

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

JOSÉ HENRIQUE DA SILVA BRAZ

**Uma análise dos dados de queimada do
INPE no Brasil (preliminar)**

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

Orientador: Prof. Dr. Lucas M. Schnorr

Porto Alegre
2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Prof^a. Patricia Pranke

Pró-Reitora de Graduação: Prof^a. Cíntia Inês Boll

Diretora do Instituto de Informática: Prof^a. Carla Maria Dal Sasso Freitas

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

Bibliotecário-chefe do Instituto de Informática: Alexsander Borges Ribeiro

*“If I have seen farther than others,
it is because I stood on the shoulders of giants.”*

— SIR ISAAC NEWTON

AGRADECIMENTOS

Agradeço ao L^AT_EX por não ter vírus de macro...

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	6
LISTA DE FIGURAS	7
LISTA DE TABELAS	8
RESUMO	9
ABSTRACT	10
1 INTRODUÇÃO	11
2 CONCEITOS BÁSICOS E VISÃO GERAL DOS DADOS.....	13
2.1 O monitoramentos das queimadas no Brasil.....	13
2.2 Os Satélites.....	14
2.3 Detecção de focos de queimadas e área queimada	17
2.4 Uma visão geral dos dados	18
3 TRABALHOS RELACIONADOS	21
4 METODOLOGIA	22
4.1 Visão geral da metodologia	22
4.2 Mapeamento dos focos para áreas.....	23
4.3 Separação e avaliação de quadrantes.....	25
4.4 Cálculo da área queimada.....	27
5 IMPLEMENTAÇÃO DO MÉTODO	29
5.1 Coleta dos dados.....	29
5.2 Implementação do método	30
6 RESULTADOS E DISCUSSÃO.....	33
6.1 Aplicações do método	33
7 CONSIDERAÇÕES FINAIS	34
REFERÊNCIAS.....	35

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface (Interface de Programação de Aplicação)
CSV	Comma Separated Values (valores separados por vírgulas).
GMT	Greenwich Mean Time
INPE	Instituto Nacional de Pesquisas Espaciais
IBGE	Instituto Brasileiro de Geografia e Estatística
URL	Uniform Resource Locator (Localizador Uniforme de Recursos)
NOAA	National Oceanic and Atmosphere Administration
MODIS	Moderate Resolution Imaging Spectroradiometer
GOES	Geostationary Operational Environmental Satellite
AVHRR	Advanced Very High Resolution Radiometer

LISTA DE FIGURAS

Figura 1.1	Emissões de gases do efeito estufa per capita de 1990 até 2020 (tCO ₂ e/capita).	11
Figura 2.1	Órbita dos satélites no dia 10 de agosto de 2022.	17
Figura 2.2	Relação do montante dos dados por satélite.	19
Figura 2.3	Amostragem por tempo de cada satélite.	19
Figura 2.4		20
Figura 4.1	Diagrama do método.	22
Figura 4.2	Comparação entre pontos e áreas dos focos detectados pelo satélite Suomi NPP.	24
Figura 4.3	Separação entre quadrantes.	25
Figura 4.4	Funções embutidas para cálculo de área.	28
Figura 4.5	Exemplo da aplicação das funções embutidas.	28
Figura 5.1	Cores escolhidas para cada satélite.	32
Figura 5.2	Diferença da avaliação para diferentes tamanhos de quadrantes.	32

LISTA DE TABELAS

Tabela 2.1	Significado de cada coluna dos dados de queimada do INPE.	14
Tabela 2.2	Características dos satélites usados pelo INPE.	15

RESUMO

Este documento é um exemplo de como formatar documentos para o Instituto de Informática da UFRGS usando as classes \LaTeX disponibilizadas pelo UTUG. Ao mesmo tempo, pode servir de consulta para comandos mais genéricos. *O texto do resumo não deve conter mais do que 500 palavras.*

Palavras-chave: Formatação eletrônica de documentos. \LaTeX . ABNT. UFRGS.

