURIGAN; RATTER, 2016; FIDELIS et al., 2018).

Besides, such initiatives have been made aiming to increase the adoption of sustainable practices and systems for agriculture in degraded or abandoned areas, for instance, through the training of technicians, farmers, and students in low carbon agricultural practices, reducing the use of fire. Furthermore, adaptive fire management actions were implemented considering its social, ecological, and economic importance through the hiring and training of fire fighters at different government levels, and the deployment of two operational bases for preventing and fighting forest fires (MMA, 2009; MCTIC, 2017).

A prominent example of a governmental policy is the Action Plan for the Prevention and Control of Deforestation and Forest Fires in the Cerrado (PPCerrado) (MMA, 2009). Created in 2010, PPCerrado is an instrument of the National Policy on Climate Change - PNMC (Law No. 12,187/2009). Aiming to reduce deforestation, forest degradation, and incidence of forest fires in the Cerrado biome, it presents three main action axes: i) Monitoring and Control; ii) Protected Areas and Land Use Planning; and iii) Fostering Sustainable Activities. As a consequence of its implementation, progress has been achieved from 2010 to 2015. For instance, a systematic monitoring of the changes in LULC has been enhanced through the TerraClass Cerrado project by means of a solid methodological protocol, providing information for decision makers and public policy (MMA, 2015).

Currently, the regulation of the fire use for ecological purposes is also present in the Law for Protection of Native Vegetation (Law 12.651/2012) and the Brazilian Integrated Fire Management Policy Bill (PL 11.276).

2.2 Fires

2.2.1 The nature of fire

Explaining the fire taxonomy is essential in order to have a deeper knowledge about the nature of the fire and its further developments. In the context of the science of nature, the Aristotle's four causes, also known as the four (be)causes or the four whys MacLennan (2018), were selected for a broader understanding of the fire phenomenon (Table 2.1). The four causes are: i) material cause: that out of which something comes to be (what something is made of); ii) formal cause: the form (essence, shape, or appearance); iii) efficient cause: the starting-point (agent of the event); and iv) final cause: for the sake of which an event comes out (end/goal of the event) (REECE, 2019).

Table 2.1 - Aristotle's four causes of fire.

Cause	Definition	Fire cause
Material	that out of which something comes to be (what something is made of)	Chemical reaction of oxygen, heat, and fuel
Formal	the form (essence, shape, or appearance)	Depends on the sensor. It can be seen through visible flames according to the human eyes, and heat release by remote sensing sensors on board satellites
Efficient	the starting-point (agent of the event)	Natural and human induced
Final	for the sake of which an event comes out (end/goal of the event)	Natural: discharge of electricity Human induced: land management and hunting

Source: Author's own elaboration.

The material cause of fires is the chemical reaction of oxygen, heat (ignition temperature), and fuel, which is commonly known as the *fire triangle*. Once the combustion chain reaction is also a requirement for the fire, this concept can also be incorporated as the so called *fire tetrahedron*. For the combustion, it is necessary that the fuel is heated enough to achieve its ignition temperature.

In order to support the comprehension of fire incidence on biomes (wildfires), the fire triangle was also adapted, for instance, formed by topography, fuel (vegetation), and air mass (COUNTRYMAN, 1972). In Table 2.2, we developed an adaptation for the here called wildfire triangle, which was based on different authors (Table 2.2) (MORITZ et al., 2005; FALK et al., 2007; PARISIEN; MORITZ, 2009; MCKENZIE et al., 2011; GOMES et al., 2018). Such elements influence the fire aspects (fire behaviour) (CRUTZEN; GOLDAMMER, 1993), and it presents a very high spatiotemporal variability (MCKENZIE et al., 2011; LEHMANN et al., 2014).

Table 2.2 - Fire and wildfire material cause elements, and their influential factors.

Fire triangle elements	**Wildfire triangle elements**	**Example**
Heat/ Ignition	Topography (landform)	Topographic effects and influences on fire spread (Slope, aspect, elevation)
Fuel	Vegetation (landscape)	Fuel biomass condition, postfire recovery, and distribution (fuel sizes and chemical composition)
Oxygen	Air mass (microclimate, weather, climate)	Temperature, wind, relative humidity, cloud cover, precipitation, air stability

Source: Based on Batchelder and Hirt (1966), Countryman (1972), Moritz et al. (2005), Falk et al. (2007), Parisien and Moritz (2009), McKenzie et al. (2011), and Gomes et al. (2018).

The formal cause of the fire is the flames, whose matter consists mainly of hot gases. In the process of combustion, there is the release of heat, light, and other reaction products (gases) in the form of flames. However, fire appearance can also vary according to the remote sensing sensor. While the human eyes see fires through the visible flames, remote sensing sensors on board satellites can be sensitive for instance to the heat release, which could support to identify the fire phenomenon based on the shape of a time series dataset, deeply described in the next sections.

There are two fire efficient causes along the Brazilian savannas: natural and human induced, which have been coexisting in the region for more than 10,000 years (MIRANDA et al., 2010; PIVELLO, 2011). Fire signs date before the human establishment in the region (FERRAZ-VICENTINI, 1999; NASCIMENTO, 2001; MIRANDA et al., 2009),

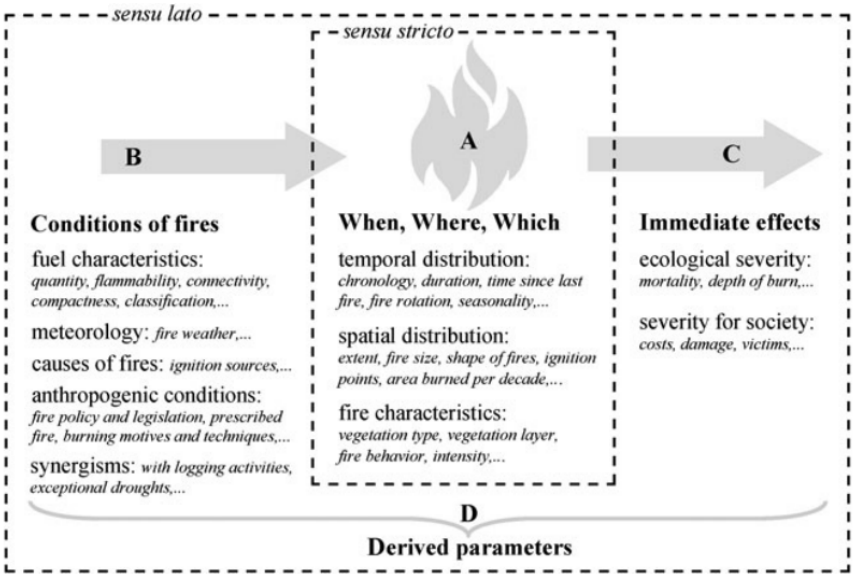
indicating the presence of natural fires.

The last but not least, the final cause of fires in Cerrado is also divided according to the agent of the fire. Natural fires are caused by rains with thunderstorms and lightning (RAMOS-NETO; PIVELLO, 2000), the final cause is the discharge of electricity. While the human induced fires are used for hunting by indigenous (ANDERSON; POSEY, 1987; LEEUWENBERG; SALIMON, 1999; MELO, 2004), and it is easiest and the cheapest way of land management, once it enables the boosting of fresh grass growth for cattle ranching, and can also be used to open new agricultural areas (ROSS, 1996; MISTRY, 1998; RAMOS-NETO; PIVELLO, 2000; NASCIMENTO, 2001; MIRANDA et al., 2002; KLINK; MACHADO, 2005; PIVELLO, 2011).

2.2.2 Fire regimes and behaviours

The concept of fire regime is used to designate an ensemble of fire-related parameters, which is divided into i) *sensu stricto* - when, where and which fire characteristics (Figure 2.3-A); and ii) *sensu lato* - conditions of the fire occurrence, which determines the characteristics of fire events, and immediate effects/impacts (Figure 2.3-B and C). Used in specific conditions, parameters derived from the aforementioned categories can be combined in order to develop methods for fire monitoring and modelling (KREBS et al., 2010) (Figure 2.3-D).

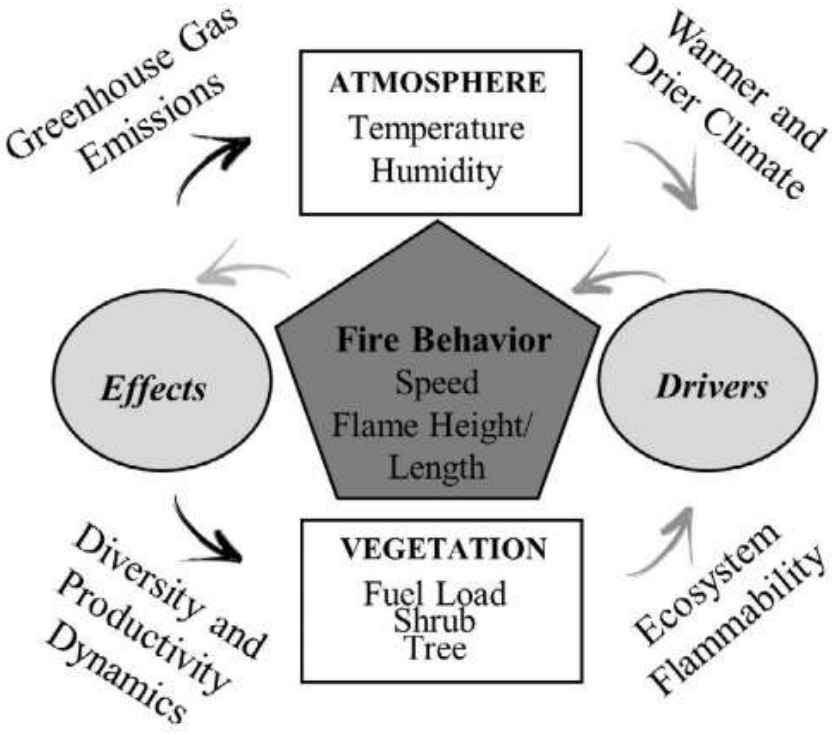
Figure 2.3 - Fire regime concepts.



SOURCE: Krebs et al. (2010).

Even though there are attempts to segregate the fire regime elements (Figure 2.3), it is not a trivial task. The whole mechanism of conditions for the fire occurrence (material cause elements also known as drivers), immediate impacts (effects), and specific fire behaviours (when, where and which) are integrated directly or not (Figure 2.4). In this manner, according to the drivers, fire can present different behaviours and immediate effects (COUNTRYMAN, 1972; KREBS et al., 2010). The generated impacts influence the atmosphere and land cover, which also influences the drivers.

Figure 2.4 - A holistic view of a fire regime: conditions for the fire occurrence (drivers), immediate impacts (effects), and specific fire aspects (fire behaviours).



SOURCE: Gomes et al. (2018).

Understanding and describing the fire behaviours and their effects are necessary for fire detection and management. Generally, the fire behaviours and their effects can be spatial, temporal and also spatiotemporal (Table 2.3). Among them, the fire severity does not present a single unit, but several possibilities, according to the impacted element and the unit of the analysis measurement. For instance, it can be related to the degree of the occurred change in the mortality of the overstory vegetation, which can refer to different characteristics such as the basal area, canopy cover, and forest types (SINGH et al., 2018).

Table 2.3 - Fire behaviours and effects description.

Dimension	**Behaviour/ effect**	**Description**	**Unit of measure/ symbol**
Spatial	Extent/size	Total scale of the fire or burned area	km^2
	Severity	Degree of change occurred (e.g. mortality of overstory vegetation)	-
	Spatial pattern / variability	Spatial heterogeneity of the fire/burned area	km^2, pixels, %, patches
Temporal	Frequency/ interval/ recurrence	How often on average fire returns to a given place (fire return interval)	Number of fires d^{-1}
	Duration/ residence time	How persistent the event is	d
	Seasonality/ timing	When the fire occurrs along a year or in relation to meteorological and phenological events (temporal pattern)	-
Spatio-temporal	Magnitude/ intensity	Amount of energy released by a flaming front (heat released)	Watts (W), kW, MW
	Expansion	Area burned by a fire for a day	km^2 d^{-1}
	Speed	Expansion (km^2 d^{-1}) divided by the length of the fire line (km) on the same day	km^2 d^{-1}

Source: Based on Agee (1998), Falk et al. (2007), Myers and Rodríguez-Trejo (2009), Archibald et al. (2012) and Andela et al. (2019).

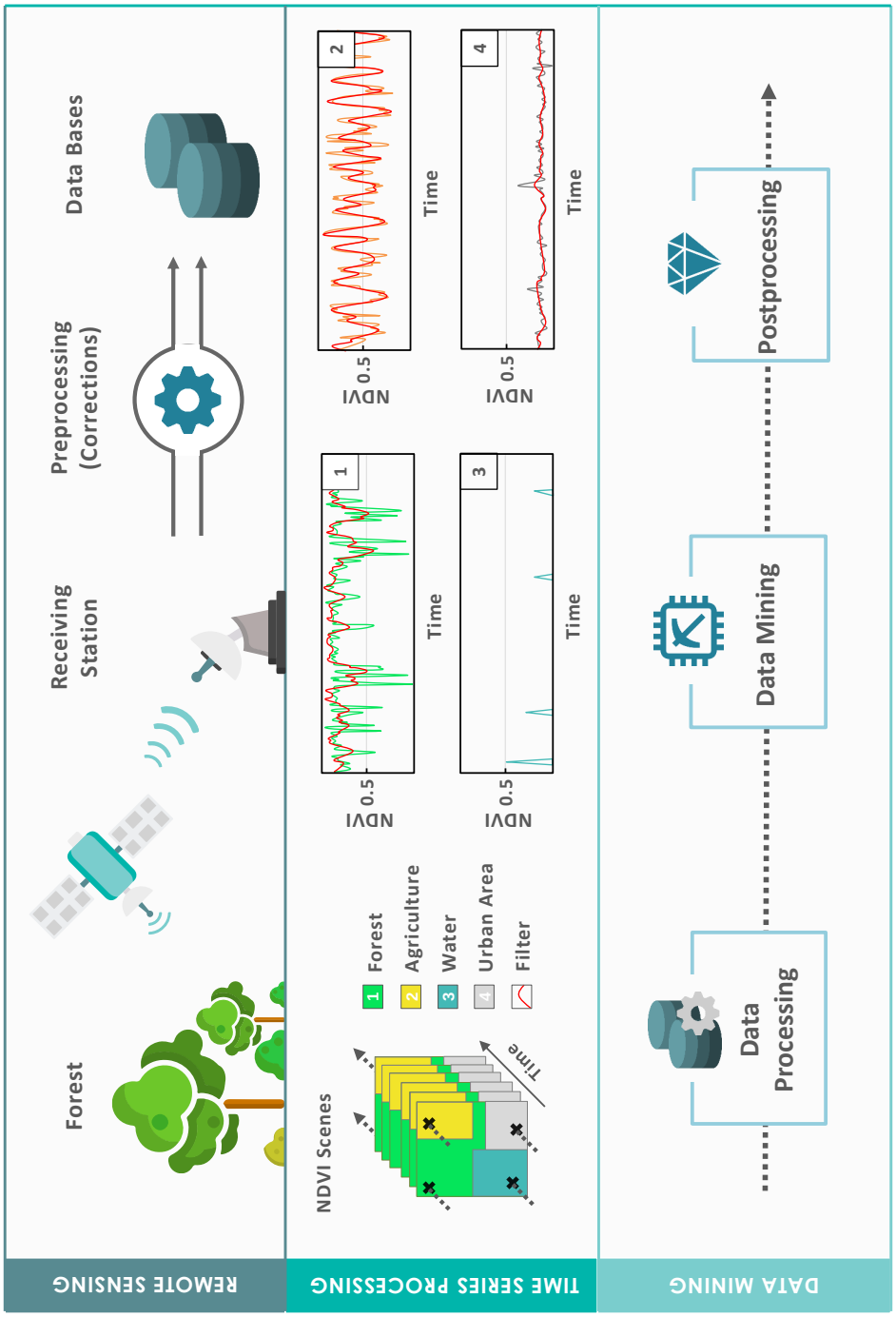
The pattern (variability) is closed related to landscape ecology metrics, whose units can be measured in square kilometers (km^2), pixels, percentage (%), and patches. GeoDMA, for instance, is a system of image analysis that supports the spectral and spatial features segmentation, as well as landscape-based features (KÖRTING et al., 2013). The magnitude (intensity) of a fire is related to the amount of heat released

per unit time. Once the fuel pursues chemical energy, when it is burnt the energy is transformed into thermal energy, and it can be measured by means of the heat release rate (HRR) (HADDEN, 2020). While in Table 2.3 we have the fire behaviours mainly horizontally distributed on the space, there is also the fire type as a vertical spatial distribution, which can be classified as surface fire (ground fuels), crown fire (aerial fuels), ground fire (organic soil fuel), or some combination of them (SCOTT et al., 2013).

2.2.3 Fire monitoring

The current knowledge about fire behaviour is mainly based on field works using controlled fires, restricted to small areas compared to natural environments (MIRANDA et al., 2010; GOMES et al., 2018). Nonetheless, fire processes are complex and present a spatiotemporal variability (DWYER et al., 2000), requiring a broad, real and long-term burning events analysis. The development of a feasible and suitable approach to support a large scale fire management, such as the FM presented in this thesis, required the integration of three main areas: i) Remote Sensing (RS) data, which is acquired by means of sensors on board satellites (JOYCE et al., 2009; LIBONATI et al., 2015; MATAVELI et al., 2017) (Section 2.3); ii) Time Series and Data Cube development (Section 2.4; and iii) Pattern recognition (Section 2.5) (Figure 2.5).

Figure 2.5 - Remote Sensing Time Series Data Mining.



SOURCE: Author's own elaboration.

2.3 Remote Sensing for fire management

In general, a RS sensor has four different resolutions: i) spectral - refers to the number and size of the bands/channels; ii) spatial - represents the measure of the smallest separation between two objects; iii) temporal - means the frequency of data acquisition; iv) radiometric - the sensitivity to detect differences in the signal strength (JENSEN, 2015). RS data present different taxonomies, for example, a sensor with a spatial resolution refiner than 10 m is considered a high spatial resolution sensor, between 10 and 50 m, is medium, and more than 50 m is low (EHLERS et al., 2002).

Furthermore, some mathematical operations engender synthetic images with proportional responses to the occurrence of certain targets in nature. In some cases, such procedure generates *spectral indices*, which may even boost the potential of the RS data (BANNARI et al., 1995).

Based on such RS dataset settings and on the phenomena, there is a plenty of possible uses of RS data. For instance, different sensors provide spectral bands with capacity to support fire studies. Detecting and monitoring fire events through the RS dataset are generally based on two main approaches: Active Fire (AF) products (*hot spots*) and Burned Area (BA) mapping (ANDERSON et al., 2005).

However, it is worth mentioning the strong trade-off between temporal and spatial resolutions. On the one hand, studies focused on small regions it is indicated refined spatial resolution data. On the other hand, for a more broad and regional fire analysis, coarse spatial resolution data are normally used since finer temporal accuracy is required for fire behavior studies. Consequently, a range of techniques has been developed for geosynchronous (temporal resolution of 60 min or less) and polar orbiting meteorological satellites (temporal resolution of 1 to 2 days) (GIGLIO et al., 1999), at the expense of the spatial resolution. Even though over larger areas coarse spatial resolution sensors are mainly used to fire activity analysis (CHUVIECO et al., 2019), they can also be used for BA estimation (OLIVA; SCHROEDER, 2015).

2.3.1 Active Fire detection

The AF data are normally based on middle infrared (3.9 μm) and thermal (10.7 μm) bands from sensors on board geosynchronous (temporal resolution of 60 minutes or less), and polar orbiting satellites (temporal resolution of 1-2 days), presenting a coarse spatial resolution (250 m, 2 km) (GIGLIO et al., 1999; WEAVER et al., 2004).

Even though a fire occupies just a fraction of the pixel, it can increase the brightness in the entire pixel, which indicates that low spatial resolution data are also suitable for this task. However, the flux of radiance should be sufficient to be detected, but not so intense to cause saturation in the pixel (ROBINSON, 1991).

Also around 4 μm, Fire Radiative Power (FRP) can be obtained, which is associated to the fuel consumption and smoke emission rates (WOOSTER et al., 2005; XU et al., 2010), thus a measure of fire intensity and destructive power (CALLE; CASANOVA, 2008; GIGLIO et al., 2008). If sufficiently fast, the FRP data could be useful for fire management, once it could be used to characterize the frontal strength of the fire (ZHUKOV et al., 2006). Meanwhile, the fire light can be detected on nocturnal data through the visible wavelengths (CHUVIECO; KASISCHKE, 2007), although it would be limited by the time interval. In summary, the AF products are able to detect and characterize current fire spots, while FRP products are indicatives of intensity.

RS data have been widely used to identify and monitor fire products at different spatial scales (JUSTICE et al., 2002; SCHROEDER et al., 2016). However, according to the user's necessity, it is important to take into account the pros and cons of each application (MORGAN et al., 2001). RS products with coarse spatial but with the finer temporal resolution are indicated to better comprehend fire behaviour through AF datasets, such as Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectrometer (MODIS), Visible Infrared Imaging Radiometer Suite (VIIRS), and Advanced Baseline Imager (ABI) (Table 2.4). The meteorology satellite Himawari-8 carries the geostationary sensor, Advanced Himawari Imager (AHI), and presents similarities with the ABI. Nonetheless, Himawari-8 provides data only from the western Pacific. Because of that, we do not present here further details.

Table 2.4 - Main current coarse spatial resolution sensors for fire analysis: Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectrometer (MODIS), Visible Infrared Imaging Radiometer Suite (VIIRS), and Advanced Baseline Imager (ABI).

Sensor (satellite)	**Resolution**		
	Spectral (number of bands)	**Spatial (m)**	**Temporal**
AVHRR (NOAA-15 and 19)	6	1,100	1-2 days
MODIS (Aqua and Terra)	36	250-1,000	1-2 days
VIIRS (Suomi-NPP)	22	375-750	1-2 days
ABI (GOES-16)	16	2,000	05-10 minutes

Source: Author's own elaboration.

As reported by Csiszar et al. (2005), the monitoring of AF through a systematic satellite approach began in the 1980s with the sensor AVHRR. Currently, the optical multispectral sensor AVHRR-3 is aboard NOAA (National Oceanic and Atmospheric Administration) polar orbiting satellites (NOAA-15 and NOAA-19), and presents six spectral bands, with a spatial resolution of about 1.1 km at nadir and 1-2 days temporal resolution (NOAA, 2014).

Launched a year after the AVHRR, MODIS sensor was the first instrument with specific band characteristics for fire detection. Onboard NASA's Terra and Aqua satellites, the MODIS sensor has provided global fire data for over a decade. It is a sun-synchronous orbit sensor, with 36 bands, and spatial resolution of 250 m, 500 m, and 1 km, and temporal resolution from 1 to 2 days (JUSTICE et al., 2002). The smallest fire detected by MODIS is around 50 m^2 (GIGLIO et al., 2018). MODIS is already working with Collection 6, which aims to address Collection 5 limitations such as false AF detections. With 1-km spatial resolution, Collection 6 is driven mainly by regional differences and fire sizes (GIGLIO et al., 2016).

VIIRS sensor is onboard two different satellites, the Suomi National Polar-orbiting Partnership (S-NPP) and NOAA-20. Its AF products were designed based on the previous MODIS Fire Thermal Anomalies algorithm in order to support data continuity (SCHROEDER; GIGLIO, 2018). VIIRS, with a 375 m spatial resolution, has already been validated by different studies proving to be superior in detecting small AF when compared with MODIS (CAO et al., 2013; CSISZAR et al., 2014; SCHROEDER et al., 2014). Besides, it presents 22 spectral bands and a temporal resolution of 1-2 days.

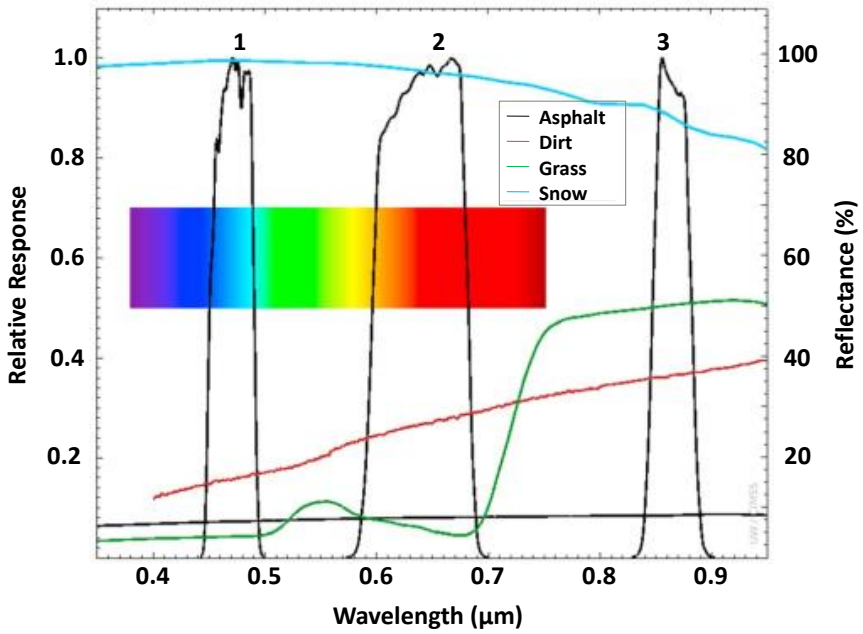
The Geostationary Operational Environmental Satellite system (GOES) is a joint effort of the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) and comprehends a constellation of satellites. GOES-16, launched in November, 2016, is the first satellite from GOES-R Series. The multispectral imager instrument ABI is onboard GOES-16, and it presents 16 spectral bands (Table 2.5), with a nominal spatial resolution of 2 km at nadir, and a full disk image every 10 minutes over North and South Americas (SCHMIT et al., 2017; SCHMIT et al., 2018). That means 144 daily remote sensing imagery acquisition (≈8 GB).

GOES-16 ABI also presents a wide range of products[4] and uses. Some of the spectral bands are more likely to support the AF detection and monitoring than others. The description of the GOES-16 ABI bands is presented as follow.

Band 01 (B01) provides information for smoke and aerosols monitoring (LINDSTROM et al., 2017a). Band 02 (B02) pursues the finest spatial resolution of all ABI bands. As such, it is used to identify features of small-scale, including boundaries and clouds (LINDSTROM et al., 2017g), while Band 03 (B03), known as *Veggie*, has the potential to detect burned areas through the computation of vegetation indices (LINDSTROM et al., 2017h). Considering a visible rainbow spectrum, it is possible to allocate the B01 and B02, and the location of the B03, referred as the *vegetation* band, presenting different potential according to the analyzed target, as shown in Figure 2.6. Although B03 is not centered in the *green* (0.55 μm), it is possible to use the three bands to generate true color RGB images, with the Veggie band instead of the green band, or by combining the B01, B02, and B03 to generate the lacking band (BAH et al., 2018).

[4]The whole list of products as well as quick guides are freely available respectively at: `www.ncdc.noaa.gov/data-access/satellite-data/goes-r-series-satellites`; and `http://rammb.cira.colostate.edu/training/visit/quick_guides/`

Figure 2.6 - Bands 01, 02, and 03 GOES-16 ABI spectral response along with the reflectance spectra for different features, asphalt, dirt, grass, and snow.

