

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

JOSÉ HENRIQUE DA SILVA BRAZ

**AQ-FA: Um método para cálculo da área  
queimada a partir da detecção de focos de  
incêndio no Brasil usando dados oriundos  
de múltiplos satélites**

Monografia apresentada como requisito parcial  
para a obtenção do grau de Bacharel em Ciência  
da Computação

Orientador: Prof. Dr. Lucas Mello Schnorr

Porto Alegre  
2023

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Prof<sup>a</sup>. Patricia Pranke

Pró-Reitora de Graduação: Prof<sup>a</sup>. Cíntia Inês Boll

Diretora do Instituto de Informática: Prof<sup>a</sup>. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Marcelo Walter

Bibliotecário-chefe do Instituto de Informática: Alexsander Borges Ribeiro

## **AGRADECIMENTOS**

## SUMÁRIO

<b>LISTA DE ABREVIATURAS E SIGLAS .....</b>	<b>5</b>
<b>LISTA DE FIGURAS .....</b>	<b>6</b>
<b>LISTA DE TABELAS .....</b>	<b>8</b>
<b>RESUMO .....</b>	<b>9</b>
<b>ABSTRACT .....</b>	<b>10</b>
<b>1 INTRODUÇÃO .....</b>	<b>11</b>
<b>2 CONCEITOS BÁSICOS E TRABALHOS RELACIONADOS .....</b>	<b>15</b>
<b>2.1 O monitoramento das queimadas no Brasil .....</b>	<b>15</b>
<b>2.2 Os Satélites e Sensores .....</b>	<b>17</b>
<b>2.3 Uma visão geral dos dados .....</b>	<b>22</b>
<b>2.4 Detecção de focos ativos e área queimada.....</b>	<b>27</b>
<b>2.5 Trabalhos Relacionados.....</b>	<b>34</b>
<b>3 METODOLOGIA .....</b>	<b>37</b>
<b>3.1 Visão geral da metodologia .....</b>	<b>37</b>
<b>3.2 Coleta e pré-processamento dos dados .....</b>	<b>39</b>
<b>3.3 Mapeamento dos focos para áreas.....</b>	<b>42</b>
<b>3.4 Separação e avaliação de quadrantes.....</b>	<b>43</b>
<b>3.5 Cálculo da área queimada.....</b>	<b>46</b>
<b>3.6 Implementação do método .....</b>	<b>48</b>
<b>4 RESULTADOS E DISCUSSÃO .....</b>	<b>51</b>
<b>4.1 Metodologia da avaliação .....</b>	<b>51</b>
<b>4.2 Validação do AQ-FA .....</b>	<b>55</b>
<b>4.3 Resultados apresentados .....</b>	<b>63</b>
<b>5 CONSIDERAÇÕES FINAIS .....</b>	<b>68</b>
<b>REFERÊNCIAS .....</b>	<b>71</b>

## LISTA DE ABREVIATURAS E SIGLAS

ABI	<i>Advanced Baseline Imager</i>
API	<i>Application Programming Interface</i> (Interface de Programação de Aplicação)
AQ-FA	Área Queimada por Focos Ativos
ATRS	<i>Along Track Scanning Radiometer</i>
AVHRR	<i>Advanced Very High Resolution Radiometer</i>
CSV	<i>Comma Separated Values</i> (valores separados por vírgulas)
EUMETSAT	<i>European Organisation for the Exploitation of Meteorological Satellites</i>
ESA	<i>European Space Agency</i>
GMT	<i>Greenwich Mean Time</i>
GOES	<i>Geostationary Operational Environmental Satellite</i>
INPE	Instituto Nacional de Pesquisas Espaciais
IBGE	Instituto Brasileiro de Geografia e Estatística
JAXA	<i>Japan Aerospace Exploration Agency</i>
URL	<i>Uniform Resource Locator</i> (Localizador Uniforme de Recursos)
NASA	<i>National Aeronautics and Space Administration</i>
NOAA	<i>National Oceanic and Atmosphere Administration</i>
MSG	<i>Meteosat Second Generation</i>
METOP	<i>Meteorological Operational Satellite</i>
MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i>
TRMM	<i>Tropical Rainfall Measuring Mission</i>
WRS	<i>Worldwide Reference System</i>
VIIRS	<i>Visible Infrared Imaging Radiometer Suite</i>

## LISTA DE FIGURAS

Figura 1.1 Comparação entre uma imagem do satélite TERRA (a direita) e os focos ativos (a esquerda), ambas no dia 8 de março de 2020 na região do Pantanal. ....	12
Figura 2.1 Órbita dos satélites no dia 10 de agosto de 2022.....	18
Figura 2.2 Órbitas-ponto Landsat que cobrem o território brasileiro com base no WRS-2.....	23
Figura 2.3 Detecções por satélite, sensor, bioma e região durante toda a série histórica	24
Figura 2.4 Amostragem por tempo de cada satélite.....	25
Figura 2.5 Somatório dos focos do satélite de referência agrupado por mês.....	25
Figura 2.6 Relação de focos por área geográfica do país por município (superior esquerdo); por estado (superior direito); por região (inferior esquerdo); e por bioma (inferior direito) .....	26
Figura 2.7 Variação da reflectância (%), no eixo y, em relação à frequência da luz eletromagnética emitida por uma floresta, no eixo x. A cor verde representa a reflectância de uma vegetação não queimada, a cor vermelha representa a reflectância do fogo ativo na parte inferior da floresta (no solo), a cor preta representa um incêndio na copa das árvores e a cor cinza representa a reflectância de uma vegetação queimada (carvão). ....	31
Figura 3.1 Diagrama do método. A entrada principal são os dados de focos de queimadas e a saída principal é a estimativa de áreas queimadas. Abaixo os resultados produzidos com a aplicação em um determinado local para cada etapa. A imagem abaixo representa a aplicação do método no sudoeste do Pará, durante os dias 1 a 3 de setembro de 2022. ....	38
Figura 3.2 Cores escolhidas para cada satélite.....	42
Figura 3.3 Comparação entre pontos e áreas dos focos detectados pelo satélite Suomi NPP. ....	43
Figura 3.4 Separação entre quadrantes. ....	44
Figura 3.5 Funções embutidas para cálculo de área.....	47
Figura 3.6 Exemplo da aplicação das funções embutidas. ....	47
Figura 3.7 Diferença da avaliação para quatro tamanhos de quadrantes. ....	50
Figura 4.1 Redução da resolução do produto AQ30m (a esquerda) para uma resolução de aproximadamente 330m usando diferentes limiares. A segunda imagem contém todos os quadrantes com valores maiores que 0. A terceira imagem mostra quadrantes que tem pelo menos metade de sua área coberta. A última imagem mostra apenas quadrantes totalmente cobertos pelo polígono original.....	54
Figura 4.2 Redução da resolução do produto AQ30m (a esquerda) para uma resolução de aproximadamente 330m sem perda significativa de área queimada. Cada quadrante na imagem da direita representa a porcentagem do quadrante que está classificado como queimado. ....	55
Figura 4.3 Box Plot representando seis distribuições, no eixo y: Acurácia Geral (OA), Coeficiente de Dice (DC), Erros de comissão (CE), Erros de Omissão (OE), Viés (B) e Diferença de área queimada. Cada distribuição foi gerada com as funções área queimada, no eixo x: linear(2.5, 15), exponencial(2.5, 15, 2), exponencial(2.5, 15, 3) e percentil(65, 85, 2.5, 15).....	58
Figura 4.4 Comparação entre a porcentagem queimada indicada pelo modelo AQ-FA (eixo y) e o AQ30m (eixo x) para quadrantes de tamanho 15 km × 15 km. ....	59

Figura 4.5 Aplicação do AQ-FA na órbita-ponto 221/70 no período de 27 de agosto a 10 de setembro de 2017. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA. Quanto mais vermelho escuro, maior o percentual de área queimada.....	61
Figura 4.6 Aplicação do AQ-FA na órbita-ponto 221/67 no período de 3 de julho a 19 de julho de 2021. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA.....	62
Figura 4.7 Aplicação do AQ-FA na órbita-ponto 221/74 no período de 30 de julho a 15 de agosto de 2016. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA.....	63
Figura 4.8 Comparação mensal do total de área queimada, representado no eixo y, a partir dos resultados do AQ-FA e do AQ1km ao longo dos anos de 2008 a 2022, apresentado no eixo x, abrangendo o território brasileiro.....	64
Figura 4.9 Variação da Área Queimada por Órbita-Ponto de 2008 a 2022. A escala de cores está em $\text{km}^2$ , com tons de azul indicando uma diminuição da área queimada em relação ao AQ1km e tons de vermelho indicando um aumento. ....	67

## **LISTA DE TABELAS**

Tabela 2.1 Significado de cada coluna dos dados de queimada do INPE .....	17
Tabela 2.2 Características dos satélites usados pelo INPE. ....	19
Tabela 2.3 Nível de maturidade dos produtos .....	28
Tabela 2.4 Nível de processamento dos dados dos produtos. ....	28
Tabela 2.5 Versões resumidas do Collection para o sensor MODIS.....	29
Tabela 2.6 Resumo dos produtos de área queimada. ....	31
Tabela 4.1 Tabela de Contingência. ....	51
Tabela 4.2 Regiões e período de tempo da validação.....	60

## RESUMO

O Brasil é o quarto maior emissor de gases de efeito estufa por habitantes segundo o ranking da UNEP de 2022, muito devido às queimadas e ao desmatamento da vegetação. Este trabalho desenvolve uma metodologia capaz de estimar a área queimada a partir de focos ativos oriundos do Programa Queimadas do INPE. A metodologia consiste em transformar focos, que são pontos no espaço, em áreas de cobertura do incêndio, de acordo com a resolução do satélite que o detectou. O espaço é dividido em quadrantes e cada quadrante é avaliado de acordo com equações que utilizam as áreas de cobertura dos focos. Por último, cada avaliação de quadrante passa por uma função de normalização, resultando na porcentagem de área queimada estimada para o quadrante. O processo apresentado recebe o nome de AQ-FA, sigla para Área Queimada por Focos Ativos. O método é parametrizável, permitindo alterar a resolução dos resultados, que neste estudo é definida como 300m. A avaliação utiliza o produto AQ30m (INPE) como referência, aplicado em 20 órbitas-ponto e diferentes períodos de tempo, apresentando altos níveis de acurácia. Os resultados são obtidos a partir da aplicação mensal do método para todo o território nacional, entre os anos de 2008 até 2022. Ao comparar os resultados com o AQ1km (INPE), há um indicativo de aumento na área queimada no Brasil, contrariando o conhecimento prévio. A diferença entre os dois produtos fica mais evidente no bioma amazônico, que apresenta mais áreas queimadas. Toda a implementação e validação é desenvolvida utilizando a linguagem Python e Jupyter Notebook, para fins de reproduzibilidade. Todos os dados e o código estão disponíveis de forma pública e em conformidade com a ciência aberta.

**Palavras-chave:** Área queimada. Análise de dados. Fogo. INPE. Ciência aberta.

# **AQ-FA: A method for calculating the burned area from the detection of fire outbreaks in Brazil using data from multiple satellites**

## **ABSTRACT**

Brazil is the fourth largest emitter of greenhouse gases per inhabitant according to the 2022 UNEP ranking, largely due to forest fires and deforestation. This document develops a methodology capable of estimating the burned area from active fires arising from INPE's Programa Queimadas. The methodology consists of transforming the active fires, which are points in space, into fire coverage areas, according to the resolution of the satellite that detected it. The space is divided into quadrants and each quadrant is evaluated according to equations that use the coverage areas of the foci. Finally, each quadrant evaluation goes through a normalization function, resulting in the estimated percentage of burned area for the quadrant. The process presented is called AQ-FA, an acronym for Area Burned by Active Fire (in Portuguese Área Queimada por Focos Ativos). The method is parameterizable, allowing to change the resolution of the results, which in this study is defined as 300m. The validation uses the AQ30m product (INPE) as a reference, applied in 20 path-rows and different time periods, presenting high levels of accuracy. The results are obtained from the monthly application of the method for the entire national territory, between the years 2008 to 2022. When comparing the results with the AQ1km (INPE), there is an indication of an increase in the burned area in Brazil, contrary to knowledge prior. The difference between the two products is more evident in the Amazon biome, which has more burned areas. All implementation and validation are developed using the Python language and Jupyter Notebook, for reproducibility purposes. All data and code are publicly available and open science compliant.

**Keywords:** Burned area. Data analysis. Fire. INPE. Open science.

## 1 INTRODUÇÃO

O fogo é um elemento que está presente há milênios no território que hoje é o Brasil, desde queimadas controladas pelo povo indígena Kayapó no cerrado para plantio ou caça, até incêndios iniciados por combustão espontânea em períodos de seca no sul da Amazônia. O uso do fogo pelos indígenas era controlado, levando em conta o clima atual e a vegetação a ser queimada, e restrito apenas a um período do ano, com o intuito de reduzir pragas e ajudar nas plantações (LEONEL, 2000).

Hoje as queimadas que mais chamam atenção estão diretamente ligadas ao processo de desmatamento e manejo de áreas agrícolas para o cultivo da monocultura de soja. O fogo também é a prática mais barata e rápida para limpar áreas inteiras para a pecuária bovina. Commodities agrícolas e carne bovina movem a economia do Brasil, que é o maior exportador desses produtos, e aumentam a pressão para o desmatamento de novas áreas na Amazônia (FUCHS, 2020).

O Brasil ocupa a quarta posição no ranking de nações que mais emitem gases de efeito estufa por habitantes, segundo dados da *United Nations Environment Programme UNEP* (2022). De acordo com o estudo, o valor absoluto se manteve estável desde 2010, e atingiu seu pico entre os anos de 2003 e 2004. O que melhor explica a alta posição do Brasil são as queimadas e o desmatamento da vegetação nativa. Entre os dez municípios que mais poluem no Brasil, oito deles estão localizados no bioma da Amazônia.

Nesse sentido, o Instituto Nacional de Pesquisas Espaciais (INPE) mantém o Programa Queimadas, que é um conjunto de ferramentas que envolve a detecção em tempo real dos focos ativos de queimadas. Esse monitoramento se dá a partir de processamento de imagens de satélites, utilizando metodologias estabelecidas capazes de identificar regiões com alta probabilidade de estarem queimando. A Figura 1.1 ilustra uma comparação entre uma imagem capturada pelo satélite TERRA, que foi corrigida em termos de cores, e os focos ativos detectados na mesma região e período de tempo pelo INPE. Os focos são representados por pontos coloridos de acordo com o satélite que possibilitou a detecção do foco. Na imagem à direita, é possível identificar pelo menos dois agrupamentos de queimadas, que também são destacados, por vários satélites, na imagem à esquerda.

Os focos ativos de queimada dão uma ideia do que está acontecendo no presente momento, mas não são uma métrica ideal para estimar o impacto ambiental do fogo na vegetação. Para dar uma visão do impacto real, existem métodos para calcular a área efetivamente queimada a partir do processamento de imagens de satélites, chamados de

Figura 1.1 – Comparação entre uma imagem do satélite TERRA (a direita) e os focos ativos (a esquerda), ambas no dia 8 de março de 2020 na região do Pantanal.



Fonte: O Autor, agregando imagem publicada pela NASA e dados de focos ativos do INPE

“produtos”. O processamento mais tradicional envolve a análise de índices de vegetação que tentam identificar regiões de vegetação queimada (PEREIRA, 1999). Novas versões desses produtos utilizam uma abordagem híbrida, integrando também os focos ativos para identificar regiões de atividade do fogo e com isso apresentam no geral menores taxas de erros (BOSCHETTI et al., 2019).

Além da existência de produtos de áreas queimadas de aplicação global, como o MCD64A1 (GIGLIO et al., 2018) e FireCCI51 (LIZUNDIA-LOIOLA et al., 2020), o INPE mantém dois produtos, nomeadamente AQ30m (MELCHIORI et al., 2014) e AQ1km (LIBONATI et al., 2015), para o cálculo da área queimada, especialmente dedicadas aos biomas do Brasil. Enquanto o AQ30m é um produto de alta resolução (30m) e taxas de erros pequenas, o AQ1km tem precisão de 1km e erros próximos a 30% em relação ao AQ30m (INPE, 2023b). Ambos produtos se completam, uma vez que o primeiro é calculado apenas para algumas regiões do Brasil por imagens a cada 16 dias. O segundo gera dados mensais sobre áreas queimadas de todo o país a partir de imagens de satélites diárias diárias.

O tema do cálculo de área queimada ainda está muito ativo atualmente devido à complexidade e importância do assunto. Muitas metodologias foram criadas e aprimoradas utilizando dados de sensoriamento remoto. Para validá-las, vários trabalhos foram desenvolvidos através da comparação do produto com uma referência com baixos níveis de erro (BOSCHETTI; STEHMAN; ROY, 2016; BOSCHETTI et al., 2019). Nesse sen-

tido, vários trabalhos demonstram que há muita variação entre os resultados dos produtos globais envolvendo as queimadas (BOSCHETTI et al., 2004; HUMBER et al., 2019). Com isso, não há um método definitivo e livre de erros para cálculo da área queimada.

Este trabalho tem como objetivo principal desenvolver uma metodologia para o cálculo de área queimada aplicada a todo o território nacional, utilizando focos ativos de incêndio. Adicionalmente, este estudo visa avaliar e analisar os resultados desse método a fim de identificar possíveis discrepâncias comparando-se com produtos já existentes. Nesse contexto, pretende-se também responder se a área queimada anualmente está em tendência de queda ou alta e identificar quais biomas e regiões são mais impactadas pelo fogo.

A análise acerca dos dados de focos ativos de queimadas no contexto do cálculo de áreas queimada se justifica por meio de estudos desenvolvidos. Segundo Giglio et al. (2010), a detecção de áreas queimadas pode se beneficiar da fusão de observações de incêndios ativos de vários sensores. Isso é respaldado pelo fato de que focos ativos apresentam maior contraste com a vegetação e por isso podem ser detectados com maior precisão (GIGLIO; SCHROEDER; JUSTICE, 2016). Por outro lado, devido ao intervalo entre as passagens dos satélites polares, detectam apenas uma fração das queimadas totais (HANTSON et al., 2013; GIGLIO et al., 2009). Desta forma, agregar dados de vários satélites, incluindo satélites geoestacionários, tende a mitigar esse problema.

A fim de atingir estes objetivos, a metodologia criada consiste em separar o território em pequenos quadrantes e avaliá-los utilizando algumas regras que priorizem a diversidade de satélites e a quantidade de focos dentro do quadrante. Após isso, os valores são normalizados em um intervalo de 0 a 1, representando a porcentagem da área do quadrante é classificada como queimada. A metodologia, inspirado nos nomes dos produtos do INPE, é chamada de AQ-FA (Área Queimada a partir de Focos Ativos).

Os dados analisados neste trabalho foram obtidos a partir do DBQueimadas, Banco de Dados de Queimadas <[www.inpe.br/queimadas/bdqueimadas](http://www.inpe.br/queimadas/bdqueimadas)>, que é uma ferramenta que faz parte do Programa Queimadas desenvolvido pelo INPE e acessível de forma aberta por meio da web. O sistema conta com mais de 300 milhões de pontos, provenientes de vários satélites (SETZER; MORELLI; SOUZA, 2019). Apesar de ter dados desde 1998, para esse trabalho o escopo de tempo dos resultados é limitado ao início de 2008 até o final de 2022. Este período de 15 anos joga luz aos dados mais recentes e não torna o cálculo dos resultados um processo muito demorado.

O trabalho pretende estimular mais estudos na área, deixando um arcabouço técnico e metodológico capaz de facilitar a análise desse tipo de dados. O desenvolvimento de todo o trabalho segue os princípios da reproduzibilidade dos resultados, por meio de cadernos de anotação (Jupyter Notebooks). Assim como os dados do Programa Queimadas, todos os dados e códigos desenvolvidos neste trabalho podem ser acessados de forma pública a partir do repositório Git hospedado em <<https://github.com/josebraz/INPE-Queimadas>>, seguindo os princípios da ciência aberta e reproduzível (LEGRAND, 2019).

Este documento estrutura os capítulos da seguinte forma. O Capítulo 2 apresenta uma revisão sobre os conceitos básicos envolvendo os dados de focos ativos, os satélites e seus sensores. Apresenta também metodologias consolidadas para a detecção de focos ativos e o cálculo de área queimada a partir de imagens de satélites, bem como alguns trabalhos inovadores no assunto. O Capítulo 3 explica com detalhes a metodologia proposta para o cálculo de área queimada, bem como questões da implementação propriamente dita e da coleta dos dados utilizados. O Capítulo 4 apresenta a validação do AQ-FA e a análise dos resultados obtidos aplicando em todo território nacional. O Capítulo 5 conclui o documento, propondo algumas possíveis melhorias e trabalhos futuros.

## 2 CONCEITOS BÁSICOS E TRABALHOS RELACIONADOS

Neste Capítulo, serão apresentados conceitos importantes para a compreensão de todo trabalho e dos resultados obtidos com o novo método proposto. Na primeira Seção, são formalizadas definições relacionadas às queimadas e à forma como são monitoradas no Brasil. Na Seção seguinte, é realizada uma sumarização dos principais satélites e sensores utilizados pelo INPE e suas características. Após isso, é feito uma análise dos dados dos focos ativos do INPE para demonstrar uma visão geral das principais tendências e gráficos. Em seguida, é abordado como as imagens geradas pelos satélites são utilizadas para a detecção de focos ativos e cálculo da área queimada. Por fim, se discute alguns trabalhos relacionados com o cálculo de área queimada em contraste com o presente trabalho.

### 2.1 O monitoramento das queimadas no Brasil

Um foco de queima, também conhecido como “Fire Pixel”, é a detecção de fogo na vegetação por uma imagem de satélite. A área de abrangência dessa detecção é do tamanho de um pixel da imagem, por isso está relacionada diretamente com a resolução do sensor do satélite. Por exemplo, se um sensor tem resolução de aproximadamente 1 km, a área coberta pelo pixel equivale a 1 km no nadir (no equador) (INPE, 2023b). A resolução do sensor, na verdade, representa uma estimativa do tamanho dos pixels, pois eles podem sofrer distorções relacionadas a inclinação e distância do satélite em relação ao ponto medido.