Using L^AT_EX to Prepare Documents at II/UFRGS

ABSTRACT

This document is an example on how to prepare documents at II/UFRGS using the L^AT_EX classes provided by the UTUG. At the same time, it may serve as a guide for general-purpose commands. *The text in the abstract should not contain more than 500 words.*

Keywords: Electronic document preparation. L^AT_EX. ABNT. UFRGS.

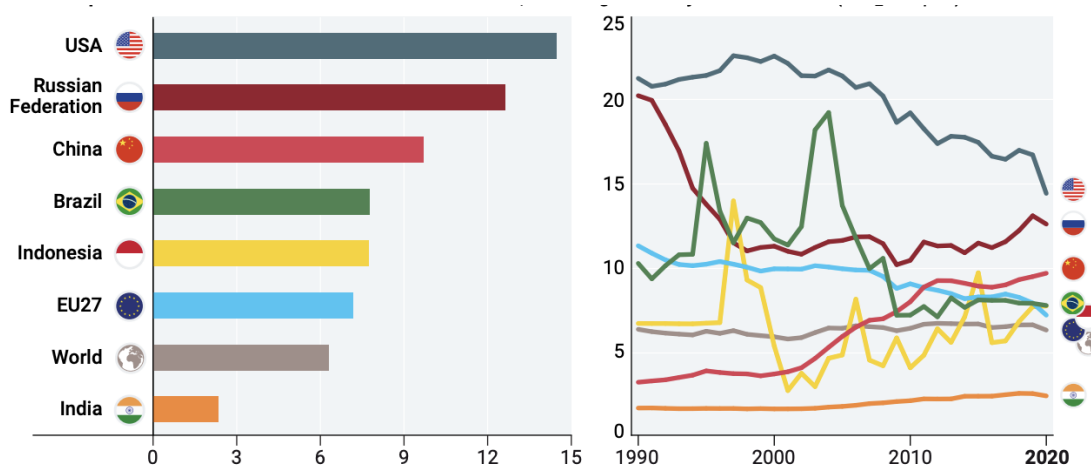
1 INTRODUÇÃO

O fogo é uma tecnologia 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). [P0. O que é uma queimada/fogo]

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). [P1. As queimadas hoje]

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) de 2022. De acordo com o estudo, o valor absoluto se manteve estável desde 2010, e atingiu seu pico por volta dos anos de 2003 a 2004. Assim como a Indonésia, o que melhor explica a alta posição do Brasil neste ranking são as queimadas e o desmatamento da vegetação nativa. Olhando para os municípios do país, dos dez que mais poluem, oito deles estão localizados no bioma da amazônia e não possuem atividades industriais que justificariam esse valor. [P2. queimadas e efeito estufa no Brasil]

Figura 1.1 – Emissões de gases do efeito estufa per capita de 1990 até 2020 (tCO₂e/capita).



Fonte: Emissions Gap Report 2022: The Closing Window

Este trabalho se dedica a estudar e apresentar de forma concisa os dados de focos de queimadas disponibilizados pelo Instituto Nacional de Pesquisas Espaciais (INPE). O principal objetivo é tornar fácil o entendimento desses dados gerados a partir de imagens de satélites, sem a necessidade de um conhecimento prévio das técnicas de ciência de dados e sensoriamento remoto. O escopo de tempo das análises é limitado ao início de 1998, ano que iniciou a base aberta de queimadas, até o final de 2022. [P3. O que é o trabalho em si]

Os dados analisados neste trabalho foram obtidos a partir do DBQueimadas, Banco de Dados de Queimadas <www.inpe.br/queimadas/bdqueimadas>, que é um sistema desenvolvido pelo INPE e acessível de forma aberta por meio da web. Conta com mais de 300 milhões de pontos coletados desde o ano de 1998, provenientes de vários satélites (SETZER; MORELLI; SOUZA, 2019). Ao disponibilizar os dados das queimadas o instituto possibilita que a sociedade retribua com pesquisas e fomenta novas abordagens ao problema das queimadas no Brasil, como é o caso deste trabalho. [P4. Fonte dos dados usados]