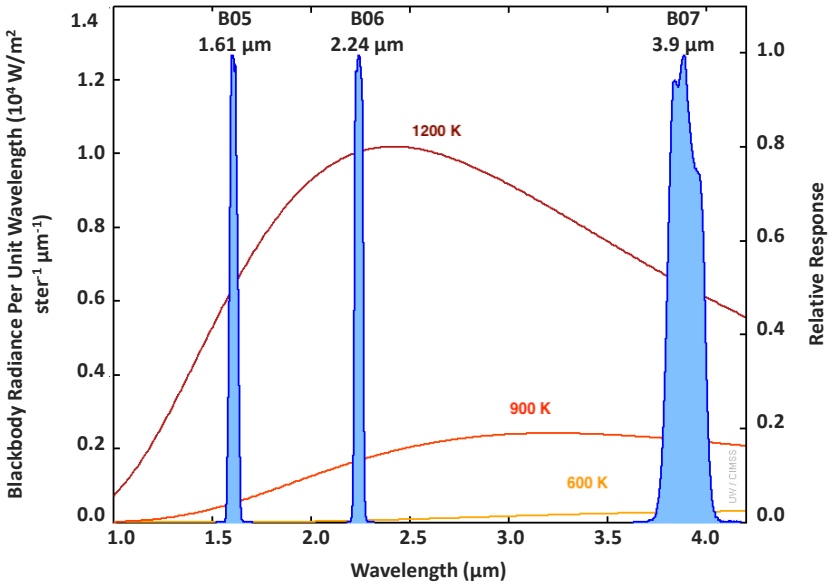


SOURCE: Adapted from Bah et al. (2018).

The Cirrus band (B04) detects thin cirrus clouds during the daytime (LINDSTROM et al., 2017i). Furthermore, in Band 05 (B05) during the daytime, liquid water clouds are bright, and ice clouds are darker, supporting the inference of the cloud phase. Its nighttime application is regarding the detection of very hot fires (LINDSTROM et al., 2017j). Moreover, during the day, it is also possible to compare B05 and B02 to discern AF in areas free of clouds (SCHMIT et al., 2018). Band 06 (B06) is used in conjunction with other bands for a couple of applications, including estimating cloud particle sizes, to create cloud masking and to detect hot fires in the absence of clouds (LINDSTROM et al., 2017k). Band 07 (B07) also can be used along a whole day. At night, it can identify fog and low clouds, and during the daytime among others, to detect AF and to estimate the temperature of sea-surface. Due to bands 05, 06, and 07 potential to detect AF (Figure 2.7), a product known as *Fire Temperature RGB* (FT-RGB) can be generated through a RGB composition (R7;G6;B5). While the True Color RGB can be used to show the fire smokes and BA, the FT-RGB indicates AF. However, there are some limitations: clouds may hide fire signals, and dry regions may present false positive AF (NOAA AND NASA, 2018), once there are

spatio-temporal color variations (SEAMAN et al., 2017; SCHMIDT, 2019).

Figure 2.7 - Spectral Response of Bands 5, 6 and 7 from GOES-16 with the different hot target spectral behaviours.



SOURCE: Adapted from Bachmeier (2019).

The Bands 08 (B08), 09 (B09), and 10 (B10) are the water vapor bands on the ABI, and its primary use is atmospheric feature identification (BACHMEIER et al., 2017c; BACHMEIER et al., 2017d; BACHMEIER et al., 2017a). The Band 11 (B11) is important for volcanic activity monitoring (LINDSTROM et al., 2017b), while the Band 12 (B12) in combination with other bands shows the dynamics of the atmosphere near the tropopause (LINDSTROM et al., 2017c), which is the boundary between the troposphere and the stratosphere (HAQQ-MISRA et al., 2011). Band 13 (B13) is used among others to improve the atmospheric moisture corrections and to identify and classify atmospheric features (BACHMEIER et al., 2017b). Moreover, the Band 14 (B14) is used in a range of products such as to support the analysis of land surface Temperature, and fire (LINDSTROM et al., 2017d). Also, part of different products, including Cloud Top Properties, and AF characterization, the Band 15 (B15) is used to identify moisture and dust (LINDSTROM et al., 2017e). Finally, the Band 16 (B16) is used due to its potential to highlight tropopause aspects, applied thus in a range

of products, including cloud mask (LINDSTROM et al., 2017f).

Table 2.5 - GOES-16 spectral bands.

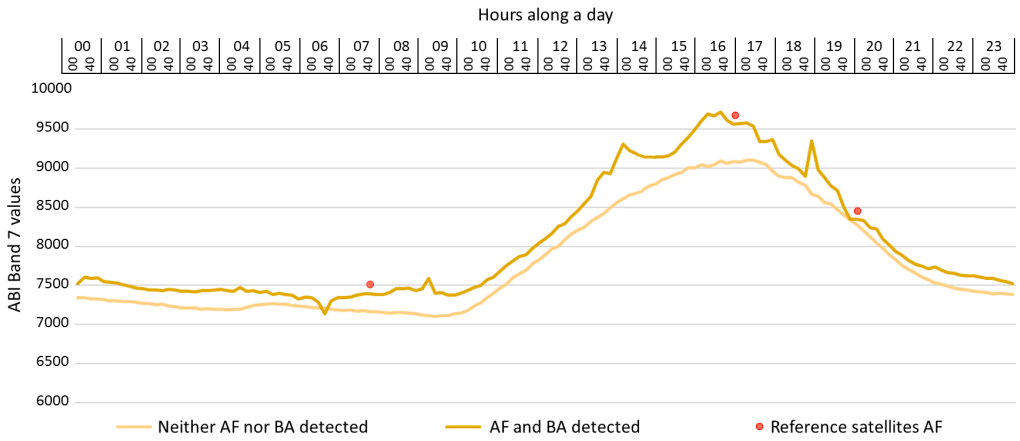
Band Number	**Central Wavelength (μm)**	**Type**	**Nickname**	**Spatial Resolution (km)**
01	0.47	Visible	Blue	1
02	0.64	Visible	Red	0.5
03	0.86	Near-Infrared	Veggie	1
04	1.37	Near-Infrared	Cirrus	2
05	1.6	Near-Infrared	Snow/Ice	1
06	2.2	Near-Infrared	Cloud particle size	2
07	3.9	Infrared	Shortwave window	2
08	6.2	Infrared	Upper-level water vapor	2
09	6.9	Infrared	Midlevel water vapor	2
10	7.3	Infrared	Lower-level water vapor	2
11	8.4	Infrared	Cloud-top phase	2
12	9.6	Infrared	Ozone	2
13	10.3	Infrared	"Clean" longwave window	2
14	11.2	Infrared	Longwave window	2
15	12.3	Infrared	"Dirty" longwave window	2
16	13.3	Infrared	CO_2 longwave	2

Source: NOAA and NASA (2022a).

Among the spectral bands, the ABI Band 7 (3.90 μm) is the most indicated for AF detection, once its short wavelength is more sensitive to the hottest part of a pixel (NOAA AND NASA, 2021). Nonetheless, two of the main limitations of such dataset are: i) small fires can be overlooked, and ii) solar reflectance can influence the ABI Band 7 values (NOAA AND NASA, 2021). Figure 2.8 shows an example of

the ABI Band 7 on a given day with and without the presence of fire, based on the reference satellites and on the manually mapped BA.

Figure 2.8 - ABI Band 7 values for two different pixel locations along a day with and without the detection of AF and BA.



AF: Active Fire; BA: Burned Area.

SOURCE: Author's own elaboration.

The bands used by the *Programa Queimadas* are around 0.6 μm, 3.9 μm, and 11.0 μm, yet the first and last wavelengths are used to remove noises during the day. The global Wild Fire Automated Biomass Burning Algorithm (WF_ABBA) uses the bands around 0.6 μm (optional), 3.9 μm, 11.0 μm (optional), and 12.0 μm (HOFFMAN et al., 2011). Once the AF is detected, notifications are sent to the fire public managers, such as firefighters. An example of such bulletins is found in Python programming language at LINDLEY et al. (2016).

Comparatively, while MODIS can identify smaller fires due to the finer spatial resolution, ABI may pursue fewer false alarms because of the high temporal information available (SCHMIDT et al., 2010), being thus suitable for more large fires (ROBINSON, 1991). According to the literature, detecting a fire through the ABI dataset is possible if the fire incidence is intense, a difference between the middle infrared and thermal bands of about 10°-15°C, or wide enough, 2% of the pixel area (0.08 km^2) (WEAVER et al., 2004). In this manner, low spatial resolution data are also suitable for AF detection, once a small fraction of the pixel with fire can already increase the brightness of the entire pixel. Although the ABI sensor presents such potential with an ultra-high temporal resolution, the use of this sensor in the litera-

ture is yet scarce in the regional scale, probably due to its coarse spatial resolution and high complexity for data processing and analysis (PLETSCH et al., 2019a).

2.3.2 Burned area mapping

In the 1980s, the first BA products were based on medium resolution sensors (CHUVIECO et al., 2019), such as the Landsat Series. The BA mapping is mainly detected based on visible and infrared wavelength bands from near-polar and sun-synchronous orbit satellites (temporal resolution of 8-16 days), and with a medium spatial resolution (20-30 m) (CHUVIECO et al., 2019). The same approach could also be used to detect fire smoke, but the direct use of plumes is not so as usual and efficient as the others approaches. In the case of BA, the detection is based on two main procedures, the charcoal (char) deposition, and on the effects of the vegetation spectral response, fire scar. The first one is quickly vanished due to wind and rainfall, and the later one is more persistent, resisting for a couple of days or weeks in tropical grasslands, and years in boreal ecosystems (PEREIRA et al., 1997; ROY et al., 2002; CHUVIECO et al., 2019).

With 30 m spatial resolution, and temporal resolution of about 16 days, Landsat satellites are commonly used for fire studies, but they are usually limited to a local scale. Another example of medium spatial resolution is the satellites Sentinel-2 A and B, presenting a temporal resolution of 5 days (ESA, 2015). Due to their similarities, studies have been developed aiming to integrate Landsat and Sentinel-2 datasets in order to refine the temporal resolution without jeopardizing the spatial resolution (ROY et al., 2019), and also to extrapolate methods from one sensor to another, such as the automatic BA mapping, which uses a dependent threshold of the image statistics to detect BA (WOŹNIAK; ALEKSANSDROWICZ, 2019).

Even though Sentinel-2 data does not provide the exact time of the fire occurrence, for BA mapping, it is one of the most suitable free available dataset, since it presents a 10-20 m spatial resolution and a revisit time of five days. For a BA mapping by means of Sentinel-2 imagery at Sentinel Hub viewer (available at: `https://www.sentinel-hub.com/`), the false color composite shortwave infrared (SWIR), RGB (B12, B8A, B04) is the most suitable, once it enables the fire damage mapping (Sentinel Hub, 2021) (Table 2.6).

Table 2.6 - Sentinel-2 bands most indicated for BA mapping.

Bands	Central Wavelength (μm)	Resolution (m)
B4 - Red	0.665	10
B8A - Narrow NIR	0.865	20
B11 - SWIR1	1.610	20
B12 - SWIR2	2.190	20

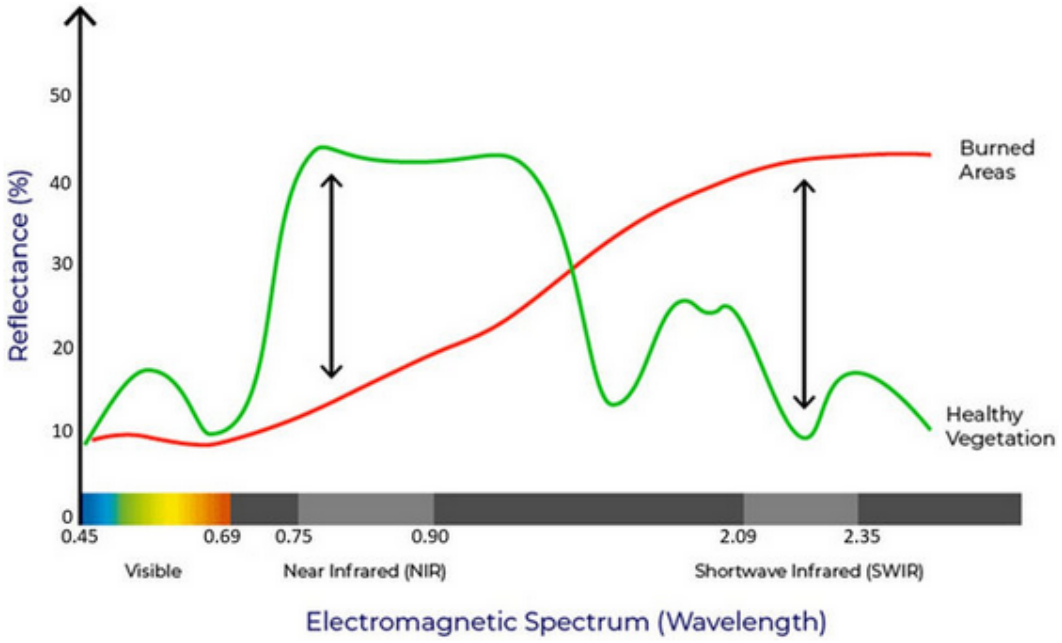
Source: Adapted from Sentinel Hub (2021).

Finally, the Planet imagery is also remarkable and could be used for BA mapping. Operated by Planet, the PlanetScope is a constellation of approximately 130 satellites. Providing daily images from the entire land surface of the Earth, its spatial resolution is approximately 3 meters (PLANET DEVELOPERS, 2022). Planet imagery access is possible by means of commercial agreements.

Aiming to highlight the affected areas, BA mapping is possible through some approaches, that may include the dimension of time (See Section 2.4), and Spectral Index (SI) (Table 2.7), which are mathematical operations performed on RS imagery (BANNARI et al., 1995; KEY; BENSON, 2005). SI with such focus normally use both near (NIR) and short-wave (SWIR) infrared bands (Table 2.6), where NIR presents a strong reflectance decrease after the burning process, while the dryness results in an increase in the SWIR reflectance (CHUVIECO et al., 2019), as indicated by reflectance response of healthy vegetation and burned areas presented in Figure 2.9. Besides, according to VAN DIJK et al. (2021), the SWIR bands 11 and 12 from Sentinel-2 have also proven to be effective to distinguish between unburned and BA.

The Normalized Difference Vegetation Index (NDVI) (Table 2.7), which ranges from [-1.0, +1.0], is a SI that presents information regarding the presence of live green vegetation in RS images, and was first developed based on the Landsat dataset (ROUSE et al., 1974). The Burned Area Index (BAI), Mid-Infrared Burn Index (MIRBI), and Normalized Burn Ratio (NBR) aims to highlight burned areas. BAI was tested on Landsat TM and NOAA images (CHUVIECO et al., 2002), while MIRBI, on MODIS (Moderate Resolution Imaging Spectroradiometer) data and on Landsat TM, they presented satisfatory results in savannas (TRIGG; FLASSE, 2001). NBR was also designed for the Landsat dataset, as well as the Normalized Burn Ratio Thermal (NBRT). The NBRT particularity is the use of the thermal band, but at the time

Figure 2.9 - Spectral response of healthy vegetation and burned areas.



SOURCE: USDA (2022).

of its development, Landsat thermal band ranged [10.40 - 12.50 μm], and Landsat 8 already presents two finer thermal bands.

A new SI for BA detection is the NBR+, which was developed based on Sentinel-2 imagery (ALCARAS et al., 2022). Presenting exceptional results, the NBR+ has the advantage of excluding part of false positives, which is mainly common due to the presence of clouds or water bodies. The SI results also has the potential to be combined and refined in order to highlight phenomena (PLETSCH et al., 2019b). Even though there is a range of SI, some challenges remain, such as the variations between the so called burned and unburned areas (CHUVIECO et al., 2019).

Table 2.7 - Main spectral indices used for burned area mapping. ρ: surface reflectance factor; NIR: near-infrared band; SWIR: short-wave infrared band; LSWIR: longer short-wave infrared band; STIR: scaled brightness temperature of the thermal band (TIR) divided by 10,000.

Spectral Index	Acronym	Equation	Reference
Normalized Difference Vegetation Index	NDVI	$\frac{\rho NIR - \rho Red}{\rho NIR + \rho Red}$	Rouse et al. (1974)
Burned Area Index	BAI	$\frac{1}{(0.1 + \rho Red)^2 + (0.06 + \rho NIR)}$	Chuvieco et al. (2002)
Mid-Infrared Burn Index	MIRBI	$10 * \rho LSWIR - 9.8 * \rho SWIR + 2$	Trigg and Flasse (2001)
Normalized Burn Ratio	NBR	$\frac{\rho NIR - \rho SWIR}{\rho NIR + \rho SWIR}$	Key and Benson (2005)
Normalized Burn Ratio Thermal	NBRT	$\frac{\rho NIR - (\rho SWIR * STIR)}{\rho NIR + (\rho SWIR * STIR)}$	Holden et al. (2005)
Normalized Burn Ratio Plus	NBR+	$NBR+ = \frac{(SWIR2 - Narrow\ NIR - Green - Blue)}{(SWIR2 + Narrow\ NIR + Green + Blue)}$	Alcaras et al. (2022)

Source: Author's own elaboration.

2.4 Time series and data cube development

At a first moment, RS data was observed mainly spatially, but with the advance in computer algorithms and the amount of available data, researchers identified that it was also possible to take into account the time as an extra dimension. In this manner, values ordered in time could enable the identification of trends, seasonal patterns and even predict future values. For that, when necessary and according to the study goal, certain Time Series (TS) processing is required, such as the harmonization when dealing with different RS imagery sources, and calibration when data is not yet calibrated. However, the creation of a Data Cube (DC) may lead to an information loss during the TS processing, for instance, by means of the use of filtering algorithms (APPEL; PEBESMA, 2019).

A TS can be composed of different RS data, such as raw digital numbers, reflectance

values (KUENZER et al., 2015), or other subproducts as spectral indices. RS vegetation indices, such as NDVI, are broadly used to analyze phenological metrics (also known as phenometrics), which refers to the observation of seasonal pattern of vegetation changes (REED et al., 2009). Due to the Big Data challenges, DC infrastructures are considered an important tool able to manage and share a great amount of data (NATIVI et al., 2017).

Generating a TS is possible by staking the RS data in a chronological order, and the product is known as Data Cube (DC) (Figure 2.5). According to Bovolo et al. (2018), multitemporal data analysis can be classified as bi-temporal and TS data, regarding the use of a few and a dense quantity of data, respectively. In this context, after the RS data acquisition (Figure 2.5), the next steps required to use RS datasets for fire management are the Time Series (TS) processing and Data Cube (DC) development.

The mathematical models used in TS are known as *deterministic* and *stochastic*. In the first one, the exact calculation of an element trajectory along the time is possible. Considering that unknown variables may influence the data, natural phenomena are rarely totally deterministic. On the other hand, when it is not possible to exactly predict the behaviour of an object, probability (stochastic) models are indicated (BOX et al., 2015). Due to the RS nature, we will consider hereafter the TS dataset as stochastic, requiring thus probabilistic analysis.

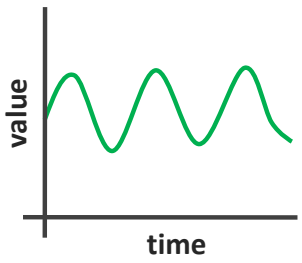
The step of obtaining usable information and analysing data patterns in DC is possible through Data Mining (DM) techniques, which include automatic or semi-automatic mechanisms of searching, discovering, and extracting relevant information in a database (HAN et al., 2006; WITTEN et al., 2016) using the integration of multiple approaches from Statistics and Machine Learning. For that, it is essential to analyze the TS, which is according to Dodge (2008), the process of making the mathematical descriptions and estimation of the four TS components:

- Trend - general tendency of data over a long period;
- Seasonal variations - changes that occur in a regular and periodic manner;
- Cyclical fluctuations - periodical changes but not with seasonal variations. Often used for business cycles in economic data, for instance, prosperity, recession and depression;
- Irregular variations (remainder) - random variations (unforeseen and unpredictable).

In case the mean, variance and covariance do not vary in a TS, it is called stationarity (Figure 2.10), and if it does, non-stationary (Figure 2.11). To predict phenomena, it is indicated the use of stationarity TS. For that, it is necessary to perform a "de-trending" (removal of long-term linear trend), and/or a "de-seasonalisation" (removal of seasonal variations) approach on the non-stationary TS, whose two main elements are trend and seasonality. In some cases, smoothing the TS is already enough. In others, it is necessary to apply decomposition techniques, which refers to computing differences in a particular time lag, and modeling the components to remove trend and seasonality, respectively (JAIN, 2016).

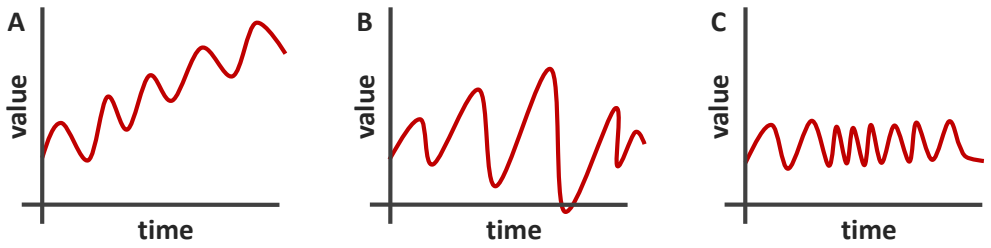
Finally, the decomposition of the TS can be additive or multiplicative. While the additive model assumes that the TS components are independent from each other and the TS is a function of the sum of its components, the multiplicative model considers that the components can influence one another and the TS is a function of the product of its components (ADHIKARI; AGRAWAL, 2013). In this manner, for the "de-trending", it would be necessary to subtract the trend estimates (additive decomposition) or to divide the series by the trend values (multiplicative decomposition).

Figure 2.10 - Example of stationary TS, with no variation of mean, variance, and covariance.



SOURCE: Singh (2018).

Figure 2.11 - Example of non-stationary TS, with a variation of: A. Mean; B. Variance; C. Covariance.



SOURCE: Singh (2018).

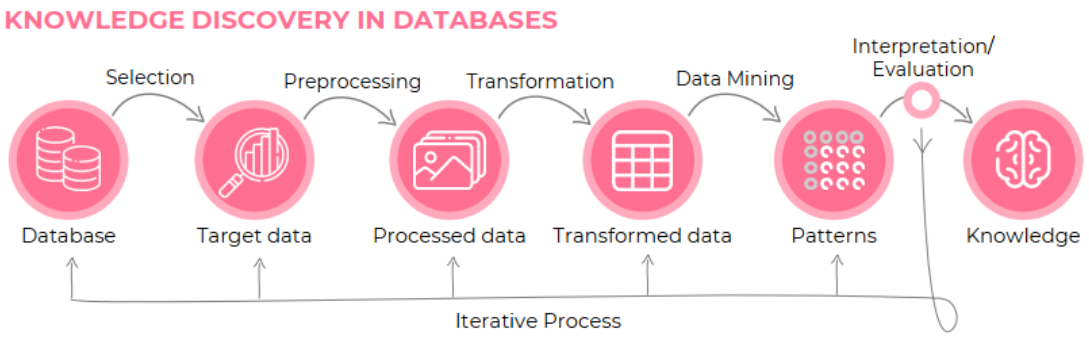
2.5 Pattern recognition

A pattern is a set of measurements that is able to describe a phenomenon. In this manner, in order to identify certain phenomena, it is essential to recognize patterns. By means of that, we can have a better comprehension of regions, which is imperative for a sustainable development and preservation. A popular object of study for several decades, the pattern recognition techniques for RS imagery have been continuously developed. However, the increasing resolution of data, expansions of data volume and complexity are yet the main open challenges. The same is also true for Cerrado, where not only the RS dataset but also geospatial methods based on DM and ML are essential to process and extract useful information from the data (FONSECA et al., 2021).

2.5.1 Data science: Discovering knowledge from databases

Data science is a broad field of study that includes all the processes related to data projects, whether the data is structured (i.e., Excel files - tabular format with relationship between the different rows and columns), semi-structured (i.e. XML files), or unstructured (i.e., RS imagery). It takes advantage of Big Data and the myriad of other field studies and approaches. The Knowledge Discovery from Databases (KDD) (Figure 2.12) is a subset of Data Science that aims to find useful information in a dataset and to use the information to recognize hidden patterns.

Figure 2.12 - Knowledge Discovery from Databases (KDD).



SOURCE: Adapted from Fayyad et al. (1996), Han et al. (2006), Yeung and Hall (2007).

KDD is a pipeline for extracting knowledge from large datasets and presents seven iterative steps (FAYYAD et al., 1996; HAN et al., 2006; YEUNG; HALL, 2007):

- Data Cleaning: to detect and treat noises and inconsistent data, such as missing values, which should be either omitted or replaced. Specific mathematical treatments are indicated according to the analysis purpose;
- Data Integration: multiple data sources may be combined and stored in a data warehouse;
- Data Selection: relevant data is retrieved from the database for analysis;
- Data Transformation: data are transformed and consolidated into appropriate structures for modelling, for instance, by means of aggregation operations. In this step, it is also possible to apply data reduction to obtain a smaller data representation without losing the data integrity. In some cases, the Data Transformation can occur before the Data Selection;
- Data Mining: methods are applied to extract data patterns;
- Pattern Evaluation: evaluation of the DM results;
- Knowledge Representation: visualization and communication of the results.

Data Cleaning and Data Integration can be considered as part of the preprocessing step. After that, Data Selection and Transformation are necessary to comprehend

and prepare the data to apply suitable statistical tools. After those steps, there is the DM application, and the post-processing, which includes the Pattern Evaluation and Knowledge Representation (Figure 2.5). In this manner, Data Mining (DM) is one of the seven aforementioned steps. Nonetheless, DM is commonly used as a synonym for KDD. Due to its broad use, in this thesis review we adopted the term Data Mining for the whole process of KDD.

DM is performed through the integration of multiple approaches including Statistics and Machine Learning (ML). According to Fayyad et al. (1996), broadly, there are six main primary DM methods for data description and prediction:

- Classification: a function that classifies the data;
- Regression: a function that discovers the functional relationship between variables;
- Clustering: identify clusters to describe the data;
- Summarization: finding a compact description for a subset of data;
- Dependency Modeling: description of significant dependencies between variables;
- Change and Deviation Detection: based on previously measured or normative values, it is the process of discovering the most significant changes in the data.

Due to the complexity of applying DM to discover knowledge from databases, Cross Industry Standard Process for Data Mining (CRISP-DM) framework was developed aiming to make large DM projects faster, more reliable, repeatable, manageable and at the same time with reduced time and costs (WIRTH; HIPP, 2000). Basically, CRISP-DM incorporated KDD generating six phases (Figure 2.13). The sequence of the phases is neither strict nor it is linear.

Figure 2.13 - Cross Industry Standard Process for Data Mining CRISP-DM.

SOURCE: Adapted from Wirth and Hipp (2000).