É importante destacar que o foco de queima por si só não representa de forma precisa o que está acontecendo na região, mas é um indicador valioso da intensidade e extensão do fogo. Poucos focos ou uma área queimada pequena não necessariamente refletem a intensidade da degradação da vegetação e o impacto ambiental. Áreas de floresta densa, mesmo que pequenas, podem ser gravemente afetadas, resultando em perdas significativas de fauna e flora, incluindo espécies exóticas e em risco de extinção.

Um incêndio, diferente de um foco de queima, pode durar dias e queimar uma grande extensão de terra, o que provavelmente resulta na detecção de vários focos de queima. O número de focos de queima está diretamente relacionado à extensão total queimada e pode ser utilizado para comparações e análises (GIGLIO et al., 2010). A área queimada, por outro lado, representa o resultado deixado pelo incêndio, que também pode ser estimada a partir de imagens de satélites. Existem vários métodos para detectar focos

ativos e calcular a área queimada a partir dos dados brutos de satélites, alguns deles são abordados no decorrer do documento.

Para monitorar e proteger o meio ambiente do Brasil das queimadas, o INPE desenvolve o programa de Monitoramento de Queimadas e Incêndios Florestais (<<https://queimadas.dgi.inpe.br/queimadas/portal>>). O programa é um conjunto de ferramentas abertas, muitas delas acessíveis via Web, desenvolvidas pelo time de tecnologia do INPE que possibilitam obter e visualizar dados sobre os focos de queimadas, risco de fogo, área queimada, focos em áreas de preservação e propriedades rurais, entre outras.

O programa também conta com o Banco de Dados de Queimadas (DBQueimadas), um site capaz de gerar mapas, tabelas, gráficos e exportar os dados sobre as queimadas no Brasil aplicando diferentes filtros (SETZER; MORELLI; SOUZA, 2019). O DBQueimadas é um excelente caso de como os dados abertos podem ajudar a sociedade, pois qualquer pessoa pode ter acesso aos dados e desenvolver trabalhos científicos, fomentando a ciência e melhorando o monitoramento. É a partir do DBQueimadas que este trabalho obteve boa parte dos dados necessários para o desenvolvimento.

Além do DBQueimadas, o INPE também disponibiliza para visualização e download, por meio da Divisão de Geração de Imagens (DGI) <[www.dgi.inpe.br/catalogo/](http://www.dgi.inpe.br/catalogo/)>, algumas imagens inteiras geradas pelos satélites que o próprio DGI captura e processa. Para obter essas imagens brutas são necessárias antenas especiais que ficam em centros de recepção de dados. Com esse propósito, a DGI possui duas Estações de Recepção e Gravação (ERG) - a primeira em Cachoeira Paulista (SP) e uma mais recente em Cuiabá (MT). Na estação de SP, é feito o processamento de mais de 200 imagens de diversos satélites todos os dias, extraíndo os dados de focos ativos de queimadas que alimentam o DBQueimadas (Divisão de Geração de Imagens, 2023).

É em posse dessas imagens que o INPE aplica algoritmos de detecção de focos de queima, abordado na Seção 2.4. No caso da detecção ser positiva, a posição exata (latitude e longitude) e a hora que a imagem foi gerada são adicionadas aos dados como uma nova linha e disponibilizados pelo DBQueimadas. O INPE ainda coloca junto com as coordenadas da detecção mais alguns dados como risco de fogo, poder do fogo, precipitação e dados referentes a região do foco. A lista completa da estrutura dos dados pode ser vista na Tabela 2.1, nela a coluna Atributo indica o nome de cada coluna nos dados, seguido do tipo e uma breve descrição (INPE, 2023b).

Tabela 2.1 – Significado de cada coluna dos dados de queimada do INPE.

Atributo	Tipo	Descrição
Id	string	Identificador único registrado no banco
Latitude	double	Graus decimais da latitude do centro do pixel de fogo ativo (valores de 90.0000 até -90.0000)
Longitude	double	Graus decimais da longitude do centro do pixel de fogo ativo (valores de 180.0000 até -180.0000)
DataHora	string	Data a hora da passagem do satélite no fuso horário de Greenwich (GMT)
Municipio	string	Nome do município, de acordo com os dados do IBGE 2000
Estado	string	Nome do estado
País	string	Nome do país
Bioma	string	Nome do bioma brasileiro, de acordo com dados do IBGE 2004 (para outros países o campo fica vazio)
Precipitação	double	Valor a precipitação do dia até o horário da medida (-999 para valores inválidos)
DiasSCh	integer	Dias sem chuva até a data da medida (-999 para valores inválidos)
RiscoFog	double	Valor do risco de fogo previsto naquele dia (-999 para valores inválidos)
FRP	double	Fire Radiative Power, MW (megawatts)

Fonte: INPE (2023b)

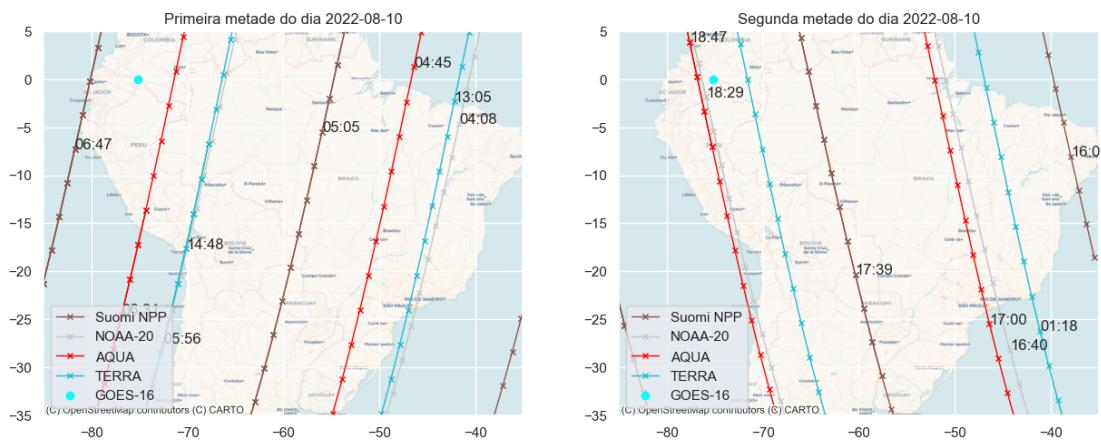
## 2.2 Os Satélites e Sensores

O INPE processa atualmente dados de diversos satélites, cada um com características distintas. Esses satélites abrangem uma ampla gama de órbitas e altitudes, desde satélites geoestacionários, como o GOES-12, localizado a 29.400 km de distância da superfície (SETZER; YOSHIDA, 2004), até satélites em órbitas polares, que ficam entre 700 e 900 km de altura. Cada um deles é equipado com diferentes sensores que são utilizados para diversos propósitos, como observação da vegetação terrestre, nuvens, oceanos e monitoramento atmosférico.

Cada satélite pode estar equipado com um sensor imageador que captura imagens com características distintas. Esses sensores são capazes de captar não apenas a luz visível, no comprimento de onda de 400 nm a 700 nm, mas também o infravermelho, de 780 nm a 1 mm. As medições realizadas pelos sensores são divididas em canais, que diferem em termos de resolução espacial - a escala geográfica representada por cada pixel da imagem - e resolução espectral, ou seja, o intervalo de comprimento de onda abrangido.

Os satélites também podem ser classificados de acordo com o tipo de órbita, que pode ser polar ou geoestacionária. Os satélites polares percorrem toda a terra, com uma altitude baixa e alta velocidade, passando várias vezes pelos polos norte e sul em um único dia. Já os satélites geoestacionários ajustam sua órbita para sempre parecerem estacionários em relação a um ponto fixo na Terra. Esses satélites ficam em altitudes mais altas devido à sua velocidade mais baixa, que é igual à rotação da Terra. A Figura 2.1, gerada a partir de dados do site Celestrak (<https://celestrak.org/>), mostra a passagem de alguns satélites polares no Brasil no dia 10 de agosto de 2022. Também é possível ver o satélite geoestacionário GOES-16, localizado ao norte do perú, representado com um ponto no gráfico.

Figura 2.1 – Órbita dos satélites no dia 10 de agosto de 2022.



Fonte: O Autor

A Tabela 2.2 contém um resumo dos satélites utilizados pelo INPE desde o início da série histórica até o final de 2022 (Embrapa Territorial, 2023). Na Tabela, são apresentados o nome do satélite, o nome do sensor imageador, a resolução espacial do sensor, o tipo de órbita do satélite, o ano em que o satélite entrou em operação e, por último, os dois horários aproximados da passagem de um satélite polar no território do Brasil. Os horários exatos dessa passagem variam de acordo com a excentricidade da órbita do satélite o que reflete no chamado “Periodo de revisita”, que indica quantos dias o satélite leva para repetir a mesma trajetória.

Atualmente, os satélites que estão em pleno funcionamento são: NOAA-20, NOAA-19, NOAA-18, GOES-16, Suomi NPP, AQUA, TERRA, MSG-03, METOP-B e METOP-C. Os demais satélites, como NOAA-17, NOAA-16, NOAA-15, TRMM, NOAA-14,

Tabela 2.2 – Características dos satélites usados pelo INPE.

<b>Nome</b>	<b>Sensor</b>	<b>Res. esp.</b>	<b>Órbita</b>	<b>Lançado</b>	<b>Passagem no Brasil</b>
NOAA-12	AVHRR	1100m	Polar	1991	2h / 15h
ERS-1	ATSR-1	1000m	Polar	1991	Variadas
NOAA-14	AVHRR	1100m	Polar	1994	9h / 21h
ERS-2	ATSR-2	1000m	Polar	1995	Variadas
TRMM	VIRS	2000m	Polar	1997	Variadas
NOAA-15	AVHRR-3	1100m	Polar	1998	5h / 17h
TERRA	MODIS	1000m	Polar	1999	11h / 23h
NOAA-16	AVHRR-3	1100m	Polar	2000	9h / 21h
AQUA	MODIS	1000m	Polar	2002	2h / 14h
NOAA-17	AVHRR-3	1100m	Polar	2002	9h / 21h
NOAA-18	AVHRR-3	1100m	Polar	2005	9h / 21h
NOAA-19	AVHRR-3	1100m	Polar	2009	2h / 14h
Suomi NPP	VIIRS	500m	Polar	2011	2h / 14h
METOP-B	AVHRR-3	1100m	Polar	2012	9h / 21h
NOAA-20	VIIRS	500m	Polar	2017	2h / 14h
METOP-C	AVHRR-3	1100m	Polar	2018	9h / 21h
GOES-08	GOES I-M	4000m	Geoest.	1994	Não se aplica
GOES-10	GOES I-M	4000m	Geoest.	1997	Não se aplica
GOES-12	GOES I-M	4000m	Geoest.	2001	Não se aplica
MSG-02	SEVIRI	3000m	Geoest.	2005	Não se aplica
GOES-13	GOES I-M	4000m	Geoest.	2006	Não se aplica
MSG-03	SEVIRI	3000m	Geoest.	2012	Não se aplica
GOES-16	ABI	2000m	Geoest.	2016	Não se aplica

Fonte: O Autor com base em Embrapa Territorial (2023)

NOAA-12, GOES-13, MSG-02, GOES-12, GOES-10 e GOES-08, deixaram de operar em diferentes momentos devido a problemas técnicos ou ao fim de sua vida útil.

Os satélites TERRA e AQUA, ambos com o sensor Moderate Resolution Imaging Spectroradiometer (MODIS), foram lançados em 1999 e 2002 respectivamente. São de origem norte americana e foram desenvolvidos em parceria com o Japão, Canadá e Brasil. O sensor MODIS possui 36 canais e resolução espectral que varia de 250 m a 1 km em diferentes espectros do infravermelho e luz visível. É especialmente capaz de detectar mudanças no uso e cobertura da terra bem como queimadas e atividades vulcânicas (LATORRE et al., 2003).

Os satélites da série National Oceanic and Atmospheric Administration (NOAA) são operados pela agência americana de mesmo nome. Embora tenha havido vários satélites NOAA ao longo do tempo, atualmente estão em operação o NOAA-20, NOAA-19 e NOAA-18. Esses satélites desempenham um papel fundamental no fornecimento de dados para a previsão do tempo e o monitoramento da vegetação.

Além dos satélites NOAA, em 2011, a NOAA em parceria com a National Aeronautics and Space Administration (NASA) lançou o satélite Suomi NPP (Suomi National Polar-orbiting Partnership). O Suomi NPP tem como objetivo principal obter observações ambientais e meteorológicas avançadas da Terra. É equipado com o sensor Visible Infrared Imaging Radiometer Suite (VIIRS), que também está presente no NOAA-20. Esses dois satélites, Suomi NPP e NOAA-20, são os satélites com mais alta resolução espacial entre todos os processados pelo INPE, equivalente a 250m.

O conjunto de satélites Geostationary Operational Environmental Satellite (GOES) é operado pela NASA e controlado também pelo NOAA. As primeiras versões desses satélites geoestacionários foram equipadas com o sensor GOES I-M (Imager Radiometer e Vertical Sounder) de resolução espacial equivalente a 4 km. A partir do GOES-16 o sensor foi substituído pelo Advanced Baseline Imager (ABI), capaz de produzir imagens com resolução 4 vezes melhores. Esse conjunto de satélites foi e ainda é muito usado nas previsões metereológicas nos países do continente americano.

O INPE processa as imagens dos satélites da série Meteorological Operational Satellite (METOP), incluindo o METOP-B e o METOP-C, bem como os satélites da série Meteosat Second Generation (MSG), como o MSG-02 (Meteosat-9) e o MSG-03 (Meteosat-10). Ambas missões projetadas em parceria com a European Space Agency (ESA) e a European Organisation for the Exploitation of Meteorological Satellites (EU-METSAT). Esses satélites têm como objetivo a previsão do tempo, o monitoramento

climático e o acompanhamento de desastres naturais. O METOP-C é o último satélite planejado para a série METOP, completando assim o conjunto de satélites METOP-A, METOP-B e METOP-C.

O instituto também processou dados dos satélites ERS-1 e ERS-2, equipados com o sensor Along Track Scanning Radiometer (ATRS) da ESA (Embrapa Territorial, 2023). Esses satélites, de órbita polar, são especializados em medir com precisão a temperatura da terra e dos oceanos. Os dados coletados por esses satélites são utilizados por cientistas para detectar mudanças climáticas e vegetação, incluindo eventos como queimadas, em todo o planeta. O ATRS possui duas versões, sendo que o primeiro (ATSR-1) possui 4 canais e o segundo (ATSR-2) foi aprimorado com um canal adicional para capturar a luz visível, possibilitando o monitoramento da vegetação.

Também estão presentes nos dados, em menor escala, os focos de queimadas obtidos pelo satélite Tropical Rainfall Measuring Mission (TRMM). O TRMM foi uma missão conjunta entre a NASA e a Japan Aerospace Exploration Agency (JAXA), que teve início em 1997 e oficialmente encerrou em 2015. O principal objetivo dessa missão foi estudar a distribuição de chuvas e tempestades, bem como suas influências no clima global.

## O satélite de referência

Para estabelecer uma série temporal consistente e permitir a análise de tendências ao longo de vários anos dos focos de queimadas detectados em diferentes regiões, o INPE definiu um satélite de referência. Esse satélite de referência é escolhido com base em critérios específicos. É importante que sua órbita cubra satisfatoriamente a área do país, sem distorcer os dados de forma significativa. Além disso, é desejável que os sensores do satélite tenham resoluções adequadas, ou seja, não muito baixas, para garantir uma análise precisa dos focos de queimadas. (INPE, 2023b)

No período de 01 de junho de 1998 a 03 de julho de 2002, o satélite de referência utilizado foi o NOAA-12, com passagem no final da tarde. Após esse período, passou-se a utilizar o satélite AQUA, com passagem à tarde (identificado nos dados como AQUA\_M-T). O AQUA tem operado além da data prevista de encerramento e provavelmente deve ser descontinuado em breve. Quando isso ocorrer, a previsão é que o satélite Suomi NPP deve ser o novo satélite de referência (INPE, 2023b). Sempre que há uma mudança de

satélite de referência, a série histórica de quantidade de focos de queimada precisa ser ajustada devido a diferença entre os satélites.

### **Satélites de média resolução**

São considerados satélites de média resolução aqueles que possuem um sensor capaz de produzir imagens em que um pixel equivale de 10 a 60 metros (Embrapa Territorial, 2023). A principal missão que emprega esses satélites é a Landsat que lançou 9 satélites e teve início em 1972. Atualmente, apenas os satélites Landsat-8 e Landsat-9 estão ativos. Esses satélites tem período de revisita de 16 dias, ou seja, a cada 16 dias produzem dados da mesma região.

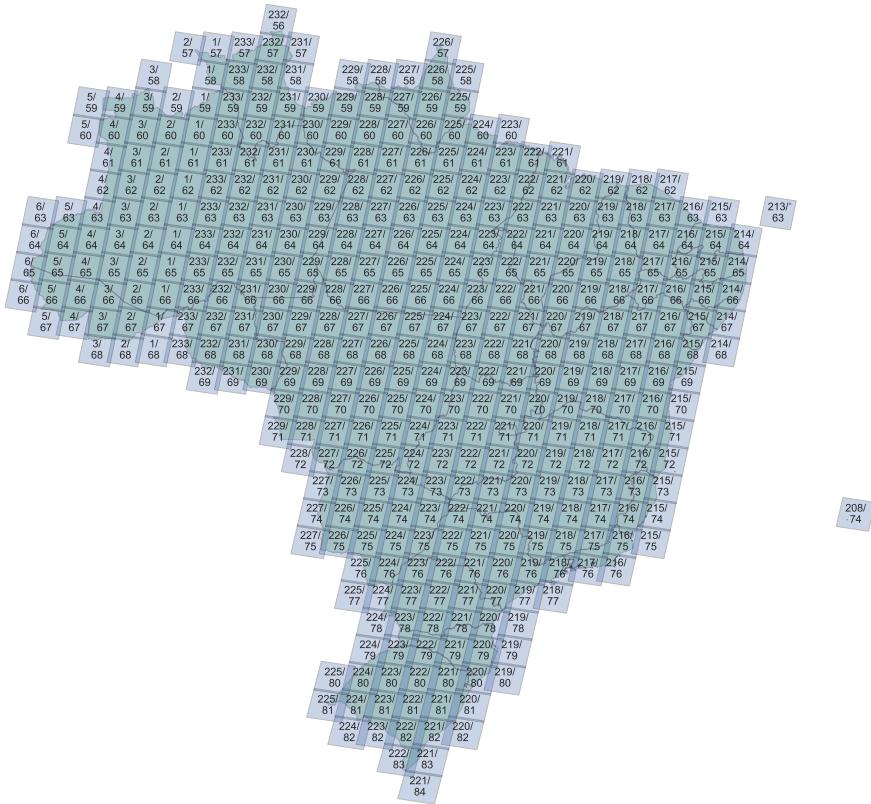
A missão Lansat estabeleceu um sistema de órbita-ponto que divide o planeta em vários quadrantes, chamo de *Worldwide Reference System* (WRS). A partir do Landsat-4 o WRS ganhou uma versão 2 (WRS-2) para se adequar ao fato de que os satélites passaram de 18 para 16 dias de tempo de revisita. O Brasil é coberto por 390 quadrantes, como na Figura 2.2, sendo marcados com o número da órbita e sua linha correspondente, chamado no inglês de *path/row*.

### **2.3 Uma visão geral dos dados**

Nesta seção, é realizada uma análise inicial dos dados disponíveis sobre focos de queimada, com o objetivo de extrair algumas informações relevantes. É importante, no entanto, ter cuidado na escolha dos satélites a serem utilizados na análise. Para algumas análises, caso sejam usados todos os satélites disponíveis, pode ocorrer a contagem de um mesmo foco várias vezes para diferentes satélites. Ou ainda, a contagem do mesmo foco em passagens diurnas e noturnas de um mesmo satélite polar. Para solucionar esse problema, em casos em que a quantidade absoluta de focos importa, será utilizado o satélite AQUA com passagem à tarde (AQUA\_M-T), que é o satélite de referência do INPE atualmente. Dessa forma, é possível evitar a contagem duplicada de focos de queimada e garantir a precisão das informações analisadas.

A análise começa, na Figura 2.3, com uma visão preliminar sobre as características dos focos detectados durante toda a série história, para todos os satélites. Aqui é importante destacar que vários focos podem representar uma única queimada, como

Figura 2.2 – Órbitas-ponto Landsat que cobrem o território brasileiro com base no WRS-2.

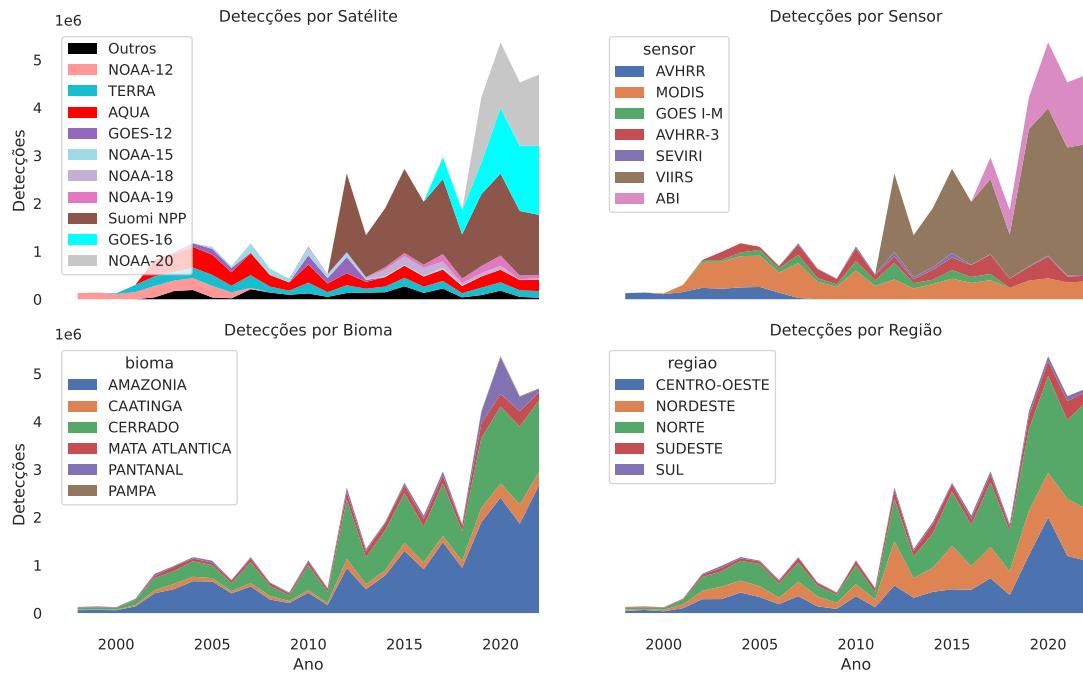


Fonte: O Autor

mencionado anteriormente, e também não há relação necessariamente direta com a área queimada. Nota-se que a série começa basicamente com os dados do satélite NOAA-12, que era o satélite de referência até 2002, e teve dados até 2007, ano que foi desativado. TERRA e AQUA foram os únicos satélites presentes no início e que até hoje tem relevância nos dados. A partir de 2012, com a entrada do satélite Suomi NPP, os dados mais que duplicaram devido a capacidade do sensor VIIRS a bordo deste satélite e depois, em 2019, junto ao satélite NOAA-20 também. A partir de 2017 o satélite GOES-16 entra em ação e, mesmo sendo geoestacionário (com menor capacidade de detecção dos focos), passou a ter relevância.