... apontaram que a detecção de áreas queimadas pode se beneficiar da fusão de observações de incêndios ativos de vários sensores. (GIGLIO et al., 2010) mas incêndios ativos podem ficar omissos (GIGLIO et al., 2009)

Durante o decorrer do documento são apresentadas diversas figuras, a maioria de construção do próprio autor, a fim de instigar a intuição do leitor para o tópico que está sendo abordado. De início, será abordado questões mais teóricas envolvendo características dos satélites, suas produções de imagens e como são usadas para detectar um foco ativo de queimada. Após isso, [P5. Estrutura do documento]

2 CONCEITOS BÁSICOS E VISÃO GERAL DOS DADOS

Neste capítulo, serão apresentados conceitos importantes para a compreensão do trabalho. Primeiramente, serão formalizadas definições relacionadas às queimadas e à forma como são monitoradas no Brasil. Em seguida, será realizada uma sumarização dos principais satélites utilizados pelo INPE e suas características. Além disso, será abordado como as imagens geradas pelos satélites são utilizadas para a detecção de focos ativos. Por fim, apresentaremos uma análise dos dados para demonstrar uma visão geral das principais tendências e gráficos.

2.1 O monitoramentos das queimadas no Brasil

[P0. definir uma queimada, focos detectados e área queimada]

[P0. Diferença entre foco detectado e área queimada]

[P1. Falar um pouco do INPE e suas divisões]

Dentro do site é possível gerar mapas, tabelas, gráficos e exportar os dados sobre as queimadas no Brasil aplicando diferentes filtros. Todo o programa foi desenvolvido com ferramentas abertas, muitas delas criadas pelo próprio time de tecnologia da informação do INPE (SETZER; MORELLI; SOUZA, 2019). [P2. Falamos sobre o programa DBQueimadas]

O Banco de Dados de Queimadas é um excelente caso de como os dados abertos podem ajudar a sociedade. 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/>, algumas imagens inteiras geradas pelos satélites que o próprio DGI captura e processa. [P3. importancia do dados abertos para a sociedade]

Para obter as imagens brutas dos satélites 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, extraindo os dados de focos ativos de queimadas que alimentam o DBQueimadas. (Divisão de Geração de Imagens, 2023) [P4. Papel do DGI]

É em posse dessas imagens brutas que o INPE aplica algoritmos de detecção de focos de queimadas. 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 das colunas pode ser vista na Tabela 2.1 (INPE, 2023). [P5. O que é uma observação nos dados]

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)
Município	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: O Autor com base em INPE (2023)

[P7. Como é feito o calculo de área queimada pelo INPE]

2.2 Os Satélites

O INPE atualmente processa dados de vários satélites com características distintas entre si. Estão presentes nos dados desde satélites geoestacionários, como o GOES-12, que está a 29.400 km de distância da superfície (SETZER; YOSHIDA, 2004), até satélites com órbitas polares, entre 700 a 900 km de altura.[P0. Visao geral dos satelites]

Abaixo segue um resumo dos satélites usados pelos INPE desde o início da série histórica até final de 2022 (Embrapa Territorial, 2023). Os que estão em funcionamento pleno atualmente são: NOAA-20, NOAA-19, NOAA-18, GOES-16, Suomi NPP, AQUA, TERRA, MSG-03, METOP-B e METOP-C. [P1. Falar sobre os principais]

Tabela 2.2 – Características dos satélites usados pelo INPE.