Following, the six phases of CRISP-DM are described based on Wirth and Hipp (2000):

- Business Understanding: firstly, it is necessary to understand the DM problem that needs to be solved, the project objectives, requirements from a business and strategy perspective, and a preliminary plan to achieve the objectives;
- Data Understanding: this phase comprehends the initial data collection and all the activities to get familiar with the data, first insights, and hypotheses. Business Understanding and Data Understanding have a close connection, since the formulation of the DM problem and the creation of a preliminary action plan requires some understanding of the dataset;

- Data Preparation: this phase includes all the activities to create the final dataset/DC, including the KDD phases: data cleaning, data integration, data selection, and data transformation for modeling, such as adding labels and splitting data into training and testing set;

- Modeling: in order to perform DM and recognize patterns in large datasets, Machine Learning (ML) models are commonly used and indicated, once algorithms are created to perform complex tasks. Thus, in this phase, different modeling techniques based on ML and Statistics can be selected, applied and assessed. Finally, there is a close relation between Data Preparation and Modeling, once different models may require certain data structures;

- Evaluation: based on the Business Understanding/issue to be solved, in this phase, the results are interpreted and evaluated. The evaluation is not only technical, but also theoretical: *the developed model really resolves the problem?*. In this manner, there is a review process and discussions in order to create a list of possible actions for the next steps;

- Deployment: the actions that need to be carried out in order to make use of the developed model. It can be the creation of reports, the development of web applications (i.e. interfaces to allow users to consume the model, APIs - Application Programming Interface) as well as the implementation of repeatable DM process, for instance, for AF detection. In some cases, it is necessary even to publish scientific articles in order to gain reliability and communicate the findings to a larger public.

2.5.2 Machine Learning

Machine Learning is a subset of Artificial Intelligence (AI), which refers to computer systems able to imitate somehow human behaviours. By means of complex algorithms, it is possible to train machines to process large data, learn patterns and deliver results. ML is based on heuristics, enabling scientists to build models according to empirical processes, encapsulating phenomena in mathematical expressions (HAN et al., 2006; YEUNG; HALL, 2007; CLARKE et al., 2009; GORUNESCU, 2011).

In order to learn patterns, ML algorithms adjust their own internal parameters during the learning process. Before that, pre-configured parameters, also known as "hyperparameters", should also be inputted to guide the learning process. Because

hyperparameters are not the same for every problem, and if not well set can cause a great impact on the ML model performance, there are different hyperparameter optimization strategies (ELGELDAWI et al., 2021). For instance, grid search is a simple application for an exhaustive search of hyperparameters adjustments in a defined subset. Nonetheless, its computing cost increases exponentially.

Three of the most popular hyperparameters optimization algorithms (LIASHCHYNSKYI; LIASHCHYNSKYI, 2019) are:

- Grid Search: makes an exhaustive search of hyperparameters adjustments in a defined subset of the hyperparameters space;
- Random Search: selects combinations of hyperparameters by means of random selection;
- Genetic Algorithm: simulating the process of natural selection, and using mechanisms that resemble biological evolution, this algorithm makes different small adjustments by sequentially selecting, combining, and varying hyperparameters.

The ML models can learn patterns through unsupervised or supervised learning. Whereas the unsupervised learning does not require labelled data, the supervised learning does. In this manner, once the labelled dataset is created, it is indicated to randomly split the data into:

- Training Set: data used to train the model. It should be as representative as possible and without bias. Otherwise, the error will be propagated;
- Validation Set: it is common to create multiple models by trying different algorithms and combinations of hyperparameters. In this manner, the Validation Set is used to evaluate the models' performance;
- Test Set: after selecting the model with the best combination of hyperparameters, it is performed a final evaluation of the model, but by means of an unbiased dataset, the test set.

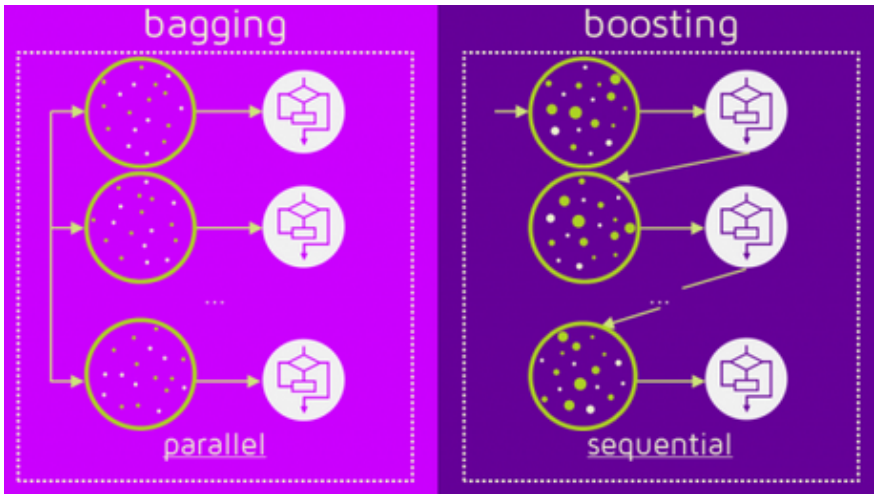
In this manner, the model learns repeatedly about the data behaviour and adjust itself for a defined purpose by means of the training dataset. Because of that, it is necessary an effort to obtain a large number of high-quality training samples (MAXWELL

et al., 2018). Once the model is built, the validation dataset is used to evaluate how well the model makes predictions, providing helpful information even to optimize the model's hyperparameters. If the model performance is not satisfactory, the model can be trained again. Finally, the test data is used to provide a final real-world check of an unseen dataset to confirm that the ML algorithm was trained effectively and can make accurate predictions. In other words, the test data is used to determine the model's true performance, which is essential before the model deployment.

There is a long list of ML algorithms, with different pros and cons. Because there is no theory beforehand about the algorithms' prediction performance in certain studies, it is indicated to experiment multiple methods to find the best one for the problem (MAXWELL et al., 2018). Some important supervised algorithms for fire management are: Random Forest, Gradient Boosting, Logistic Regression (JAIN et al., 2020; BOT; BORGES, 2022), and Deep Learning.

Random forest is an ensemble of decision tree ML algorithms, an extension of bootstrap aggregation (Bagging), which combines the predictions from many decision trees into a 'forest' (BREIMAN, 2001). That means, that RF combines the output of multiple decision trees to reach a single result. The combination of the tree predictors can be used to predict, classify or cluster events. Boosting, on the other hand, describes a strategy that uses a sequential additive model to combine a set of weak learners to make a strong learner, usually through decision threes (JAIN et al., 2020). In this manner, while the training process is parallel for Bagging, it is sequential for Boosting, as presented in Figure 2.14 (GARRIDO, 2016). Faster and with a higher performance when compared to the traditional Gradient Boosting, the Extreme Gradient Boosting (XGBoost) is a tree-based ensemble ML algorithm that continuously minimizes the bias error aiming to produce a new optimized model (FRIEDMAN, 2001; AMPOMAH et al., 2020).

Figure 2.14 - Training process for Bagging (parallel) and Boosting (sequential) approaches.



SOURCE: Garrido (2016).

Logistic Regression (LR) is an analysis method indicated for binary outcomes, such as the presence or absence of fires. LR is able to predicts a dependent data variable by analyzing the relationship between one or more existing independent variables. As such, LR allows the analysis of the explanatory power of the independent variables (i.e., temperature brightness) on the response variable (i.e., fires) through the analysis of the regression coefficients of those independent variables (LEGENDRE; LEGENDRE, 1998). Finally, Deep Learning (DL) refers to those algorithms with a brain-like logical structure. There are many DL models, such as the Convolutional Neural Networks (CNN) and the Recursive Neural Network (RNN). While CNN is commonly used for image processing, segmentation, and classification, RNN was designed to process and interpret temporal (or sequential) data.

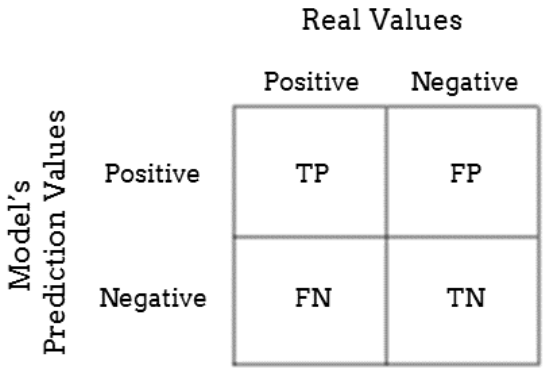
2.5.3 Model evaluation

Different statistical metrics can be applied to make a final model evaluation. Besides the traditional confusion matrix, when the main issue is related to a binary classification (i.e., fire / not fire), the indicated statistical metrics are: prevalence, accuracy rate, sensitivity, specificity, positive predictive value, and negative predictive value (detailed at Shreffler and Huecker (2022)).

The confusion matrix (Figure 2.15) is the comparison between the predicted value

and the real value, which is normally tabled as: True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN). Where the predicted value by the model is described as Positive and Negative, and the actual values in "real-world" as True and False.

Figure 2.15 - Example of a confusion matrix. Where: True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN).



SOURCE: Adapted from Kohavi and Provost (1998).

- TP: The model predicted positive and it's true in real-world;
- TN: The model predicted negative and it's true in real-world;
- FP (Type 1 Error): The model predicted positive and it's false in real-world;
- FN (Type 2 Error): The model predicted negative and it's false in real-world.

The prevalence is simply the proportion of the phenomenon's presence in the population. This metric is important, once it shows how common the phenomenon is. For instance, in a study about AF detection, when the dataset presents rare cases of AF in the region, the model performance can be affected by the data imbalance (HE; GARCIA, 2009; MAXWELL et al., 2018). It is stated by:

Prevalence = cases with AF / total population

The accuracy rate is related to the overall accuracy, and is expressed as:

Accuracy rate = number of cases correctly classified / total population

Sensitivity (or Recall) is the ability of the model to yield a positive result for a subject that presents AF. In other words, the ability to correctly classify is essential, and is calculated as:

Sensitivity = TP / (TP + FN)

Specificity is the ability of the model to yield a negative result for a subject that does not present AF. The equation for specificity is the following:

Specificity = TN / (TN + FP)

The positive predictive value (or Precision) determines: out of all the AF detections, how many are TP. While negative predictive value is: out of all the absence of AF, how many are TN. The equations are:

Positive predictive value = TP / (TP + FP)

Negative predictive value = TN / (TN + FN)

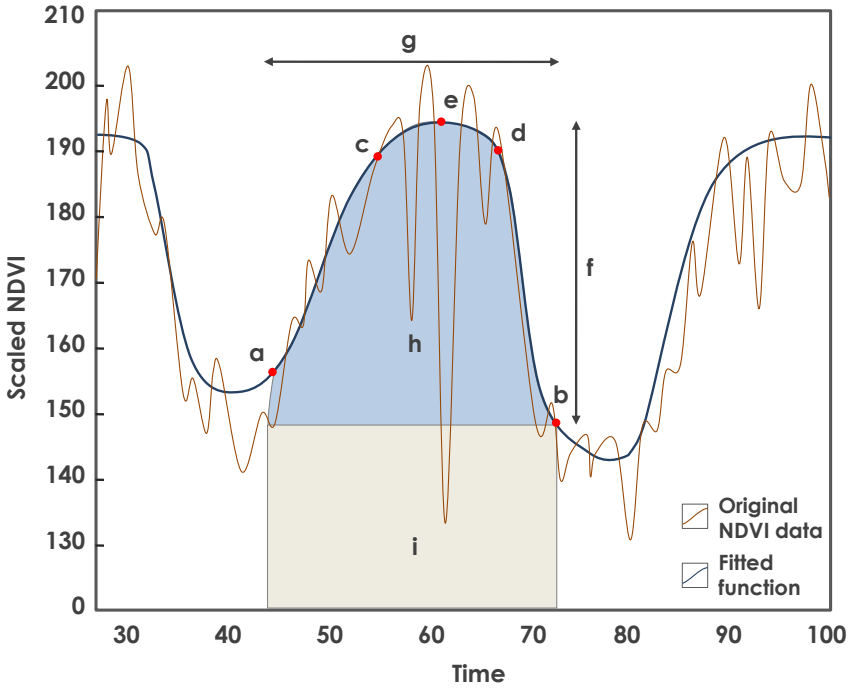
2.6 Traditional methods for pattern recognition in Remote Sensing Time Series

Due to the amount of available data, most of the ML approaches for TS were developed for Landsat series, with temporal resolution at the maximum of twice a month, and analysis along years. In such case, the long term (trend curve) components could be of interest to climate scientists, while seasonal studies could be related to agriculture growth analysis (KUENZER et al., 2015; BOVOLO et al., 2018).

A remarkable example of that is the TIMESAT system, developed by (JÖNSSON; EKLUNDH, 2004). It aims to smooth curves on dense TS in order to extract seasonal parameters (JONSSON; EKLUNDH, 2002). According to the authors, TIMESAT is based on three different least-squares methods, whose preliminary definition of seasonality could be uni or bi-modal combined with growing season approximate timings. Due to the high presence of noise, it is difficult to identify the number of annual seasons based on a single year's data. In this manner, the determination of the number of seasons is based on a model function, and the analysis of the surrounding years data to reduce mistakes.

Aiming to smooth the data and suppressing disturbances/noises, Savitzky-Golay filtering is used. The Savitzky-Golay filtering is based on different steps, and as a result it presents a smoothed curve. Nonetheless, according to the parameters used, the filtering may affect the detection of abrupt changes. Finally, seasonal data are extracted according to the season characteristics (Figure 2.16). The increased and decreased value are characterized as the beginning (Figure 2.16-a) and the end (Figure 2.16-b) of the season, respectively. The mid of the season (e) is estimated between 90% of the fitted function left (c) and right (d). The difference between the peak value and the average of right and left minimum values indicates the amplitude (f). Over the growing season, there are two integrals, which represent the active vegetation (h), and the total vegetation production (i). For evergreen areas, active vegetation may be really small.

Figure 2.16 - Examples of Timesat parameters. Where: a) beginning of the season; b) end of the season; c) left 90% level; d) right 90% level; e) peak; f) amplitude; g) length of season; h) integral over growing season giving area between fitted function and the average of left and right minimum values; i) integral over growing season giving area between fitted function and zero level.

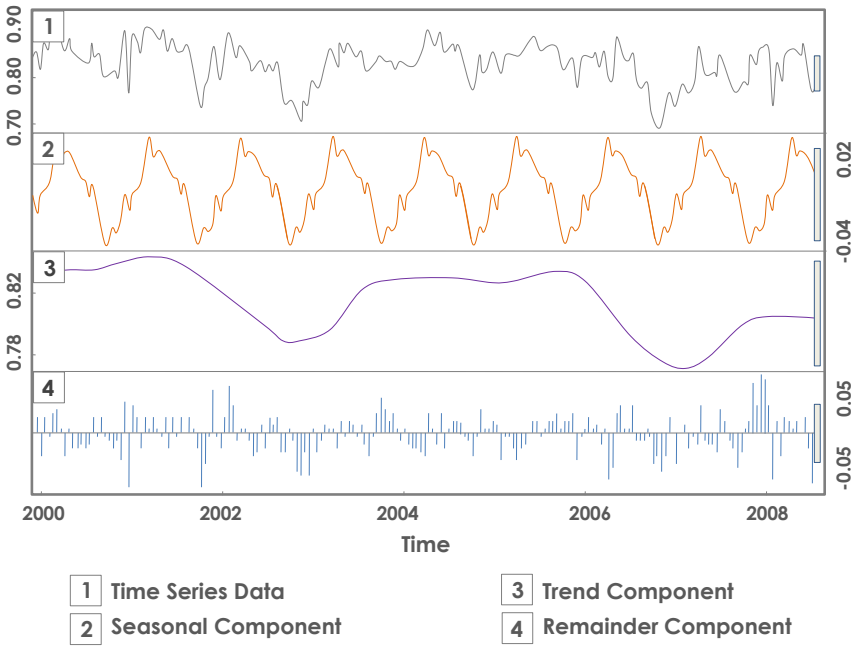


SOURCE: Adapted from Jönsson and Eklundh (2004).

The Breaks For Additive Seasonal and Trend (BFAST) (Figure 2.17) (VERBESSELT et al., 2010) is another example of an algorithm that decomposes TS. BFAST assumes for that a linear trend and harmonic season in the components model. In this manner, the ordinary least squares residual moving sum (OLS-MOSUM) statistical test is used to determine whether the breakpoints are occurring, the number, and its position in the TS.

BFAST is able to detect changes with NDVI magnitudes greater than > 0.1 in TS with different levels of noise and seasonal amplitudes. It was designed to be generic enough to comprehend a range of data types without the requirement to normalize the dataset, select a reference period, threshold, or change trajectory. If the ordinary least squares approach does not indicate significant change ($P < 0.05$), the breaking point is not identified. Besides, BFAST considers an additive model. In case of a multiplicative TS, BFAST can be used if it is converted to additive by taking a log of the TS (ABBES; FARAH, 2017).

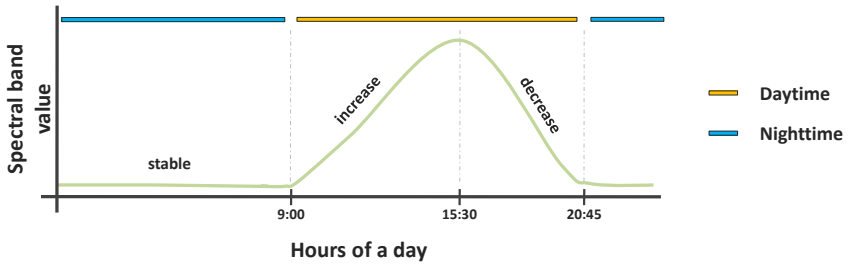
Figure 2.17 - Example of BFAST application on a 16-day NDVI TS of a pine plantation. In order to support comparisons, the bars on the right hand side of the plot show the same data range.



SOURCE: Adapted from Verbesselt et al. (2010).

In this perspective, the use of an ultra high temporal resolution as the ABI sensor, with a new image every 10 minutes, opens a completely new perspective and a paradigm shift is required. The general trend of ABI sensor TS could be the values gradually shifting along a month, due to the amount of available data. The seasonality is related to the seasonal differences and could also be considered along a year. The cyclical movements could be the days and nights variations. It is worth noting that such day and nighttime fluctuations would be cyclic in the cases of the ABI bands that depend on the reflected visible solar radiation, as the visible and near-infrared bands, presenting thus three main phases: stable, increase and decrease values (Figure 2.18). Unexpected variations refer to random movements, such as noises.

Figure 2.18 - Example of a three phase visible solar radiation of a daytime dependable ABI band.



SOURCE: Author's own elaboration.

2.7 Recent supporting studies for fire management

Several studies have been developed aiming to detect and monitor fires. An extensive research trends analysis of studies using RS to retrieve fire information is found in Chuvieco et al. (2020), Barmpoutis et al. (2020) and Jain et al. (2020). Rostami et al. (2022) used a deep Convolutional Neural Network (CNN) "MultiScale-Net" for AF detection in Landsat-8 datasets at the pixel level worldwide. Besides an innovative Active Fire Index (AFI) for AF detection through Landsat-8, a combination of bands SWIR2, SWIR1 and Blue, the study achieved a F1-score of more than 90%. Although the study presents such a high accuracy, it is important to highlight that Landsat-8 temporal resolution is 16 days, and in the best case scenario, 8 days. As already explained, for a suitable fire management, it is imperative the use of RS

dataset in NRT.

Aiming to detect early fires, a recent study used a deep learning framework (based on Gated Recurrent Units - GRU), GOES-16 and GOES-17 datasets in diferent LULC (ZHAO; BAN, 2022). The training dataset was based on US, in four large wildfires in California in 2020, and the model was tested in thirty eight different areas along US, Canada and Brazil. The bands used were Band 7 (sensitive to burning processes) and Band 14, due to its good contrast to the Band 7. After that, a normalized difference equation was applied. The band 15 was also used as a cloud and smoke mask. The results showed that the proposed method has a great ability to detect AF. It could detect AF earlier or similar to the VIIRS dataset. Finally, the authors also compared the deep learning approach with two classical ML algorithms: the Support Vector Machine (SVM) and the Random Forest (RF). Both ML algorithms were trained using the same training dataset and tested over three regions. In this comparison, SVM F1 score varied from 0.12 to 0.63, RF varied from 0.17 to 0.69, and the deep learning approach, from 0.18 to 0.71.

Aiming to characterize fire, precipitation and vegetation condition regimes and to establish spatial patterns in Cerrado from 2002 to 2015, Mataveli et al. (2018) used different RS data: i) TRMM monthly precipitation product (3B43); ii) MODIS fire products, AF, BA and FRP; iii) the global MODIS 1km Normalized Difference Vegetation Index (MOD13A3) to determine the Vegetation Condition Index (VCI), which is useful for evaluating fire incidence danger; iv) MODIS MCD12Q1 land use and land cover type product. After a boxplot analysis of the monthly total hotspots, monthly total BA, monthly average precipitation and monthly average VCI, the algorithm BFAST was applied to identify trends in the four TS. For the spatial statistical analysis, the Pearson's correlation coefficient (R) was used to show the linear relationship between the datasets. As result, the article presents the fire frequency, extension, seasonality, and variability along the biome. About 70% of the AF were identified along the physiognomy of savanna formations. Besides, possibly there is an inverse relationship between the incidence of fire and vegetation and precipitation.

Alvarado et al. (2017) analyzed the history of the fire incidence along 31 years (1984-2014) in the region of Serra do Cipó in Minas Gerais state. The dataset comprehends daily precipitation, burn scars based on Landsat images visually interpreted, and MODIS products (MCD45A1v5, MYD14, and MOD13Q1, which represent BA, EVI, and AF, respectively). In order to comprehend the extension, pattern, frequency, du-

ration, seasonality, and synergism of fire with the rainfall, the authors created some hypotheses about the relation between rainfall and BA. For that, a simple linear regression model was used to test the relationship between BA and mean annual rainfall. According to the authors, the strongest predictor of BA was drought during the ignition season, with an increasing fire occurrence during the driest periods. That means, the annual rainfall volume was weakly and negatively correlated with BA. Additionally, there is a lower fire frequency in higher density woody vegetation.

The estimates of AF, BA, and emissions are important variables to evaluating interannual fire variability. In this way, this set of variables was used by Chen et al. (2013) to analyze the fire incidence in the South America between 2001 and 2012. Using also fire persistence, deforestation and precipitation datasets, the study found out that there is a positive trend of fire activity at the leading edge of the deforestation frontier. Besides, the AF detections in evergreen forest and savanna biomes of South America covaried from year to year.

With a dataset with spatial resolution varying from 25 m to 9 km, and temporal resolution of hourly to days (ERA5-Land imagery, MODIS, FIRMS, etc.), Apostolakis et al. (2021) implemented a RF classifier to predict AF in Greece, and showed promising results, with precision varying from 77% to 94%, which could be used to support fire management in the region. Furthermore, among the used features, NDVI presented one of the most important data used.

2.8 First results of AF detection and monitoring along Cerrado by means of ABI bands

Since there was no complete model that describes near real-time AF in Cerrado, we investigated the potential of GOES-16 ABI TS through the known Fire Temperature RGB (FT-RGB) composition bands as a support for near real-time AF detection and characterization in Cerrado (PLETSCH et al., 2019). The premise used was that aiming to detect fires, GOES-16 has additional ABI bands in the near- and shortwave infrared (Table 2.8). Using the bands 7, 6, and 5 in a RGB composition (R7;G6;B5), hot spots and fires are highlighted in red, orange, yellow or white, as the fire gets hotter and the pixels become saturated, and according to the fire size. Green and blue shades can be related to ice and water clouds, respectively. Besides, the True Color RGB composition can be used to show the fire smokes (NOAA AND NASA, 2018).

Some limitations of the application are the presence of clouds, which can hide fire

signals, the possible false *red* fires in more dry regions (NOAA AND NASA, 2018), and the color variations along days, seasons, and localization (SEAMAN et al., 2017; SCHMIDT, 2019). Finally, due to the possibility of a more frequent data acquisition (LINDLEY et al., 2016), the incorporation of TS analysis in FT-RGB bands was a novel approach by itself in the work of Pletsch et al. (2019).

Table 2.8 - GOES-16 spectral bands used in a FT-RGB composition, R7;G6;B5.

Band	**Band central wavelength (μ m)**	**Contribution to a saturated pixel**	**Fire Temperature**
7	3.9	Hot land surface	Low
6	2.2	Small ice/ water particles	Medium
5	1.6	Water clouds	High

Source: Adapted from NOAA and NASA (2018).

Based on the concept of near (<3.0 km), and far (>6.0 km) from AF pixels, and using VIIRS and MODIS AF, two main questions guided the analysis: 1) *How does the GOES-16 ABI TS behave in near and far from AF environments?*; 2) *How does the GOES-16 FT-RGB compositions behave in near and far from AF environments?* To answer that, Pletsch et al. (2019) assessed the FT-RGB bands behaviour in a study area of about 3,000 km^2 located in Tocantins State during one day (October 24th, 2018) in the Cerrado biome. The answers (extracted from the original scientific article) are detailed below.

2.8.1 How does the GOES-16 TS behave in near and far from AF environments?

Bands 5 (B05) and 6 (B06) do not present information from 9:00 pm to 8:00 am due to the lack of daylight, while band 7 (B07) does. Even though, this period was not analyzed in B07, once there was no True Color composition to support the analysis (Figure 2.19).

Figure 2.19 - TS spectral behaviour of the average values of the GOES-16 FT-RGB bands near (<3.0 km) and far (>6.0 km) from AF pixels - October 24th, 2018. A highlight is available during the acquisition of AF data (2:00 pm - 5:00 pm). A. Band 5 (B05) along a whole day; B. B05 from 2:00 pm to 5:59 pm - interval with detected AF; C. Band 6 (B06) along a whole day; D. B06 from 2:00 pm to 5:59 pm - interval with detected AF; E. Band 7 (B07) along a whole day; F. B07 from 2:00 pm to 5:59 pm - interval with detected AF.

SOURCE: Pletsch et al. (2019).

The B05 TS curve presented a quite similar behaviour in near (NAF) and far from AF (FAF) spectral curves along the day, with peaks around 10:00 am and 4:00 pm (Figure 2.19-A). From 8:00 am to about 12:00 pm, the area presented dense clouds, which may have affected the band curve in this period. During the time interval between 2:00 pm and 4:00 pm, the NAF values were higher (Figure 2.19-B),

which may indicate that the presence of AF increases the values in this band.

TS of B06 data also presents an analogous pattern to the B05, with peaks in both curves, NAF and FAF, around 10:00 am and 4:00 pm (Figure 2.19-C). The first one can be also due to the presence of clouds. The second peak as well as B05, NAF values are superior. At 5:00 pm, as well as in B05, there is a decrease in the values, probably due to the sunset and the inherit sensor characteristics (Figure 2.19-D). Both bands can be physically related to aerosol particle size, and especially the B06 primary use was hot spot detection at the emission of temperatures greater than 600 K.

Differently from the aforementioned bands, the B07 values pursue another pattern along the time, and it was not possible to identify the influence of the clouds around 10:00 am. B07 TS presented just a soft peak curve from 9:00 am to 9:00 pm (Figure 2.19-E), which can be explained considering this band contains a daytime solar reflectance component. Along it, it seems that the curves NAF and FAF are overlapped. Nonetheless, a closer analysis between 2:00 pm and 5:59 pm shows that the curves are slightly different, and the NAF values are again higher (Figure 2.19-F).