Quanto as detecções por bioma, nota-se que no decorrer dos anos a proporção entre os focos em cada bioma se manteve basicamente constante, com excessão do período de 2019 a 2022. Nesse período, o bioma Pantanal quebrou a tendência de ter poucos focos detectados, apresentando um aumento de 600% e justificado por um período de seca prolongado (MIRANDA; MARTINHO; CARVALHO, 2021). Nesse mesmo pe-

Figura 2.3 – Detecções por satélite, sensor, bioma e região durante toda a série histórica



Fonte: O Autor

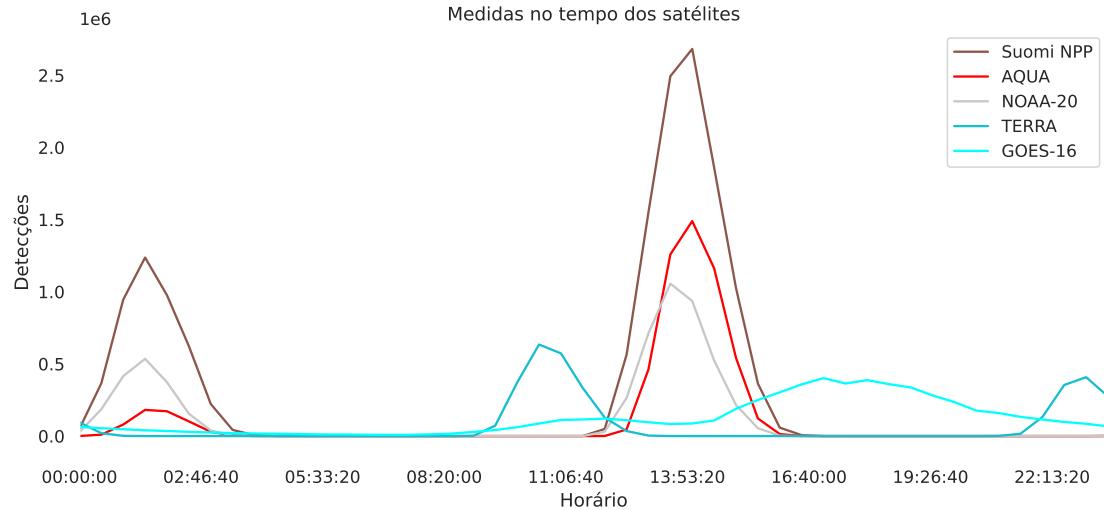
íodo, também notou-se um aumento nas detecções, ainda que bem menor, no bioma da Mata Atlântica.

Saber em quais momentos os satélites passam também é importante para a análise. Os satélites polares passam duas vezes por dia no Brasil, variando o local exato da passagem de acordo com as características de sua órbita. Pela Figura 2.4 é possível observar esse comportamento também nos dados. As quatro primeiras linhas, que representam dados gerados por satélites polares, apresentam dois picos durante um período de 24 horas. Já para o caso dos satélites geoestacionários (GOES-16) não se observou o mesmo padrão, apresentando mais dados entre o período da tarde e início da noite, provavelmente por ser um intervalo de tempo mais quente e seco, mais propício para o fogo (NEPSTAD et al., 2007).

Ao realizar uma análise quantitativa dos focos de queimadas, considerando apenas o satélite de referência, é possível identificar uma sazonalidade clara na Figura 2.5. Os períodos do ano com maior número de detecções estão concentrados entre julho e setembro, com um pico máximo em 2007. Por outro lado, entre dezembro e março, as detecções diminuem significativamente, representando menos de 10% do total de focos no ano.

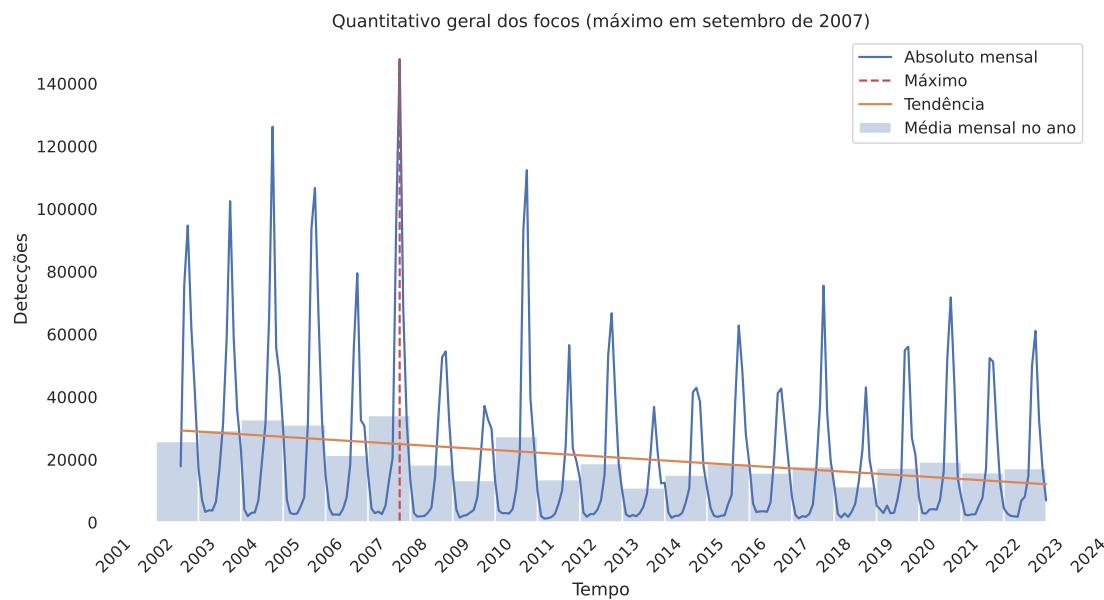
Esse padrão sazonal foi explicado em um estudo realizado por Martins et al. (2020). O trabalho analisou os dados de focos de queimada do satélite AQUA, obtidos

Figura 2.4 – Amostragem por tempo de cada satélite.



Fonte: O Autor

Figura 2.5 – Somatório dos focos do satélite de referência agrupado por mês

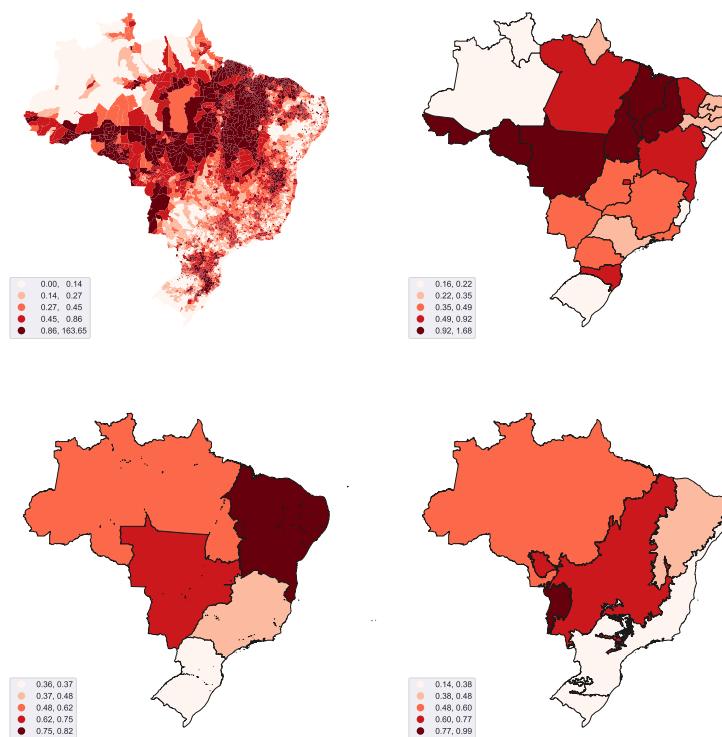


Fonte: O Autor

pelo INPE, para o período de 2003 a 2018. Os resultados indicaram que a precipitação exerce uma influência direta na quantidade e localização dos focos no território nacional. Durante o período chuvoso, de dezembro a março, o Brasil é afetado pela Zona de Convergência Intertropical e pela Zona de Convergência do Atlântico Sul. Quando esses fenômenos estão menos intensos, a região nordeste entra em um período de seca, e à medida que a seca se intensifica, as queimadas se tornam mais frequentes e intensas.

No geral, observa-se uma tendência de queda nos focos de queimada. Tanto os períodos de baixa quanto os períodos de alta apresentam uma diminuição nas ocorrências. Essa tendência é fortemente influenciada pelos anos de 2009, 2011 e 2013, que registraram o menor número de focos, bem como pelos anos anteriores, de 2003 a 2005 e 2007, com ocorrências mais elevadas. É importante destacar que nos últimos anos, a média anual dentro de cada ano rompeu essa tendência, o que pode ser um sinal de alerta do aumento das queimadas.

Figura 2.6 – Relação de focos por área geográfica do país por município (superior esquerdo); por estado (superior direito); por região (inferior esquerdo); e por bioma (inferior direito)



Fonte: O Autor

A Figura 2.6 é fruto da relação dos dados de focos de queimadas com dados geográficos do país proveniente do IBGE. Nela é apresentado a quantidade de focos por  $km^2$  para cada município, estado, região e bioma do país desde o início dos dados do satélite de referência. A quantificação não é fixa para todos dos mapas, por isso cada mapa possui uma legenda que fica no canto inferior. Cada linha da legenda representa o valor inicial (exclusivo, exceto no primeiro) e o valor final (inclusivo) do intervalo.

O bioma Pantanal, apesar de apresentar queimadas mais recentes, já se destaca como o bioma que teve mais focos detectados, quase um foco por  $km^2$ . Logo depois vem o Cerrado, bioma de interesse de vários trabalhos que envolvem as queimadas, com a

maior concentração de focos na parte superior de bioma, onde se intersecta com a região Norte e Nordeste. Os biomas Pampa, Mata Atlântica e Caatinga apresentam a menor proporção de focos por  $km^2$ , provavelmente por já estarem muito degradados pelo processo de urbanização que começou na costa do país.

Pará, Rondônia e Acre são estados que apresentaram bastante focos e são cobertos quase completamente pelo bioma da amazônia. Por outro lado, estados nessa mesma situação, como Amazonas, Roraima e Amapá tem poucos focos. Muitas cidades do nordeste do Rio Grande do Sul e Santa Catarina também se destacaram, apesar da região Sul inteira não apresentar valores significativos.

Por fim, com essa exploração inicial dos dados de focos do INPE, ficou claro que o tema ainda é bem relevante atualmente. Só no ano de 2022 foram detectados mais de 4.5 milhões de focos e mais da metade desses no bioma da Amazônia. O tema das queimadas no país é algo que precisa de atenção e de políticas públicas direcionadas às áreas que risco. A importância do tema, somado a alta cobertura de satélites produzindo dados públicos sobre os focos diariamente, fomenta o desenvolvimento de novos métodos de análise automática desses dados.

## **2.4 Detecção de focos ativos e área queimada**

Nesta seção é abordada de forma mais detalhada como os dados brutos dos sensores dos satélites são usados de fato para encontrar focos ativos de queimadas e como o cálculo da área queimada é realizado atualmente. O texto começa com uma explicação de como funciona o processamento dos dados brutos e depois apresenta o método mais consolidado atualmente de detecção de focos ativos. Em seguida, são apresentados quatro métodos diferentes para o cálculo da área queimada, sendo dois mantidos pelo INPE, outro pela NASA e outro pela ESA.

Os dados brutos dos sensores dos satélites são processados usando algoritmos específicos e produzem resultados chamados de “produtos”. Nesse processo, os algoritmos corrigem distorções, eliminam ruídos e convertem os valores para dados que podem ser representados visualmente. Os produtos podem ter várias finalidades, como detecção de níveis de  $CO_2$ , radiância do solo, anomalias térmicas, entre outros, além de servirem como entrada para outros produtos. Os produtos podem ser classificados pelo seu nível de maturidade, como apresentado na Tabela 2.3, e pelo nível de processamento dos dados, como listado na Tabela 2.4.

Tabela 2.3 – Nível de maturidade dos produtos

Nível	Descrição
Beta	Apenas deve ser usado para testes
Provisório	O produto está em melhoria e segue sendo validado, não há garantias de qualidade
Validado	O produto tem alta qualidade e pode ser usado para estudos rigorosos, como de mudanças climáticas. É subdividido em 4 estágios de validação. Nessa fase, o produto ainda pode seguir melhorando para atingir estágios mais altos de validação

Fonte: O Autor com base em (NASA, 2023)

Tabela 2.4 – Nível de processamento dos dados dos produtos.

Nível	Descrição
Nível 0	É o mais próximo dos dados brutos, com tratamento apenas de erros gerados pela comunicação dos dados, como dados duplicados
Nível 1	São dados georeferenciados com correções geométricas em relação a superfície da terra. Contém dados como radiância e reflectâncias. São subdivididos entre 1A, 1B e 1C
Nível 2	São dados referentes a geofísica da terra, como características da superfície terrestre, dados atmosféricos, níveis de poluentes, entre outros. São subdivididos entre 2A e 2B
Nível 3	São dados de nível 2 organizados no espaço e agrupados em períodos de tempo fixo (diário, mensal, anual), permitindo uma análise em escala global.
Nível 4	Envolve o uso dos dados em modelos ou uso deles em conjuntos com outros dados ou produtos, produzindo estimativas mais aprimoradas

Fonte: O Autor com base em (EARTHDATA, 2023)

A evolução dos produtos gerados a partir do sensor MODIS é contínua à medida que a série temporal se expande. Isso resulta em diferentes versões do conjunto de produtos, denominado “Collection” e seguido por um número que define a versão. Por exemplo, a versão mais recente atualmente é a Collection 6.1, que também pode ser referida de forma simplificada como C6.1. Durante o desenvolvimento de uma nova versão, a versão anterior continua disponível enquanto os cálculos e processamentos são atualizados para os diferentes produtos. Isso permite que os usuários tenham acesso contínuo aos dados e resultados até que a nova versão esteja completa. A Tabela 2.5 apresenta um resumo das principais mudanças que ocorreram de uma versão para outra do Collection, começando pela versão que surgiu no mesmo ano do lançamento do satélite AQUA.

Tabela 2.5 – Versões resumidas do Collection para o sensor MODIS.

Versão	Criação	Fim	Descrição
4	2002	2006	Primeira versão que atingiu 3 estágios de validação
5	2005	2016	Primeira grande coleção científica, que foi amplamente distribuída.
5.1	2008	2016	Atualização fundamental para o Produto L2 de Aerossol (04_L2) e o Produto L2 de Nuvem (06_L2) - juntamente com todos os produtos L3 para incorporar as atualizações do L2.
6	2013	2018	A Coleção 6 incluiu muitas novas atualizações científicas e melhorias.
6.1	2017	Atual	Corrigiu problemas nos dados de entrada do Nível-1B (L1B) - com algumas novas melhorias nos produtos do Nível-2 (L2) e Nível-3 (L3) adicionadas.

Fonte: Traduzido e modificado a partir de (PRODUCTS, 2023)

## Produtos de focos ativos

Os algoritmos que geram produtos de focos ativos, geralmente usam diferentes canais dos sensores dos satélites, entre a luz visível e o infravermelho, e podem ter comportamentos distintos dependendo se a imagem foi gerada à noite ou de dia. Podem também usar limiares dinâmicos, de acordo com a região do planeta, que são calculados, com base em uma espécie de média das temperaturas nas regiões próximas ao longo dos dias. Além disso, é comum a aplicação de máscaras para eliminar regiões submersas, costeiras, desérticas e que estavam nubladas na hora da passagem do satélite empregado.

Para simplificar o trabalho, apresentamos apenas o método empregado para a detecção de focos ativos oficial da NASA, a partir do sensor MODIS (GIGLIO; SCHROEDER; JUSTICE, 2016), que faz parte de Collection 6.1. O método recebe como entrada os produtos Nível 1B de radiância dos dois satélites (MOD021KM/MYD021KM) e os produtos Nível 1A (MOD03/MYD03) que identificam regiões de água, regiões costeiras e de terra. O resultado são dois produtos de Nível 2 e de resolução espacial 1 km, MOD14 para o satélite Terra e MYD14 para o satélite Aqua.

O processamento começa identificando e eliminando as regiões costeiras dos dados de reflectância, a fim de evitar confundir áreas terrestres mais quentes com áreas de água. Em seguida, ocorre a filtragem das regiões com cobertura de nuvens, utilizando limites atualizados, com cuidado para não excluir regiões de fumaça proveniente de incêndios ativos. As regiões remanescentes passam por um processo de avaliação para determinar potenciais píxels de fogo ativo. Esse processo envolve o teste da reflectância

em relação a limiares dinâmicos, que são constantemente atualizados e ajustados para diferentes regiões do planeta, no caso das áreas terrestres. Para regiões onde não há dados suficientes para calcular os limiares dinâmicos ou para áreas de água, é utilizado um limiar fixo igual ao do Collection 5.

O algoritmo passa então a calcular médias de reflectância em uma grade centrada no pixel potencial de fogo ativo. Com base nessas médias locais, os pixels que atendem aos diversos testes relacionados a essas médias são selecionados. Os pixels que não possuem dados suficientes na vizinhança e não passaram nos testes são rotulados como “sem dado”. Em seguida, o algoritmo entra em uma fase de descarte dos pixels preliminarmente classificados como fogo ativo, a fim de reduzir os falsos positivos. Durante essa fase de descarte, os pixels são eliminados se forem identificados reflexos solares maiores do que o aceitável ou se estiverem próximos a regiões costeiras, tanto para pixels na água quanto na terra. Além disso, outros filtros são aplicados exclusivamente para pixels em terra: se eles pertencem a regiões desérticas ou se estão em regiões de clareiras florestais.

O INPE empregava seu próprio algoritmo de detecção de focos de queimadas para o sensor MODIS, o qual produzia dados de maior confiabilidade (INPE, 2023b). No entanto, a partir de 2017, o instituto adotou também os algoritmos do Collection 6 e reprocessou toda a sua base de dados de focos de queimadas. Isso marcou o início da chamada Base 2 de queimadas, que é totalmente compatível com os produtos de queimadas da NASA (INPE, 2023b). Anteriormente, quando o Collection 5 estava em uso, eram observados falsos positivos em clareiras florestais e falsos negativos para grandes queimadas que estavam obscurecidas por fumaça densa (SCHROEDER et al., 2008), o que justificava o desenvolvimento de um algoritmo próprio pelo INPE para o Brasil. Como esse algoritmo não é mais utilizado e seus resultados foram substituídos na base de dados do INPE, ele não é abordado neste trabalho.

## **Produtos de área queimada**

Nesta parte, são apresentados alguns dos produtos mais consolidados para o cálculo de área queimada no Brasil e no mundo: AQ1Km, AQ30m, MCD64A1 e FireCCI51. Um resumo desses quatro produtos pode ser encontrado na Tabela 2.6. A tabela fornece informações do nome do produto, o ano de início de uso, a resolução espacial e temporal, os satélites utilizados no cálculo e a principal referência teórica que embasam esses mé-

todos. Além desses produtos, existem muitos outros trabalhos que abordam o cálculo de área queimada, e alguns deles são mencionados na Seção 2.5.

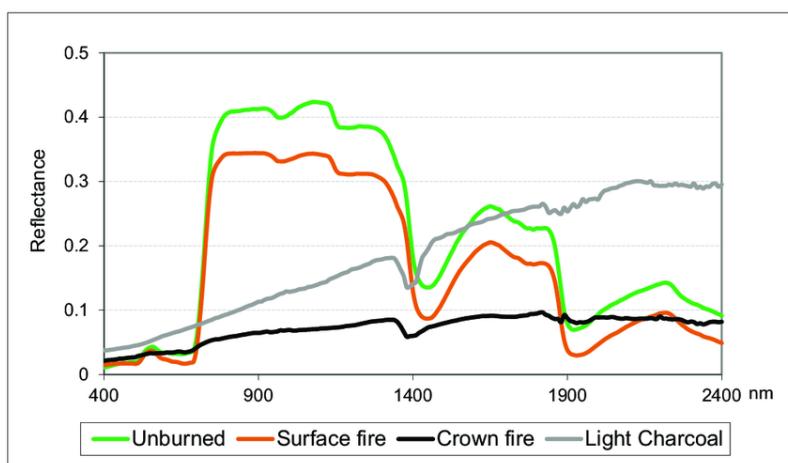
Tabela 2.6 – Resumo dos produtos de área queimada.

Nome	Ano	Res. esp.	Res. tem.	Satélites	Embassamento teórico
AQ1km	2015	1 km	mensal	Aqua/Terra	Libonati et al. (2015)
AQ30m	2014	30 m	quinzenal	Landsat-5,7,8	Melchiori et al. (2014)
MCD64A1	2016	500 m	mensal	Aqua/Terra	Giglio et al. (2018)
FireCCIS1	2018	250 m	mensal	Aqua/Terra	Lizundia-Loiola et al. (2020)

Fonte: O Autor

Os algoritmos utilizados para calcular a área queimada empregam índices que são sensíveis à vegetação queimada, aplicados aos produtos de reflectância dos satélites. Atualmente, os mais utilizados envolvem uma combinação das bandas do infravermelho próximo (NIR:  $0,78 - 1,2\mu\text{m}$ ) e do infravermelho de onda curta (SWIR:  $1,2 - 3,0\mu\text{m}$ ), enquanto o uso do infravermelho de onda média (MWIR:  $3 - 8\mu\text{m}$ ) é menos comum. Estudos indicam que, em uma área queimada, a reflectância do NIR é reduzida e a reflectância do SWIR é aumentada, devido à queima da vegetação e à secura do ambiente, respectivamente (CHUVIECO et al., 2019). Esse comportamento é ilustrado na Figura 2.7.