Nome	Sensor	Resolução esp.	Órbita	Lançamento	Passagem
METOP-C	AVHRR-3	1100m	Polar	2018	21h
NOAA-20	VIIRS	500m	Polar	2017	2h / 14h
METOP-B	AVHRR-3	1100m	Polar	2012	21h
Suomi NPP	VIIRS	500m	Polar	2011	2h / 14h
NOAA-19	AVHRR-3	1100m	Polar	2009	2h / 14h
NOAA-18	AVHRR-3	1100m	Polar	2005	Variadas
AQUA	MODIS	1000m	Polar	2002	2h / 14h
NOAA-17	AVHRR-3	1100m	Polar	2002	21h
NOAA-16	AVHRR-3	1100m	Polar	2000	Variadas
TERRA	MODIS	1000m	Polar	1999	11h / 23h
NOAA-15	AVHRR-3	1100m	Polar	1998	5h / 17h
TRMM	VIRS	2000m	Polar	1997	Variadas
NOAA-14	AVHRR	1100m	Polar	1994	21h
NOAA-12	AVHRR	1100m	Polar	1991	2h / 15h
GOES-16	ABI	2000m	Geoest.	2016	Não se aplica
MSG-03	SEVIRI	3000m	Geoest.	2012	Não se aplica
GOES-13	GOES I-M	4000m	Geoest.	2006	Não se aplica
MSG-02	SEVIRI	3000m	Geoest.	2005	Não se aplica
GOES-12	GOES I-M	4000m	Geoest.	2001	Não se aplica
GOES-10	GOES I-M	4000m	Geoest.	1997	Não se aplica
GOES-08	GOES I-M	4000m	Geoest.	1994	Não se aplica

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

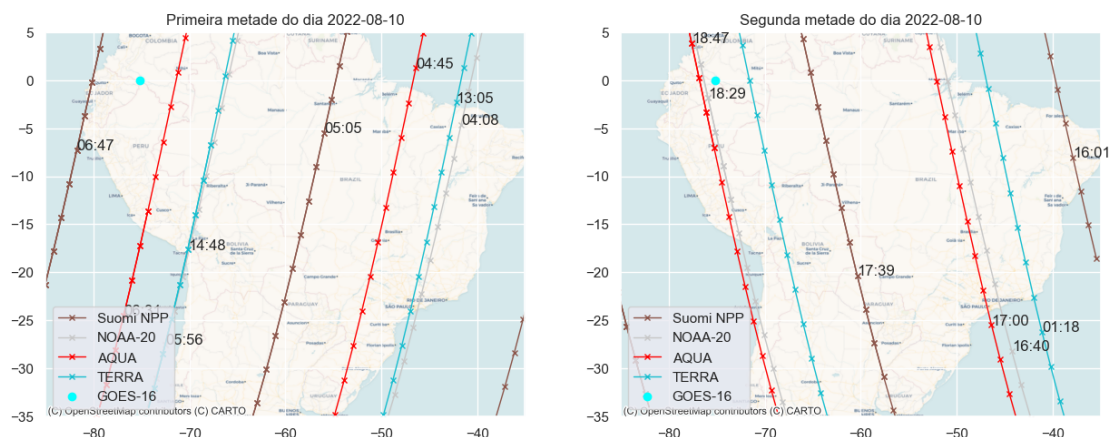
Cada satélite pode ter um sensor imageador, que gera imagens, com características distintas. Nelas são captadas imagens não só no comprimento de onda da luz visível (de 400nm a 700nm), mas também no infravermelho (de 780nm a 1mm). Suas medições são divididas em canais, que variam em resolução espacial e espectral (intervalo de comprimento de onda). Geralmente o primeiro canal é dedicado à luz visível, entre o laranja e o vermelho, e com a maior resolução espacial possível para o sensor. Os outros canais utilizam diferentes intervalos do infravermelho e luz visível. [P2. visão geral dos sensores e porque geram dados diferentes]

Com todas essas diferenças entre os satélites foi necessário estabelecer um satélite base, que ficou conhecido como Satélite de Referência. Ele é usado para estabelecer uma série temporal e permitir análise de tendência durante vários anos dos focos detectados para diferentes regiões. Este balizador precisa cobrir a área do país de forma satisfatória, ou seja, sua órbita não deve distorcer os dados no geral. Além disso, resoluções dos sensores muito baixas, maiores de 1 km, tornam a análise dos focos pouca precisa. [P3. Satélite de referência]