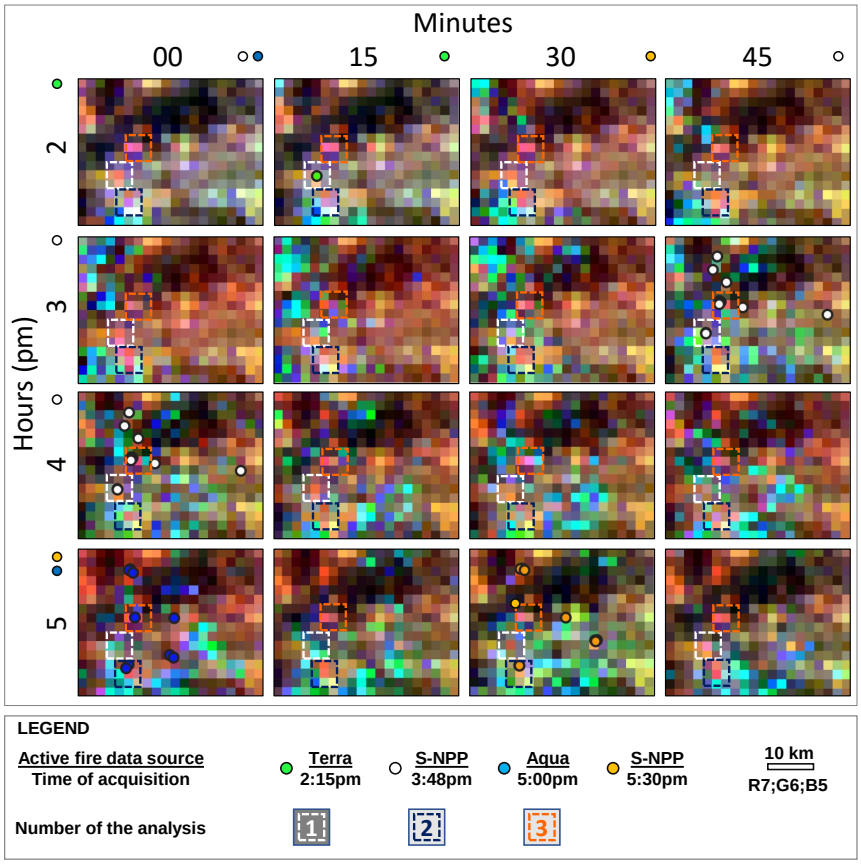
Such results could support the development of a reliable process able to classify GOES-16 based on the spectral-temporal characteristics. However, the use of nearer pixels (<3 km) may distinguish the results of NAF and FAF even better. Furthermore, the aforementioned bands main advantage is in RGB composition, but GOES-16 presents 16 different bands. As such, it would be suitable to develop an approach, as an index that can use the full potential of the integrated bands. For instance, the use of a middle infrared band (B07) may be even boosted when integrated with the thermal infrared band (Band 14 in the case of GOES-16). Moreover, the analysis of FAF temporal patterns could benefit the identification of anomalies in TS that can be related to the incidence of fires.

2.8.2 How do the GOES-16 FT-RGB compositions visually behave near and far from AF environments?

According to the FT-RGB quick guide (NOAA AND NASA, 2018), colors derived from the composition can be related to different targets. In theory, near black color is more related to water/snow/night, shades of blue to water clouds, green to ice clouds, purples/pinks to clear land, red to *warm* fire, and shades of brown to burn scars. However, for the same target the presented color may vary in space and time.

In this study, the True Color RGB composition enabled the identification of just three different targets in Cerrado: natural vegetation, cloud, and cloud shadow. In general, the presence of clouds was highly accurate with the pixels in shades of blue and green. The FT-RGB could identify a small water course in the north of the scene mainly from 2:00 pm to 2:45 pm (Figure 2.20). As the size of the water course is too small, possibly there are contributions from other factors for the amount of near black pixels.

Figure 2.20 - GOES-16 FT-RGB compositions near (<3.0 km) and far from (>6.0 km) AF environments along October 24th, 2018, from 2:00 pm to 5:45 pm.



SOURCE: Pletsch et al. (2019).

Due to the complexity of the TS, we are going to present in detail three different analyses. In the first one (Figure 2.20, white dotted square), it is possible to notice at 2:00 pm the presence of a red pixel in the center of the square, which is the

source of an AF detected by the satellites Terra at 2:15 pm, and S-NPP at 3:48 pm. Except for the presence of cloud or cloud shadow, this pixel remained with a hot color palette during the whole analyzed interval.

The second analysis (Figure 2.20, dark blue dotted square) also presented a pixel in the center with shades of red, orange, pink, and purple. It was identified AF in this pixel or in the adjacent pixel above at 5:00 pm and 5:30 pm, by the satellites Aqua and S-NPP, respectively. As it is located in the extreme southwest of the area, a region that presents clouds along the afternoon, the result may be influenced by this target. The last analysis (Figure 2.20, orange dotted square) is regarding two adjacent pixels in the center of the square, whose color along the period was also red, orange, pink, and purple. It was identified AF by the satellites S-NPP at 3:48 pm, Aqua at 5:00 pm, and in the adjacent pixel above by the S-NPP at 5:30 pm.

Due to the spatio-temporal variations of the colors in the FT-RGB composition, it would be necessary to assess the correlation between the targets and the possible correspondent band values in Cerrado. Furthermore, the presence of clouds could have influenced the analysis of the red pixels and the AF spots. Such a process would be crucial to identify which values of the RGB could endorse the presence of fire.

2.8.3 GOES-16 FT-RGB findings

In Pletsch et al. (2019), we analyzed the use of GOES-16 TS to characterize near real-time AF in Cerrado. As a result, they identified a certain pattern along the 24 hours TS from bands 05, 06, and 07. However, the TS curve analysis was influenced by the presence of clouds from 8:00 am to 12:00 pm. During the presence of AF (2:00 pm - 5:59 pm), the NAF values tend to be slightly higher than FAF, which may indicate that techniques of digital image processing could strength the differences and separate both groups to improve the AF detection.

Regarding the FT-RGB composition, they visually identified a certain relation of: i) shades of red with AF; ii) shades of blue and green with clouds; iii) the color black with the presence of water; iv) shades of brown with clear sky. For a deeper analysis of AF, the presence of water clouds was a hindrance to better relate the color with the target. Moreover, they identified as a limitation of the technique the minimum and maximum RGB values, which presents a spatio-temporal variation, and may influence the final color. Finally, the ABI Band 7 also showed promising results in order to detect hot land surfaces, and it was highly indicated for AF detection and monitoring.

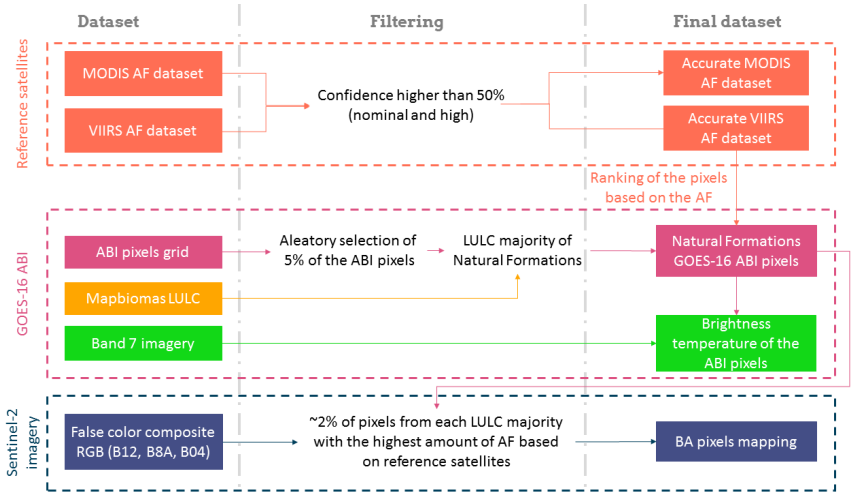
3 METHODOLOGY

The methodology was developed based on the CRISP-DM (Figure 2.13) framework presented in Section 2.5.1. Firstly, the used Data is described in Section 3.1. The Business Understanding step, which is related to understand the DM problem that needs to be solved, was already presented in Section 2, especially regarding Fires in Cerrado (Section 2.1.1). The Data Understanding process was performed by means of exploratory statistical and visual data analysis, and guided the steps required for the Data Preparation and Modeling (Section 3.2). The Evaluation step is presented in Sections 4 and 5, and the possible steps and studies prior to the Deployment are presented as Recommendations for Future Work in Section 6.2.

3.1 Data

The dataset used in this thesis is composed of AF products from the reference satellites (MODIS and VIIRS), GOES-16 ABI (Band 7), and Sentinel-2 (Bands 4, 8A, and 12) imagery, and Mapbiomas LULC mapping (Figure 3.1). Planet imagery (Bands 1, 2, and 3) was used only to support the visual pixel analysis (Section 4.5).

Figure 3.1 - Dataset used and the filtering process of the data.



AF: Active Fire; LULC: Land Used and Land Cover.

SOURCE: Author's own elaboration.

Among the products, MODIS and VIIRS are the most prominent and are consid-

ered references to identify AF (Figure 3.1, red dotted rectangle). The AF datasets, MODIS and VIIRS, were acquired from the Fire Information for Resource Management System (available at https://firms2.modaps.eosdis.nasa.gov/, accessed on 20^{th} November 2020) for August, 2019 and filtered for our study area and confidence level, higher than 50% for MODIS, and nominal and high confidence for VIIRS because false alarms are particularly undesirable. The data were used both to indicate the most representative burned ABI Band 7 pixels in order to train the FM model and to assess the FM performance.

Besides, we also used the Band 7 imagery ABI pixel grid raster and brightness temperature for the whole month of August, 2019, over MATOPIBA (≈240 GB) (Figure 3.1, pink dotted rectangle). Due to such big data, firstly, we randomly selected 5% of the ABI pixel grid to work with. Afterward, we used the 2019 Brazilian Annual LULC Mapping Project (MapBiomas), the 4^{th} collection (available at https://mapbiomas.org/, accessed on 10^{th} October 2020), to filter the random grid only for those with a majority of LULC natural formations once it presented the most expressive number of AF throughout August, 2019. In addition, the Natural Formation also composes the most representative LULC with 70% of MATOPIBA territory: 16% Natural Forest (NF), 43% Savanna Formation (SF), and 11% Grasslands (Gr). For the selected areas, we extracted brightness temperature from Band 7 imagery for the whole month of August, 2019.

Based on VIIRS and MODIS data, we selected 2% of the ABI filtered pixel grid with the highest AF recurrence for the manual BA mapping using Sentinel-2 imagery (Figure 3.1, blue dotted rectangle). In addition, the pixels were equally distributed based on the three LULC natural formations in order to identify areas more prone to fire and, therefore, better support our ML training and modeling.

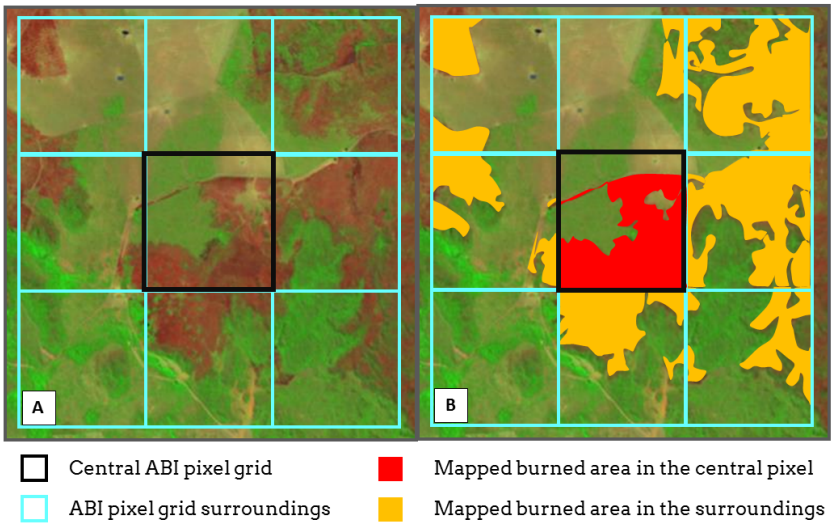
The BA mapping was performed considering both: (i) the inside of the central ABI pixel grid and (ii) the surroundings of the central pixel, as one of the assumptions of this study is that the presence of BA in the surrounding pixels may influence the brightness temperature of the ABI Band 7 central pixel. In this manner, for the BA mapping, we used Sentinel-2 imagery at Sentinel Hub viewer (available at: https://www.sentinel-hub.com/, accessed on 25^{th} January 2022). We also used the false color composite shortwave infrared (SWIR), RGB (B12, B8A, B04), as it enables fire damage mapping (Sentinel Hub, 2021). An example of such a process is available in Figure 3.2.

Even though Sentinel-2 data does not provide the exact time of a fire occurrence, it

is one of the most suitable satellites since it presents a 10–20 m spatial resolution and a five-day revisit time. Due to the satellite temporal resolution, the BA mapping was quantified by the ABI pixel grid and by date and only in cloud and cloud shadow free areas. The use of such data allowed us to design the FM in a way that it could detect AF, whose impact can be seen on Sentinel-2 imagery by means of burned areas.

Finally, for specific results analysis, when necessary and available, Planet imagery (Bands 1, 2, and 3) was used to visually support fire comprehension.

Figure 3.2 - Example of burned area mapping in the central ABI pixel grid and surroundings based on Sentinel-2 imagery. RGB (B12, B8A, B04).



(**A**) Remote sensing image without BA mapping. (**B**) Remote sensing image with BA mapping.

SOURCE: Author's own elaboration.

3.2 Methods: data preparation and modeling

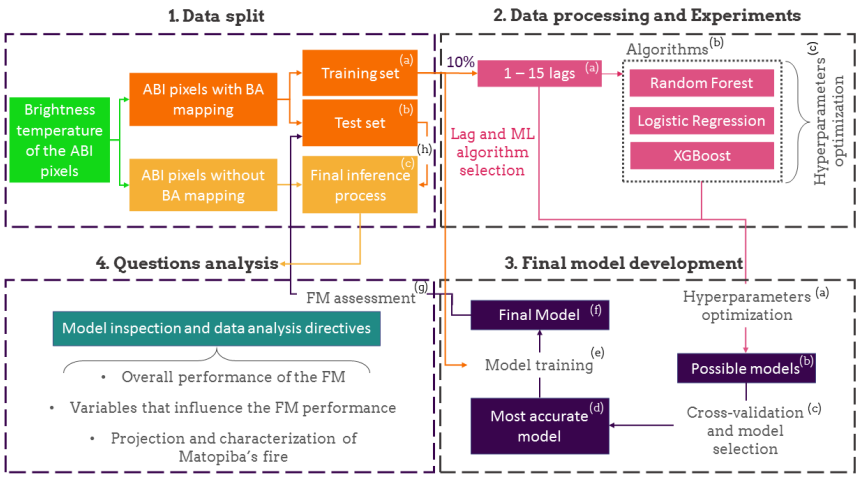
As indicated by KDD, after that data cleaning, integration and selection, we transformed the data to create a DC, where different algorithms were performed (available at: https://github.com/MikhaAJSP).

For the DC, firstly, the ABI dataset was pre-processed, including the transformation

from NetCDF to TIF dataset format, reprojection to WGS 84 (EPSG: 4326), and data cropping to the extension of the study area. Due to the technical improvements that ABI presents over the previous GOES imagers, including calibration and navigation (KALLURI et al., 2018), no further pre-processing was necessary. Furthermore, once the used dataset comprehends only 30 days, the TS was considered stationary, which means that there are no trend or seasonality influences. Hence, no TS decomposition was required. Finally, the remaining data was also integrated into a data frame, generating thus the DC.

Once the DC was prepared, we were able to perform the next steps of the methodology: data split, data processing and experiments, FM development, and FM assessment by means of important question analysis (Figure 3.3).

Figure 3.3 - Methods for Data Preparation, Modeling, and Evaluation divided into four main steps: Data Split, Data Processing and Experiments, FM Development, and Questions Analysis.



XGBoost: Extreme Gradient Boosting; FM: Final Model.

SOURCE: Author's own elaboration.

3.2.1 Data split

For the ML processes, we used the areas with BA mapping, equally distributed among the three Natural Formation LULC. For the training set, 94 of the brightness temperature pixels were selected, from which 25% of the observations were used

for the validation set (Figure 3.3(1a)). Besides, 40 pixels were used for the test set (Figure 3.3(1b)). It was divided once each pixel presented more than 4000 GOES-16 observations due to its ultrahigh temporal resolution. The remaining 2291 pixels were used for the final inference process (Figure 3.3(1c)).

3.2.2 Data processing and experiments

The data processing workflow used in this study starts with the identification of the amount of historical data (lag) the FM requires before a fire event to make accurate AF classifications (Figure 3.3(2a)). To do so, we applied the data normalization known as standard score (z-score), which requires not only the last brightness temperature but also a historical time series, from which the last brightness temperature can be compared and analyzed. In this manner, we would like to know *how distant the last value is from the historical time series average*. Based on an empirical analysis, the premise is that the average pixel value is the absence of fire, positive values are related to fires, and negative values, to the presence of clouds. In this thesis, the unit of the historical time series lag is a day, where 1 lag represents the last 144 ABI Band 7 observations, and 15 lags represent the last 2160 observations. It is important to highlight that the greater the amount of data, the more computational power is required.

Integrated with the lag analysis, we also conducted experiments with the three different ML algorithms, aiming to identify the most suitable combination for the FM development (Figure 3.3(2b)). The ML algorithms used were Random Forest (RF), Logistic Regression (LR), and Extreme Gradient Boosting (XGBoost). They were selected based on their performance and literature recurrence as techniques for fire management and decision-making processes (JAIN et al., 2020; BOT; BORGES, 2022).

Because GOES-16 ABI presents such an ultrahigh temporal resolution dataset, for the FM analysis, there were two main hindrances: the necessity to (i) create an approach able to compare datasets with different temporal resolutions (GOES-16 ABI, MODIS/VIIRS, and Sentinel-2); (ii) develop ways to support firefighters prioritization planning to maximize the efficiency of the response team. In this manner, for the FM performance assessment, we analyzed not only a single AF detection (naive) but also consecutive AF indications as well.

3.2.2.1 Algorithms and hyperparameters optimization

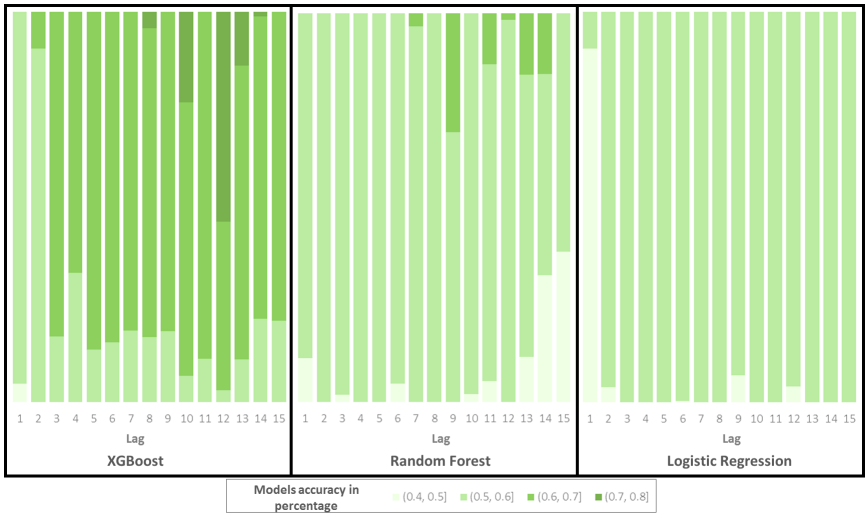
By using RF, XGBoost, and LR, we aimed to develop a model trained by the historical time series in order to correctly identify AF by means of a NRT dataset. For the experiments, we used 10% of the training set and created a z-score based on historical time series of 1 to 15 lags. For each lag, we applied each ML algorithm.

Aiming more fast and effective results, we developed our own hyperparameter optimization strategy. For that, we developed two base approaches: i) adjustment of all the hyperparameters; ii) adjustment of only one of the hyperparameters based on the best hyperparameters combinations until that moment. Both were used alternatively, whose prioritization was based on the incremental performance gain and processing time cost. Due to the number of ML algorithm hyperparameters, we also developed and applied an optimization step for each model and lag (days), where the automatic hyperparameter adjustments were run for at least three hours on the computer and with a minimum of 100 attempts per model (Figure 3.3(2c)). The possibilities of lag and algorithm combinations reached almost 10,000 models.

3.2.2.2 Lag and Machine Learning algorithm selection

According to Milanović et al. (2021), RF models were more efficient than LR for forest fire probability mapping, in our study, XGBoost presented an even higher performance. RF and LR hardly achieved an accuracy of 60–70%, whereas XGBoost achieved an accuracy of 70–80% (Figure 3.4).

Figure 3.4 - Lag comparison over 15 days and overall Machine Learning models performance.



SOURCE: Author's own elaboration.

In XGBoost lag accuracy, lags 12 and 13 had the best results. Although lag 12 presented the greatest number of accurate models, 70–80%, we selected lag 13 as it presented more accurate models than those of lag 12—around 80%. In other words, using lag 13 means that for every new piece of ABI data, the spectral distance from the mean of the last 13 days is analyzed in standard deviation units in order to confirm if it is above or below the local pattern, where positive z-score values are generally related to fires and negative z-score values to clouds.

Once the algorithm XGBoost and lag 13 were selected, the FM was developed and applied to the whole dataset in order to answer the proposed questions.

3.2.3 Final model development and assessment

For the FM, we aimed to optimize its hyperparameters until its saturation by means of 25% of the training data (Figure 3.3(3a)). This process generated about 2,150 models (Figure 3.3(3b)). Afterward, 75% of the remaining training data and the test set were used to assess the models in a process known as cross-validation (Figure 3.3(3c)).

The most accurate model was then selected (Figure 3.3(3d)) and trained based on

the training set (Figure 3.3(3e)), which resulted in the FM (Figure 3.3(3f)). We then assessed the FM with the test set (Figure 3.3(3g)) and applied it to the pixels without BA mapping for the inference process (Figure 3.3(3h)). Due to the temporal resolution difference of the GOES ABI dataset (10 min), reference satellites (12 h), and the BA mapping (5 days), we assessed the FM accuracy in the test set considering both: a single indication of AF and a certain sequence of AF. The consecutive AF indications represent a more persistent fire in TP cases. For fire management, it is important to comprehend the presence of fire throughout the territory (single detection). Nonetheless, it is also essential to understand how persistent a fire is in order to direct efforts to where it is most needed (sequence of consecutive AF detection). In addition, it is vital to highlight that as a consequence of the used dataset, only fires whose BA impact can be seen on Sentinel-2 false color composition can be detected by the FM.

For both single and sequential AF detections, the analysis was conducted from 14 to 31 August 2019. As lag 13 was selected, the first 13 days of August were only used to compose the historical data required for the z-score. The accuracy analysis for the single detection approach (named hereafter as 'naive') aimed to identify, based on the previous 13 days' z-score, if the brightness temperature found on the 14th day would present a fire or not. The same was performed for the consecutive AF detections; however, we considered a certain number of consecutive AF indications as a prediction. An attention was given to 15 and 125 consecutive AF detections.

Evaluating FM performance is a complex task because all of the data used have different temporal resolutions. In this manner, for the overall FM performance analysis by LULC and BA mapping, we only considered the dataset on the days with BA mapping. In addition, for the FM performance evaluation considering a consecutive sequence of AF indications, we analyzed the previous days of a BA mapping in order to identify the presence of AF indication from FM and the reference satellites.

4 RESULTS

4.1 Overall performance of the FM

The FM applied to the test set resulted in an accuracy rate of 78.9% (Table 4.1). In our analysis, the probability of an FM being right when it points out an AF detection is around 87%, and when it indicates a non-fire, around 70%.

In order to understand if LULC plays an important role in FM performance, we also analyzed the FM for each LULC natural formation: Natural Forest (NF), Savanna Formation (SF), and Grassland (Gr) (Table 4.1). We observed that the percentage of fire prevalence among the classes varies up to 50%. Nonetheless, the overall accuracy rate ranged between 70% and $\approx$ 90%, which indicates that the FM performance is versatile among the three natural formations.

Furthermore, to explore the weaknesses and strengths in specific situations, it is important to take into account the rate of FM accuracy in detecting AF when there is fire (sensitivity) and not detecting AF when the fire is absent (specificity). In this manner, it is possible to have greater clarity on the FM classifications' interpretability and the required further improvements in the FM.

The FM performance in the different LULC (Table 4.1) can be grouped into two: (i) Natural Forest (NF), with high sensitivity (more than 90%) and low specificity (58.9%); and (ii) Savanna Formation (SF) and Grasslands (Gr), with low sensitivity (58% and 54%, respectively) and high specificity (72% and 77%, respectively). In NF, the FM has a tendency to identify practically all the AF activities; however, it can indicate more AF than there really are, whereas in SF and Gr, we have the opposite process. Even though the FM presents such specificities, its performance in any LULC presents an overall accuracy rate higher than 70%, which means that the FM classifications will be right in at least 70% of the cases.

Compared with NF and Gr, SF presents higher false positives ($\approx$11%). Such result could be related to (i) the high heterogeneity of the physiognomies presented in this natural formation that embraces areas with defined tree and shrub-herbaceous stratum, and (ii) the absence of BA on the Sentinel-2 imagery due to the fast grassland vegetation recovery. Finally, Gr presented the highest true negatives and the lowest false positive among the natural formations, probably due to the predominance of herbaceous-shrub species.

Table 4.1 - Overall and Land Use and Land Cover (LULC) final model prediction assessment metrics. NF: Natural Forest; SF: Savanna Formation; Gr: Grassland; BA: Burned Area. True positives, false positives, false negatives and true negatives. Results in absolute numbers and percentage.

Metrics	Overall FM Assessment	FM Assessment by LULC		
		NF	SF	Gr
True positives (real: fire, predicted: fire)	6607 (40.60%)	3906 (82.09%)	1691 (24.98%)	1010 (21.28%)
False negatives (real: fire, predicted: non-fire)	2468 (15.17%)	419 (08.81%)	1190 (17.58%)	859 (18.10%)
False positives (real: non-fire, predicted: fire)	971 (05.96%)	178 (03.74%)	763 (11.27%)	30 (00.63%)
True negatives (real: non-fire, predicted: non-fire)	6228 (38.27%)	255 (05.36%)	3125 (46.17%)	2848 (60.00%)
Fire prevalence on test data	55.8%	90.9%	42.6%	39.4%
Accuracy rate	78.9%	87.5%	71.1%	81.3%
Sensitivity	72.8%	90.3%	58.7%	54.0%
Specificity	86.5%	58.9%	80.4%	99.0%
Positive Predictive Value	87.2%	95.6%	68.9%	97.1%
Negative Predictive Value	71.6%	37.8%	72.4%	76.8%

4.2 FM performance regarding burned areas mapping

The size of the BA does not influence the FM accuracy (Table 4.2a). Actually, smaller BAs (0.01–0.1 km^2) presented about 10% higher true positives than those of bigger BAs (>1.0 km^2), which can be explained by the following reasons: (i) most of the BA data are smaller than 1.0 km^2, and as a consequence, there is a great number of representative samples of smaller BA proportions within the pixels to train the models; (ii) fires of different proportions may present different z-score patterns, and due to the great number of small BA samples (<1.0 km^2), the FM is probably more focused on this dimension of fire. In other words, the FM can be more accurate at predicting AF at the beginning of the fire phenomenon; (iii) larger BAs (>1.0 km^2) can be generated by means of fire of small proportions burning for a longer period of time, which would not necessarily sensitize the FM.