Figura 2.7 – Variação da reflectância (%), no eixo y, em relação à frequência da luz eletromagnética emitida por uma floresta, no eixo x. A cor verde representa a reflectância de uma vegetação não queimada, a cor vermelha representa a reflectância do fogo ativo na parte inferior da floresta (no solo), a cor preta representa um incêndio na copa das árvores e a cor cinza representa a reflectância de uma vegetação queimada (carvão).



Fonte: Extraído de Chuvieco et al. (2019) com base nos modelos Prospect + Geosail

O INPE utiliza dois produtos principais para a detecção de áreas queimadas: o AQ30m e o AQ1km. O produto AQ1km, desenvolvido em parceria com o Laboratório de

Aplicações de Satélites Ambientais (LASA) do Departamento de Meteorologia da UFRJ, possui uma resolução espacial de 1 km e fornece dados diários agrupados em períodos mensais (INPE, 2023a). Por outro lado, o produto AQ30m possui uma resolução espacial de 30 metros e fornece dados quinzenais. Ambos os produtos são complementares e permitem uma análise abrangente das áreas queimadas. O AQ1km é adequado para avaliações mais amplas, como a determinação da porcentagem de área queimada em todo o Brasil durante um mês específico. Já o AQ30m fornece dados locais mais precisos, embora muitas regiões do Brasil não são processadas ou os dados não estão disponíveis publicamente.

O produto AQ1Km está em fase provisória de desenvolvimento (conforme Tabela 2.3). É derivado de dados coletados pelo sensor MODIS, utilizando simultaneamente os satélites Aqua e Terra, além de incorporar informações de focos de queimada do INPE de diferentes satélites (LIBONATI et al., 2015). Resumidamente, o algoritmo do AQ1km inicia filtrando os dados de reflectância com base em critérios de cobertura de nuvens e posição espacial para minimizar distorções. Em seguida, os dados são agrupados por mês para preencher eventuais lacunas diárias. O algoritmo identifica as áreas denominadas “HotSpots” (pontos quentes), que correspondem a regiões com atividade de fogo durante o período analisado. Posteriormente, é calculado um índice de vegetação para essas áreas, utilizando dados de reflectância no infravermelho próximo (NIR) e no infravermelho de onda média (MWIR), conforme proposto em Libonati et al. (2011), comparando-o com o mesmo índice do mês anterior. Por fim, são identificadas regiões próximas às áreas classificadas como queimadas no passo anterior, mas que não apresentaram um indicador suficientemente forte, podendo indicar que ocorreu uma queimada parcial ou que a intensidade do fogo foi baixa naquela localidade.

O produto AQ30m é gerado a partir de imagens dos satélites Landsat-5, Landsat-7 e Landsat-8, utilizando a diferença entre os índices de vegetação NDVI (Normalized Difference Vegetation Index), NBRL (Normalized Burn Ratio) e BAI (Burned Area Index) entre imagens consecutivas (MELCHIORI et al., 2014). Esses dados passam por uma filtragem com base em limiares definidos por especialistas e, em seguida, são suavizados usando um filtro de mediana 3x3 para eliminar valores discrepantes. O passo final envolve a construção de polígonos com base nos valores calculados e a remoção de polígonos muito pequenos. A implementação desse produto foi feita em Python, utilizando algumas ferramentas de processamento de dados que o presente trabalho também utiliza (mais detalhes na Seção 3.6).

No entanto, o processo atualmente não é totalmente automatizado e requer uma avaliação final de um especialista do instituto antes da divulgação. Isso é necessário pois a degradação da vegetação pode ter ocorrido por outro fator além da queimada, como desmatamento, colheita ou preparação do solo. Quando os especialistas não tem certeza se a área destacada foi originada por uma queimada, a classificação é de não-queimada. Em Júnior et al. (2019), são apresentadas algumas técnicas de aprendizado de máquina aplicadas na avaliação automática desse produto, obtendo resultados promissores (ver a Seção 2.5).

A NASA também mantém um produto específico para o cálculo da área queimada, chamado de MCD64A1, utilizando também os dados dos dois satélites com sensor MODIS (GIGLIO et al., 2018), de resolução espacial de 500m e temporal de um mês. Diferente da versão anterior, a MCD45 (C5), esse algoritmo combina as detecções de fogos ativos de 1 km do Collection 6.1 (ver Seção 2.4) com um índice de vegetação sensível a queimadas, construído a partir dos canais 5 e 7 do SWIR do sensor MODIS, capturando a reflectância característica do solo em áreas queimadas. A combinação dos dados de fogos ativos e do índice de vegetação permite que o algoritmo se adapte às diferentes regiões e ecossistemas em todo o mundo. Os artefatos produzidos pelo produto podem ser baixados gratuitamente para cada mês a partir de novembro do ano de 2000 pelo site da Earthdata (<<https://urs.earthdata.nasa.gov>>).

O algoritmo inicia selecionando os dados diários de reflectância válidos, ou seja, dados que não contêm nuvens, fogo ativo, estão localizados em terra e possuem valores dentro do intervalo de 0 a 1. Em seguida, calcula o Índice de Vegetação (IV) para cada célula da grade. Posteriormente, o algoritmo calcula a diferença estatística do IV (S), comparando uma janela de dados de 8 dias antes e 8 dias depois do período avaliado. As regiões que não possuem dados confiáveis dentro dessas janelas de tempo são marcadas como não classificadas. Os dados passam por uma classificação inicial utilizando as estatísticas de S e outras informações de probabilidade de queima. Por fim, o algoritmo analisa as regiões vizinhas às que foram previamente classificadas como queimadas, aplicando uma fórmula probabilística que utiliza dados de focos ativos nessas regiões.

A ESA desenvolveu o FireCCI51, um produto avançado para o cálculo de áreas queimadas (LIZUNDIA-LOIOLA et al., 2020). Esse produto representa uma evolução em relação aos produtos anteriores, FireCCI41 e FireCCI50. Assim como o MCD64A1, o FireCCI51 combina detecções de focos ativos e dados de reflectância NIR do sensor MODIS, porém com uma resolução mais refinada de 250m. O algoritmo do FireCCI51 é

dividido em duas fases principais. Na primeira fase, chamada de “semente”, são coletados os pontos com alta probabilidade de terem sido queimados. Esses pontos são selecionados com base nas detecções de anomalias térmicas (focos ativos) e na diminuição do NIR, que é um indicativo de área queimada (PEREIRA, 1999). Em seguida, o algoritmo passa para a fase de “crescimento”, na qual são aplicados limiares locais a partir das sementes para identificar completamente as áreas queimadas. Posteriormente, é realizada uma poda para eliminar áreas que cresceram muito além das sementes iniciais, utilizando heurísticas que visam reduzir os falsos positivos.

## 2.5 Trabalhos Relacionados

Muitos outros trabalhos antes deste versaram sobre os dados de queimadas e, em menor escala, especificamente sobre o cálculo das áreas queimadas. Além dos trabalhos fundamentais, mencionados na Seção 2.4, que deram origem aos produtos AQ1Km, AQ30m, MCD64A1 e FireCCI51, realizados por Libonati et al. (2015), Melchiori et al. (2014), Giglio et al. (2018) e Lizundia-Loiola et al. (2020), respectivamente, são apresentados outros estudos que também inovaram no tema. São discutidos e apresentados os resultados obtidos por esses estudos e, ao final, são abordadas questões práticas e as características que tornam o presente trabalho único em relação aos demais.

Júnior et al. (2019) se dedica a eliminar a necessidade de uma avaliação humana no produto AQ30m, para o cálculo de áreas queimadas a partir de imagens do satélite Landsat-8. O documento aborda os resultados do treinamento de cinco modelos de aprendizagem de máquina: K-Nearest Neighbors (kNN), Decision Trees (DT), Random Forests (RF), Neural Networks (NN) e Support Vector Machines (SVM). São usados quatro dados de entradas para esse modelo: quantidade de queimadas anteriores próximas a área analisada (i), focos ativos de queima de diferentes satélites extraídos da base de dados do INPE (ii), valor do índice Mid-Infrared Burn Index (MIRBI) (iii) e valor do índice Normalized Difference Water Index (NDWI) (iv).

Os modelos foram treinados com 9 imagens no período de 16/03/17 a 24/09/17 na região órbita-ponto 223/067. A validação foi feita com uma imagem em 10/10/17 e obteve resultados melhores nos modelos Random Forests e Redes Neurais. A acurácia geral ficou em torno de 96% para esses modelos. O modelos de K-Nearest Neighbors e Decision Trees ficaram entre 80% e 90% de acurácia. O pior modelo, o Support Vector Machines, obteve acurácia de pouco mais de 50%. O trabalho conclui que os modelos

podem atingir boas acuráciais, mas como são treinados usando modelos de classificação pesada (*hard classification*), ao final o resultado é sempre uma classificação binária entre “Queimado” e “Não queimado”. Isso simplifica muito o treinamento dos modelos, mas pode não ser ideal por agrupar áreas de desmatamento, colheita ou preparação do solo em uma classe errada.

Pereira et al. (2017) usa um modelo de aprendizagem de máquina chamado de Máquina de Vetores de Suporte de Classe Única (OC-SVDD) para o mapeamento de áreas queimadas no Cerrado brasileiro com uma resolução de 300m. O método precisa apenas de dados da classe “Queimada”, o que simplifica sua automatização. Para alimentar o modelo, foram usados dados pré-processados de focos ativos detectados pelo sensor VIIRS (resolução de 375m) representando a classe “Queimada”. Junto aos focos, também foram usados dados quizenais de reflectância do satélite PROBA-V (missão da ESA descontínua em 2020), com o intuito de detectar regiões queimadas que não foram indicadas pelos dados dos focos ativos, o que reduz o erro de omissão. Devido ao uso dos dados do PROBA-V o método recebeu o nome de AQM-PROBA.

O método foi avaliado contra o produto AQ30m e comparando os resultados com o MCD64A1 aplicados em 13 órbitas-ponto Landsat. A abordagem de avaliação é a mesma usada em Libonati et al. (2015) e neste presente trabalho (mais detalhes na Seção 4.1). O AQM-PROBA apresentou em média 30% de erros de omissão e 22% de erros de comissão. Comparando com o MCD64A1, o AQM-PROBA indica mais queimadas verdadeiras, mas também aumenta a taxa de falsos positivos. A conclusão do trabalho é que o método AQM-PROBA produz resultados mais próximos da referência AQ30m comparado com o MCD64A1, mesmo que tanto o AQM-PROBA quanto o MCD64A1 subestimam a área queimada de referência.

## **Comparação e discussão**

Quanto aos produtos AQ1Km e MCD64A1, ambos utilizam dados do sensor MODIS para monitorar a reflectância da vegetação a fim de detectar áreas queimadas, embora tenham utilizado índices de vegetação diferentes. Os resultados apresentados pelo produto AQ1km em comparação com o MCD64A1 indicam muito mais áreas queimadas, mas apresentam uma menor acurácia geral. Em contraste ao MCD64A1, o produto AQ1km inova e se beneficia dos dados de diferentes satélites para definir as áreas de interesse e calcular o índice de vegetação. No entanto, ainda sofre das mesmas limitações do

sensor MODIS, dependendo apenas de dois satélites, AQUA e TERRA, para o resultado final.

O produto AQ30m, embora tenha uma resolução espacial mais alta, possui algumas limitações em relação à sua disponibilidade e cobertura. A dependência de intervenção humana para a publicação final dos resultados pode atrasar a divulgação periódica dos dados no site do DGI, dependendo da disponibilidade da equipe do INPE. Além disso, a disponibilidade dos dados do satélite Landsat-8 nem sempre é garantida, podendo estar faltando ou corrompidos, o que impossibilita a derivação do produto. Outro aspecto é que a cobertura do produto não abrange todo o território nacional, ou pelo menos essa informação não é divulgada, cobrindo apenas algumas órbitas-ponto que incluem parte do Cerrado, parte de Rondônia e a região central da Bahia.

O trabalho de Júnior et al. (2019) é interessante pois aborda o treinamento de diferentes modelos de aprendizado de máquina. No entanto, para tornar os resultados mais robustos e representativos, seria necessário incluir mais regiões de treinamento e uma avaliação mais diversificada. Embora o estudo tenha alcançado uma alta acurácia geral, acima de 96%, é importante destacar que essa métrica pode não indicar alta precisão devido ao desequilíbrio de classes. Nesse caso, a classe “Não Queimada” é muito mais comum do que a classe “Queimada”, fazendo com que modelos que simplesmente classifiquem todas as áreas como não queimadas ainda apresentem uma acurácia alta.

No caso do presente trabalho, não são utilizados os dados de produtos de reflectância dos satélites, como os produtos de infravermelho usados nos cálculos dos índices de vegetação dos trabalhos anteriormente mencionados. Em vez disso, são utilizados apenas os dados de focos ativos provenientes de diferentes satélites processados pelo INPE. Isso simplifica o processo, uma vez que as considerações relacionadas a diferentes regiões e biomas do planeta já foram resolvidas no processamento dos focos ativos. Além disso, essa abordagem é mais resistente a falta de dados, pois se um ou mais satélites falharem, o método ainda será capaz de operar com os dados dos satélites restantes.

No próximo capítulo é apresentado a metodologia empregada para estimar a área queimada apenas com focos ativos. A Seção 3.6 discute algumas questões práticas da sua implementação e de como os dados utilizados foram obtidos. Por fim, no Capítulo 4, é feito a avaliação do método, comparando-o com o produto AQ30m e apresentando métricas animadoras.

## 3 METODOLOGIA

Neste capítulo, é apresentado o desenvolvimento do método utilizado para calcular as áreas queimadas a partir de focos ativos. Inicia-se com uma visão geral do método e, em seguida, são abordadas detalhadamente cada etapa do processo. Questões práticas e de implementação são discutidas no final deste capítulo, na Seção 3.6. Por fim, a aplicação do método bem como os resultados obtidos são abordados no Capítulo 4.

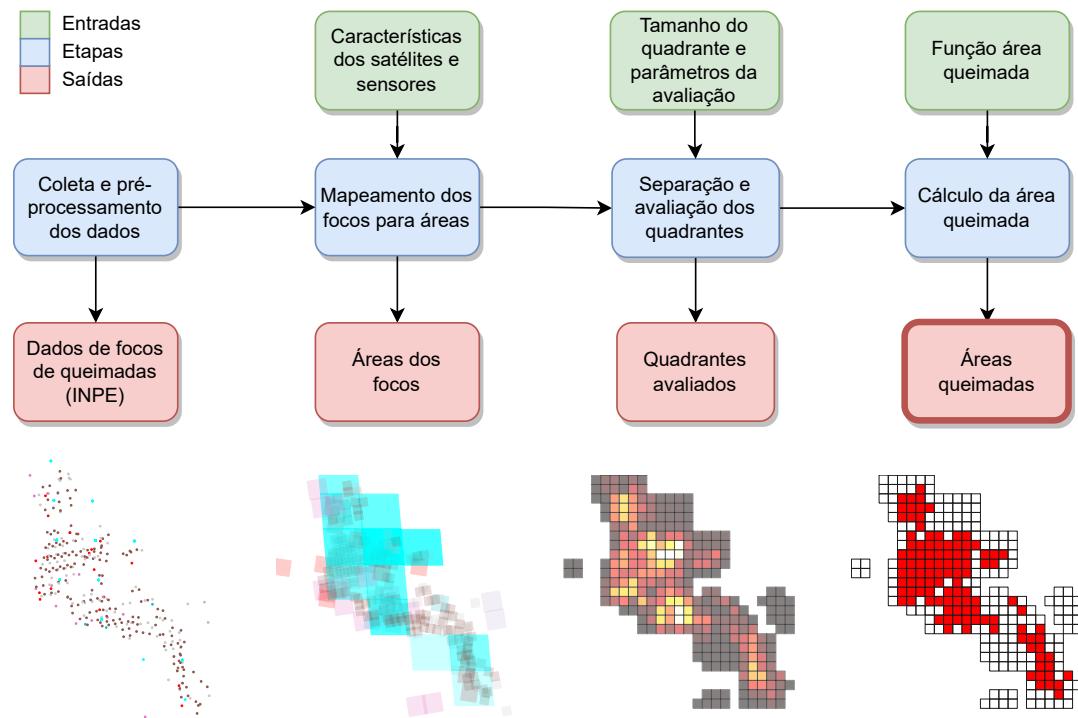
### 3.1 Visão geral da metodologia

A metodologia desenvolvida visa calcular a área de vegetação queimada no Brasil por meio da análise dos dados de focos de queimadas disponibilizados pelo INPE, juntamente com as características dos diferentes satélites e sensores. O método proposto recebeu o nome de AQ-FA, sigla para Área Queimada por Focos Ativos, por ser a principal informação utilizada para estimar a área queimada. A premissa fundamental é que um foco de queimada detectado resulta em uma área queimada. Além disso, considera-se que a quantidade de focos detectados em uma determinada região, em um intervalo de tempo, está diretamente relacionada com a área efetivamente queimada na região.

O método completo, conforme ilustrado na Figura 3.1, é composto por quatro etapas principais (cartões em azul): coleta e pré-processamento dos dados, mapeamento dos focos para áreas, separação e avaliação dos quadrantes, e cálculo da área queimada. Cada etapa pode receber dados de entrada e gera uma saída que, com exceção da última etapa, é utilizada como entrada para a etapa seguinte. Os cartões em verde representam as entradas de cada etapa, que incluem dados de focos ativos do período de tempo de interesse, características dos satélites e sensores, e parâmetros de configuração do método. O território e o período de tempo a serem avaliados são definidos por meio do filtro dos dados de focos ativos. Por fim, os cartões em vermelho representam as saídas de cada etapa, exemplificadas na parte inferior do diagrama.

A primeira etapa consiste em obter os dados da base de dados do INPE e prepará-los para serem utilizados no método. Essa parte é executada apenas uma vez, salvando os dados de todo o Brasil para serem utilizados nas próximas execuções do método. É nessa etapa também, que os dados são filtrados para o período de tempo e região de interesse para a aplicação do método.

Figura 3.1 – Diagrama do método. A entrada principal são os dados de focos de queimadas e a saída principal é a estimativa de áreas queimadas. Abaixo os resultados produzidos com a aplicação em um determinado local para cada etapa. A imagem abaixo representa a aplicação do método no sudoeste do Pará, durante os dias 1 a 3 de setembro de 2022.



Fonte: O Autor

A segunda etapa consiste na transformação dos pontos de foco em áreas que representam a extensão de cobertura do foco. Nessa etapa, são utilizados dados de focos de queimadas fornecidos pelo INPE, previamente filtrados para o período de tempo de interesse. A saída dessa etapa são as áreas correspondentes a cada foco. Essas áreas estão diretamente relacionadas às características da órbita do satélite e à resolução do sensor que capturou a imagem original, conforme abordado na Seção 2.4. Portanto, as informações sobre os satélites e sensores também são necessárias como entrada para essa etapa.

Na terceira etapa, ocorre a divisão do espaço em quadrantes de tamanho fixo e sua avaliação. Nesse passo, são utilizadas as áreas dos focos calculadas na etapa anterior, juntamente com o tamanho do quadrante, que pode ser definido pelo pesquisador. A saída consiste nos quadrantes associados a valores que representam a intensidade das queimadas ocorridas dentro de cada quadrante, avaliados por meio de uma função de avaliação. Essa função de avaliação leva em consideração a diversidade dos dados e a quantidade de áreas de foco que intersectam cada quadrante.

Na quarta e última etapa, é realizado o cálculo das áreas queimadas, produzindo o resultado final do método. Essa etapa recebe como entrada os quadrantes avaliados no passo anterior e uma função de avaliação da área queimada, que pode ser definida pelo pesquisador. A função é aplicada a todos os quadrantes, resultando em um valor que representa a porcentagem provável de que a área do quadrante tenha sido queimada.

A Figura 3.1 também ilustra a aplicação completa do método em uma área específica localizada no sudoeste do Pará, durante os dias 1 a 3 de setembro de 2022. Na primeira imagem, são mostrados os focos de queimada, sem nenhum tipo de pré-processamento. Na segunda imagem, os pontos são transformados em áreas. Em seguida, o espaço é dividido em quadrantes avaliados. Na última imagem, o cálculo da área queimada considera todas as avaliações acima do valor 5 como quadrante totalmente queimado, resultando em uma área queimada de  $33,32\text{km}^2$ .

O método é flexível e altamente configurável, podendo assim gerar resultados bem diferentes dependendo de suas entradas, ainda que para o mesmo conjunto de dados de focos. O pesquisador pode testar combinações de entradas e avaliar os resultados usando algum comparador (*benchmark*). Um exemplo seria comparar as saídas do método com os resultado obtidos pelo produto AQ1km (abordado na Seção 2.4).

### **3.2 Coleta e pré-processamento dos dados**

Uma parte fundamental do processo foi coletar os dados do site DBQueimadas, que é a fonte primária usada para cálculo das áreas queimadas. Para exportar os dados usando o navegador, é necessário preencher um formulário com os campos de data inicial, data final e um endereço de e-mail. O intervalo de tempo máximo para o pedido é de 1 ano. Também é possível aplicar filtros ainda mais específicos, como continente, país, estado, município, satélite, bioma e unidades de conservação/terras indígenas. Após clicar em “Exportar”, uma mensagem contendo um link de download é enviada para o e-mail informado, e o arquivo disponibilizado é um CSV compactado em formato zip.

Embora este processo seja simples, não seria prático baixar todos os dados do Brasil manualmente sem sobrecarregar os servidores no INPE. Por isso, foi necessário entender como a solicitação dos dados é processada pelo site e, assim, automatizar o download. Foi identificado que o site faz uma requisição GET para a API do DBQueimadas, localizada em <<https://queimadas.dgi.inpe.br/queimadas/exportacaobdq/exportar>>, passando

os parâmetros codificados em JSON na URL, incluindo os filtros, e-mail e formato de arquivo desejado.