De 01/junho/1998 a 03/julho/2002 o satélite de referência utilizado foi o NOAA-12 com passagem no final da tarde. Depois desse período passou-se a utilizar o AQUA com passagem à tarde (chamado nos dados de AQUA_M-T). O satélite AQUA já ultrapassou a data prevista de encerrar o funcionamento em muitos anos e será descontinuado em breve, quando isso acontecer o satélite Suomi NPP será o novo balizador (INPE, 2023). [P3. Satélite de referência]

Para finalizar, podemos observar na Figura 2.1 a passagem dos satélites no dia 10 de agosto de 2022, gerada a partir de dados do <<https://celestrak.org/>>. Os satélites podem levar vários dias para passar no mesmo local devido a características de sua órbita e podemos ter uma estimativa razoavelmente precisa de sua trajetória ao longo do tempo. Também é possível ver o satélite geoestacionário GOES-16, no norte do Peru, representado com um ponto no gráfico.

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



Fonte: O Autor

2.3 Detecção de focos de queimadas e área queimada

O algoritmo para identificação de focos ativos é específico para cada tipo de sensor. No geral, eles 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, deserticas e que estavam nubladas na hora da passagem. [P2. Visão geral dos algoritmos para detecção]

Para o caso dos satélites TERRA e AQUA, que têm o sensor MODIS, o instituto mantinha seu próprio método de detecção, que produzia dados de maior confiabilidade (INPE, 2023). A partir de 2017 o INPE migrou toda a base de dados para o "Collection 6", algoritmo aplicado pela *National Aeronautics and Space Administration* (NASA), marcando o início da chamada Base 2 de queimadas. Anteriormente, a NASA empregava o Collection 5, que gerava falsos positivos em clareiras florestais e falsos negativos para grandes queimadas obscurecidas por fumaça densa. (SCHROEDER et al., 2008). [P3. Algoritmo empregado pelo INPE]

O collection 6 (GIGLIO; SCHROEDER; JUSTICE, 2016)

Para detectar áreas queimadas, o instituto atualmente usa o produto AQ1km, desenvolvido em parceria com o Laboratório de Aplicações de Satélites Ambientais (LASA)

(ÁREA. . . , 2023) que ainda está em fase Provisória (ainda pode sofrer mudanças e não foi validada completamente). O produto é aplicado nos dados do sensor MODIS e, desta forma, usa os satélites AQUA e TERRA de forma concomitante (LIBONATI et al., 2015). Uma vez que o AQUA foi lançado apenas em 2002, o produto só pode ser usado para dados a partir de 2003.

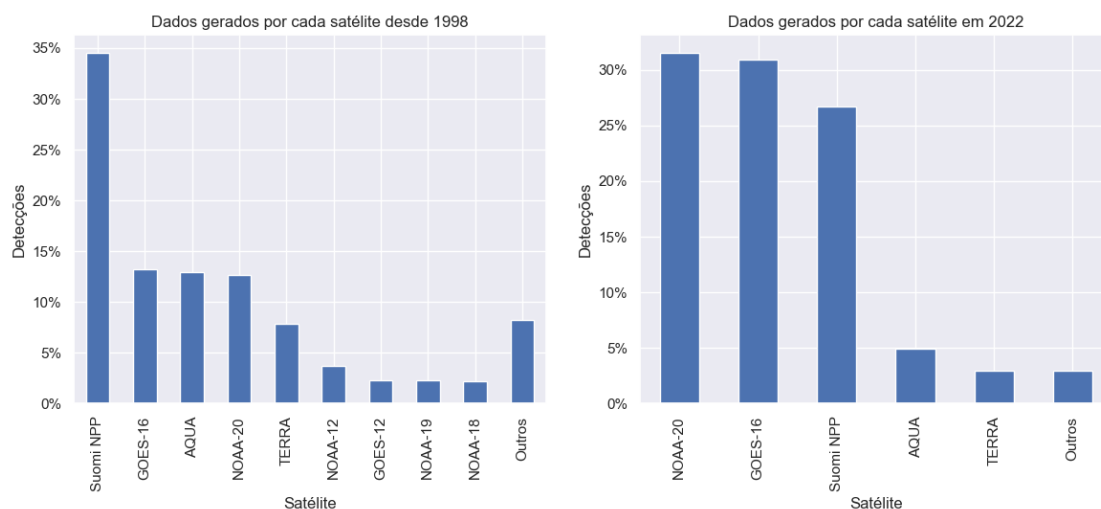
O AQ1km (LIBONATI et al., 2015)