Table 4.2b shows that when there is a larger BA mapping in the surroundings (>1.0 km^2), the FM result presents a higher false negative in the central pixel. In addition, BA mapping in the surroundings larger than 0.1 km^2 already negatively affects the true positive values in the central pixel.

Table 4.2 - Final model prediction accuracy considering burned areas. F: Fire; NF: Non Fire.

a)

		FM accuracy according to BA Mapping in the central pixel (km^2)							
		0 - 0.01		0.01 - 0.1		0.1 - 1.0		>1.0	
	Classification	F	NF	F	NF	F	NF	F	NF
BA Mapping	F	0.00%	0.00%	77.10%	22.90%	71.20%	28.80%	67.80%	32.20%
	NF	13.50%	86.50%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

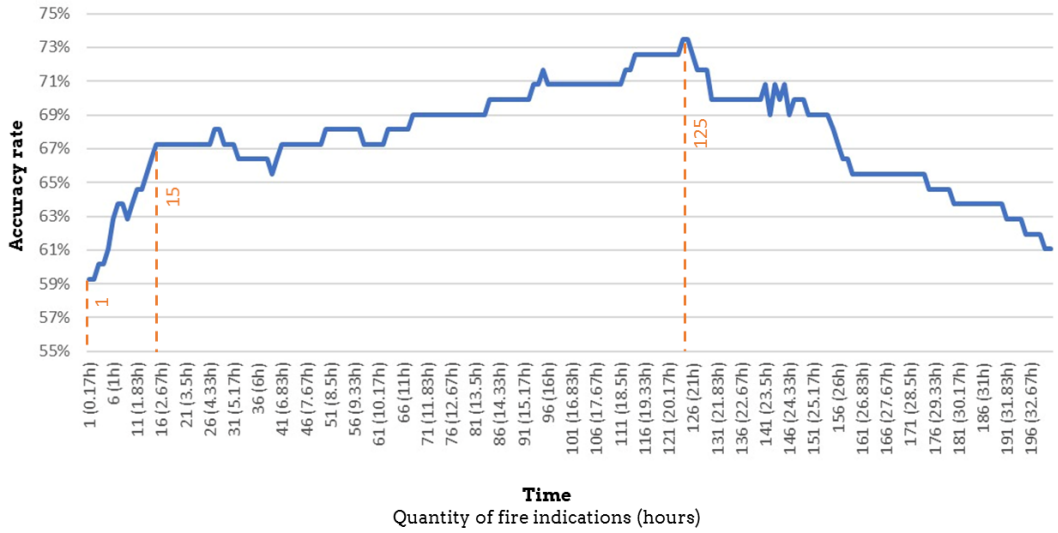
b)

		FM accuracy according to BA Mapping in the surroundings (km^2)							
		0 - 0.01		0.01 - 0.1		0.1 - 1.0		>1.0	
	Classification	F	NF	F	NF	F	NF	F	NF
BA Mapping	F	0.00%	4.00%	82.00%	10.00%	31.00%	13.00%	47.00%	28.00%
	NF	12.00%	84.00%	0.00%	8.00%	9.00%	47.00%	4.00%	21.00%

4.3 What is the FM potential when considering a consecutive sequence of positive predictions?

Aiming to explore ABI temporal resolution of 10 min, we analyzed the FM accuracy considering a single detection (naive) and a sequence of 15 and 125 consecutive AF detections (Figure 4.1). While the accuracy rate of the naive approach is 56.6%, 15 consecutive AF detections (after 2.5 h) is 67.3%, and 125 (after $\approx$ 20 h) achieves an accuracy peak of 73.4%. From then on, the increase in consecutive AF detections does not improve fire detection performance. Such a fact can also be associated with the low number of samples with more than 20 h of consecutive fire indications.

Figure 4.1 - FM consecutive AF prediction assessment metrics.



SOURCE: Author's own elaboration.

Comparatively, the results show that with the increase in consecutive AF detections, there are slightly lower true positives, around 6%, but significantly higher true negatives, more than 20% (Table 4.3). Moreover, it also presents lower false positive cases, from ≈39% of the naive approach to ≈16% of the 125 consecutive AF. Roughly, until the 125 consecutive AF detections, more detections result in better overall accuracy metrics, but the time required to identify more fire indications can be decisive for firefighters in real life. Furthermore, fires with a shorter lifetime are more likely to be unseen when a longer consecutive AF detection approach is considered.

The accuracy rate of the reference satellites is almost 71% and roughly half of the fires are correctly detected. In addition, reference satellites (MODIS and VIIRS) rarely commit false positives, less than 3%, yet their true positives are lower than those of the FM. In comparison with the reference satellites, in the three presented approaches, FM has a higher sensitivity and a lower specificity. In addition, the 125 consecutive AF detection has a higher accuracy rate. However, it is noteworthy that due to our methodology, the reference satellites have an advantage, as the data were already filtered, and low confidence detections were removed.

According to the BA manual mapping, almost 50% of the 125 consecutive AF detections from the FM and the reference satellites are correct and in agreement, and almost 5% are incorrectly classified by both (Figure 4.4). The main difference is regarding the errors. While the 125 consecutive AF detections approach sees more fires than there are (false positives is almost 16%), the reference satellites are more restrictive and point to fewer fires than there really are (false negative is 23%), probably due to the filtering process of high confidence AF from the reference satellites. In this context, the FM can be considered an important improvement over the reference satellites, not because it is more accurate, but because it presents a high agreement with traditional methods, not to mention its ultrahigh temporal resolution of 10 min, which could be integrated with the already consolidated reference data in order to provide NRT fire detection in the MATOPIBA region.

Table 4.3 - Reference satellites (MODIS and VIIRS) and Final Model (FM) prediction assessment metrics by 1 (naive), 15, and 125 consecutive Active Fire (AF) detections. Results in absolute numbers and percentages.

Metrics	Reference Satellites	Consecutive AF Detection		
		Naive	15	125
True positives (real: fire, predicted: fire)	32 (28.32%)	58 (51.33%)	56 (49.56%)	51 (45.13%)
False negatives (real: fire, predicted: non-fire)	30 (26.55%)	4 (3.54%)	6 (5.31%)	11 (9.73%)
False positives (real: non-fire, predicted: fire)	3 (2.65%)	45 (39.82%)	31 (27.43%)	19 (16.82%)
True negatives (real: non-fire, predicted: non-fire)	48 (42.48%)	6 (5.31%)	20 (17.70%)	32 (28.32%)
Fire prevalence on test data	55.76%	55.76%	55.76%	55.76%
Accuracy rate	70.80%	56.64%	67.26%	73.45%
Sensitivity	51.61%	93.55%	90.32%	82.26%
Specificity	94.12%	11.76%	39.22%	62.75%
Positive Predictive Value	91.43%	56.31%	64.37%	72.86%
Negative Predictive Value	61.54%	60.00%	76.92%	74.42%

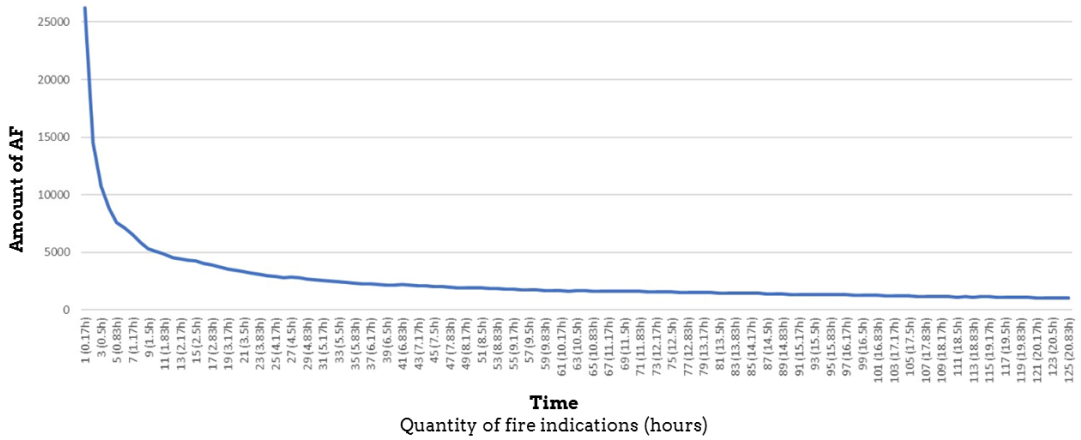
Table 4.4 - Agreement between the Final Model (FM) and the reference satellites (MODIS and VIIRS). True Positive (TP); True Negative (TN); False Positive (FP); False Negatives (FN).

		Reference satellites			
		TP	**FN**	**FP**	**TN**
125 consecutive AF detections	**TP**	22.10%	23.00%	0.00%	0.00%
	FN	6.20%	3.50%	0.00%	0.00%
	FP	0.00%	0.00%	0.90%	15.90%
	TN	0.00%	0.00%	1.90%	26.50%

4.4 Fire reality in the remaining data over MATOPIBA

We applied the FM consecutive AF detections to the remaining 5% of the dataset in MATOPIBA territory throughout August, 2019. Thus, the fire prediction varied according to the selected FM approach. If we considered the naive approach, we would have more than 26,000 AF. For the 5 consecutive AF detections, it would represent more than 4200, and for the 125 consecutive AF, 1042 detections (Figure 4.2).

Figure 4.2 - Number of active fires detected according to the number of consecutive AF indicated by the FM.



SOURCE: Author's own elaboration.

For the same area and time interval, according to the confusion matrix from the

reference satellites, the total number of AF would reach 1209. Comparing the fire prediction based on the reference satellites and the 125 consecutive AF detections, the difference between the results was only 167 AF in more than 2400 of the analyzed pixels. Although fire characteristics (quantity, velocity and persistence) can be a hindrance when comparing AF of different sources and temporal resolutions, both aforementioned approaches presented similar results.

4.5 Detailed pixel and fire behaviour analysis

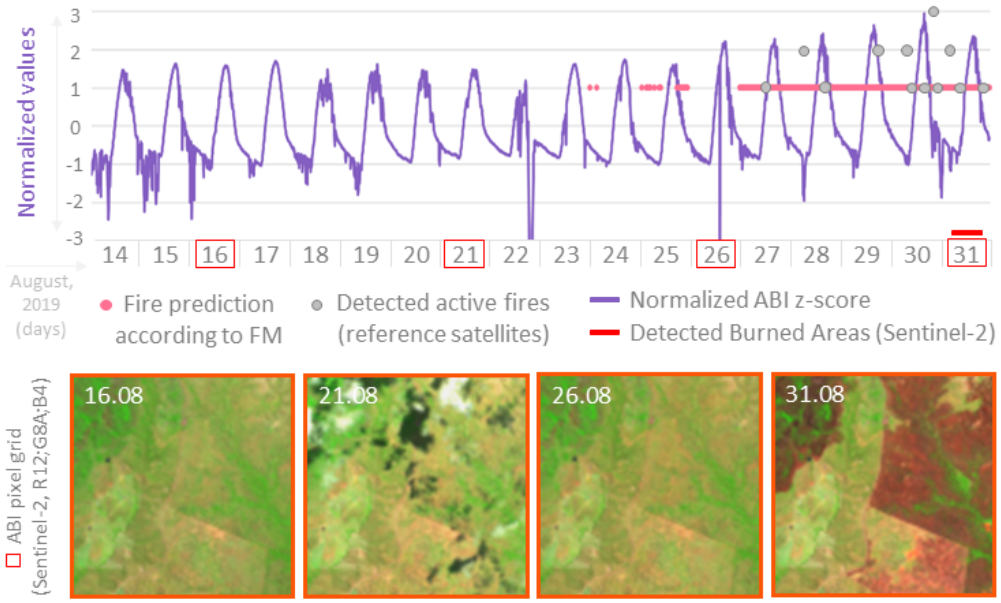
To support the results comprehension, we performed a detailed pixel and fire analysis, presented in Table A.1. The whole list of detailed pixel analysis is presented in Appendix. When there was more than one fire incidence in the same pixel along August, we divided the pixel analysis into sub-analysis (A, B and so on).

Table 4.5 - Detailed pixel analysis. FM: Final Modell RefSat: Reference Satellites. NA: Not Applicable. d: day.

ID analysis	Fire firstly detected by	Fire duration (days)	Fire extent (km^2)	Expansion (km^2 d^{-1})	Days of burning (August)	Figure
01	FM	4	2	0.5	26-31	4.3
02-A	RefSat	5	0.52	0.10	14-18	4.4
02-B	FM	1	0.12	0.12	22	4.4
02-C	RefSat	1	0.52	0.52	25	4.4
03	FM	2	0.08	0.04	30-31	4.5
04	FM	4	0.28	0.07	26-29	4.6
05	RefSat	2	2.6	1.3	15-16	4.7
06	FM	4	3.3	0.83	24-27	4.8
07	RefSat	2	2.8	1.4	20-22	4.9
08	RefSat	4	3.2	0.8	19-22	4.10

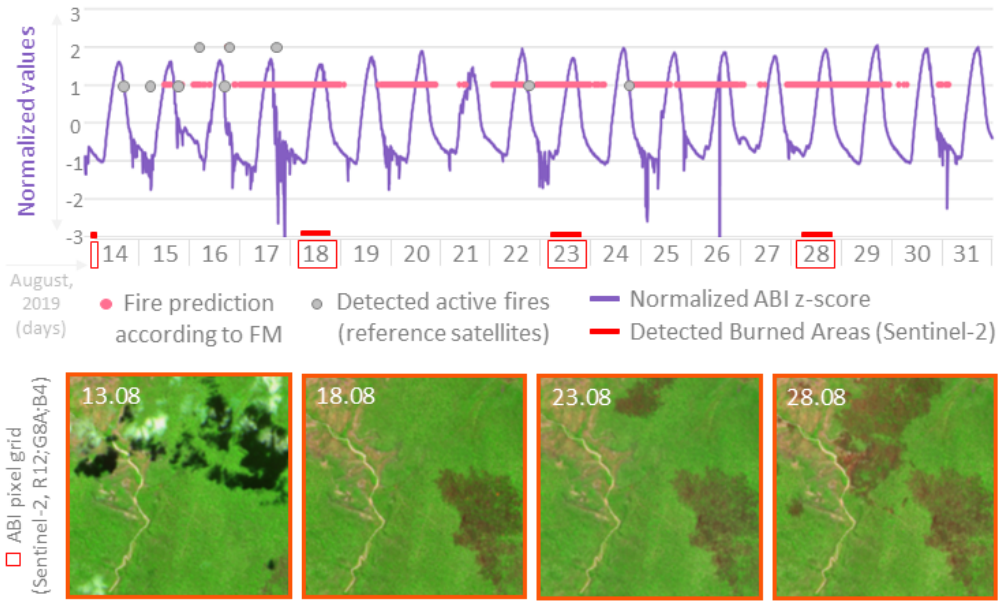
In the ID analysis 01 (Figure 4.3), the fire behaviour achieved an expansion of about 0.5 km^2 d^{-1} with a fire extent of about 2 km^2. Presenting a very high agreement between FM and the Reference Satellites, FM could identify fire firstly. Whereas, in the ID analysis 02 (Figure 4.4), two of the three fire events were firstly detected by the Reference Satellites (Table A.1, ID analysis 02-A and 02-C). One important issue is related to the continuous FM fire indication after a more intense fire, which could indicate that FM can be sensitive not only to the presence of AF but also to a more severe anomaly such as BA.

Figure 4.3 - Detailed analysis ID 01.



SOURCE: Author's own elaboration.

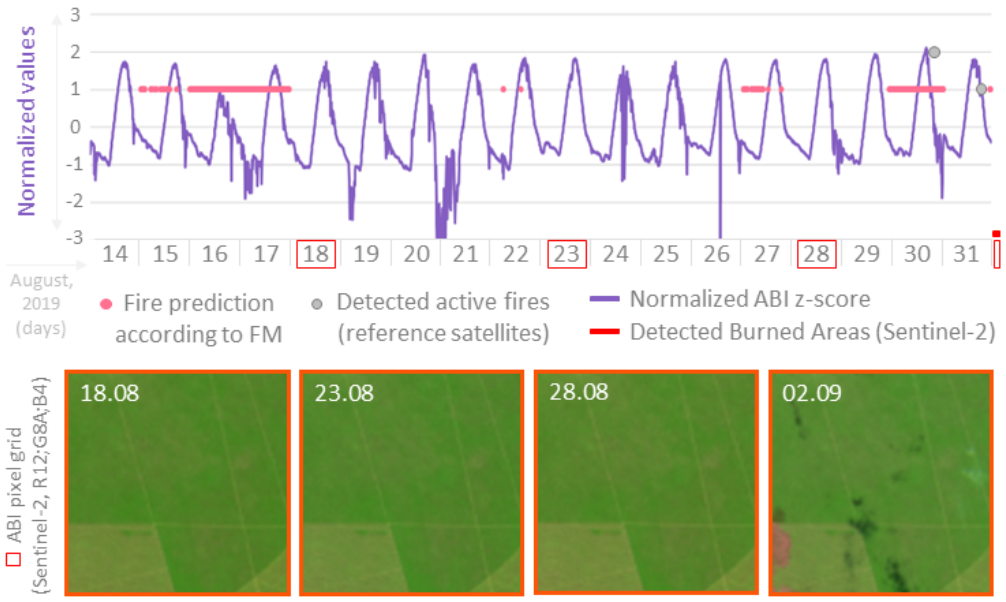
Figure 4.4 - Detailed analysis ID 02.



SOURCE: Author's own elaboration.

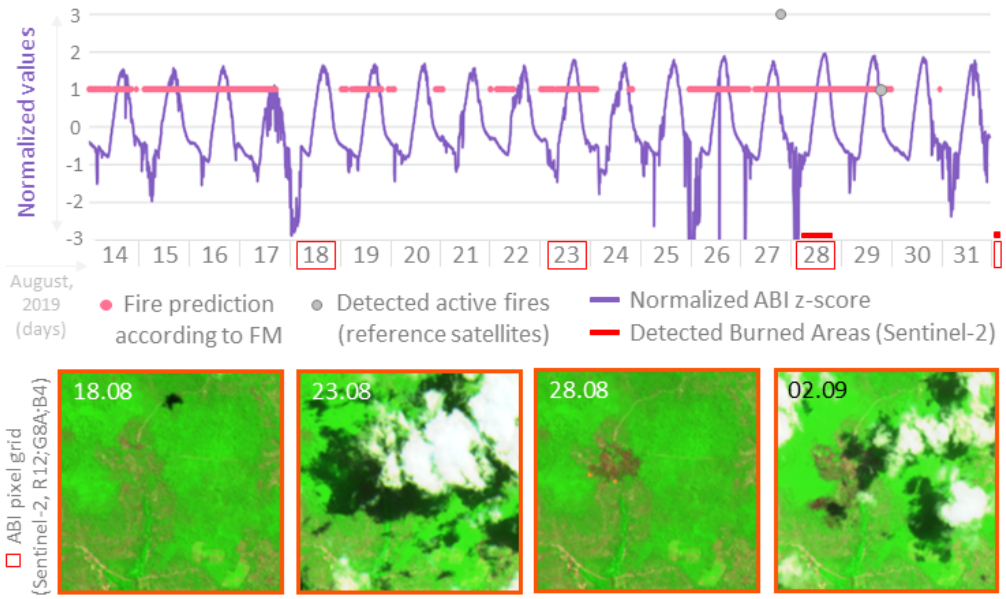
FM and the Reference Satellites also presented agreement at the end of August in the analysis 03 (Figure 4.5), even though the fire extent was only 0.08 km^2 and expansion of 0.04 km^2 d^{-1}. However, FM presented false positive values on 16^{th} and 17^{th}. The same confusion is possible to notice in the analysis 04 (Figure 4.6). It is also possible to notice the gap of information access due to the presence of clouds on 23^{th}. On 28^{th}, because we noticed the AF on the ground, we considered that the fire started between 25^{th} and 26^{th} with duration until 29^{th}, according to FM and the Reference Satellites.

Figure 4.5 - Detailed analysis ID 03.



SOURCE: Author's own elaboration.

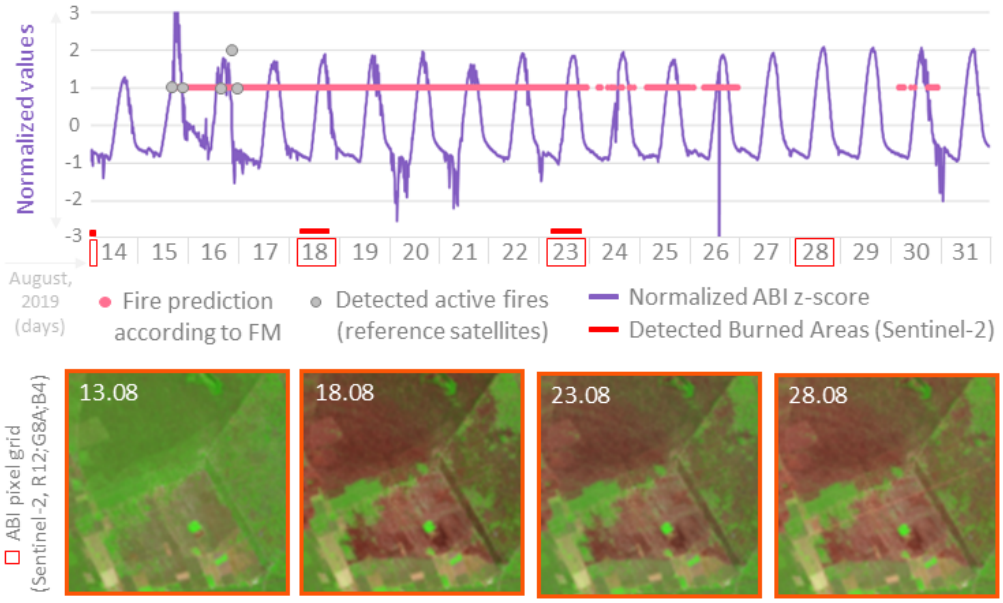
Figure 4.6 - Detailed analysis ID 04.



SOURCE: Author's own elaboration.

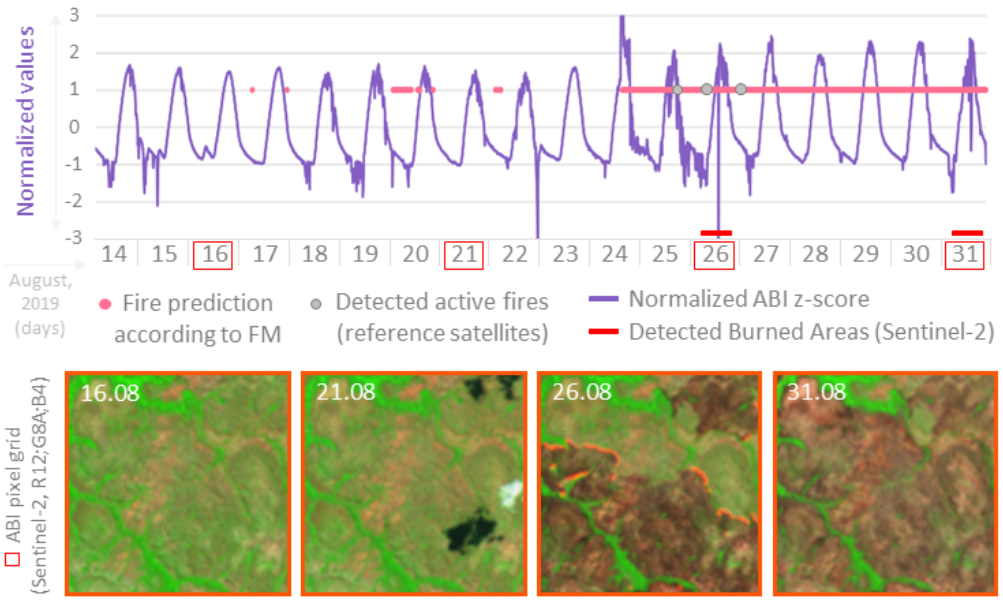
In the ID analysis 05 (Figure 4.7), 06 (Figure 4.8), 07 (Figure 4.9) and 08 (Figure 4.10), we notice an agreement between the FM and the Reference Satellites detections, but as well as in the analysis 02, there is continuous FM fire indication after the intense fire. Supported by Planet imagery, we could identify that the fire duration in the analysis 06 was of 4 days and not 8 as indicated by FM. Looking with more detail for the analysis 07, we also found out that the fire in the area was lasts only for 2 days, while FM detected AF for at least 10 days. In the analysis 08, the fire duration was about 4 days, but the FM indicated AF for around 9 days.

Figure 4.7 - Detailed analysis ID 05.



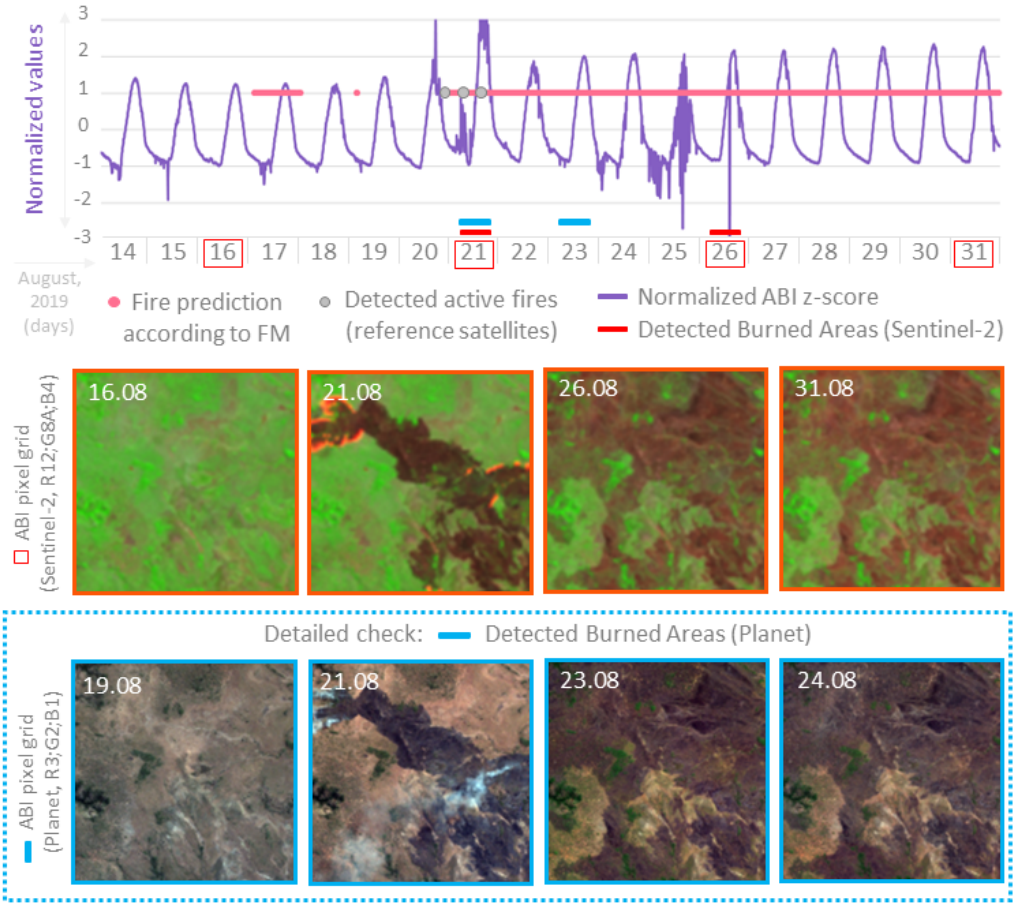
SOURCE: Author's own elaboration.