Para automatizar o processo, foi desenvolvido um *script* em Python que solicita os dados de cada 30 dias, totalizando 300 requisições desde 1998 até 2022. Para não sobrecarregar os servidores do INPE, foi adicionado uma espera de um minuto entre as requisições. Além disso, foi implementado um sistema de registro de requisições em arquivo, que armazenou o estado de cada requisição, evitando duplicadas e possibilitando a retentativa das requisições que falharam por algum motivo. O programa foi considerado concluído apenas quando todas as requisições contidas no registro estavam marcadas como concluídas.

Para concluir o processo, ainda era necessário fazer o download dos arquivos por meio do link enviado por e-mail. Para isso, foi utilizado o *Google Scripts*, uma ferramenta que possibilita escrever programas simples, utilizando uma linguagem semelhante a JavaScript, com integração aos serviços do Google (como o Gmail). O *script* era executado com um intervalo de dois minutos, varrendo todos os e-mails novos provenientes do DBQueimadas. Com ele, foi possível extrair o link de cada mensagem e, finalmente, salvar o arquivo de forma automatizada.

Os dois *scripts* trabalharam de forma concomitantemente, funcionando como uma espécie de produtor e consumidor distribuído. Enquanto um requisitava os dados do INPE, o outro vasculhava os e-mails e salvava o arquivo em uma pasta específica. Quando o primeiro *script* identificava (por meio do nome do arquivo) que a resposta já estava salva, marcava a requisição como concluída no registro de requisições. Caso uma requisição permanecesse por mais de 30 minutos no estado pendente, o *script* faz uma nova requisição aos servidores do INPE.

Todo o processo de investigação e recuperação dos dados levou cerca de uma semana. Todos os arquivos baixados ocupam pouco mais de 4 GBytes de armazenamento em disco e, juntos, somam exatamente 43.782.758 linhas. Ao final, eles foram recompactados em um único arquivo zip (450 megabytes) que está disponível para download em <<https://bit.ly/3IgHIXH>>, independentemente dos servidores do INPE.

Também foi utilizado os dados públicos territoriais do Brasil provenientes do Instituto Brasileiro de Geografia e Estatística (IBGE) para gerar gráficos delimitados em municípios, unidades federativas e biomas. O processo de aquisição desses arquivos se deu diretamente no site do IBGE. Todos os arquivos baixados estão no formato *Shapefile*, que é o formato responsável por armazenar dados vetoriais geográficos.

## Pré-processamento dos dados

Após a etapa de coleta de dados, essas informações foram carregadas no Python e algumas otimizações nas estruturas foram realizadas. Os 300 arquivos das queimadas foram lidos pelo *Pandas* e concatenados em um único *DataFrame* (estrutura semelhante a uma tabela). Foi necessário converter o fuso horário das datas, que estavam no formato UTC, para o fuso horário de Brasília, a fim de facilitar a compreensão dos gráficos para os usuários brasileiros. Além disso, as colunas de texto, como município, estado, país e satélite, foram convertidas em categorias, que são uma forma de enumeração disponível no *Pandas*. Essa conversão ajudou a reduzir o consumo de memória, uma vez essas colunas possuíam muitos valores repetidos, ocupando aproximadamente 2 GBytes em memória.

Para otimizar o processamento e a leitura dos dados, também foi utilizada a biblioteca *Dask* especializada no processamento paralelo e distribuído. Essa biblioteca oferece a capacidade de instanciar um cluster computacional lógico sobre um cluster físico local ou remoto, permitindo o envio de tarefas completas para serem processadas pelo cluster. Além disso, o *Dask* possibilita a leitura dos arquivos de forma paralela, o que resulta em uma redução significativa no tempo de carregamento dos dados. Após o carregamento completo do *DataFrame* a partir dos arquivos CSV originais durante a primeira execução, foi criado um arquivo no formato *H5* que contém os mesmos dados, mas com uma estrutura otimizada para permitir uma leitura mais rápida. Nas execuções subsequentes do método, ao invés de carregar novamente os arquivos CSV, o programa faz a leitura desse único arquivo *H5*, resultando em um carregamento ainda mais rápido dos dados.

Durante o pré-processamento dos dados, foram identificadas regras de nomenclatura especiais para alguns satélites. No caso dos satélites AQUA (AQUA\_M-T e AQUA\_M-M) e TERRA (TERRA\_M-T e TERRA\_M-M), a última letra indica o período do dia em que ocorre a passagem do satélite, sendo “M” para Manhã e “T” para Tarde. Outros satélites, como Suomi NPP, NOAA-19, NOAA-18, NOAA-16, NOAA-15 e NOAA-12, podem apresentar a última letra do nome como “D” para Diurno. Com base nessa compreensão das regras de nomenclatura, foi possível criar uma nova coluna que fornece o nome simplificado dos satélites, facilitando algumas análises. Além disso, os satélites com sensor ATRS (ERS-1 e ERS-2) e o satélite TRMM foram removidos dos dados a fim de simplificar a implementação, assim como no produto AQ1km (LIBONATI et al., 2015). Em comparação com a coluna original dos satélites, que tinha 32 valores

possíveis, a nova coluna contém apenas 20 valores possíveis. Para cada satélite, foi es-  
colhida uma cor única, conforme apresentado na Figura 3.2, que é utilizada em todos os  
gráficos apresentados neste documento.

Figura 3.2 – Cores escolhidas para cada satélite.



Fonte: O Autor

### 3.3 Mapeamento dos focos para áreas

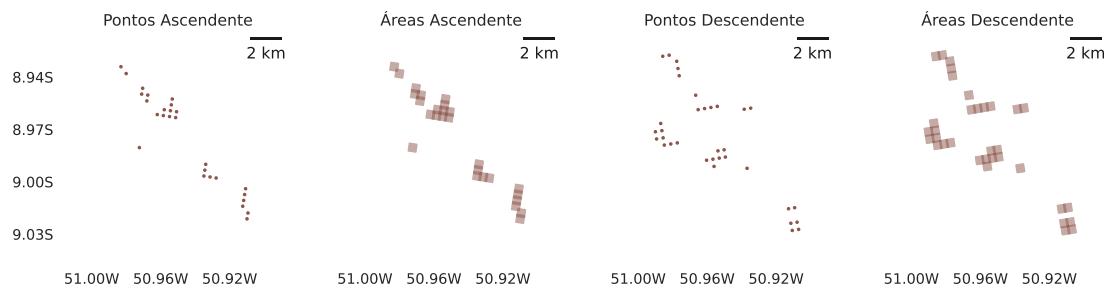
Quando um foco de queimada é detectado pelo INPE a partir de um satélite, o foco representa uma área quadrada aproximadamente do tamanho da resolução do sensor que capturou a imagem original. Por exemplo, um foco detectado a partir de imagens do satélite AQUA, que utiliza o sensor MODIS, representa uma área de  $\approx 1\text{km}^2$ . Para um satélite com o sensor menos preciso, como o GOES-13, que utiliza o sensor GOES I-M com resolução espacial de  $4\text{km}$ , a área representada seria 16 vezes maior, indicando uma menor precisão.

O cálculo exato da área coberta pelo foco deve levar em conta, além da resolução do sensor, as distorções provenientes dos seguintes fatores: diferença de localização entre o foco e o satélite, e as características da órbita do satélite. No primeiro fator, quanto maior a distância entre os dois pontos, maior será a distorção em relação a área de cobertura do sensor. Para um satélite geoestacionário, por exemplo, essa distância é sempre muito grande, devido a sua órbita com altura de  $36\text{km}$  (Embrapa Territorial, 2023).

No segundo fator, a inclinação dos satélites determinam como é a rotação da área coberta. Os satélites que orbitam a Terra em órbitas polares possuem uma determinada inclinação que lhes permitem cobrir diferentes áreas de todo o planeta durante sua trajetória (Figura 2.1). Quando o satélite está se movendo em uma trajetória ascendente, ou seja, do sul para o norte, a área coberta pelo sensor deve ser rotacionada no sentido horário. Por outro lado, se a trajetória for descendente, do norte para o sul, a rotação deve ser no sentido anti-horário.

A Figura 3.3 apresenta dados coletados pelo satélite Suomi NPP no dia dois de setembro de 2022. O primeiro par de imagens da esquerda para a direita corresponde à órbita ascendente do satélite, enquanto o segundo par de imagens corresponde à órbita descendente. Observa-se que, na primeira e terceira imagens, os pontos dos focos de queimada estão alinhados, mas ligeiramente rotacionados em alguns graus, coincidindo com a inclinação do satélite. Na segunda e quarta imagens, os pontos são convertidos em áreas a partir da aplicação desta etapa do método. Eles passam a representar a cobertura territorial do foco e encaixam perfeitamente entre seus vizinhos. Neste documento, a área coberta pelo foco será chamada de medição.

Figura 3.3 – Comparação entre pontos e áreas dos focos detectados pelo satélite Suomi NPP.



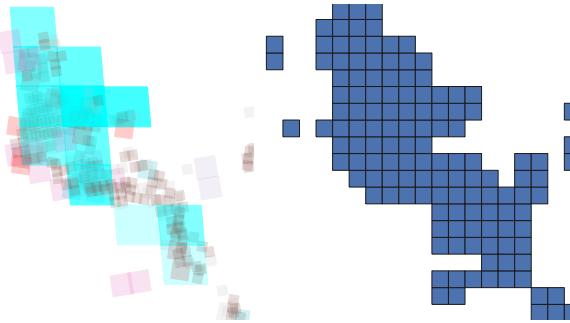
Fonte: O Autor

### 3.4 Separação e avaliação de quadrantes

A separação entre quadrantes, neste trabalho definido como um polígono quadrado e não rotacionado, é a forma de discretizar os dados do espaço, que são contínuos. Os quadrantes abstraem os detalhes das medições dos diferentes satélites, com diferentes áreas e orientações. Como resultado, ilustrado na Figura 3.4, temos um conjunto de quadrantes, todos do mesmo tamanho, que formam uma grade regular. Por fim, esses quadrantes são avaliados individualmente usando operações que serão descritas a seguir.

O tamanho do quadrante é uma entrada dessa etapa, chamada de `quadrant_width`, e pode ser definida pelo pesquisador. Esse parâmetro é expresso em graus e significa quantos graus de longitude e latitude o quadrante vai ocupar. Normalmente o valor deve ser pequeno, a fim de não reduzir muito a resolução da solução, sendo recomendado em torno de 0,004 a 0,005 graus. Esse valor coincide com o tamanho da maior resolução de sensor presente nos dados, que é o VIIRS de 375m.

Figura 3.4 – Separação entre quadrantes.



Fonte: O Autor

Com a distribuição dos quadrantes definida, é necessário atribuir a cada um deles um valor que represente a intensidade de focos dentro do quadrante. A avaliação dos quadrantes deve penalizar quadrantes que apresentam poucas medições de diferentes satélites. Essa abordagem justifica-se pelo fato de que ajuda a reduzir os ruídos nas medições e a identificar queimadas mais intensas. Além disso, diferentes satélites realizam medições em horários distintos (ver Figura 2.4), o que indica uma queimada mais prolongada. Em ambos os casos, queimadas mais intensas e duradouras sinalizam um maior potencial de o fogo ter atingido todo o quadrante.

Sendo  $q$  o quadrante a ser avaliado;  $M$  um conjunto de todas as medições (área coberta pelo foco); a operação  $area(p)$  retorna a área de um polígono  $p$  qualquer (que pode ser um quadrante ou uma medição); a operação  $unique\_satellite(Q)$  retorna todos os satélites diferentes do conjunto  $Q$ ; a constante  $min\_area$  é a área mínima de uma medição; a constante  $threshold\_satellite$  é o número mínimo de satélites para não ser penalizado. A definição formal da avaliação é dada por:

Para cada quadrante ( $q$ ), é realizado o cálculo da interseção com cada medição ( $m$ ), resultando em um conjunto de polígonos ( $Qm$ ) contidos em  $q$ :

$$Qm = \{m \cap q \mid m \in M \text{ and } m \cap q \neq \emptyset\} \quad (3.1)$$

Em seguida, é realizada uma filtragem no conjunto  $Qm$ , removendo todos os polígonos que não possuem uma determinada área mínima, resultando em  $Qm'$ :

$$Qm' = \{m \in Qm \mid area(m) \geq min\_area\} \quad (3.2)$$

A partir de  $Qm'$ , é extraído o número de satélites diferentes presentes nesse conjunto, que é denominado  $Us$ :

$$Us = \{m \in unique\_satelite(Qm')\} \quad (3.3)$$

Além disso, é calculada a soma das áreas de  $Qm'$  e dividida pela área total de  $q$ , resultando em  $ia$ . Ou seja, quantas vezes o somatório das áreas de  $Qm'$  cabem dentro de  $q$ :

$$ia = \left( \sum_m^{Qm'} area(m) \right) \div area(q) \quad (3.4)$$

Penalização de quadrantes que tenham poucos satélites diferentes, o valor de  $c$  fica entre 0 e 1.

$$c = 1 - min \left( 1, \frac{|Us|}{threshold\_satellite} \right) \quad (3.5)$$

Finalmente, a avaliação final do quadrante, denominada  $aq$ , é obtida pela expressão:

$$aq = |Us|^2 + ia - ia \cdot c \quad (3.6)$$

A fórmula final 3.6 atribui um peso quadrático à quantidade de satélites diferentes presente no quadrante  $q$ . Isso garante que, à medida que a diversidade de satélites aumenta dentro do quadrante, o valor de  $aq$  aumente de forma exponencial e se destaque dos demais quadrantes com menos diversidade. O resultado é somado com  $ai$ , que representa quantas vezes a área das medições somadas cabem no quadrante  $q$ . Esse termo impede que medições com pouca interseção dentro do quadrante tenham uma influência desproporcional no resultado final.

Além disso, há uma penalização aplicada aos quadrantes que apresentam poucos satélites diferentes. Quando a quantidade de satélites diferentes em um quadrante é maior do que um limite pré-estabelecido, chamado de `threshold_satellite`, não há penalizações, ou seja, o valor de  $c$  é igual a 0. Por outro lado, quando o quadrante não atinge o limite, uma parcela de  $ai$  é descontada do resultado final.

Essa etapa é considerada concluída quando todos os quadrantes estão avaliados seguindo as equações apresentadas. Desta forma, a avaliação de um quadrante não depende de outras avaliações de seus vizinhos ou outra forma de dependência de dados, o que possibilita a avaliação paralela dos quadrantes.

### 3.5 Cálculo da área queimada

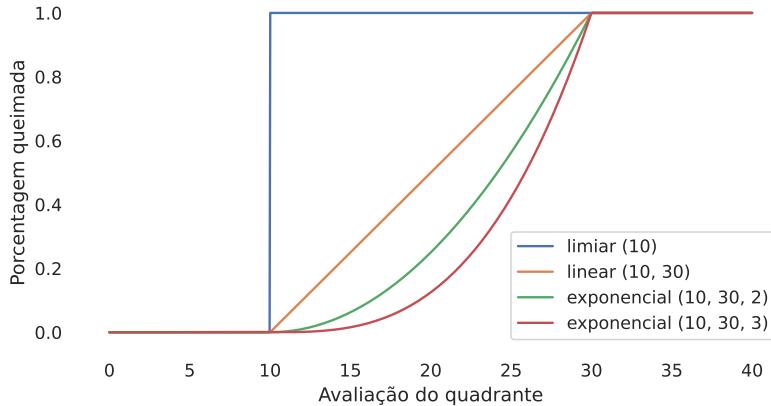
Finalmente, com os quadrantes avaliados, é possível estimar a área queimada para cada quadrante. É importante que o método de cálculo seja flexível, permitindo que o pesquisador teste diferentes métodos. Uma forma de fazer isso é atribuir um número que represente a porcentagem de área queimada dentro do quadrante e, em seguida, multiplicá-lo pela área total do quadrante para obter a área queimada dentro do quadrante. Após calcular a área queimada de cada quadrante, é necessário somar todos esses valores para obter a área total queimada.

Nesse sentido, o papel do pesquisador é definir uma função `eval(v, vs)` que receba o valor do quadrante `v`, calculado na etapa anterior, e retorne um número real entre zero e um. Essa função também pode receber como parâmetro a lista de todos os valores de avaliação dos quadrantes `vs`, que podem ser usados para normalizar o cálculo ou realizar uma análise estatística. Por exemplo, o pesquisador pode definir uma função que indique que o quadrante só deve ser considerado totalmente queimado se o valor do quadrante for maior que o percentil 85, ou seja, maior que pelo menos 85% dos valores.

Para facilitar o trabalho do pesquisador, a implementação pode fornecer funções embutidas comuns que definem a função `eval(v, vs)`. Na Figura 3.5, são apresentadas algumas possibilidades de definições para essa função. A função mais simples é chamada de `limiar(threshold)`, que estabelece que toda a área do quadrante deve ser considerada queimada se a avaliação for maior ou igual a `threshold` e nenhuma área deve ser considerada queimada se não alcançar esse valor. Outra função simples é a `linear(vmin, vmax)`, que faz o valor da área queimada crescer de forma linear dentro dos valores `vmin` e `vmax`. A última função é a `exponencial(vmin, vmax, exponent)`, semelhante à `linear`, mas elevando o resultado a `exponent`, o que faz o valor crescer mais lentamente no início e de forma mais acentuada no final do intervalo.

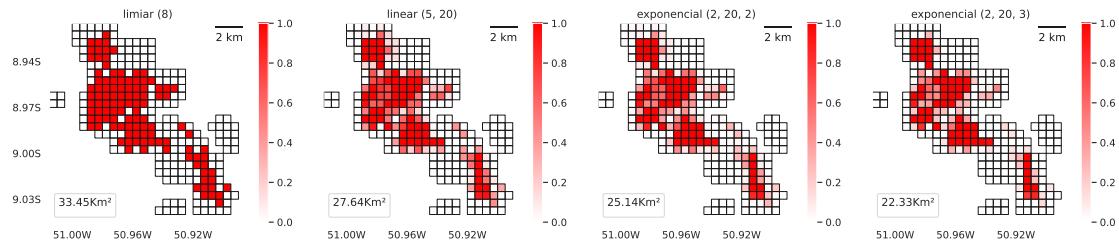
A Figura 3.6 ilustra claramente como diferentes funções e parâmetros podem afetar significativamente os resultados na determinação da área queimada. Cada gráfico mostra a área queimada total calculada usando a função de área queimada indicada no título. Observa-se que a função `limiar` identifica mais áreas como queimadas, pois trata de forma igual os quadrantes com diferentes níveis de intensidade. No entanto, as funções exponenciais e linear parecem mais próximas da realidade, já que consideram apenas uma porcentagem da área dos quadrantes próximos aos agrupamentos de alta intensidade como queimados.

Figura 3.5 – Funções embutidas para cálculo de área.



Fonte: O Autor

Figura 3.6 – Exemplo da aplicação das funções embutidas.



Fonte: O Autor

Também são possíveis funções mais complexas que utilizam análises estatísticas da distribuição, como a função `percentil` (`pmin`, `pmax`, `vmin`, `vmax`). Essa função define um percentil mínimo (`pmin`) e um percentil máximo (`pmax`) como os limiares mínimos e máximos para considerar o quadrante como parcialmente queimado, assim como na função `linear`. Valores abaixo do limiar mínimo são considerados não queimados, enquanto valores acima do limiar máximo são considerados totalmente queimados. Para evitar limiares discrepantes, também são utilizados limites fixos `vmin`, `vmax` e a média desses limiares, chamada de `vmean` ( $(pmin + pmax)/2$ ). O valor do limiar mínimo é restrigido ao intervalo `vmin` e `vmean`, enquanto o limiar máximo deve ficar entre `vmean` e `vmax`.

Desta forma, a função ajusta seus limiares de acordo com o conjunto de dados, se comportando de forma adequada para diferentes tipos de distribuição e ainda mantém uma margem de segurança para evitar classificar áreas utilizando valores discrepantes. Funções que utilizam análises estatísticas da distribuição produzem resultados melhores

que as com limiares fixos. Uma avaliação mais detalhada sobre a escolha dessas funções é apresentada na Seção 4.2.

### 3.6 Implementação do método

Após a formalização do método de forma abstrata, apresenta-se a implementação de fato do método. Para toda a implementação e análise dos dados, foram utilizadas diversas ferramentas do ecossistema Python para *Data Science*, tais como *NumPy*, *Pandas* e *Matplotlib*. Além disso, foram empregadas bibliotecas específicas para análise de dados geográficos, como *GeoPandas*, *Pysal*, *Xarray* e *Shapely*. Para garantir a reproduibilidade das execuções e a organização do código, foi adotado o Jupyter Notebook. Todos o código da implementação do método, da elaboração do documento e demais artefatos gerados durante o projeto podem ser encontrados em <<https://github.com/josebraz/INPE-Queimadas>>, disponibilizados sob a licença MIT.

A biblioteca *GeoPandas*, amplamente utilizada na implementação, é uma extensão da biblioteca *Pandas* que adiciona uma coluna especial chamada “geometry” no *DataFrame* do *Pandas*, chamado de *GeoDataFrame*. Essa coluna permite armazenar as coordenadas e o formato geográfico dos dados em um sistema de coordenadas pré-definido. Além disso, o *GeoPandas* tem integração com a biblioteca *Shapely*, que é usada para realizar cálculos e operações em estruturas geométricas espaciais. Com o uso dessas bibliotecas, é possível executar operações avançadas baseadas na teoria dos conjuntos em elementos geométricos, como união, interseção, diferença, entre outras, além de calcular a área de qualquer polígono que represente uma área no espaço.

Para implementar a etapa de mapeamento dos focos para áreas (Seção 3.3), foi necessário obter as informações de posição, altitude e inclinação dos satélites, a fim de aplicar os ajustes necessários às áreas resultantes. Para os satélites geoestacionários, as coordenadas de posição e altitude são fixas e foram codificadas diretamente no código. No entanto, as posições dos satélites polares são variáveis ao longo do tempo. Para determinar a posição aproximada em um determinado momento, utilizou-se a biblioteca *pyorbital*, que implementa os cálculos necessários.