Figure 4.8 - Detailed analysis ID 06.



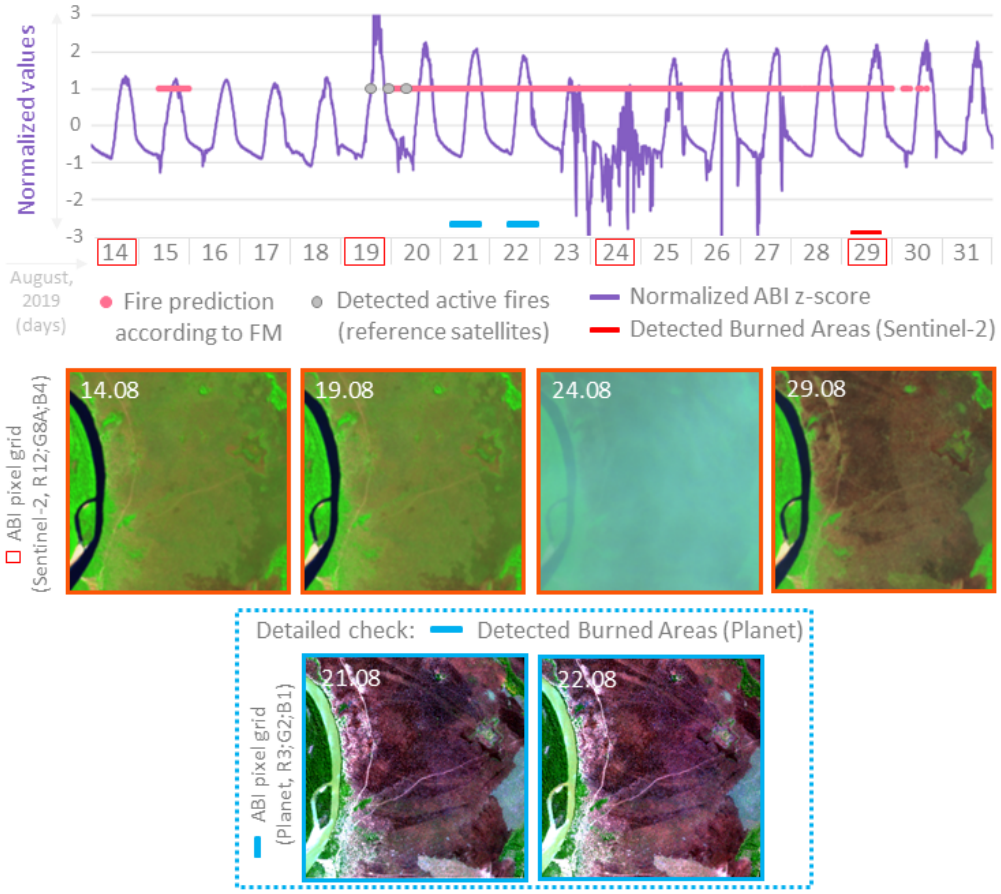
SOURCE: Author's own elaboration.

Figure 4.9 - Detailed analysis ID 07.



SOURCE: Author's own elaboration.

Figure 4.10 - Detailed analysis ID 08.



SOURCE: Author's own elaboration.

5 DISCUSSION

As presented in Table 4.1, FM proved to be versatile among the three analyzed natural formations, NF, SF, and Gr, but with specificities in each one of them. Although the overall accuracy rate of FM is high, the FM single detection accuracy can be negatively influenced by BA greater than 1 km^2 in the central pixel and in its surroundings. Consequently, the greatest potential of this approach is when the fire is in its initial phase. In addition, considering that the fire intensity needs to be high enough to sensitize the ABI sensor, and the human-induced fires are frequent and intense (PIVELLO et al., 2021), most of the FMs potential is also regarding this kind of fire.

In order to visually compare FM and the Reference Satellites in different scenarios, we performed a detailed pixel analysis (Section 4.5). It was possible to notice a certain agreement between FM and the references satellites AF detections, where both were suitable for AF detection at the same time that they can complement each other. In such occasions, either the FM or the Reference Satellites would firstly detect AF or even detect it when the other would not. FM presents generally two sensitive points: i) occasional presence of FM false positives; and ii) when the surface presents an AF followed by a severe BA, FM continues to detect AF, probably due to the Machine Learning training process used. In this manner, FM indicates not only AF but also burning anomalies on the ground. Finally, among the analyzed pixels, the fire duration ranged from 1 to 5 days, burning up to 3 km^2 and maximum expansion of 1.3 (km^2 d^{-1}).

Because fire is a dynamic phenomenon, NRT datasets are the most recommended for AF studies. However, ultrahigh temporal resolution data such as GOES-16 ABI is thus far poorly explored in this field. Higa et al. (2022) proposed an approach based on object detection methods to map AF in the Brazilian Pantanal biome. For that, the authors used deep learning (a subset of ML-based on neural networks) and CBERS 4A (China Brazil Earth Resources Satellite) imagery. After extensive experiments and the generation of 150 models, the study achieved a high precision, more than 80%. Nonetheless, CBERS 4A presents a spatial resolution of 55m and a five-day revisit time. A similar approach was also developed by (PEREIRA et al., 2021), where deep learning techniques were used for active fire detection. The final model achieved a precision of more than 87%; however, the authors used the Landsat-8 imagery, with a revisit time of 16 days. Considering that different sensors present idiosyncrasies, we can also notice an opportunity to harmonize multi-sensors for AF

studies.

Understanding fire behavior is essential for fire management. As resources are scarce, and proper allocation of firefighters is essential for firefighting success, a more refined AF detection and monitoring are imperative. Because of that, we developed the FM consecutive AF. It presented a higher accuracy, reaching its peak after around 20 h (125 consecutive AF). Even though there is an important trade-off between the consecutive AF and time, the 125 consecutive AF presented a number of true positives almost as accurate as those of the reference satellites. Therefore, incorporating the approach of consecutive AF detection, for instance, at INPE's Fire Monitoring Program, is vital for proper firefighting and management. Such an approach could provide not only a single AF detection but also a better comprehension of the NRT fire behaviours. Moreover, different AF sources could also be integrated to boost the confidence of one another.

In this study, the FM was developed based on the XGBoost ML model and also considering the z-score of the last 13 days. However, once it is implemented to support fire management, retraining the FM is recommended throughout the year due to the seasonal variability. Furthermore, since a pixel is the smallest unit of analysis, the FM could also be trained for other biomes. As such, not only MATOPIBA but also other areas could benefit from the FM predictions. Further studies are needed to improve the FM in order to reduce weaknesses, such as the false positives found in SF. Finally, the FM could also be improved by using a fire prediction confidence rate instead of the binary prediction (true or false). Additionally, the integration of the FM with other models, for instance, the integration of FM results with the fuel load dynamics and fire spread probability (OLIVEIRA et al., 2021), could also better guide firefighters in allocating resources efficiently where they are most needed.

The combination of factors such as the removal of natural vegetation and the inadequate soil management by means of recurrent human-induced fires has already been proven to contribute to soil degradation in MATOPIBA (VIEIRA et al., 2021). Because of such impacts, different laws and initiatives have been developed in order to protect natural tropical biomes. In Brazil, the "Zero Fire" policy aimed to ban fires. However, the advance of science in the 1970s resulted in changes in fire management discussions, and from the 2000s on, it became more evident that such policy was inefficient in protecting fire-dependent biomes, including Cerrado (PIVELLO et al., 2021). Hence, other more updated strategies have been created, such as the Law for Protection of Native Vegetation (Law 12.651/2012) and the Brazilian Integrated

Fire Management Policy Bill (PL 11.276), which include in their regulation the use of fire for ecological purposes.

Although some initiatives already exist, there is a lack of studies in areas with high fire frequency in Cerrado, as already indicated by Arruda et al. (2018), and little can be achieved without deeper knowledge about fire behavior in the region. Consequently, the NRT dataset and ML approaches, such as the FM, are crucial in supporting fire management.

MATOPIBA may lose approximately 120,000 km^2 of natural formations to anthropogenic uses before 2050 (VIEIRA et al., 2021). That being the case, fire management and new agricultural practices in MATOPIBA are fundamental not only to preserving local biodiversity but also to guaranteeing food security and avoiding its associated impacts on the national economy. However, contrary to what is needed, environmental management and research have suffered budget cuts by the Brazilian government in recent years (SCHMIDT; ELOY, 2020).

To avoid further impacts, it is of utmost importance to have financial support for infrastructure as well as human resources for environmental monitoring and research development, where techniques with cutting-edge technology, such as the FM, can be developed and applied for better national fire management. While budgets are scarce, the integration among share- and stakeholders is inefficient, and public policies remain only on paper, and the fire phenomenon persists as an open issue in Brazil.

6 FINAL CONSIDERATIONS

6.1 Conclusions

Active fire mapping based on geostationary satellites is already proven to be an important source of information for wildfire management and surveillance initiaties (WICKRAMASINGHE et al., 2018). In this thesis, we developed the Final Model, the first Machine Learning algorithm able to detect AF in near real-time in the MATOPIBA region. In addition, Final Model can also be considered a major improvement over the reference satellites for a couple of reasons: the Final Model is versatile and can be used not only considering a single detection but also consecutive AF detections while retaining a high overall accuracy rate. Such process is able to support an expanded comprehension of fire behavior (e.g., duration and expansion) and prioritize the daily activities of firefighters. In regions so extensive as MATOPIBA and with low resources for environmental management, such prioritization is essential.

Although further advances and studies are required, Final Model and the reference satellites could even be integrated, providing an even more accurate AF detection and monitoring in the region; for instance, those fires that are already burning and were also detected by VIIRS/MODIS should have a priority.

Finally, because consistent fire policies are urged for Cerrado conservation (DURIGAN; RATTER, 2016), and objective regulations require a better comprehension of the fire scenario, Final Model can be an important tool for providing detailed information about the fire behaviour in the region.

6.2 Recommendations for future work

This thesis was developed based o CRISP-DM framework. In this manner, once Final Model was evaluated and published (PLETSCH et al., 2022), it is indicated its diffusing for a large public as well as technical training about its use and application in real life for fire management.

Besides the diffusing, future studies could be focused on data, modeling and applications:

- It is important to obtain a larger number of high-quality BA training samples to train more realistic and updated models;

- Incorporation of new and harmonized data, including the integration of Band 7 with other bands, smoke/fire plumes, NDVI, and precipitation, once there is an inverse relationship between the incidence of fire and vegetation and precipitation Mataveli et al. (2018);

- The use of statistical metrics, such as minimum, maximum, and standard deviation, from different bands could also be used instead of the whole time series. This process could provide a less computational process requirements;

- Consider the fire susceptibility as a support data in the model, where more susceptible areas could have more weight. For instance, the fuel moisture content is also important for fire management, once it can support firefighters strategies (CHUVIECO et al., 2020). In this manner, Remote Sensing dataset related to local moisture could also be used to create an algorithm able to perform fire risk maps incorporated to Final Model results;

- Comprehend in more detail the ABI temporal behaviour for different targets;

- Because Machine Learning is still an active area of research in Remote Sensing, different algorithms have been continuously developed (MAXWELL et al., 2018). As such, we indicate the use of CRISP-DM for new algorithms, which could provide improved Machine Learning performance;

- We used binary classification, as such it would be interesting to measure the probabilistic relationships between variables;

- Train the model for the whole Brazil and not only for MATOPIBA;

- Studies are required to understand: After how much time it is indicated to retrain the model?

- Exploration of more fire behaviours by means of Final Model, such as fire direction;

- Compare the results with similar AF datasets.

REFERENCES

ABBES, A. B.; FARAH, I. R. Prediction changes for nonstationary multi-temporal satellite images using hmm. In: FAIZ, S.; MAHMOUDI, K. (Ed.). **Handbook of research on geographic information systems applications and advancements**. [S.l.]: IGI Global, 2017. p. 387–406. 48

ABDI, A. M.; BRANDT, M.; ABEL, C.; FENSHOLT, R. Satellite remote sensing of savannas: Current status and emerging opportunities. **Journal of Remote Sensing**, v. 2022, 2022. 2

ABREU, R. C.; HOFFMANN, W. A.; VASCONCELOS, H. L.; PILON, N. A.; ROSSATTO, D. R.; DURIGAN, G. The biodiversity cost of carbon sequestration in tropical savanna. **Science Advances**, v. 3, n. 8, p. e1701284, 2017. 1, 12

ADHIKARI, R.; AGRAWAL, R. K. **An introductory study on time series modeling and forecasting**. 2013. [S.l.]: Lap, 2013. 36

AGEE, J. K. The landscape ecology of western forest fire regimes. **Northwest Science**, v. 72, n. 17, p. 24–34, 1998. 13, 20

ALCARAS, E.; COSTANTINO, D.; GUASTAFERRO, F.; PARENTE, C.; PEPE, M. Normalized burn ratio plus (nbr+): a new index for sentinel-2 imagery. **Remote Sensing**, v. 14, n. 7, p. 1727, 2022. 33, 34

ALENCAR, A.; MOUTINHO, P.; ARRUDA, V.; SILVÉRIO, D. **The Amazon in flames: fire and deforestation in 2019—and what's to come in 2020**. [S.l.]: IPAM Technical Note, 3, 1–14, 2020. Available from: <https://ipam.org.br/wp-content/uploads/2020/04/NT3-Fire-2019.pdf>. Access on: 03 Feb. 2022. 1

ALVARADO, S. T.; FORNAZARI, T.; CÓSTOLA, A.; MORELLATO, L. P. C.; SILVA, T. S. F. Drivers of fire occurrence in a mountainous Brazilian Cerrado savanna: tracking long-term fire regimes using remote sensing. **Ecological Indicators**, v. 78, p. 270–281, 2017. 50

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; MORAES, G. de; LEONARDO, J.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013. 7

AMPOMAH, E. K.; QIN, Z.; NYAME, G. Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. **Information**, v. 11, n. 6, p. 332, 2020. 43

ANDELA, N.; MORTON, D. C.; GIGLIO, L.; PAUGAM, R.; CHEN, Y.; HANTSON, S.; WERF, G. R.; RANDERSON, J. T. The global fire atlas of individual fire size, duration, speed and direction. **Earth System Science Data**, v. 11, n. 2, p. 529–552, 2019. 20

ANDERSON, A. B.; POSEY, D. A. Reflorestamento indígena. **Ciência Hoje**, v. 6, n. 31, p. 44–50, 1987. 12, 17

ANDERSON, L. O. et al. Detecção de cicatrizes de áreas queimadas baseada no modelo linear de mistura espectral e imagens índice de vegetação utilizando dados multitemporais do sensor modis/terra no estado do Mato Grosso, Amazônia brasileira. **Acta Amazonica**, v. 35, n. 4, p. 445–456, 2005. 23

APOSTOLAKIS, A.; GIRTSOU, S.; KONTOES, C.; PAPOUTSIS, I.; TSOUTSOS, M. Implementation of a random forest classifier to examine wildfire predictive modelling in greece using diachronically collected fire occurrence and fire mapping data. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA MODELING, 27., 2021, Prague, Czech Republic. **Proceedings...** Czech Republic: Springer, 2021. p. 318–329. 51

APPEL, M.; PEBESMA, E. On-demand processing of data cubes from satellite image collections with the gdalcubes library. **Data**, v. 4, n. 3, p. 92, 2019. 34

ARAGÃO, L. E. O. E. C. de; Silva Junior, C. H. L.; ANDERSON, L. O. **O desafio do Brasil para conter o desmatamento e as queimadas na Amazônia durante a pandemia por COVID-19 em 2020: implicações ambientais, sociais e sua governança**. São José dos Campos: INPE, 2020. 34 p. Doi 10.13140/RG.2.2.17256.49921. 1

ARCHIBALD, S.; STAVER, A. C.; LEVIN, S. A. Evolution of human-driven fire regimes in Africa. **Proceedings of the National Academy of Sciences**, v. 109, n. 3, p. 847–852, 2012. 13, 20

ARRUDA, F. V. d.; SOUSA, D. G. d.; TERESA, F. B.; PRADO, V. H. M. d.; CUNHA, H. F. d.; IZZO, T. J. Trends and gaps of the scientific literature about the effects of fire on brazilian Cerrado. **Biota Neotropica**, v. 18, 2018. 81

BACHMEIER, S. **Fire at a Jim Beam bourbon warehouse in Kentucky**. 2019. Available from: <https://cimss.ssec.wisc.edu/goes/blog/archives/33679>. Access on: 09 Nov. 2019. 28

BACHMEIER, S.; SCHMIT, T.; GERTH, J. **Quick guide: ABI Band 10 (7.3 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band10.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 13 (10.3 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band13.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 8 (6.2 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band08.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 9 (6.9 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band09.pdf>. Access on: 07 Nov. 2019. 28

BAH, M.; GUNSHOR, M.; SCHMIT, T. Generation of goes-16 true color imagery without a green band. **Earth and Space Science**, v. 5, n. 9, p. 549–558, 2018. 26, 27

BANNARI, A.; MORIN, D.; BONN, F.; HUETE, A. A review of vegetation indices. **Remote Sensing Reviews**, v. 13, n. 1-2, p. 95–120, 1995. 23, 32

BARMPOUTIS, P.; PAPAIOANNOU, P.; DIMITROPOULOS, K.; GRAMMALIDIS, N. A review on early forest fire detection systems using optical remote sensing. **Sensors**, v. 20, n. 22, p. 6442, 2020. 49

BATCHELDER, R. B.; HIRT, H. F. **Fire in tropical forests and grasslands**. [S.l.: s.n.], 1966. 16

BENCHERIF, H. et al. Investigating the long-range transport of aerosol plumes following the Amazon fires (august 2019): a multi-instrumental approach from ground-based and satellite observations. **Remote Sensing**, v. 12, n. 22, p. 3846, 2020. 1

BEUCHLE, R.; GRECCHI, R. C.; SHIMABUKURO, Y. E.; SELIGER, R.; EVA, H. D.; SANO, E.; ACHARD, F. Land cover changes in the Brazilian Cerrado and

Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. **Applied Geography**, v. 58, p. 116–127, mar 2015. ISSN 0143-6228. 11, 13

BOND, W. J.; WOODWARD, F. I.; MIDGLEY, G. F. The global distribution of ecosystems in a world without fire. **New Phytologist**, v. 165, n. 2, p. 525–538, 2005. 1

BOT, K.; BORGES, J. G. A systematic review of applications of machine learning techniques for wildfire management decision support. **Inventions**, v. 7, n. 1, p. 15, 2022. 43, 61

BOVOLO, F.; BRUZZONE, L.; SOLANO-CORREA, Y. Multitemporal analysis of remotely sensed image data. **Comprehensive Remote Sensing**, p. 156–185, jan 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780124095489103380>. 35, 46

BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2015. 35

BRASIL. MINISTÉRIO DE CIÊNCIA, TECNOLOGIA E INOVAÇÕES. **Second biennial update report of Brazil: to the United Nations framework convention on climate change**. [S.l.]: Brasília: MCTIC, 2017. 14

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001. 43

CALLE, A.; CASANOVA, J.-L. Forest fires and remote sensing. In: COSKUN, H. G.; CIGIZOGLU, H. K.; MAKTAV, M. D. (Ed.). **Integration of information for environmental security**. [S.l.]: Springer, 2008. p. 247–290. 24

CAO, C.; XIONG, J.; BLONSKI, S.; LIU, Q.; UPRETY, S.; SHAO, X.; BAI, Y.; WENG, F. Suomi npp viirs sensor data record verification, validation, and long-term performance monitoring. **Journal of Geophysical Research: Atmospheres**, v. 118, n. 20, p. 11–664, 2013. 25

CASTRO, E. A. D.; KAUFFMAN, J. B. Ecosystem structure in the brazilian Cerrado: a vegetation gradient of aboveground biomass, root mass and consumption by fire. **Journal of Tropical Ecology**, v. 14, n. 3, p. 263–283, 1998. 8

CHEN, Y.; MORTON, D. C.; JIN, Y.; COLLATZ, G. J.; KASIBHATLA, P. S.; WERF, G. R. van der; DEFRIES, R. S.; RANDERSON, J. T. Long-term trends and interannual variability of forest, savanna and agricultural fires in South America. **Carbon Management**, v. 4, n. 6, p. 617–638, 2013. 51

CHUVIECO, E.; AGUADO, I.; SALAS, J.; GARCÍA, M.; YEBRA, M.; OLIVA, P. Satellite remote sensing contributions to wildland fire science and management. **Current Forestry Reports**, v. 6, n. 2, p. 81–96, 2020. 3, 49, 84

CHUVIECO, E.; KASISCHKE, E. S. Remote sensing information for fire management and fire effects assessment. **Journal of Geophysical Research: Biogeosciences**, v. 112, n. G1, 2007. 24

CHUVIECO, E.; MARTIN, M. P.; PALACIOS, A. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. **International Journal of Remote Sensing**, v. 23, n. 23, p. 5103–5110, 2002. 32, 34

CHUVIECO, E. et al. Historical background and current developments for mapping burned area from satellite earth observation. **Remote Sensing of Environment**, v. 225, p. 45–64, 2019. 23, 31, 32, 33

CIANCIARUSO, M. V.; BATALHA, M. A.; SILVA, I. A. da. Seasonal variation of a hyperseasonal Cerrado, Emas National Park, central Brazil. **Flora-Morphology, Distribution, Functional Ecology of Plants**, v. 200, n. 4, p. 345–353, 2005. 8

CLARKE, B.; FOKOUE, E.; ZHANG, H. H. **Principles and theory for data mining and machine learning**. [S.l.]: Springer Science & Business Media, 2009. 41

COUNTRYMAN, C. M. **Fire environment concept**. [S.l.]: Pacific Southwest Forest and Range Experiment Station, 1972. 16, 18

COUTINHO, L. M. Fire in the ecology of the brazilian Cerrado. In: GOLDAMMER, J. G. (Ed.). **Fire in the tropical biota**. [S.l.]: Springer, 1990. p. 82–105. 11, 12, 13

CRUTZEN, P.; GOLDAMMER, J. Fire in the environment: The ecological, atmospheric, and climatic importance of vegetation fire. In: CRUTZEN, P.; GOLDAMMER, J. (Ed.). **Dahlem Workshop Reports – Environmental Sciences**. Chichester: John Wileys & Sons, 1993. p. 400. 16

CSISZAR, I.; DENIS, L.; GIGLIO, L.; JUSTICE, C. O.; HEWSON, J. Global fire activity from two years of modis data. **International Journal of Wildland Fire**, v. 14, n. 2, p. 117–130, 2005. 25

CSISZAR, I.; SCHROEDER, W.; GIGLIO, L.; ELLICOTT, E.; VADREVU, K. P.; JUSTICE, C. O.; WIND, B. Active fires from the suomi npp visible infrared imaging radiometer suite: product status and first evaluation results. **Journal of Geophysical Research: Atmospheres**, v. 119, n. 2, p. 803–816, 2014. 25

DAMASCO, G.; FONTES, C.; FRANÇOSO, R.; HAIDAR, R. The Cerrado biome: a forgotten biodiversity hotspot. **Frontiers for Young Minds**, Frontiers Media SA, v. 6, 2018. 2

DIAS, B.; MIRANDA, H. O projeto fogo. In: MIRANDA, H. (Ed.). **Efeitos do regime do fogo sobre a estrutura de comunidades de Cerrado: resultados do Projeto Fogo**. Brasília: IBAMA, 2010. p. 15–22. 11, 12

DODGE, Y. **The concise encyclopedia of statistics**. [S.l.]: Springer Science & Business Media, 2008. 35

DURIGAN, G.; RATTER, J. A. The need for a consistent fire policy for Cerrado conservation. **Journal of Applied Ecology**, v. 53, n. 1, p. 11–15, 2016. 12, 14, 83

DWYER, E.; PINNOCK, S.; GRÉGOIRE, J.-M.; PEREIRA, J. Global spatial and temporal distribution of vegetation fire as determined from satellite observations. **International Journal of Remote Sensing**, v. 21, n. 6-7, p. 1289–1302, 2000. 21

EHLERS, M.; JANOWSKY, R.; GAEHLER, M. New remote sensing concepts for environmental monitoring. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Remote sensing for environmental monitoring, GIS applications, and geology**. [S.l.], 2002. v. 4545, p. 1–12. 23

ELGELDAWI, E.; SAYED, A.; GALAL, A. R.; ZAKI, A. M. Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis. **Informatics**, v. 8, n. 4, p. 79, 2021. 42

EUROPEAN SPACE AGENCY-(ESA). **Sentinel-2 user handbook**. 2015. Available from: <https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook>. Access on: 30 Oct. 2019. 31

FALK, D. A.; MILLER, C.; MCKENZIE, D.; BLACK, A. E. Cross-scale analysis of fire regimes. **Ecosystems**, v. 10, n. 5, p. 809–823, 2007. 16, 20

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. Knowledge discovery and data mining: towards a unifying framework. In: KDD'96: SECOND INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 2., 1996, Portland Oregon. **Proceedings...** Portland Oregon: AAAI Press, 1996. p. 82–88. 38, 39

FEARNSIDE, P. M. Soybean cultivation as a threat to the environment in Brazil. **Environmental Conservation**, v. 28, n. 1, p. 23–38, 2001. 11

FERRAZ-VICENTINI, K. **História do fogo no Cerrado: uma análise palinológica**. 208 p. Tese (Doutorado em Ecologia) — Universidade de Brasília, Brasília, DF, 1999. 12, 16

FERREIRA, D. H. L.; BADINGER, A.; TENDOLINI, J. C. Distribuições de tendências sazonais de temperatura média e precipitação nos biomas brasileiros. **Revista Brasileira de Meteorologia**, v. 33, n. 1, p. 97–113, 2018. 7

FIDELIS, A. Is fire always the "bad guy"? **Flora**, v. 268, p. 151611, 2020. ISSN 0367-2530. 1

FIDELIS, A.; ALVARADO, S. T.; BARRADAS, A. C. S.; PIVELLO, V. R. The year 2017: megafires and management in the Cerrado. **Fire**, v. 1, n. 3, 2018. ISSN 2571-6255. 1, 13, 14

FONSECA, L. M.; KöRTING, T. S.; BENDINI, H. do N.; GIROLAMO-NETO, C. D.; NEVES, A. K.; SOARES, A. R.; TAQUARY, E. C.; MARETTO, R. V. Pattern recognition and remote sensing techniques applied to land use and land cover mapping in the brazilian savannah. **Pattern Recognition Letters**, v. 148, p. 54–60, 2021. ISSN 0167-8655. 37

FRANÇOSO, R. D.; BRANDÃO, R.; NOGUEIRA, C. C.; SALMONA, Y. B.; MACHADO, R. B.; COLLI, G. R. Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. **Natureza e Conservacao**, v. 13, n. 1, p. 35–40, 2015. ISSN 16790073. 11

FRANKE, J.; BARRADAS, A. C. S.; BORGES, M. A.; COSTA, M. M.; DIAS, P. A.; HOFFMANN, A. A.; OROZCO FILHO, J. C.; MELCHIORI, A. E.; SIEGERT, F. Fuel load mapping in the Brazilian Cerrado in support of integrated fire management. **Remote Sensing of Environment**, v. 217, p. 221 – 232, 2018.

ISSN 0034-4257. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425718303936>. 14

FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. **Annals of Statistics**, p. 1189–1232, 2001. 43

FRIZZO, T. L.; BONIZARIO, C.; BORGES, M. P.; VASCONCELOS, H. Uma revisão dos efeitos do fogo sobre a fauna de formações savânicas do Brasil. **Oecologia Australis**, v. 15, n. 2, p. 365–379, 2011. 13

GARRIDO, A. P. **What is the difference between bagging and boosting?** [S.l.]: Quantdare, 2016. Available from: <https://quantdare.com/what-is-the-difference-between-bagging-and-boosting/>. Access on: 01 Aug. 2022. 43, 44

GIBBS, H. K.; RAUSCH, L.; MUNGER, J.; SCHELLY, I.; MORTON, D. C.; NOOJIPADY, P.; SOARES-FILHO, B.; BARRETO, P.; MICOL, L.; WALKER, N. F. Brazil's soy moratorium. **Science**, v. 347, n. 6220, p. 377–378, 2015. 11

GIGLIO, L.; CSISZAR, I.; RESTÁS, Á.; MORISETTE, J. T.; SCHROEDER, W.; MORTON, D.; JUSTICE, C. O. Active fire detection and characterization with the advanced spaceborne thermal emission and reflection radiometer (aster). **Remote Sensing of Environment**, v. 112, n. 6, p. 3055–3063, 2008. 24

GIGLIO, L.; KENDALL, J.; JUSTICE, C. Evaluation of global fire detection algorithms using simulated avhrr infrared data. **International Journal of Remote Sensing**, v. 20, n. 10, p. 1947–1985, 1999. 23

GIGLIO, L.; SCHROEDER, W.; HALL, J. V.; JUSTICE, C. O. **Modis collection 6 active fire product user's guide revision B**. Washington, DC, USA: NASA, 2018. 25

GIGLIO, L.; SCHROEDER, W.; JUSTICE, C. O. The collection 6 modis active fire detection algorithm and fire products. **Remote Sensing of Environment**, v. 178, p. 31–41, 2016. 3, 25

GOMES, L.; MIRANDA, H. S.; BUSTAMANTE, M. M. da C. How can we advance the knowledge on the behavior and effects of fire in the Cerrado biome? **Forest Ecology and Management**, v. 417, p. 281–290, 2018. 16, 19, 21

GORUNESCU, F. **Data mining: concepts, models and techniques**. [S.l.]: Springer Science & Business Media, 2011. 41

HADDEN, R. M. Heat release rate. In: MANZELLO, S. L. (Ed.). **Encyclopedia of wildfires and wildland-Urban interface (WUI) fires**. Cham: Springer International Publishing, 2020. p. 603–610. ISBN 978-3-319-52090-2. 21

HALL, J. V.; ZHANG, R.; SCHROEDER, W.; HUANG, C.; GIGLIO, L. Validation of GOES-16 ABI and MSG SEVIRI active fire products. **International Journal of Applied Earth Observation and Geoinformation**, v. 83, p. 101928, 2019. 3

HAN, J.; KAMBER, M.; MINING, D. **Data mining: concepts and techniques**. [S.l.]: Elsevier, 2006. 35, 38, 41

HAQQ-MISRA, J.; LEE, S.; FRIERSON, D. M. Tropopause structure and the role of eddies. **Journal of the Atmospheric Sciences**, v. 68, n. 12, p. 2930–2944, 2011. 28

HE, H.; GARCIA, E. A. Learning from imbalanced data. **IEEE Transactions on Knowledge and Data Engineering**, v. 21, n. 9, p. 1263–1284, 2009. 45

HIGA, L. et al. Active fire mapping on brazilian pantanal based on deep learning and cbers 04a imagery. **Remote Sensing**, v. 14, n. 3, p. 688, 2022. 79

HOFFMAN, J.; SCHMIDT, C.; PRINS, E.; BRUNNER, J. The goes-r abi wild fire automated biomass burning algorithm. In: AGU FALL MEETING, 2011, San Francisco, CA. **Abstracts...** Washington, DC: American Geophysical Union, 2011. p. 1636. Abstract IN53C-1636. 30

HOLDEN, Z.; SMITH, A.; MORGAN, P.; ROLLINS, M.; GESSLER, P. Evaluation of novel thermally enhanced spectral indices for mapping fire perimeters and comparisons with fire atlas data. **International Journal of Remote Sensing**, v. 26, n. 21, p. 4801–4808, 2005. 34

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS-(INPE). **FIP Cerrado: Projeto de desenvolvimento de sistemas de prevenção de incêndios florestais e monitoramento da cobertura vegetal no Cerrado brasileiro**. 2018. Available from: <www.terrabrasilis.dpi.inpe.br/app/dashboard>. Access on: 13 Jul. 2019. 11

_____. **Lançamento da "Base2" de focos de queima de vegetação**. [S.l.]: INPE, 2018. Available from: <http://queimadas.dgi.inpe.br/queimadas/portal/informacoes/

novidades/LanamentodaBase2defocosdequeimadevegetao.pdf>. Access on: 18 Nov. 2019. 11

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS-(INPE). **Programa Queimadas**. 2021. Available from: <https://queimadas.dgi.inpe.br/queimadas/portal/>. Access on: 10 Nov. 2021. 2

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS-(INPE). **TerraBrasilis, PRODES (Desmatamento)**. 2022. Available from: <http://terrabrasilis.dpi.inpe.br/>. Access on: 22 Feb. 2022. 2

IVO, I. O.; BIUDES, M. S.; VOURLITIS, G. L.; GOMES, N.; MARTIM, C. C. Effect of fires on biophysical parameters, energy balance and evapotranspiration in a protected area in the Brazilian Cerrado. **Remote Sensing Applications: Society and Environment**, p. 100342, 2020. ISSN 2352-9385. Disponível em: <https://doi.org/10.1016/j.rsase.2020.100342>. 1

JAIN, A. **A comprehensive beginner's guide to create a Time Series Forecast (with Codes in Python and R)**. [S.l.]: Analytics Vidhya: Learn everything about analytics, 2016. Available from: <https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>. Access on: 19 Nov. 2019. 36

JAIN, P.; COOGAN, S. C.; SUBRAMANIAN, S. G.; CROWLEY, M.; TAYLOR, S.; FLANNIGAN, M. D. A review of machine learning applications in wildfire science and management. **Environmental Reviews**, v. 28, n. 4, p. 478–505, 2020. 43, 49, 61

JENSEN, J. R. **Introductory digital image processing: a remote sensing perspective**. [S.l.]: Prentice Hall Press, 2015. 23

JONSSON, P.; EKLUNDH, L. Seasonality extraction by function fitting to time-series of satellite sensor data. **IEEE transactions on Geoscience and Remote Sensing**, v. 40, n. 8, p. 1824–1832, 2002. 46

JÖNSSON, P.; EKLUNDH, L. Timesat—a program for analyzing time-series of satellite sensor data. **Computers & Geosciences**, v. 30, n. 8, p. 833–845, 2004. 46, 47

JOYCE, K. E.; BELLISS, S. E.; SAMSONOV, S. V.; MCNEILL, S. J.; GLASSEY, P. J. A review of the status of satellite remote sensing and image processing

techniques for mapping natural hazards and disasters. **Progress in Physical Geography**, v. 33, n. 2, p. 183–207, 2009. 21

JUSTICE, C.; GIGLIO, L.; KORONTZI, S.; OWENS, J.; MORISETTE, J.; ROY, D.; DESCLOITRES, J.; ALLEAUME, S.; PETITCOLIN, F.; KAUFMAN, Y. The modis fire products. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 244–262, 2002. 24, 25

KALLURI, S.; ALCALA, C.; CARR, J.; GRIFFITH, P.; LEBAIR, W.; LINDSEY, D.; RACE, R.; WU, X.; ZIERK, S. From photons to pixels: processing data from the advanced baseline imager. **Remote Sensing**, v. 10, n. 2, p. 177, 2018. 60

KELLEY, D. I.; BURTON, C.; HUNTINGFORD, C.; BROWN, M. A.; WHITLEY, R.; DONG, N. Low meteorological influence found in 2019 amazonia fires. **Biogeosciences**, v. 18, n. 3, p. 787–804, 2021. 1

KEY, C.; BENSON, N. Landscape assessment (la) sampling and analysis methods. In: LUTTES, D. C. et al. (Ed.). **Firemon: fire effects monitoring and inventory systems**. [S.l.]: USDA Forest Service, 2005. p. LA 1–51. 32, 34

KLINK, C. A.; MACHADO, R. B. A conservação do Cerrado brasileiro. **Megadiversidade**, Belo Horizonte, v. 1, n. 1, p. 147–155, 2005. 1, 11, 12, 13, 17

KLINK, C. A.; MOREIRA, A. Past and current human occupation, and land use. In: OLIVEIRA, P.; MARQUES, R. (Ed.). **The Cerrados of Brazil: ecology and natural history of a neotropical savanna**. [S.l.]: Columbia University Press USA, 2002. p. 69–88. 8, 9

KOHAVI, R.; PROVOST, F. Glossary of terms. special issue of applications of machine learning and the knowledge discovery process. **Machine Learning**, v. 30, p. 271–274, 01 1998. 45

KÖRTING, T. S.; FONSECA, L. M. G.; CÂMARA, G. Geodma—geographic data mining analyst. **Computers & Geosciences**, v. 57, p. 133–145, 2013. 20

KREBS, P.; PEZZATTI, G. B.; MAZZOLENI, S.; TALBOT, L. M.; CONEDERA, M. Fire regime: history and definition of a key concept in disturbance ecology. **Theory in Biosciences**, v. 129, n. 1, p. 53–69, 2010. 17, 18

KUENZER, C.; DECH, S.; WAGNER, W. Remote sensing time series revealing land surface dynamics: status quo and the pathway ahead. In: KUENZER, C.; DECH, S.; WAGER, W. (Ed.). **Remote Sensing Time Series**. [S.l.]: Springer, 2015. p. 1–24. 35, 46

LANEY, D. et al. 3d data management: controlling data volume, velocity and variety. **META Group Research Note**, v. 6, n. 70, p. 1, 2001. 2

LASHOF, D. The contribution of biomass burning to global warming: An integrated assessment. In: LEVINE, J. (Ed.). **Global biomass burning: atmospheric, climatic, and biospheric implications**. Williamsburg, VA (United States): Massachusetts Institute of Technology Press, 1991. p. 441–444. 1

LEEUWENBERG, F.; SALIMON, M. **Os Xavantes na balança das civilizações**. Rio de Janeiro: [s.n.], 1999. 12, 17

LEGENDRE, P.; LEGENDRE, L. **Numerical ecology (Vol. 20): Developments in environmental modelling**. 2. ed. [S.l.]: Elsevier, 1998. 44

LEHMANN, C. E. et al. Savanna vegetation-fire-climate relationships differ among continents. **Science**, v. 343, n. 6170, p. 548–552, 2014. 16

LI, F.; ZHANG, X.; KONDRAGUNTA, S.; SCHMIDT, C. C.; HOLMES, C. D. A preliminary evaluation of goes-16 active fire product using Landsat-8 and VIIRS active fire data, and ground-based prescribed fire records. **Remote Sensing of Environment**, v. 237, p. 111600, 2020. 3

LIASHCHYNSKYI, P.; LIASHCHYNSKYI, P. Grid search, random search, genetic algorithm: a big comparison for nas. **arXiv preprint arXiv:1912.06059**, 2019. 42

LIBONATI, R.; DACAMARA, C.; SETZER, A.; MORELLI, F.; MELCHIORI, A. An algorithm for burned area detection in the Brazilian Cerrado using 4 μm modis imagery. **Remote Sensing**, v. 7, n. 11, p. 15782–15803, 2015. 21

LINDLEY, T.; ANDERSON, A. R.; MAHALE, V. N.; CURL, T. S.; LINE, W. E.; LINDSTROM, S. S.; BACHMEIER, A. S. Wildfire detection notifications for impact-based decision support services in Oklahoma using geostationary super rapid scan satellite imagery. **Journal of Operational Meteorology**, v. 4, n. 14, 2016. 30, 52

LINDSTROM, S.; SCHMIT, T.; GERTH, J. **Quick guide: ABI Band 1 (0.47 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band01.pdf>. Access on: 07 Nov. 2019. 26

_____. **Quick guide: ABI Band 11 (8.5 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band11.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 12 (9.6 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band12.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 14 (11.2 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band14.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 15 (12.3 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band15.pdf>. Access on: 07 Nov. 2019. 28

_____. **Quick guide: ABI Band 16 (13.3 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band16.pdf>. Access on: 07 Nov. 2019. 29

_____. **Quick guide: ABI Band 2 (0.64 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band02.pdf>. Access on: 07 Nov. 2019. 26

_____. **Quick guide: ABI Band 3 (0.86 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band03.pdf>. Access on: 07 Nov. 2019. 26

_____. **Quick guide: ABI Band 4 (1.37 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band04.pdf>. Access on: 07 Nov. 2019. 27

_____. **Quick guide: ABI Band 5 (1.61 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band05.pdf>. Access on: 07 Nov. 2019. 27

_____. **Quick guide: ABI Band 6 (2.24 μm)**. 2017. Available from: <http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band06.pdf>. Access on: 07 Nov. 2019. 27

MACHADO-SILVA, F.; LIBONATI, R.; LIMA, T. F. M. de; PEIXOTO, R. B.; FRANÇA, J. R. de A.; MAGALHÃES, M. d. A. F. M.; SANTOS, F. L. M.; RODRIGUES, J. A.; DACAMARA, C. C. Drought and fires influence the respiratory diseases hospitalizations in the amazon. **Ecological Indicators**, v. 109, p. 105817, 2020. 1

MACLENNAN, B. J. Philosophia naturalis rediviva: natural philosophy for the twenty-first century. **Philosophies**, v. 3, n. 4, p. 38–0, 2018. ISSN 2409-9287. 15

MARENGO, J. A.; JIMENEZ, J. C.; ESPINOZA, J.-C.; CUNHA, A. P.; ARAGÃO, L. E. Increased climate pressure on the agricultural frontier in the eastern Amazonia–Cerrado transition zone. **Scientific Reports**, v. 12, n. 1, p. 1–10, 2022. 2

MATAVELI, G. A. V.; SILVA, M. E. S.; PEREIRA, G.; SILVA, F. da; CARDOZO, F. S. K.; BERTANI, G.; COSTA, J. C.; RAMOS, R. de C.; SILVA, V. V. da. Analysis of fire dynamics in the brazilian savannas. **Natural Hazards and Earth System Sciences Discussions**, p. 1–27, 2017. 13, 21

MATAVELI, G. A. V.; SILVA, M. E. S.; PEREIRA, G.; CARDOZO, F. da S.; KAWAKUBO, F. S.; BERTANI, G.; COSTA, J. C.; RAMOS, R. de C.; SILVA, V. V. da. Satellite observations for describing fire patterns and climate-related fire drivers in the brazilian savannas. **Natural Hazards and Earth System Sciences**, v. 18, n. 1, p. 125, 2018. 11, 13, 50, 84

MAXWELL, A. E.; WARNER, T. A.; FANG, F. Implementation of machine-learning classification in remote sensing: an applied review. **International Journal of Remote Sensing**, v. 39, n. 9, p. 2784–2817, 2018. 43, 45, 84

MCKENZIE, D.; MILLER, C.; FALK, D. A. **The landscape ecology of fire**. [S.l.]: Springer Science & Business Media, 2011. 16

MELO, M. M. d. **A confluência entre a ecologia do fogo e o conhecimento Xavante sobre o manejo do fogo no Cerrado**. 127 p. Tese (Doutorado em Desenvolvimento Sustentável) — Universidade de Brasília, Brasília, 2004. 12, 17

MILANOVIĆ, S.; MARKOVIĆ, N.; PAMUČAR, D.; GIGOVIĆ, L.; KOSTIĆ, P.; MILANOVIĆ, S. D. Forest fire probability mapping in eastern serbia: logistic regression versus random forest method. **Forests**, v. 12, n. 1, p. 5, 2021. 62

MINISTÉRIO DO MEIO AMBIENTE-(MMA). **Plano de ação para prevenção e controle do desmatamento e das queimadas no Cerrado - PPCerrado**. Brasília: MMA, 2009. 7, 12, 14

_____. **PPCerrado – Plano de Ação para prevenção e controle do desmatamento e das queimadas no Cerrado: 2 fase (2014-2015)**. Brasília: MMA, 2014. ISBN 978-85-7738-214-9. 7, 8, 11, 12

_____. **Mapeamento do uso e cobertura da terra do Cerrado - Projeto TerraClass Cerrado 2013**. Brasília: MMA, 2015. Available from: <www.dpi.inpe.br/tccerrado/Metodologia_TCCerrado_2013.pdf>. Access on: 13 Jul. 2019. 14

_____. **Brazil's Forest reference emission level for reducing emissions from deforestation in the Cerrado biome for results-based payments for REDD+ under the United Nations Framework Convention on Climate Change**. Brasília, 2017. Working Group of Technical Experts on REDD+ created through MMA Ordinance No. 41-2014. 10

MIRANDA, E. E.; MAGALHÃES, L. A.; CARVALHO, C. A. **Nota técnica: proposta de delimitação territorial do MATOPIBA**. [S.l.]: Campinas: Embrapa, 2014. Available from: <https://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1037313>. Access on: 05 Nov. 2021. 2, 9, 11

MIRANDA, H.; NETO, W.; NEVES, B. Caracterização das queimadas de Cerrado. In: MIRANDA, H. (Ed.). **Efeitos do regime do fogo sobre a estrutura de comunidades de Cerrado: resultados do Projeto Fogo.** Brasília, DF, Brasil.: IBAMA, 2010. p. 23–33. 1, 12, 16, 21

MIRANDA, H. S.; BUSTAMANTE, M. M.; MIRANDA, A. C.; OLIVEIRA, P.; MARQUIS, R. The fire factor. In: OLIVEIRA, P.; MARQUIS, R. (Ed.). **The Cerrados of Brazil: ecology and natural history of a neotropical savanna**. New York: Columbia University Press, 2002. p. 51–68. 1, 12, 17

MIRANDA, H. S.; SATO, M. N.; NETO, W. N.; AIRES, F. S. Fires in the Cerrado, the Brazilian Savanna. In: COCHRANE, M. (Ed.). **Tropical fire ecology**. [S.l.]: Springer, 2009. p. 427–450. 12, 16

MISTRY, J. Fire in the Cerrado (Savannas) of Brazil: an ecological review. **Progress in Physical Geography**, v. 22, n. 4, p. 425–448, 1998. 12, 14, 17

MORGAN, P.; HARDY, C. C.; SWETNAM, T. W.; ROLLINS, M. G.; LONG, D. G. Mapping fire regimes across time and space: understanding coarse and fine-scale fire patterns. **International Journal of Wildland Fire**, v. 10, n. 4, p. 329–342, 2001. 13, 24

MORITZ, M. A.; MORAIS, M. E.; SUMMERELL, L. A.; CARLSON, J.; DOYLE, J. Wildfires, complexity, and highly optimized tolerance. **Proceedings of the National Academy of Sciences**, v. 102, n. 50, p. 17912–17917, 2005. 16

MYERS, R. L.; RODRÍGUEZ-TREJO, D. A. Fire in tropical pine ecosystems. In: COCHRANE, M. (Ed.). **Tropical fire ecology**. [S.l.]: Springer, 2009. p. 557–605. 20

NASCIMENTO, I. V. Cerrado: o fogo como agente ecológico. **Territorium**, n. 8, p. 25–35, 2001. 12, 13, 16, 17

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION-(NASA). **FIRMS - Fire Information for Resource Management System**. 2021. Available from: <https://firms.modaps.eosdis.nasa.gov/>. Access on: 10 Nov. 2021. 2

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION-(NOAA). **NOAA KLM users guide**. 2014. Available from: <https:\/\/www1.ncdc.noaa.gov\/pub/data\/satellite\/publications\/podguides\/N-15%20thru%20N-19/pdf/0.0%20NOAA%20KLM%20Users%20Guide.pdf>. Access on: 30 Jun. 2019. 25

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION-(NOAA) and NATIONAL AERONAUTICS AND SPACE ADMINISTRATION-(NASA). **Fire temperature RGB: quick guide**. [S.l.], 2018. Available from: <https://rammb.cira.colostate.edu/training/visit/quick_guides/Fire_Temperature_RGB.pdf>. Access on: 25 Set. 2022. 27, 51, 52, 54

_____. **Geostationary Operational Environamental Satellite - R Series: Quick guide - ABI Band 7**. 2021. Available from: <https://www.goes-r.gov/exit.html?http://cimss.ssec.wisc.edu/goes/OCLOFactSheetPDFs/ABIQuickGuide_Band07.pdf>. Access on: 09 Dec. 2021. 29

_____. **ABI bands quick information guides**. 2022. Available from: <https://www.goes-r.gov/mission/ABI-bands-quick-info.html>. Access on: 14 Oct. 2022. 29

_____. **Geostationary Operational Environamental Satellite - R Series: Instruments - Advanced Baseline Imager (ABI)**. 2022. Available from: <https://www.goes-r.gov/spacesegment/abi.html>. Access on: 14 Sep. 2022. 2

NATIVI, S.; MAZZETTI, P.; CRAGLIA, M. A view-based model of data-cube to support big earth data systems interoperability. **Big Earth Data**, v. 1, n. 1-2, p. 75–99, 2017. 35

OLIVA, P.; SCHROEDER, W. Assessment of viirs 375 m active fire detection product for direct burned area mapping. **Remote Sensing of Environment**, v. 160, p. 144–155, 2015. 23

OLIVEIRA, U.; SOARES-FILHO, B.; COSTA, W. L. de S.; GOMES, L.; BUSTAMANTE, M.; MIRANDA, H. Modeling fuel loads dynamics and fire spread probability in the Brazilian Cerrado. **Forest Ecology and Management**, v. 482, p. 118889, 2021. 80

PARISIEN, M.-A.; MORITZ, M. A. Environmental controls on the distribution of wildfire at multiple spatial scales. **Ecological Monographs**, v. 79, n. 1, p. 127–154, 2009. 16

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the köppen-geiger climate classification. **Hydrology and Earth System Sciences Discussions**, v. 4, n. 2, p. 439–473, 2007. 7

PEREIRA, G. H. de A.; FUSIOKA, A. M.; NASSU, B. T.; MINETTO, R. Active fire detection in Landsat-8 imagery: a large-scale dataset and a deep-learning study. **ISPRS Journal of Photogrammetry and Remote Sensing**, Elsevier, v. 178, p. 171–186, 2021. 79

PEREIRA, J.; CHUVIECO, E.; BEAUDOIN, A.; DESBOIS, N. **Remote sensing of burned areas: A review of remote sensing methods for the study of large wildland fires**. [S.l.]: Report of the Mega fires Project ENV-CT, Universidad de Alcalá, 1997. 127-184 p. 31

PEREIRA, P. A. A.; MARTHA, G. B.; SANTANA, C. A.; ALVES, E. The development of brazilian agriculture: future technological challenges and opportunities. **Agriculture & Food Security**, v. 1, n. 1, p. 4, 2012. 9

PIVELLO, V. R. The use of fire in the Cerrado and Amazonian rainforests of Brazil: past and present. **Fire Ecology**, v. 7, n. 1, p. 24–39, 2011. 1, 12, 14, 16, 17

PIVELLO, V. R. et al. Understanding Brazil's catastrophic fires: causes, consequences and policy needed to prevent future tragedies. **Perspectives in Ecology and Conservation**, v. 19, n. 3, p. 233–255, 2021. 1, 79, 80

PLANET DEVELOPERS. **PlanetScope: constellation and sensor overview**. 2022. Available from: <https://developers.planet.com/docs/data/planetscope/>. Access on: 10 Jul. 2022. 32

PLETSCH, M. A.; KÖRTING, T. S.; MORITA, F. C.; MORELLI, F.; BITTENCOURT, O.; VICTORINO, P. S. Using goes-16 time series to characterize near real-time active fires in Cerrado. In: BRAZILIAN SYMPOSIUM ON GEOINFORMATICS (GEOINFO), 20., 2019, São José dos Campos. **Proceedings...** São José dos Campos: INPE, 2019. p. 66–76. 2, 51, 52, 53, 55, 56

PLETSCH, M. A.; KÖRTING, T. S.; MORITA, F. C.; SILVA-JUNIOR, C. H.; ANDERSON, L. O.; ARAGÃO, L. E. Near real-time fire detection and monitoring in the matopiba region, Brazil. **Remote Sensing**, v. 14, n. 13, p. 3141, 2022. 83

PLETSCH, M. A. J. S.; KöRTING, T. S.; MORELLI, F. Trends of wildfire studies as a support to burned area monitoring. In: INTERNATIONAL WILDLAND FIRE CONFERENCE (WILDFIRE), 7., 2019, Campo Grande – MS, Brazil. **Proceedings...** [S.l.]: IBAMA, 2019. Poster. 31

PLETSCH, M. A. J. S.; PENHA, T. V.; JUNIOR, C. H. L. S.; MORELLI, F. Combination of spectral indices for burned area detection in the brazilian amazonia. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 19., 2019, Santos. **Proceedings...** São José dos Campos: INPE, 2019. 33

RAMOS-NETO, M. B.; PIVELLO, V. R. Lightning fires in a Brazilian Savanna national park: rethinking management strategies. **Environmental Management**, v. 26, n. 6, p. 675–684, 2000. 1, 12, 13, 14, 17

REECE, B. C. Aristotle's four causes of action. **Australasian Journal of Philosophy**, v. 97, n. 2, p. 213–227, 2019. 15

REED, B. C.; SCHWARTZ, M. D.; XIAO, X. Remote sensing phenology. In: NOORMETS, A. (Ed.). **Phenology of ecosystem processes: applications in global change research**. New York, NY: Springer, 2009. p. 231–246. ISBN 978-1-4419-0026-5. 35

RIBEIRO, J. F.; WALTER, B. M. T. As principais fitofisionomias do bioma Cerrado. **Cerrado: Ecologia e Flora**, v. 1, p. 151–212, 2008. 7, 8, 10

RISSI, M. N.; BAEZA, M. J.; GORGONE-BARBOSA, E.; ZUPO, T.; FIDELIS, A. Does season affect fire behaviour in the Cerrado? **International Journal of Wildland Fire**, v. 26, n. 5, p. 427–433, 2017. 1, 13, 14

ROBINSON, J. M. Fire from space: global fire evaluation using infrared remote sensing. **International Journal of Remote Sensing**, v. 12, n. 1, p. 3–24, 1991. 24, 30

ROSS, J. L. S. **Geografia do Brasil**. [S.l.]: Edusp, 1996. 12, 17

ROSTAMI, A.; SHAH-HOSSEINI, R.; ASGARI, S.; ZAREI, A.; AGHDAMI-NIA, M.; HOMAYOUNI, S. Active fire detection from Landsat-8 imagery using deep multiple kernel learning. **Remote Sensing**, v. 14, n. 4, 2022. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/14/4/992>. 49