Com a posição do satélite polar, podemos determinar a direção da trajetória observando a diferença em sua posição entre dois momentos de tempo próximos. Se a trajetória for descendente, a área é rotacionada com uma inclinação negativa; caso contrário, é rotacionada com uma inclinação positiva. Além da rotação, também é aplicado

um escalonamento na área da medição com base na distância entre a medição e o satélite. No caso do satélite GOES-16, que possui uma altitude de aproximadamente 36.000 km, o sensor possui uma resolução de cerca de 2 km no equador. No entanto, em regiões distantes de sua posição, ocorre uma distorção e a área correspondente da medição deve ser maior. Para simplificar e otimizar a implementação dessa etapa, aplicamos o escalonamento da área apenas nas medições dos satélites geoestacionários, pois são os satélites que estão mais distantes das medições.

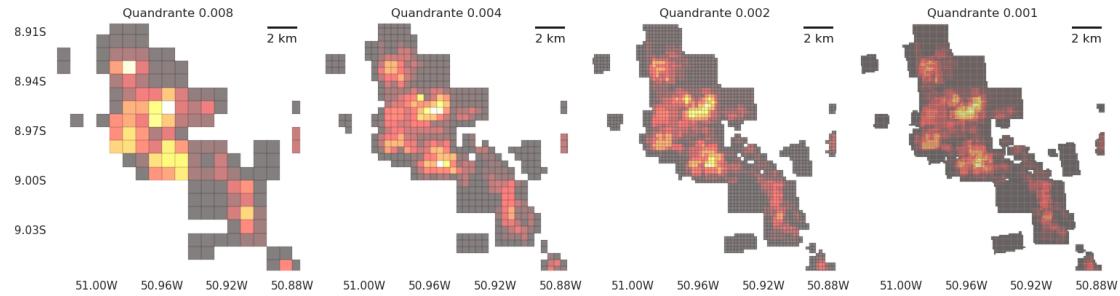
A implementação da etapa de separação e avaliação de quadrantes (Seção 3.4) fez uso da operação Junção Espacial (*Spatial Join*) suportado pela biblioteca *GeoPandas*. A Junção Espacial permite combinar diferentes conjuntos de dados com base em sua relação espacial, de maneira semelhante a uma junção em um banco de dados, porém levando em consideração a proximidade geográfica. É possível utilizar diferentes critérios de proximidade geográfica, como verificar se um ponto está dentro de um polígono ou determinar a interseção entre dois polígonos. No contexto da implementação, a Junção Espacial é utilizada para obter todas as medições que intersectam cada quadrante (interseção entre dois polígonos), que é a principal informação usada para avaliar um quadrante.

A separação em quadrantes também foi planejada para melhorar o desempenho do processamento. Ao dividir o problema em partes menores, é possível resolver cada uma de forma paralela, pois as avaliações dos quadrantes são independentes entre si. Além disso, a separação em quadrantes otimiza ainda mais o processamento da junção espacial, que se beneficia quando os quadrantes para a interseção não são muito grandes. No entanto, é importante considerar que quanto maior o quadrante utilizado, menos avaliações são necessárias, mas a precisão da área queimada também é reduzida.

Com base em experimentos, foi constatado que o uso de quadrantes muito pequenos (com menos de 200m) não aumentam significativamente a precisão dos resultados e tornam a execução muito mais demorada. Como se trata de um espaço bidimensional, a complexidade computacional da avaliação em relação à quantidade de quadrantes é  $O(n^2)$ . A Figura 3.7 mostra a diferença entre quatro tamanhos de quadrantes (em graus): 0,008; 0,004; 0,002 e 0,001.

A implementação da etapa de cálculo da área queimada (Seção 3.5) consistiu em aplicar um mapeamento dos valores dos quadrantes avaliados com a função área queimada. Depois é feito uma multiplicação desse valor com a área física abrangida pelo quadrante, calculada usando a biblioteca *pyproj* que retorna a área em metros. Por último, para saber a área queimada total foi aplicado um somatório de todos esses valores.

Figura 3.7 – Diferença da avaliação para quatro tamanhos de quadrantes.



Fonte: O Autor

O código foi estruturado basicamente com duas classes. A classe principal é chamada de `SatellitesExplore` e é responsável por ter a implementação do AQ-FA. Essa classe recebe os parâmetros de configuração, contém o cálculo da avaliação e da área queimada, bem como funções usadas para gerar gráficos dos resultados obtidos pelo modelo. A classe `SatellitesMeasureGeometry` ficou responsável por resolver o problema das áreas das medições dos satélites. Ela recebe os pontos dos focos e retorna uma estrutura `GeoDataFrame` com os polígonos representando as áreas das medições.

Após a apresentação da implementação do método, é possível utilizar o código para a validação e resultados no próximo capítulo. No processo de validação também foram desenvolvidos alguns pequenos programas para geração das imagens e análises estatísticas em um Jupyter Notebook separado, que também está disponível no repositório.

## 4 RESULTADOS E DISCUSSÃO

Neste Capítulo é apresentado a validação do AQ-FA comparando-o com o estado da arte do cálculo de área queimada. Primeiro é discutido a metodologia de avaliação escolhida, depois a avaliação é aplicada ao produto e é discutido sua qualidade. Por fim, os resultados do AQ-FA para todo o território brasileiro são apresentados e discutidos.

### 4.1 Metodologia da avaliação

Para a avaliação, usamos uma Tabela de Contingência modificada (ver Tabela 4.1) de forma a acomodar valores contínuos, representando o percentual de área queimada (*soft classification*). Esse método de avaliação foi baseado em Libonati et al. (2015), que também aplicou a mesma técnica de avaliação para avaliar o produto AQ1km. A Tabela apresenta duas classes de interesse, “Queimado” e “Não queimado”, de forma a comparar elas com um método de referência, nas colunas, e um método de interesse, nas linhas. A partir dessa comparação surgem os valores de Verdadeiro Positivo (TP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (TN).

Tabela 4.1 – Tabela de Contingência.

		Referência		
		Queimado	Não queimado	Total
Método proposto	Queimado	TP	FP	TP + FP
	Não queimado	FN	TN	FN + TN
	Total	TP + FN	FP + TN	TP + FP + FN + TN

Fonte: Libonati et al. (2015), modificado pelo autor

A versão tradicional da Tabela de Contingência, também chamada de matriz de erro ou matriz de contingência, é comumente usada para avaliar problemas de classificação pesada (*hard classification*), em que para cada objeto é atribuído uma classe única e sem classificações probabilísticas. No entanto, quando se trata de mapear a área queimada, a abordagem é um pouco mais complexa devido à natureza contínua e gradual da queima. Em vez de apenas classificar um quadrante como queimado ou não queimado, é necessário considerar a porcentagem da área efetivamente queimada dentro do quadrante. Essa abordagem é documentada em Binaghi et al. (1999) para a *soft classification* utilizando a teoria dos conjuntos disjuntos.

Seja  $p_r$  a avaliação de um quadrante pelo método de referência e  $p_p$  a avaliação pelo método proposto, utilizamos as equações definidas na Equação 4.1 para calcular os valores TP, FP, FN e TN da Tabela de Contingência. Essas equações podem ser compreendidas usando a noção intuitiva de conjuntos, em que a operação  $\min$  representa a união e a subtração representa a diferença de dois conjuntos. Desse forma, a união de  $p_r$  e  $p_p$  indica o valor em que tanto o método proposto quanto o de referência concordam, resultando nos valores de TP. O mesmo pensamento vale também para o TN. Para os valores de erros, FN e FP, o resultado é extraído a partir de uma diferença que satura em zero. O valor de FP é zero quando  $p_p \leq p_r$ , indicando que não há superestimação. Da mesma forma, o valor de FN é zero quando  $p_r \leq p_p$ , indicando que não há subestimação.

$$\begin{aligned} TP &= \min(p_r, p_p) \\ TN &= \min(1 - p_r, 1 - p_p) \\ FP &= \max(p_p - p_r, 0) \\ FN &= \max(p_r - p_p, 0) \end{aligned} \tag{4.1}$$

Com esses dados, calculamos algumas métricas que nos ajudam a entender os resultados e melhor avaliar a solução proposta. A acurácia geral (OA), definido na Equação 4.2, e o Índice de Sucesso Crítico (CSI), definido na Equação 4.3, estão relacionadas. Enquanto a primeira leva em conta todos os dados de classificação, a segunda omite a informação dos verdadeiros negativos. Isso pode ser bom para o caso das queimadas, em que a classe “Não queimado” é muito mais frequente que a classe “Queimado”. Para esses caso, pode-se entender que a métrica CSI foca na classificação correta da classe que seria “difícil” de acertar.

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \tag{4.2}$$

$$CSI = \frac{TP}{TP + FP + FN} \tag{4.3}$$

Outras duas métricas importantes relacionadas são o Erro de Omissão (OE), definido na Equação 4.4, e o Erro de Comissão (CE), definido na Equação 4.5. O OE representa a proporção de elementos que deveriam ser classificados como queimados, mas foram erroneamente classificados como não queimados, ou seja, houve uma omissão. Por outro lado, o CE representa a proporção de elementos que foram classificados

como queimados, mas que deveriam ser classificados como não queimados, em outras palavras, houve uma comissão. O valor dessas métricas fica entre 0 e 1 e quanto menor, mais preciso é o modelo.

$$OE = \frac{FN}{FN + TP} \quad (4.4)$$

$$CE = \frac{FP}{FP + TP} \quad (4.5)$$

O Viés (B) é um indicativo de qual a tendência que o modelo seguiu. Um valor menor que 1 indica que o modelo está classificando menos áreas queimadas do que deveria. Caso o valor for maior que 1, está classificando mais áreas queimadas comparado a referência. Quando a métrica vale 1 a quantidade de áreas classificadas como queimada é a mesma. Essa métrica não deve ser usada como indicativo de precisão porque, mesmo quando vale 1, pode não haver correspondência com a classificação correta de áreas queimadas do modelo com as áreas queimadas da referência.

$$B = \frac{TP + FP}{TP + FN} \quad (4.6)$$

O Coeficiente de Dice (DC), também conhecido como Índice de Sørensen–Dice, é muito usado em processamento de imagens para calcular a similaridade de duas imagens, fazendo um sobreposição de seus valores (ALBUQUERQUE et al., 2016). O valor fica entre 0 e 1, em que 1 representa uma sobreposição perfeita e valores próximos de 0 indicam uma maior taxa de erro. A equação de DC é similar a CSI, porém com um peso maior aplicado aos Verdadeiros positivos.

$$DC = \frac{2 * TP}{2 * TP + FP + FN} \quad (4.7)$$

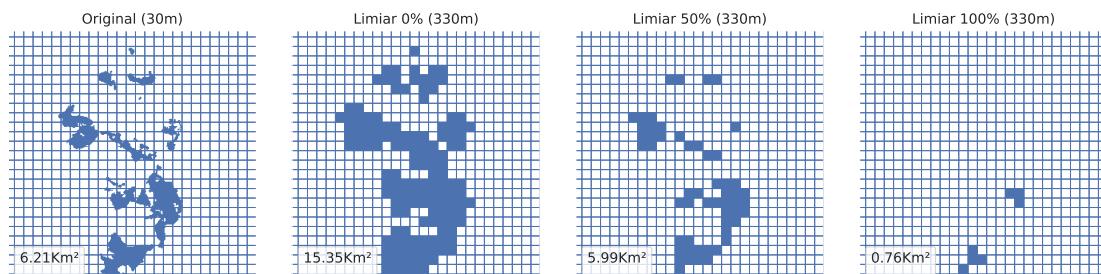
### Avaliações de produtos com resoluções diferentes

Para aplicar a Tabela de Contingência, é necessário que os dados avaliados tenham a mesma resolução e que seus pixels estejam alinhados. Diversos produtos mencionados neste documento possuem resoluções distintas, tais como AQ30m, AQ1km, MCD64A1 e FireCCI51, que têm resoluções de 30 m, 1 km, 500 m e 250 m, respectivamente. Existem várias formas de realizar essa transformação, e em Boschetti, Flasse e Brivio (2004), é

apresentado um método que minimiza os erros de comissão e omissão ao converter dados de áreas queimadas de resoluções maiores para resoluções menores.

Esse estudo classificou os quadrantes de forma binária, em “Queimado” e “Não queimado”, o que afeta o resultado total da área queimada. A Figura 4.1 ilustra a diferença no total da área queimada entre a versão original, com resolução de 30m, e três variações de limiares em resoluções maiores. Para cada quadrante, os valores foram obtidos fazendo uma interseção entre os limites do quadrante e o polígono original. Em seguida, apenas os quadrantes que apresentaram um valor acima do limiar testado foram filtrados. Ou seja, todos os quadrantes maiores que 0, 0,5 ou igual a 1 são apresentados, respectivamente, nas segunda, terceira e quarta imagens.

Figura 4.1 – Redução da resolução do produto AQ30m (a esquerda) para uma resolução de aproximadamente 330m usando diferentes limiares. A segunda imagem contém todos os quadrantes com valores maiores que 0. A terceira imagem mostra quadrantes que tem pelo menos metade de sua área coberta. A última imagem mostra apenas quadrantes totalmente cobertos pelo polígono original.

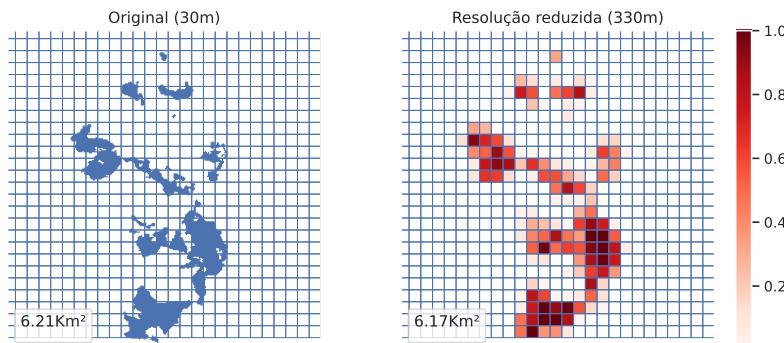


Fonte: O Autor

Outra abordagem para transformar uma resolução maior em uma menor é manter o valor contínuo resultante da interseção das extremidades do quadrante com o polígono original. Com isso, os problemas dos erros de comissão e omissão da abordagem anterior não estão presentes. Essa abordagem é ilustrada na Figura 4.2, em que a área queimada praticamente não foi alterada entre as duas imagens. A diferença foi de cerca de 40m a menos em relação à imagem original, o que pode ser atribuído ao método usado para cálculo da área espacial a partir do polígono ou a problemas de arredondamento de ponto flutuante.

Esse método contínuo é mais adequado para trabalhar com áreas queimadas, pois não distorce tanto o somatório final, que é uma métrica importante para fazer comparações entre produtos que envolvem área queimada. Além disso, o AQ-FA não requer uma classificação pesada, o que torna viável a utilização desse método de transformação. Por-

Figura 4.2 – Redução da resolução do produto AQ30m (a esquerda) para uma resolução de aproximadamente 330m sem perda significativa de área queimada. Cada quadrante na imagem da direita representa a porcentagem do quadrante que está classificado como queimado.



Fonte: O Autor

tanto, este é o método escolhido na validação do AQ-FA, que é apresentado na próxima seção.

## 4.2 Validação do AQ-FA

Nesta seção, é apresentada a validação do AQ-FA em relação ao estado da arte. Para isso, utilizou-se o produto AQ30m como referência devido à sua alta resolução e baixa taxa de erros prevista, que está abaixo de 10% (MELCHIORI et al., 2014). Além disso, é um produto desenvolvido especialmente para o Brasil e seus dados são avaliados por especialistas do INPE. Os dados do AQ30m usados foram obtidos pelo site do Programa Queimadas (<<https://queimadas.dgi.inpe.br/queimadas/aq30m>>). Cada arquivo é identificado com o nome do satélite, a data da passagem e a órbita-ponto da região coberta.

No processo de validação com o AQ30m, não é utilizado o primeiro produto processado após uma sequência de imagens não processadas. Essa abordagem foi adotada para evitar que o AQ30m incluísse áreas que foram queimadas há mais de 16 dias, que é o intervalo de revisita do satélite, mas não foram computadas nas imagens anteriores. Isso garante que o método AQ-FA seja avaliado de forma justa, pois evita discrepâncias entre os resultados do AQ30m e os dados de focos ativos usados no método para o intervalo de tempo estabelecido, que poderiam prejudicar sua performance.

Para garantir uma correta comparação entre o AQ-FA e o produto de referência AQ30m, é essencial filtrar adequadamente os dados de focos ativos para o período de

tempo correspondente. Para o satélite Landsat-8, principal satélite usado no produto, a passagem na linha do equador é aproximadamente às 10 horas da manhã no horário local e cobre a mesma região a cada 16 dias. Portanto, os dados de focos ativos precisam ser filtrados para cobrir esses 16 dias anteriores à passagem do satélite e limitados ao horário das 10 horas da manhã. Por exemplo, se o satélite passou dia 26 de agosto para uma determinada região, o intervalo de tempo dos focos ativos dever ser entre 10 de agosto até 26 de agosto, ambos às 10 horas da manhã.

Mesmo o AQ-FA possibilitando a parametrização da resolução do resultado final, seria inviável e improdutivo calcular o produto utilizando a resolução de 30m do AQ30m devido a complexidade computacional. Foi decidido transformar a resolução do AQ30m para algo em torno de 330m, utilizando a abordagem contínua apresentada na seção anterior. Como mostrado anteriormente, isso não altera significativamente a área queimada e possibilita que o resultado seja computado mais rapidamente.

Além do obstáculo das resoluções diferentes entre os produtos, um outro problema é a escolha da porção territorial a ser avaliada. Para simplificar e possibilitar a comparação entre diversos trabalhos mencionados (como em Libonati et al. (2015)), foi escolhido usar as áreas delimitadas pelo WRS-2, descritas na Figura 2.2. O produto AQ30m também é distribuído seguindo o mesmo sistema, mas com uma diferença que demanda atenção. No sistema WRS-2, as órbitas-ponto podem ter interseção com outras órbitas-ponto vizinhas. Caso isso fosse mantido no produto AQ30m, poderia haver contagens repetidas de uma mesma área queimada nas extremidades da área delimitada pela órbita-ponto. Para solucionar esse problema, foi usado uma versão modificada da grade WRS-2 sem sobreposição, obtida a partir do site do DGI (<http://www.dgi.inpe.br/documentacao/grades>) sendo esta a mesma usada no AQ30m.

## **Escolha dos parâmetros do AQ-FA**

Antes de começar a avaliação, é necessário definir os parâmetros passados ao AQ-FA. Para a validação do método e todo o restante do trabalho utilizamos parâmetros fixos definidos da seguinte forma: a constante `min_area` vale 0.1; a constante `threshold_satellite` teve valor 2 atribuído; e a constante `quadrant_width` recebeu o valor de 0.005. Essa combinação de parâmetros teve o intuito de gerar resultados com menos erros de comissão, penalizando quadrantes com dados de apenas dois satélites diferentes, o que pode indicar menos intensidade de fogo. A resolução do resul-

tado final ficou em torno de 300 m, que é uma resolução comparável aos outros produtos e ainda assim não torna o cálculo muito custoso.

Vários testes foram realizados para encontrar uma função de avaliação da área queimada (`eval(v, vs)`) que reduzisse simultaneamente os erros de omissão e comissão em diferentes regiões e períodos de tempo. Além disso, o resultado final da área queimada também não deve ser muito diferente da referência para que essa métrica seja relativamente confiável. A comparação utilizou o AQ30m como referência para as órbitas-ponto e intervalo de tempo apresentados na Tabela 4.2.

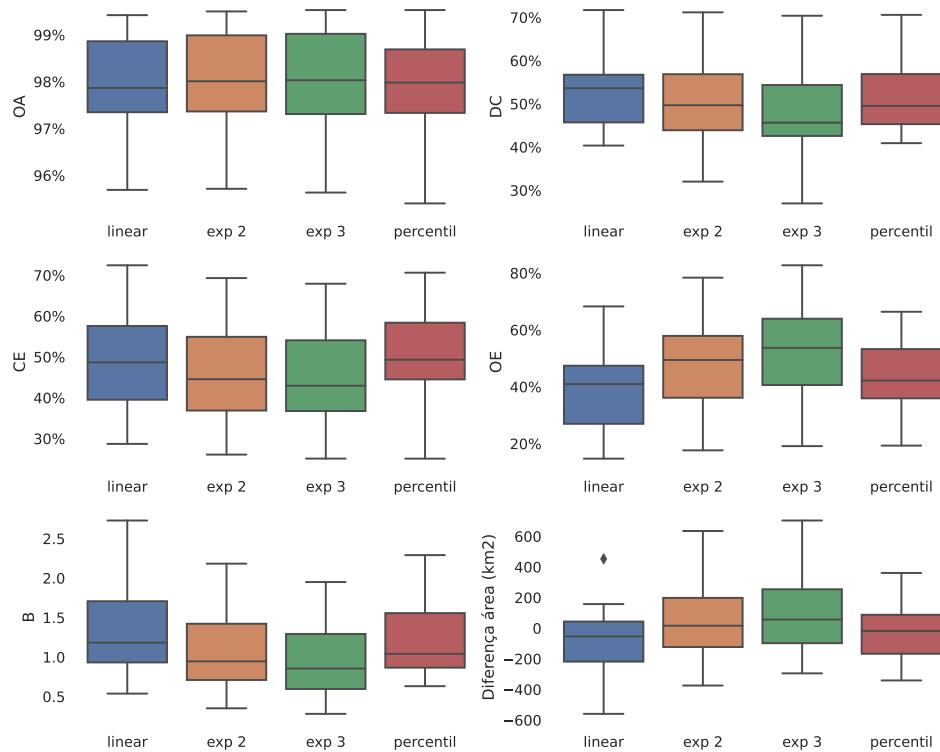
A Figura 4.3 apresenta os dados da comparação em formato de *Box Plot* para as métricas: Acurácia Geral (OA); Coeficiente de Dice (DC); Erros de comissão (CE); Erros de Omissão (OE); Viés (B); e Diferença de área queimada (referência - modelo). Aplicado para quatro funções de avaliação da área queimada: linear, exponencial quadrática, exponencial cúbica e percentil (conforme descritas na Seção 3.5). De forma mais detalhada, as quatro funções apresentadas têm os limiares inferiores `vmin` e superiores `vmax` iguais, equivalentes a 2.5 e 15, respectivamente. Esses valores foram obtidos a partir de experimentos comparando várias combinações. A função `percentil` teve também os parâmetros `pmax`, `pmin` assinalados como 65% e 85%, também obtidos a partir de testes.

A Figura 4.3 mostra que as diferentes funções área queimada apresentam dados relativamente parecidos, mas é possível tirar algumas conclusões importantes. As funções exponenciais (`exp_2` e `exp_3`) apresentaram mais erros de omissão, o que impactou no gráfico de diferença de área queimada e viés, indicando menos áreas queimadas em relação a referência. Além disso, essas funções apresentam os menores DC o que indicam uma menor precisão absoluta. Por esses motivos, as funções exponenciais não foram escolhidas.

As funções `linear` e `percentil` apresentaram os melhores resultados, com alta Acurácia Geral (OA), baixos Erros de Omissão (OE) e alto Coeficiente de Dice (DC). Lembrando que a fórmula da função `percentil` se parece com a função `linear`, mas com limiares decididos a partir de uma análise estatística dos valores. Comparando essas duas funções, a `linear` apresenta mais diferença de área queimada, tanto para mais quanto para menos. Isso se reflete no viés que, ficou entre 2.5 a 0.5 e indica que o erro pode ser de até o dobro da área queimada em relação a referência.

A função `percentil` de forma geral apresentou os melhores resultados nas diferentes métricas e por isso é a função escolhida para todo o restante do trabalho. Principalmente devido aos menores erros nos cálculos de área queimada, representados pelo viés

Figura 4.3 – Box Plot representando seis distribuições, no eixo y: Acurácia Geral (OA), Coeficiente de Dice (DC), Erros de comissão (CE), Erros de Omissão (OE) e Viés (B) e Diferença de área queimada. Cada distribuição foi gerada com as funções área queimada, no eixo x: linear(2.5, 15), exponencial(2.5, 15, 2), exponencial(2.5, 15, 3) e percentil(65, 85, 2.5, 15).



Fonte: O Autor

próximo de 1 e diferença de área queimada próximos a 0. Os Erros de Comissão (CE) foram relativamente maiores, assim como na função `linear`, provavelmente devido a baixa resolução dos satélites provocando erros nas extremidades de uma área queimada. Por esses motivos, `percentil` foi a função escolhida para todo trabalho, tanto na validação quanto nos resultados.

## Análise da validação

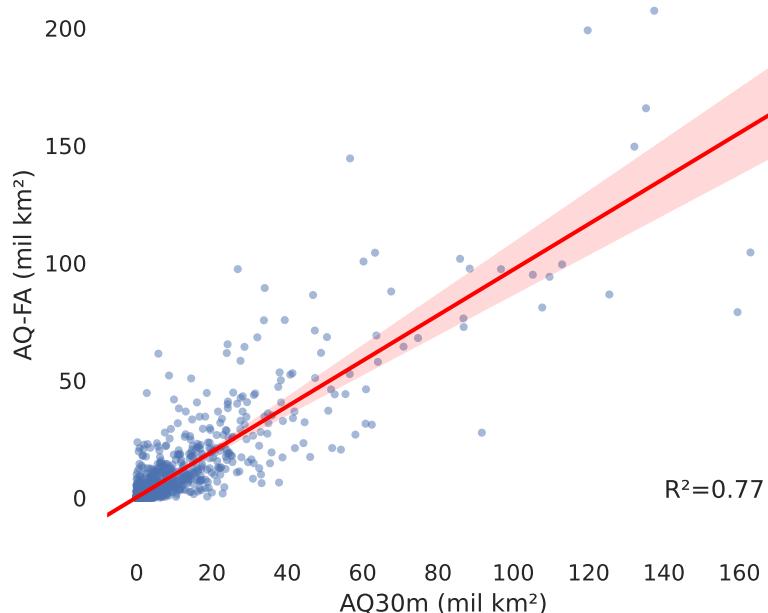
A Tabela 4.2 apresenta os resultados da comparação entre o AQ30m e o AQ-FA para várias órbitas-ponto, nomeadamente: 226/071, 225/069, 224/069, 222/074, 222/067, 221/074, 221/073, 221/070, 221/067, 220/074 e 220/065, abrangendo diferentes períodos de tempo. Essas regiões foram selecionadas manualmente tentando abranger a maior parte do território nacional e diferentes períodos de tempo dentre os disponíveis no site do DGI para o produto AQ30m. O intervalo de tempo considerado para cada órbita-ponto

é especificado na tabela, juntamente com os valores estatísticos de OA, CSI, CE, OE, B e DC, bem como a área queimada estimada pelo AQ30m e pelo AQ-FA. A tabela é ordenada pelo valor do Coeficiente de Dice (DC).

Pela Tabela é possível notar que a área queimada do AQ30m e do AQ-FA são parecidas, variando cerca de 20% para baixo ou pra cima na maioria dos casos. A acurácia geral (OA) tem mediana de 98%, mas apresenta alguns valores discrepantes nas regiões com mais áreas queimadas, como o caso da linha 11 e 19. O DC de maneira geral teve bons resultados, ressaltando que a comparação foi com um produto de alta resolução, variando de 70% a 40% com mediana de cerca de 50%. Os menores DC, como esperado, foram nos testes em que o Viés ficou longe de 1, como nas linhas 17 até 20.

A Figura 4.4 apresenta a porcentagem queimada de quadrantes classificados pelo AQ-FA, no eixo y, e o AQ30m, no eixo x. Os resultados usados para gerar a imagem são os mesmos usados para a construção da Tabela 4.2, porém redistribuídos para uma resolução de aproximadamente 250 Km<sup>2</sup>. Foi aplicado uma regressão linear nos dados que apresentou  $R^2$  igual a 0.77. Esse resultado, no geral, não é muito bom pois indica alta variação entre os dois modelos. Nesse contexto, levando em conta que os dois produtos têm resoluções e metodologias bem diferentes, é um resultado aceitável.

Figura 4.4 – Comparação entre a porcentagem queimada indicada pelo modelo AQ-FA (eixo y) e o AQ30m (eixo x) para quadrantes de tamanho 15 km × 15 km.



Fonte: O Autor

Tabela 4.2 – Regiões e período de tempo da validação.

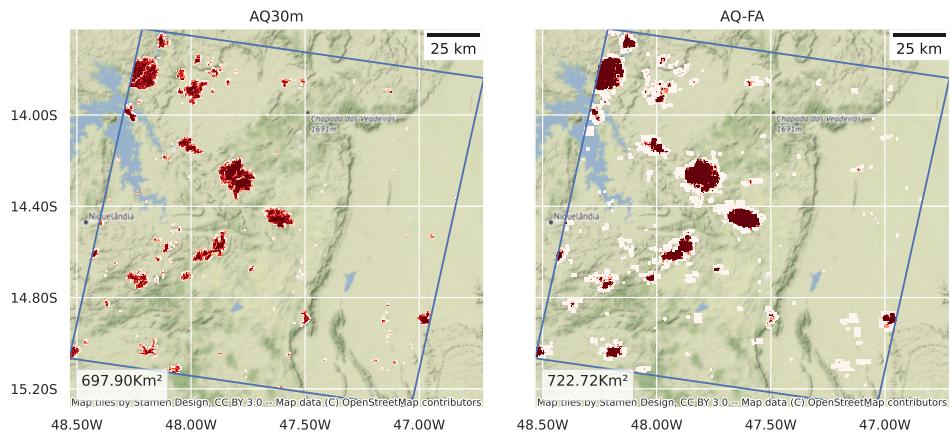
N	Órbita-ponto	Início	OA	CE	B	Área AQ30m
		Fim	CSI	OE	DC	Área AQ-FA
1	221/070	25/08/2017	98.44%	30.62%	1.04	697.9 Km <sup>2</sup>
		10/09/2017	0.55	28.15%	70.58%	722.7 Km <sup>2</sup>
2	221/070	31/08/2019	97.89%	31.37%	0.96	882.2 Km <sup>2</sup>
		16/09/2019	0.51	33.78%	67.39%	849.8 Km <sup>2</sup>
3	224/069	25/08/2021	97.31%	25.10%	0.78	1181.4 Km <sup>2</sup>
		10/09/2021	0.49	41.84%	65.47%	917.9 Km <sup>2</sup>
4	225/069	24/08/2018	99.52%	54.18%	1.76	111.1 Km <sup>2</sup>
		09/09/2018	0.41	19.31%	58.44%	195.4 Km <sup>2</sup>
5	220/074	11/09/2020	97.20%	51.36%	1.49	699.0 Km <sup>2</sup>
		27/09/2020	0.41	27.51%	58.20%	1039.9 Km <sup>2</sup>
6	221/067	31/08/2019	97.34%	28.71%	0.66	1006.1 Km <sup>2</sup>
		16/09/2019	0.39	53.26%	56.45%	659.6 Km <sup>2</sup>
7	221/067	04/08/2021	98.59%	37.67%	0.75	469.0 Km <sup>2</sup>
		20/08/2021	0.36	53.25%	53.42%	351.8 Km <sup>2</sup>
8	226/071	19/09/2019	99.10%	49.53%	1.05	240.2 Km <sup>2</sup>
		05/10/2019	0.35	47.20%	51.60%	251.3 Km <sup>2</sup>
9	222/067	07/09/2019	97.75%	54.97%	1.33	542.4 Km <sup>2</sup>
		23/09/2019	0.35	40.09%	51.41%	720.1 Km <sup>2</sup>
10	221/070	02/10/2019	97.70%	47.10%	0.88	649.6 Km <sup>2</sup>
		18/10/2019	0.33	53.38%	49.55%	571.7 Km <sup>2</sup>
11	220/065	13/08/2021	95.40%	47.69%	0.90	1317.3 Km <sup>2</sup>
		29/08/2021	0.33	53.05%	49.47%	1183.1 Km <sup>2</sup>
12	220/074	30/07/2016	98.51%	60.14%	1.56	292.9 Km <sup>2</sup>
		15/08/2016	0.32	37.97%	48.52%	456.0 Km <sup>2</sup>
13	221/067	19/07/2021	99.02%	48.69%	0.87	271.8 Km <sup>2</sup>
		04/08/2021	0.31	55.34%	47.74%	236.6 Km <sup>2</sup>
14	221/067	03/07/2021	99.06%	49.09%	0.85	257.6 Km <sup>2</sup>
		19/07/2021	0.31	56.53%	46.89%	219.9 Km <sup>2</sup>
15	226/071	28/08/2017	98.06%	62.74%	1.56	369.1 Km <sup>2</sup>
		13/09/2017	0.29	41.89%	45.40%	575.0 Km <sup>2</sup>
16	220/065	03/09/2017	96.64%	57.78%	1.14	783.1 Km <sup>2</sup>
		19/09/2017	0.29	51.79%	45.01%	894.3 Km <sup>2</sup>
17	221/073	05/09/2021	97.55%	66.23%	1.70	412.3 Km <sup>2</sup>
		21/09/2021	0.27	42.50%	42.54%	701.2 Km <sup>2</sup>
18	221/073	02/09/2020	98.96%	69.29%	2.06	151.5 Km <sup>2</sup>
		18/09/2020	0.26	36.65%	41.35%	312.4 Km <sup>2</sup>
19	222/067	29/08/2016	96.56%	46.76%	0.63	978.0 Km <sup>2</sup>
		14/09/2016	0.26	66.38%	41.21%	617.1 Km <sup>2</sup>
20	222/074	25/09/2020	98.38%	70.62%	2.29	214.6 Km <sup>2</sup>
		11/10/2020	0.26	32.69%	40.90%	491.6 Km <sup>2</sup>

Fonte: O Autor

A seguir, é apresentado uma análise mais detalhada para alguns testes a fim de entender melhor os resultados em diferentes casos. É apresentado algumas figuras comparando os dois produtos aplicados a regiões de interesse. Utiliza-se mapas provenientes do OpenStreetMap, para ressaltar as características geográficas, como elevações e regiões aquáticas, bem como nomes de cidades e regiões delimitadas. As figuras apresentam um quadrilátero que delimita a órbita-ponto analisada, fora dessa região delimitada não são mostrados dados de focos para não atrapalhar a análise.

O primeiro caso que chama mais atenção é o que apresentou melhor precisão entre todos os testes, localizado na linha 1 da Tabela 4.2. A Figura 4.5 ilustra a aplicação método AQ-FA, à direita, comparado com a referência, à esquerda, para a órbita-ponto 221/70 em um contexto de altos focos de queima. É possível notar pelo menos três grandes aglomerados de regiões queimadas nas duas imagens. Para o produto AQ30m essas regiões não aparecem completamente queimadas no seu interior, como acontece no AQ-FA. Isso pode ser explicado pela alta resolução do produto de referência e ocasiona um aumento dos erros de comissão do AQ-FA, que ficaram em torno de 30% para esse caso.

Figura 4.5 – Aplicação do AQ-FA na órbita-ponto 221/70 no período de 27 de agosto a 10 de setembro de 2017. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA. Quanto mais vermelho escuro, maior o percentual de área queimada.



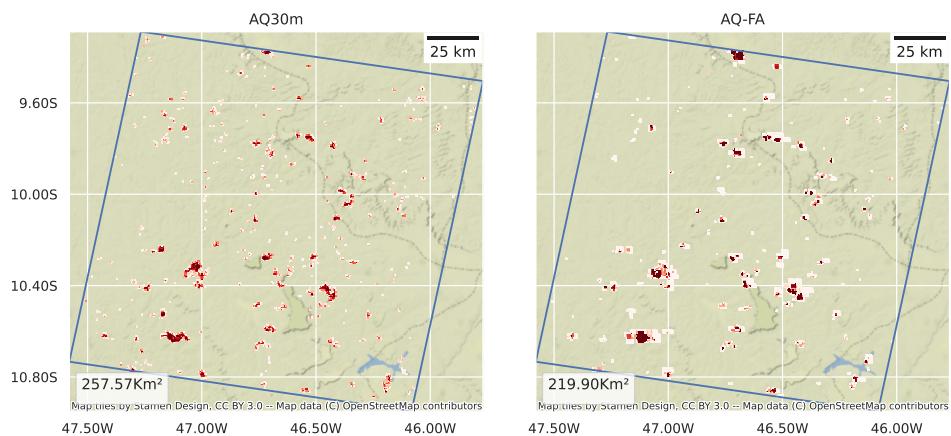
Fonte: O Autor

Regiões de queima menores ou com mais detalhes não foram bem detectadas pelo método testado, como nas regiões próximas as coordenadas 13°90'S 48°00'W e 14°70'S 48°20'W. Para esses dois casos, muitos focos ativos foram detectados, observados pelas áreas em branco próximas a essas regiões, mas não foram suficientemente fortes para indicar uma área de queima maior. Outras queimas menores também não foram bem

detectadas na parte direita da região. Isso fez com que os erros de omissão ficassem em torno de 28%.

Outro caso interessante é o da órbita-ponto 221/67, entre o período de 3 de junho de 2021 a 19 de junho de 2021, indicado na linha 15 da Tabela 4.2. Na Figura 4.6, é possível perceber que as duas imagens não são muito diferentes nas grandes cicatrizes (porções de vegetação queimada), mesmo nesse caso em que a área queimada não é tão grande em comparação com outros casos. A área queimada total no AQ-FA é cerca de 50 km menor, provavelmente devido à imprecisão na detecção das bordas das cicatrizes. Além disso, muitas regiões pequenas que foram classificadas como queimadas pelo AQ30m passaram despercebidas ou não foram suficientemente intensas para indicar uma queima total do quadrante para o método testado.

Figura 4.6 – Aplicação do AQ-FA na órbita-ponto 221/67 no período de 3 de julho a 19 de julho de 2021. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA.

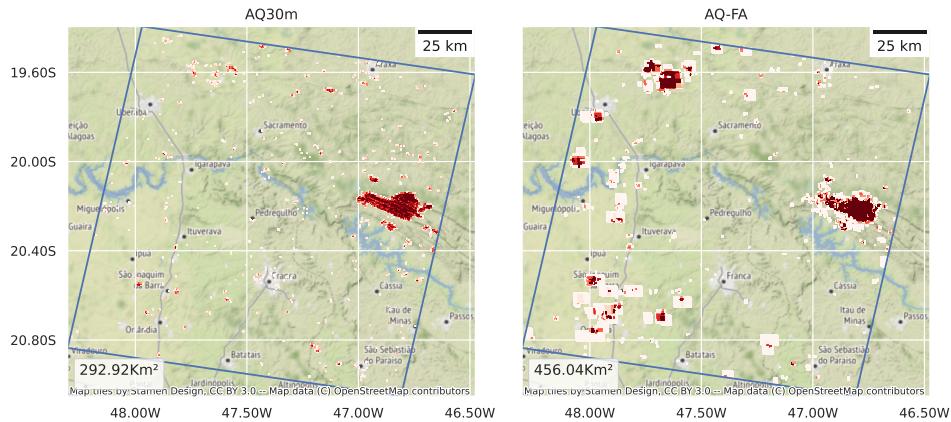


Fonte: O Autor

A Figura 4.7 apresenta principalmente duas manchas de área queimada no AQ-FA e apenas uma grande mancha no AQ30m na órbita-ponto 220/74, que fica no sudoeste de Minas Gerais, entre 30 de julho a 15 de agosto de 2016. A região próxima a coordenada 19°60'S 47°75'W teve alguns focos ativos detectados no final do mês de julho. Com isso, o rastro do fogo nessa região na passagem do satélite provavelmente já estava fraco, indicando uma área queimada pequena.

O processo de validação deixou claro que a técnica proposta é capaz de detectar com relativa precisão áreas grandes de queima, acima de 3 km<sup>2</sup>. A principal dificuldade do AQ-FA está nas bordas das cicatrizes, muito devido à baixa resolução dos sensores dos satélites usados, sendo o mais preciso o Suomi NPP e o NOAA-20, ambos com resolução de 375 m. A escolha do tamanho do quadrante como entrada para o AQ-FA, que foi de

Figura 4.7 – Aplicação do AQ-FA na órbita-ponto 221/74 no período de 30 de julho a 15 de agosto de 2016. A esquerda é o produto AQ30m (referência) e a direita é o AQ-FA.



Fonte: O Autor

aproximadamente 300 m, também pode afetar a precisão do modelo, embora torne mais complexo o processamento dos resultados.

Áreas queimadas muito pequenas, menores que 400 m, podem não ser detectadas corretamente pelo modelo. Na maioria dos casos, a área queimada é subestimada devido à falta de focos ativos. Isso pode ser resultado de queimadas de curta duração ou de baixa intensidade. Queimadas desse tipo podem ser imperceptíveis para satélites com resolução mais baixa ou ocorrer entre as passagens dos satélites.

### 4.3 Resultados apresentados

Como parte dos objetivos deste estudo e após a validação mencionada anteriormente, o método foi aplicado em todo o território nacional, abrangendo o período de 2008 a 2022. Os resultados obtidos foram comparados com o produto AQ1km, que é a referência oficial do INPE para dados de área queimada em todo o Brasil. O intuito é identificar regiões e períodos de tempo que divergem dos dados já conhecidos, a fim de incentivar novas pesquisas sobre esse tema.

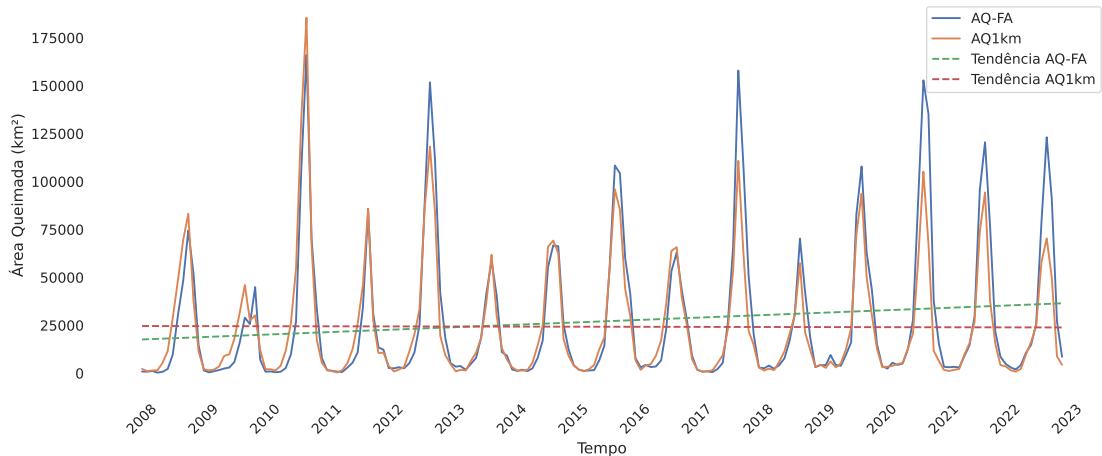
A execução do AQ-FA foi realizada em uma máquina desktop equipada com um processador Intel Core i5-9600KF Hexa-Core de 3.7GHz, 32GB de memória RAM DDR4 3200MHz e um SSD NVMe de 1TB. O sistema operacional utilizado foi o Pop!\_OS, uma distribuição Linux com o kernel na versão 5.16. Durante o processamento, foram

dedicados 5 núcleos do processador exclusivamente para essa tarefa. O tempo total de processamento para esse período demorou aproximadamente 45 horas no total.

Os dados de focos ativos fornecidos pelo INPE foram segmentados por mês e pelas órbitas-ponto que abrangem todo o território nacional. Cada órbita-ponto foi processada na íntegra dentro de um único ano. Essa abordagem otimizou o desempenho, pois permitiu manter em memória a grade correspondente à órbita-ponto sem a necessidade de recalcular-a para cada mês. Esse procedimento foi repetido para todos os anos analisados.

A Figura 4.8 apresenta uma comparação mensal da área queimada obtida pelo AQ-FA, em laranja, e o AQ1km, em azul, ao longo do período de 2008 a 2022, abrangendo todo o território nacional. Em diversos anos, nomeadamente 2008, 2009, 2011, 2013, 2014, 2016 e 2018, os resultados de área queimada foram bastante próximos. Porém, alguns anos, como 2012, 2017, 2020, 2021 e 2022, apresentaram uma diferença significativa. Além disso, como esperado, o caráter sazonal dos focos ativos se manteve para a área queimada nos dois produtos. O AQ-FA indicou uma quantidade igual ou maior de áreas queimadas em todos os anos, com exceção de 2010.

Figura 4.8 – Comparação mensal do total de área queimada, representado no eixo y, a partir dos resultados do AQ-FA e do AQ1km ao longo dos anos de 2008 a 2022, apresentado no eixo x, abrangendo o território brasileiro.