ROUSE, J.; HAAS, R.; SCHELL, J.; DEERING, D. Monitoring vegetation systems in the great plains with erts. 1974. Available from: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf>. Access on: 20 Jul. 2019. 32, 34

ROY, D.; LEWIS, P.; JUSTICE, C. Burned area mapping using multi-temporal moderate spatial resolution data—a bi-directional reflectance model-based expectation approach. **Remote Sensing of Environment**, v. 83, n. 1, p. 263 – 286, 2002. ISSN 0034-4257. 31

ROY, D. P.; HUANG, H.; BOSCHETTI, L.; GIGLIO, L.; YAN, L.; ZHANG, H. H.; LI, Z. Landsat-8 and sentinel-2 burned area mapping - a combined sensor multi-temporal change detection approach. **Remote Sensing of Environment**, v. 231, p. 111254, 2019. ISSN 0034-4257. 31

SANO, E. E.; ROSA, R.; BRITO, J. L. S.; FERREIRA, L. G. Land cover mapping of the tropical Savanna region in Brazil. **Environmental Monitoring and Assessment**, v. 166, n. 1-4, p. 113–124, 2010. ISSN 0167-6369. 7, 8

SCHMIDT, C. **GOES-R fire detection and characterization fact sheet**. [S.l.: s.n.], 06 2019. GOES-R Series Program Office. Available from: <https://www.goes-r.gov/education/docs/fs_fire.pdf>. Access on: 15 Set. 2022. 28, 52

SCHMIDT, C.; HOFFMAN, J.; PRINS, E.; LINDSTROM, S. **GOES-R advanced baseline imager (ABI) algorithm theoretical basis document for fire/hot spot characterization, version 2.0**. Silver Spring, Md: NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION-(NOAA), 2010. 3, 30

SCHMIDT, C. C. The current status of the ABI fire detection and characterization algorithm. In: FALL MEETING ABSTRACTS, 2019. **Proceedings...** [S.l.]: American Geophysical Union, 2019. p. A51F–07. 3

_____. Validation of and future plans for GOES ABI fire detection. In: FALL MEETING ABSTRACTS, 2020. **Proceedings...** [S.l.]: American Geophysical Union, 2020. p. A176–0020. 3

SCHMIDT, I. B.; ELOY, L. Fire regime in the Brazilian Savanna: recent changes, policy and management. **Flora**, v. 268, p. 151613, 2020. ISSN 0367-2530. 2, 81

SCHMIT, T.; LINDSTROM, S.; GERTH, J.; GUNSHOR, M. Applications of the 16 spectral bands on the Advanced Baseline Imager (ABI). **Journal of Operational Meteorology**, v. 06, p. 33–46, 06 2018. 26, 27

SCHMIT, T. J.; GRIFFITH, P.; GUNSHOR, M. M.; DANIELS, J. M.; GOODMAN, S. J.; LEBAIR, W. J. A closer look at the ABI on the GOES-R series. **Bulletin of the American Meteorological Society**, v. 98, n. 4, p. 681–698, 2017. 2, 26

SCHMIT, T. J.; GUNSHOR, M. M.; MENZEL, W. P.; GURKA, J. J.; LI, J.; BACHMEIER, A. S. Introducing the next-generation advanced baseline imager on GOES-R. **Bulletin of the American Meteorological Society**, v. 86, n. 8, p. 1079–1096, 2005. 2

SCHROEDER, W.; GIGLIO, L. **NASA VIIRS Land Science Investigator Processing System (SIPS) Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m & 750 m active fire products: product User's Guide Version 1.4**. Washington: NASA, 2018. Available from: <https://viirsland.gsfc.nasa.gov/PDF/VIIRS_activefire_User_Guide.pdf>. Access on: 25 Set. 2022. 25

SCHROEDER, W.; OLIVA, P.; GIGLIO, L.; CSISZAR, I. A. The new viirs 375m active fire detection data product: algorithm description and initial assessment. **Remote Sensing of Environment**, v. 143, p. 85–96, 2014. ISSN 0034-4257. 3, 25

SCHROEDER, W.; OLIVA, P.; GIGLIO, L.; QUAYLE, B.; LORENZ, E.; MORELLI, F. Active fire detection using Landsat-8/OLI data. **Remote Sensing of Environment**, v. 185, p. 210–220, 2016. 24

SCOTT, A. C.; BOWMAN, D. M.; BOND, W. J.; PYNE, S. J.; ALEXANDER, M. E. **Fire on earth: an introduction**. [S.l.]: John Wiley & Sons, 2013. 21

SEAMAN, C. J.; MILLER, S. D.; LINDSEY, D. T.; HILLGER, D. W. Jpss and goes-r multispectral imagery applications and product development at cira. In: MCWILLIAMS, G.; JAMILKOWSKI, M. L.; KALLURI, S.; SCHMIT, T. J.;

MANGO, S. A. (Ed.). **Proceedings of 13th Annual Symposium on New Generation Operational Environmental Satellite Systems**. [S.l.]: American Meteorological Society, 2017. p. 259. 28, 52

Sentinel Hub. **KSWIR - short wave infrared RGB composite**. 2021. Available from: <https://custom-scripts.sentinel-hub.com/sentinel-2/swir-rgb/>. Access on: 25 Nov. 2021. 31, 32, 58

SHREFFLER, J.; HUECKER, M. R. **Diagnostic testing accuracy: Sensitivity, specificity, predictive values and likelihood ratios**. [S.l.]: StatPearls Publishing, 2022. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK557491/>. Access on: 25 Sep. 2022. 44

SILVA, J.; FARINAS, M.; FELFILI, J.; KLINK, C. Spatial heterogeneity, land use and conservation in the Cerrado region of Brazil. **Journal of Biogeography**, v. 33, n. 3, p. 536–548, 2006. 13

SILVA, P. R. d. S.; IGNOTTI, E.; OLIVEIRA, B. F. A. de; JUNGER, W. L.; MORAIS, F.; ARTAXO, P.; HACON, S. High risk of respiratory diseases in children in the fire period in Western Amazon. **Revista de Saúde Pública**, v. 50, p. 1–11, 2016. ISSN 15188787. 1

SINGH, A. **A gentle introduction to handling a non-stationary time series in Python**. [S.l.]: Analytics Vidhya: Learn everything about analytics, 2018. Available from: <https://www.analyticsvidhya.com/blog/2018/09/non-stationary-time-series-python/>. Access on: 19 Nov. 2019. 36, 37

SINGH, K.; GAGNé, S.; MEENTEMEYER, R. 9.22 - urban forests and human well-being. In: LIANG, S. (Ed.). **Comprehensive remote sensing**. Oxford: Elsevier, 2018. p. 287–305. ISBN 978-0-12-803221-3. 19

SOTERRONI, A. C.; RAMOS, F. M.; MOSNIER, A.; FARGIONE, J.; ANDRADE, P. R.; BAUMGARTEN, L.; PIRKER, J.; OBERSTEINER, M.; KRAXNER, F.; CâMARA, G.; CARVALHO, A. X. Y.; POLASKY, S. Expanding the soy moratorium to Brazil's Cerrado. **Science Advances**, v. 5, n. 7, p. eaav7336, 2019. 2, 11

STRASSBURG, B. B. N.; BROOKS, T.; FELTRAN-BARBIERI, R.; IRIBARREM, A.; CROUZEILLES, R.; LOYOLA, R.; LATAWIEC, A. E.; FILHO, F. J. B. O.; SCARAMUZZA, C. A. D. M.; SCARANO, F. R.; SOARES-FILHO,

B.; BALMFORD, A. Moment of truth for the Cerrado hotspot. **Nature Ecology \& Evolution**, v. 1, p. 1–3, 2017. 13

TRIGG, S.; FLASSE, S. An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. **International Journal of Remote Sensing**, v. 22, n. 13, p. 2641–2647, 2001. 32, 34

UNITED STATES DEPARTMENT OF AGRICULTURE-(USDA). **Normalized Burn Ratio (NBR): comparison of the spectral response of healthy vegetation and burned areas**. 2022. Available from: <https://un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio>. Access on: 10 Oct. 2022. 33

URIARTE, M.; YACKULIC, C. B.; COOPER, T.; FLYNN, D.; CORTES, M.; CRK, T.; CULLMAN, G.; MCGINTY, M.; SIRCELY, J. Expansion of sugarcane production in São Paulo, Brazil: implications for fire occurrence and respiratory health. **Agriculture, Ecosystems and Environment**, v. 132, n. 1-2, p. 48–56, 2009. ISSN 01678809. 1

VAN DIJK, D.; SHOAIE, S.; VAN LEEUWEN, T.; VERAVERBEKE, S. Spectral signature analysis of false positive burned area detection from agricultural harvests using sentinel-2 data. **International Journal of Applied Earth Observation and Geoinformation**, v. 97, p. 102296, 2021. ISSN 1569-8432. Available from: <https://www.sciencedirect.com/science/article/pii/S0303243421000027>. Access on: 25 Set. 2022. 32

VERBESSELT, J.; HYNDMAN, R.; NEWNHAM, G.; CULVENOR, D. Detecting trend and seasonal changes in satellite image time series. **Remote sensing of Environment**, v. 114, n. 1, p. 106–115, 2010. 48

VIEIRA, R. M. d. S. P.; TOMASELLA, J.; BARBOSA, A. A.; POLIZEL, S. P.; OMETTO, J. P. H. B.; SANTOS, F. C.; FERREIRA, Y. da C.; TOLEDO, P. M. de. Land degradation mapping in the MATOPIBA region (Brazil) using remote sensing data and decision-tree analysis. **Science of The Total Environment**, v. 782, p. 146900, 2021. 80, 81

WEAVER, J. F.; LINDSEY, D.; BIKOS, D.; SCHMIDT, C. C.; PRINS, E. Fire detection using goes rapid scan imagery. **Weather and Forecasting**, v. 19, n. 3, p. 496–510, 2004. 23, 30

WERF, G. R. Van der; RANDERSON, J. T.; GIGLIO, L.; COLLATZ, G.; MU, M.; KASIBHATLA, P. S.; MORTON, D. C.; DEFRIES, R.; JIN, Y. v.; LEEUWEN, T. T. van. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). **Atmospheric Chemistry and Physics**, v. 10, n. 23, p. 11707–11735, 2010. 1

WHELAN, R. J. **The ecology of fire**. [S.l.]: Cambridge University Press, 1995. 13

WICKRAMASINGHE, C.; WALLACE, L.; REINKE, K.; JONES, S. Implementation of a new algorithm resulting in improvements in accuracy and resolution of seviri hotspot products. **Remote Sensing Letters**, v. 9, n. 9, p. 877–885, 2018. 83

WIRTH, R.; HIPP, J. Crisp-dm: Towards a standard process model for data mining. In: INTERNATIONAL CONFERENCE ON THE PRACTICAL APPLICATION OF KNOWLEDGE DISCOVERY AND DATA MINING, 4., 2000. **Proceedings...** Manchester: Practical Application Company, 2000. p. 29–39. 39, 40

WITTEN, I. H.; FRANK, E.; HALL, M. A.; PAL, C. J. **Data mining: practical machine learning tools and techniques**. [S.l.]: Morgan Kaufmann, 2016. 35

WOOSTER, M. J.; ROBERTS, G.; PERRY, G.; KAUFMAN, Y. Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release. **Journal of Geophysical Research: Atmospheres**, v. 110, n. D24, 2005. 24

WOOSTER, M. J.; ROBERTS, G. J.; GIGLIO, L.; ROY, D. P.; FREEBORN, P. H.; BOSCHETTI, L.; JUSTICE, C.; ICHOKU, C.; SCHROEDER, W.; DAVIES, D. et al. Satellite remote sensing of active fires: History and current status, applications and future requirements. **Remote Sensing of Environment**, v. 267, p. 112694, 2021. 2

WOŹNIAK, E.; ALEKSANSDROWICZ, S. Self-adjusting thresholding for burnt area detection based on optical images. **Remote Sensing**, v. 11, n. 22, p. 2669, 2019. 31

XU, W.; WOOSTER, M.; ROBERTS, G.; FREEBORN, P. New goes imager algorithms for cloud and active fire detection and fire radiative power assessment across North, South and Central America. **Remote Sensing of Environment**, v. 114, n. 9, p. 1876–1895, 2010. 24

YEUNG, A. K.; HALL, G. B. Spatial data mining and decision support systems. In: YEUNG, A. K. W.; HALL, G. B. (Ed.). **Spatial database systems: design, implementation and project management**. [S.l.]: Springer, 2007. p. 409–467. 38, 41

ZHAO, Y.; BAN, Y. Goes-r time series for early detection of wildfires with deep gru-network. **Remote Sensing**, v. 14, n. 17, p. 4347, 2022. 50

ZHUKOV, B.; LORENZ, E.; OERTEL, D.; WOOSTER, M.; ROBERTS, G. Spaceborne detection and characterization of fires during the bi-spectral infrared detection (bird) experimental small satellite mission (2001–2004). **Remote Sensing of Environment**, v. 100, n. 1, p. 29–51, 2006. 24

APPENDIX A - COMPLETE DETAILED PIXEL ANALYSIS

Below, we present the complete detailed pixel analysis (Table A.1). When there were more than one fire incidence in the same pixel along August, we divided the pixel analysis into sub-analysis (A, B and so on).

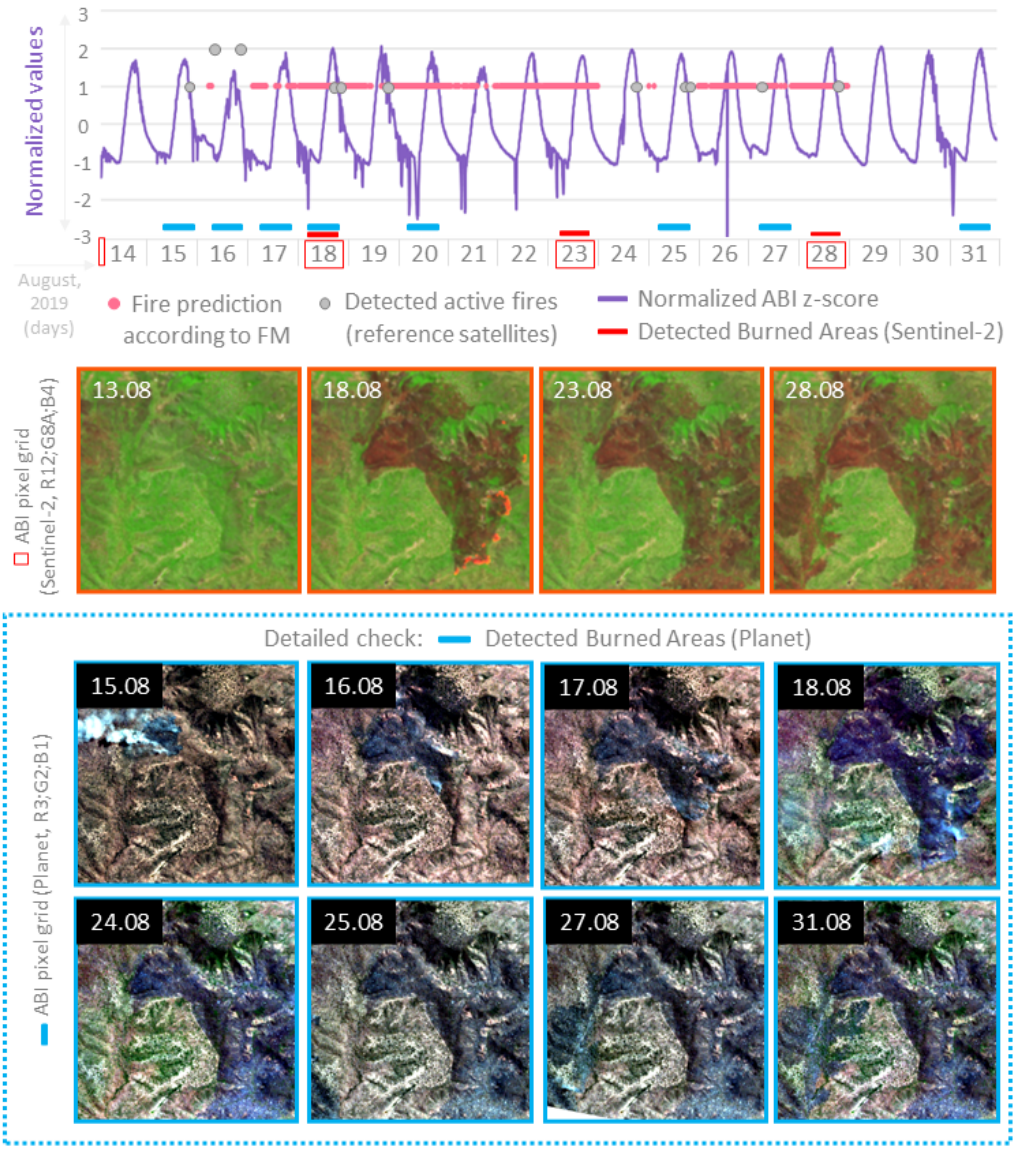
Table A.1 - Complete detailed pixel analysis. FM: Final Modell RefSat: Reference Satellites. NA: Not Applicable. d: day.

ID analysis	Fire firstly detected by	Fire duration (days)	Fire extent (km^2)	Expansion (km^2 d^{-1})	Days of burning (August)	Figure
09-A	RefSat	6	1.76	0.29	15-20	A.1
09-B	RefSat	4	0.48	0.12	24-28	A.1
10-A	FM	1	0.04	0.04	15	A.2
10-B	FM	1	0.07	0.07	21	A.2
10-C	FM	2	0.52	0.26	23-24	A.2
11	RefSat	5	2.04	0.41	24-28	A.3
12-A	FM	1	0.07	0.07	16	A.4
12-B	RefSat	1	0.04	0.04	21	A.4
12-C	FM	1	0.04	0.04	27	A.4
13	NA					A.5

For the analysis 09 (Figure A.1), FM also points out more AF than there really were (09-A), but also presented a high agreement with the Reference Satellites, specially in the analysis 09-B, which presented around 4 days of fire duration, 0.48 km^2 of fire extent and 0.12 km^2 d^{-1} of fire expansion.

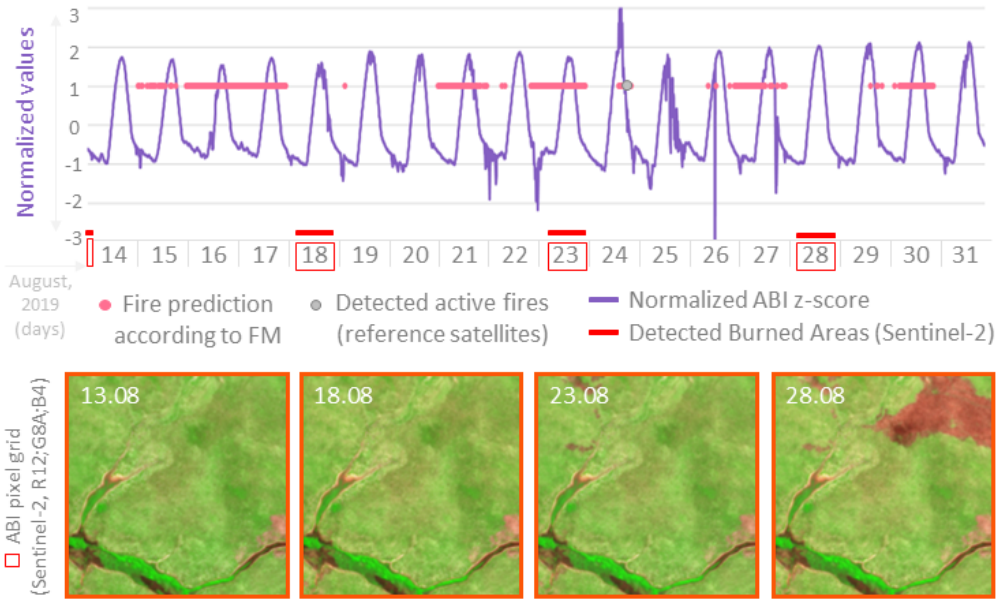
In the period of 15^{th} to 17^{th}, FM detected AF in the analysis 10-A (Figure A.2), yet only a small BA of 0.04 km^2 was detected on 18^{th}. The same also happens on 23^{th}, when a BA of 0.07 km^2 was also detected by FM (10-B). Finally, a more significant BA was found on 28^{th}, however, it is not exactly certain the fire duration, which we supposed that occurred from 23^{th} to 24^{th}. The analysis 11 (Figure A.3) presented a case, where FM wrongly detected AF from 15^{th} to 17^{th}, and only the Reference Satellites could correctly identify the AF. Once FM was mainly trained with smaller AF, it is a case where FM does not perform so well for fire with greater extent.

Figure A.1 - Detailed analysis ID 09



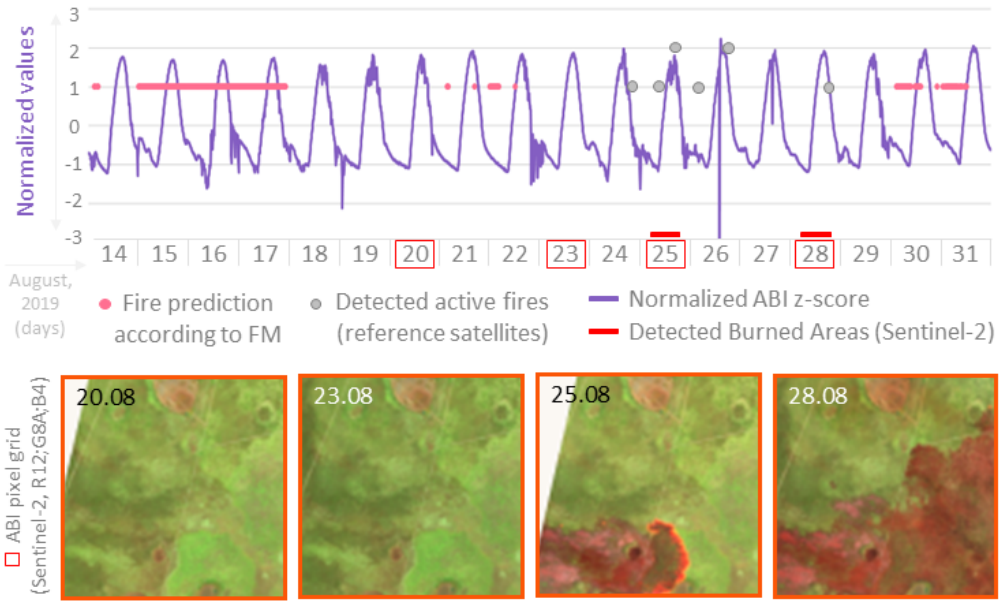
SOURCE: Author's own elaboration.

Figure A.2 - Detailed analysis ID 10.



SOURCE: Author's own elaboration.

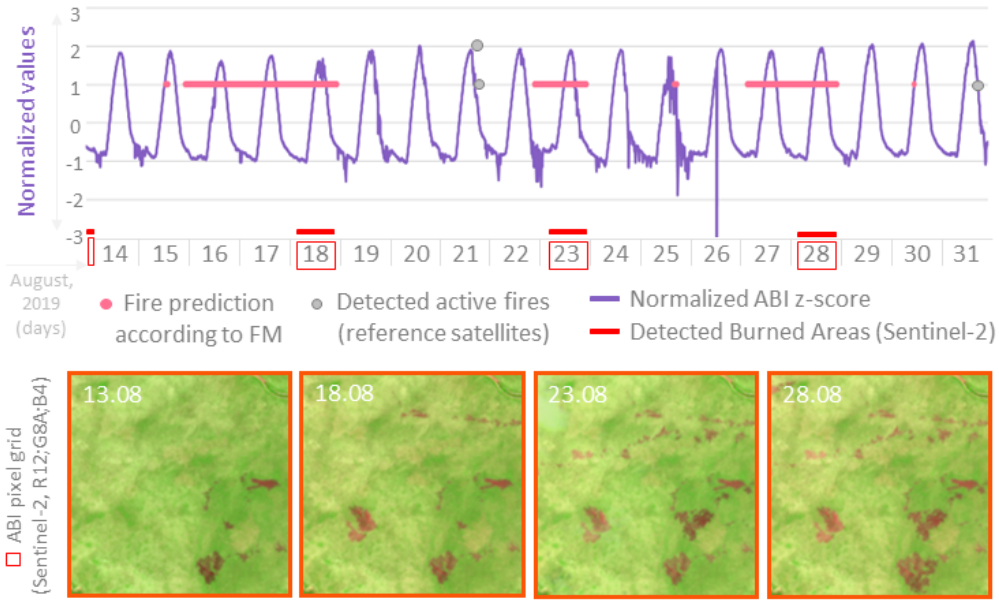
Figure A.3 - Detailed analysis ID 11.



SOURCE: Author's own elaboration.

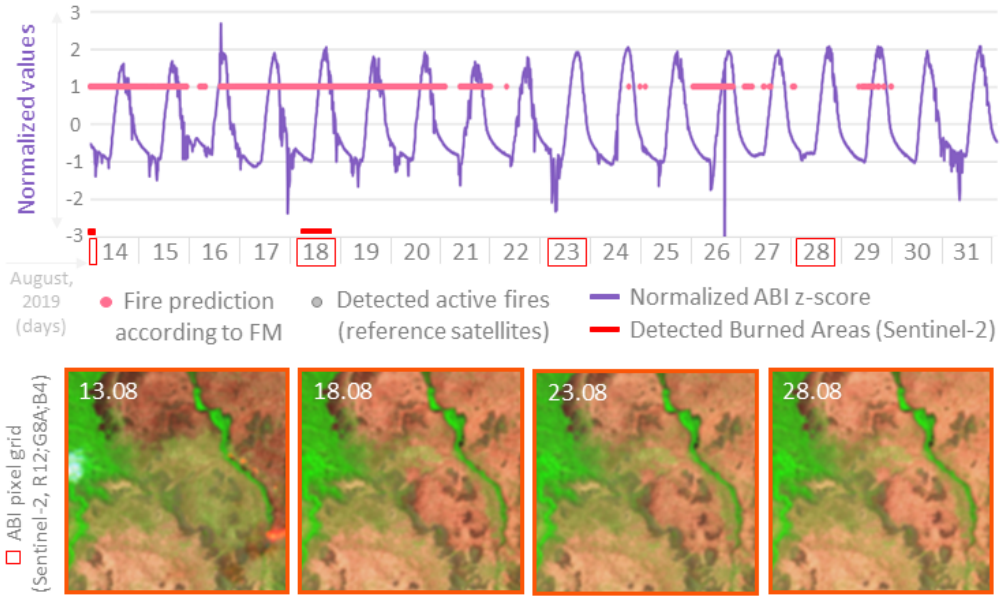
The analysis 12-A and 12-C (Figure A.4) present cases where only FM detected the small AF, 0.07 and 0.04 km^2, respectively. Whereas in the ID analysis 12-B, there is a disagreement between FM and the Reference Satellites. Because it is not possible to detect AF on 23^{th}, as pointed by FM, we considered that the Reference Satellites correctly indicated fire on 21^{th}. Finally, the analysis 13 shows an example of a detected AF on 13^{th} that was continuously indicated as fire by FM until 20^{th}.

Figure A.4 - Detailed analysis ID 12.



SOURCE: Author's own elaboration.

Figure A.5 - Detailed analysis ID 13.



SOURCE: Author's own elaboration.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.