Fonte: O Autor

O AQ-FA mostrou que, durante o período analisado, o ano de 2010 teve uma área queimada menor do que a indicada pelo AQ1km, embora a diferença entre os dois tenha sido pequena nesse caso. Os anos 2012, 2017 e 2020 também se destacaram por terem muitas áreas queimadas contabilizadas, mas com uma diferença mais notável entre os resultados do AQ-FA e do AQ1km. Entre 2018 e 2020, houve um aumento progressivo da

área queimada, e esse patamar elevado se manteve em 2021 e 2022. Contudo, para esses dois últimos anos, o AQ1km apontou uma redução na área queimada, o que não foi corroborado pelo AQ-FA. Em 2022, a maior discrepância entre os dois produtos foi registrada, com o AQ-FA indicando quase o dobro da área queimada em relação ao AQ1km.

Durante o período analisado, os dados do AQ1km demonstram uma tendência de estabilização, com uma leve diminuição. Por outro lado, os resultados do AQ-FA apontam para um aumento na área queimada. O motivo dessa diferença provavelmente é devido a discrepancia entre os últimos anos e também os anos de 2012 e 2017. Essa diferença entre os dois modelos deve ser interpretada como um alerta para os órgãos públicos relacionados à defesa ambiental, uma vez que sugere a possibilidade de um aumento real na extensão das áreas queimadas no Brasil nesse período.

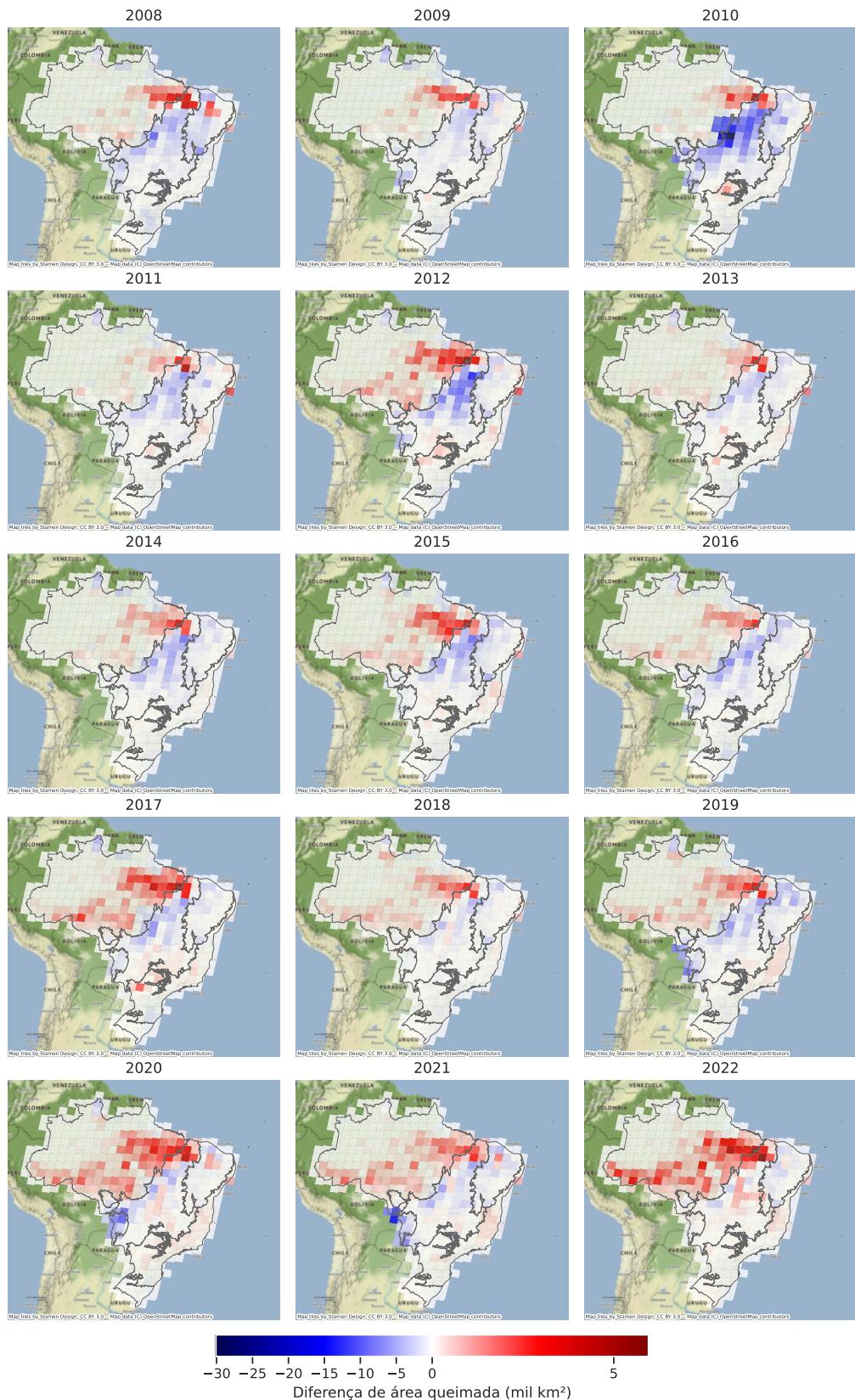
A análise quantitativa revelou que o AQ-FA, em geral, indica uma área queimada maior do que o AQ1km. No entanto, ainda não foi discutida a localização geográfica dessas diferenças. A Figura 4.9 ilustra a variação da área queimada por órbita-ponto em todo o território nacional para os anos de 2008 a 2022. A escala de cores não é linear: tons de azul, que compreendem valores entre 0 e -30 mil km<sup>2</sup>, indicam uma redução na área queimada em relação ao valor do AQ1km. Já tons de vermelho indicam um aumento e representam valores entre 0 e 6 mil km<sup>2</sup>. Quanto mais próximo do branco, menor é a diferença, o que sugere concordância entre os dois modelos.

É notável que, ao longo dos anos, o leste do bioma Amazônico e o norte do Cerrado se destacam consistentemente como aquelas com maior número de áreas queimadas em comparação com as indicações do AQ1km. Com o decorrer dos anos, essa tendência permanece no leste da Amazônia e se expande em direção ao sul e sudoeste desse bioma. O ano de 2022, em particular, apresentou a maior concentração de regiões com mais áreas queimadas, corroborando a discrepancia mencionada anteriormente e com uma predominância de ocorrências na Amazônia.

Em relação ao Cerrado, o AQ-FA apontou, na maioria dos anos, para uma menor extensão de áreas queimadas em comparação com os valores já conhecidos. Entretanto, essa discrepancia diminuiu gradualmente, a ponto de, a partir de 2020, a divergência entre os resultados dos dois modelos praticamente desaparecer. O ano de 2010, apesar de ter registrado uma alta atividade de focos ativos, não se traduziu em uma ampliação significativa da área queimada pelo modelo proposto. Nesse período, é possível observar várias órbitas-ponto desse bioma apresentando menos áreas queimadas em relação a referência.

Os biomas Pampa, Mata Atlântica e Caatinga quase não tiveram diferença na área queimada, o que pode ser explicado pela baixa atividade do fogo nesses biomas, já apresentada anteriormente na Figura 2.6. O Pantanal se manteve constante, exceto nos anos de 2020 e 2021. Nesses anos, que compreendem um longo período de seca nessa região, o AQ-FA não identificou tantas áreas queimadas como o valor conhecido.

Figura 4.9 – Variação da Área Queimada por Órbita-Ponto de 2008 a 2022. A escala de cores está em km<sup>2</sup>, com tons de azul indicando uma diminuição da área queimada em relação ao AQ1km e tons de vermelho indicando um aumento.



Fonte: O Autor

## 5 CONSIDERAÇÕES FINAIS

O AQ-FA foi validado contra os resultados do AQ30m, um produto de alta precisão mantido pelo INPE que abrange algumas regiões do país, em 20 regiões e períodos de tempo diferentes, atingindo uma acurácia geral alta e pouca diferença de área queimada entre os dois modelos. Ao comparar o método proposto com o AQ1km, também um produto do INPE mas com menor precisão, em escala nacional e 15 anos de dados recentes, os resultados indicam que o AQ-FA, de maneira geral, classifica mais regiões como queimadas. Pelo modelo, os últimos anos são caracterizados por um aumento na área queimada, principalmente na região da Amazônia.

O método implementado, seus resultados e os dados empregados estão acessíveis ao público através do link <<https://github.com/josebraz/INPE-Queimadas>>. O código responsável pela geração de todas as imagens, desenvolvido pelo autor, também está disponível. Isso facilita a reproduzibilidade dos resultados e possibilita novas validações. A disponibilização desse arcabouço técnico abre oportunidades para que um público mais amplo se envolva no tema, incentivando o desenvolvimento e aprimoramento de ferramentas que contribuam para a preservação ambiental, tanto no Brasil como no mundo.

O autor espera que esses resultados incentivem futuras pesquisas na área, especialmente na região Amazônica, que foi a área que mais apresentou aumento na área queimada de acordo com o método proposto. Contudo, é importante ressaltar que este estudo não permite afirmar que a região tenha níveis de queimada mais elevados do que os já reconhecidos. No entanto, o estudo deve, no mínimo, servir como um alerta de atenção para essa região de extrema importância para a fauna e flora do planeta.

O método proposto se beneficia quanto mais satélites são utilizados, principalmente na questão de ter múltiplas fontes de dados. Nesta forma, é interessante obter dados de focos ativos provenientes de outros satélites, mesmo que por outras bases de dados que não o INPE. Muitos satélites foram lançados nos últimos 20 anos, que aumentaram as detecções de focos ativos. Muitos outros estão planejados para os próximos anos, com sensores ainda mais potentes. Quanto mais diverso são os satélites, sensores e tipos de órbitas, mais potente em detectar áreas queimadas o método se torna. Nesse sentido, é necessário continuar o desenvolvimento do método para agregar essas novas fontes de dados, uma vez que isso não é um processo automatizado.

Por outros lado, o AQ-FA está sujeito a erros nos diversos algoritmos empregados para detecção de focos ativos, o que pode afetar a confiabilidade dos resultados. O método

não incorpora medidas para mitigar dados discrepantes ou tendenciosos de focos ativos. Além disso, não utiliza outros tipos de dados, como os dados de reflectância empregados em algoritmos de área queimada mais tradicionais. Sem isso, o método precisa inferir uma área queimada com base na intensidade do foco ativo e não propriamente a partir da área já queimada.

Outro desafio associado à utilização desse método é a seleção dos parâmetros. Neste estudo, os parâmetros foram escolhidos através de uma abordagem manual e análise estatística. No entanto, é bastante provável que esses parâmetros não sejam os melhores. Dada a ampla gama de combinações possíveis de parâmetros e funções de área queimada, uma análise completa se torna complexa sem a aplicação de técnicas automáticas. Sem isso, não é possível identificar os limites de precisão que o método pode atingir. Além disso, determinar uma amostra representativa das órbitas-ponto e períodos de tempo que melhor representam o conjunto total é uma tarefa desafiadora que também pode ser aprimorada.

Ainda há muito a desenvolver nesta área de pesquisa, em complemento a este trabalho. Por exemplo, uma vez que os parâmetros escolhidos para o AQ-FA não são ótimos, um trabalho futuro poderia focar em encontrar uma combinação que aumentasse a precisão do método, utilizando um arcabouço matemático mais sofisticado. Nesse sentido, existe a biblioteca Python Hyperopt que lança mão de heurísticas aplicadas a equações que podem variar entre um limite estabelecido. Com isso é possível testar de forma automatizada várias combinações de parâmetros, mas sem garantia de serem ótimos.

Outra possibilidade para uma melhor escolha de parâmetros, mas pode dar um pouco mais de trabalho, é reorganizar as equações descritas na Seção 3.4 para aplicá-las em um sistema de otimização combinatória. Seria necessário fixar uma função área queimada (conforme Seção 3.5) para orientar a minimização dos erros em relação a uma referência. As variáveis desse sistema seriam os parâmetros do modelo, com exceção do `quadrant_width`, que define a resolução desejada, estando sujeitas às restrições definidas pelo método proposto.

Outra importante contribuição para o tema seria a disponibilização dos resultados do AQ-FA por meio de um site interativo. Isso simplificaria o acesso a esses dados, possibilitando que pessoas sem conhecimentos em programação pudessem explorá-los facilmente. Essa abordagem é relevante, visto que ambientalistas e geógrafos, que talvez não possuam esse conhecimento, poderiam beneficiar-se consideravelmente desses dados. Um exemplo dessa ideia é o site do Programa Queimadas, que demonstra ser

uma ferramenta eficaz para promover o avanço científico e empoderar a sociedade com informações confiáveis e abertas.

Como outros produtos de satélites, trabalhos de validação sempre estão sendo desenvolvidos. Nesse sentido, é importantes que novos pesquisadores também produzam trabalhos que se dediquem a validar e a medir a precisão do AQ-FA ou de alguma outra variação de produtos para cálculo de área queimada. Ao fazer isso, todos os produtos envolvidos são beneficiados, uma vez que aumenta o nível de confiabilidade dos modelos, expõe limitações técnicas, identifica melhorias e possibilita ajustes finos nos parâmetros.

Para o Brasil, esses últimos anos foram marcados por um governo que negou a ciência e enfraqueceu a pesquisa e tecnologia do país. O governo tentou controlar a divulgação de resultados de queimadas, retirando a responsabilidade do INPE e transferindo-a para o Instituto Nacional de Meteorologia (INMET), subordinado ao Ministério da Agricultura. Tornaram-se frequentes as notícias de violações ambientais na Amazônia, como as minerações ilegais que prejudicam a sobrevivência da população indígena ribeirinha; e no cerrado, com a queima ilegal da vegetação para a prática de grilagem. Este estudo, portanto, representa também um ato político em prol da ciência no país, da livre divulgação de informações científicas e da preservação ambiental.

## REFERÊNCIAS

ALBUQUERQUE, M. A. et al. **Matemática e Estatística em Foco**. 4. ed. [S.l.]: Universidade Federal da Uberlândia, 2016. v. 2.

BINAGHI, E. et al. A fuzzy set-based accuracy assessment of soft classification. **Pattern Recognition Letters**, v. 20, n. 9, p. 935–948, 1999. ISSN 0167-8655. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167865599000616>>.

BOSCHETTI, L. et al. Lessons to be learned from the comparison of three satellite-derived biomass burning products. **Geophysical Research Letters**, v. 31, n. 21, 2004. Disponível em: <<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GL021229>>.

BOSCHETTI, L.; FLASSE, S. P.; BRIVIO, P. A. Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The pareto boundary. **Remote Sensing of Environment**, v. 91, n. 3, p. 280–292, 2004. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425704000707>>.

BOSCHETTI, L. et al. Global validation of the collection 6 modis burned area product. **Remote Sensing of Environment**, v. 235, p. 111490, 2019. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425719305097>>.

BOSCHETTI, L.; STEHMAN, S. V.; ROY, D. P. A stratified random sampling design in space and time for regional to global scale burned area product validation. **Remote Sensing of Environment**, v. 186, p. 465–478, 2016. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425716303595>>.

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. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425719300689>>.

Divisão de Geração de Imagens. **Geração de imagens**. 2023. Disponível em: <<https://www.dgi.inpe.br/>>. Acesso em: 16 de fev. 2023.

EARTHDATA. **Data Processing Levels**. 2023. Disponível em <<https://www.earthdata.nasa.gov/engage/open-data-services-and-software/data-information-policy/data-levels>>. Acesso em: 3 de jun. 2023.

Embrapa Territorial. **Satélites de Monitoramento**. 2023. Disponível em: <<https://www.embrapa.br/satelites-de-monitoramento/>>. Acesso em: 28 de jan. 2023.

FUCHS, V. B. Chinese-driven frontier expansion in the amazon: four axes of pressure caused by the growing demand for soy trade. **Civitas - Revista de Ciências Sociais**, Pontifícia Universidade Católica do Rio Grande do Sul, v. 20, n. Civitas, Rev. Ciênc. Soc., 2020 20(1), p. 16–31, Jan 2020. ISSN 1519-6089. Disponível em: <<https://doi.org/10.15448/1984-7289.2020.1.34656>>.

GIGLIO, L. et al. The collection 6 modis burned area mapping algorithm and product. **Remote Sensing of Environment**, v. 217, p. 72–85, 2018. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425718303705>>.

GIGLIO, L. et al. An active-fire based burned area mapping algorithm for the modis sensor. **Remote sensing of environment**, Elsevier, v. 113, n. 2, p. 408–420, 2009.

GIGLIO, L. et al. Assessing variability and long-term trends in burned area by merging multiple satellite fire products. **Biogeosciences**, Copernicus GmbH, v. 7, n. 3, p. 1171–1186, 2010.

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. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425716300827>>.

HANTSON, S. et al. Strengths and weaknesses of modis hotspots to characterize global fire occurrence. **Remote Sensing of Environment**, v. 131, p. 152–159, 2013. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425712004610>>.

HUMBER, M. L. et al. Spatial and temporal intercomparison of four global burned area products. **Int J Digit Earth**, Department of Geographical Sciences, University of Maryland, College Park, MD, USA.; Department of Natural Resources and Society, University of Idaho, Moscow, ID, USA.; Department of Geographical Sciences, University of Maryland, College Park, MD, USA.; Department of Geographical Sciences, University of Maryland, College Park, MD, USA., v. 12, n. 4, p. 460–484, 2019. ISSN 1753-8955 (Electronic); 1753-8947 (Print); 1753-8947 (Linking).

INPE. **Área queimada**. 2023. Disponível em: <<https://queimadas.dgi.inpe.br/queimadas/aq1km/>>. Acesso em: 24 de abr. 2023.

INPE. **Programa Queimadas Perguntas frequentes**. 2023. Disponível em: <<http://www.inpe.br/queimadas/portal/informacoes/perguntas-frequentes>>. Acesso em: 23 de jan. 2023.

JÚNIOR, C. A. dos S. et al. Classificação de áreas queimadas por machine learning usando dados de sensoriamento remoto. **Anais do XIX Simpósio Brasileiro de Sensoriamento Remoto**, p. 1784–1787, 2019.

LATORRE, M. et al. Sensor modis: Características gerais e aplicações. v. 0, 01 2003.

LEGRAND, A. **Reproducible Research: where do we stand?** 2019. Disponível em: <[https://www.inf.ufrgs.br/site/wp-content/uploads/2019/08/talk\\_19\\_10\\_08\\_Brazil.pdf](https://www.inf.ufrgs.br/site/wp-content/uploads/2019/08/talk_19_10_08_Brazil.pdf)>. Acesso em: 5 de set. 2023.

LEONEL, M. O uso do fogo: o manejo indígena e a piromania da monocultura. **Estudos Avançados**, Instituto de Estudos Avançados da Universidade de São Paulo, v. 14, n. Estud. av., 2000 14(40), p. 231–250, Sep 2000. ISSN 0103-4014. Disponível em: <<https://doi.org/10.1590/S0103-40142000000300019>>.

LIBONATI, R. et al. On a new coordinate system for improved discrimination of vegetation and burned areas using mir/nir information. **Remote Sensing of Environment**, v. 115, n. 6, p. 1464–1477, 2011. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425711000496>>.

LIBONATI, R. et al. 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.

LIZUNDIA-LOIOLA, J. et al. A spatio-temporal active-fire clustering approach for global burned area mapping at 250m from modis data. **Remote Sensing of Environment**, v. 236, p. 111493, 2020. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425719305127>>.

MARTINS, G. et al. Dinâmica espaço-temporal das queimadas no brasil no período de 2003 a 2018. **Revista Brasileira de Geografia Física**, v. 13, n. 4, p. 1558–1569, 2020.

MELCHIORI, A. E. et al. A landsat-tm/oli algorithm for burned areas in the brazilian cerrado: preliminary results. **Advances in Forest Fire Research**, Coimbra University Press Coimbra, Portugal, p. 23–30, 2014.

MIRANDA, E. de; MARTINHO, P.; CARVALHO, C. de. Dinâmica de queimadas e incêndios na bacia do rio paraguai e no pantanal (2018 e 2020). Embrapa Territorial: Campinas, 2021., 2021.

NASA. **Data Maturity Levels**. [S.I.]: NASA, 2023. Disponível em: <<https://science.nasa.gov/earth-science/earth-science-data/data-maturity-levels/>>. Acesso em: 3 de jun. 2023.

NEPSTAD, D. C. et al. Mortality of large trees and lianas following experimental drought in an amazon forest. **Ecology**, Wiley Online Library, v. 88, n. 9, p. 2259–2269, 2007.

PEREIRA, A. A. et al. Burned area mapping in the brazilian savanna using a one-class support vector machine trained by active fires. **Remote Sensing**, MDPI, v. 9, n. 11, p. 1161, 2017.

PEREIRA, J. A comparative evaluation of noaa/avhrr vegetation indexes for burned surface detection and mapping. **IEEE Transactions on Geoscience and Remote Sensing**, v. 37, n. 1, p. 217–226, 1999.

PRODUCTS, A. D. T. I. **MODIS Standard Collection Versions**. 2023. Disponível em <<https://atmosphere-imager.gsfc.nasa.gov/collections/overview>>. Acesso em: 3 de jun. 2023.

SCHROEDER, W. et al. Validation of goes and modis active fire detection products using aster and etm+ data. **Remote Sensing of Environment**, v. 112, n. 5, p. 2711–2726, 2008. ISSN 0034-4257. Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425708000102>>.

SETZER, A.; MORELLI, F.; SOUZA, J. C. O banco de dados de queimadas do inpe. **Biodiversidade Brasileira-BioBrasil**, n. 1, p. 239–239, 2019.

SETZER, A.; YOSHIDA, M. C. **Detecção de Queimadas nas Imagens do Satélite Geoestacionário GOES-12.** 2004. Disponível em: <[https://queimadas.dgi.inpe.br/~rqueimadas/documentos/relat\\_goes12\\_3\\_4.htm](https://queimadas.dgi.inpe.br/~rqueimadas/documentos/relat_goes12_3_4.htm)>. Acesso em: 1 de fev. 2023.

UNEP, U. N. E. P. **Emissions Gap Report 2022: The Closing Window — Climate crisis calls for rapid transformation of societies.** Nairobi, 2022. Disponível em: <<https://www.unep.org/emissions-gap-report-2022>>.