2.4 Uma visão geral dos dados

Nesta sessão, será 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 ou, ainda, a contagem do mesmo foco em passagens diurnas e noturnas de um mesmo satélite polar. Para solucionar esse problema, 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.

Com relação aos satélites, é possível perceber a partir da Figura 2.2 cinco satélites que mais identificaram focos de queimada em toda a série histórica, são eles: Suomi NPP, GOES-16, AQUA, NOAA-16 e TERRA. Todos eles estão ativos atualmente e com quantidades de coleta significativas em 2022. AQUA e TERRA são os mais antigos e possuem um sensor obsoleto. O GOES-16 é um satélite geoestacionário, conhecido por gerar dados de forma mais frequente, geralmente detecta apenas queimadas maiores devido a sua posição distante da Terra. Suomi NPP e NOAA-20 possuem um sensor que detecta 10 vezes mais focos que o sensor MODIS. Por estar em atividade a mais tempo, o Suomi NPP gerou mais dados que o NOAA-20.

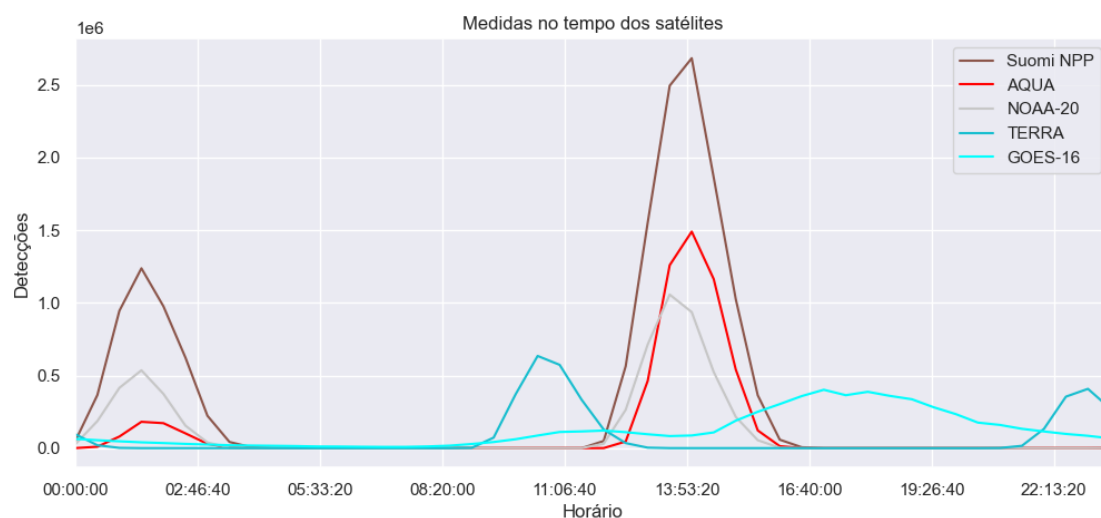
Figura 2.2 – Relação do montante dos dados por satélite.



Fonte: O Autor

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 2.3 é possível observar esse comportamento empiricamente, em que as 5 primeiras linhas, que representam dados gerados por satélites polares, apresentam dois picos durante um período de 24 horas. Já para o caso dos geoestacionários (GOES-16), que ficam fixos em relação a uma posição na Terra, não se observou o mesmo padrão.

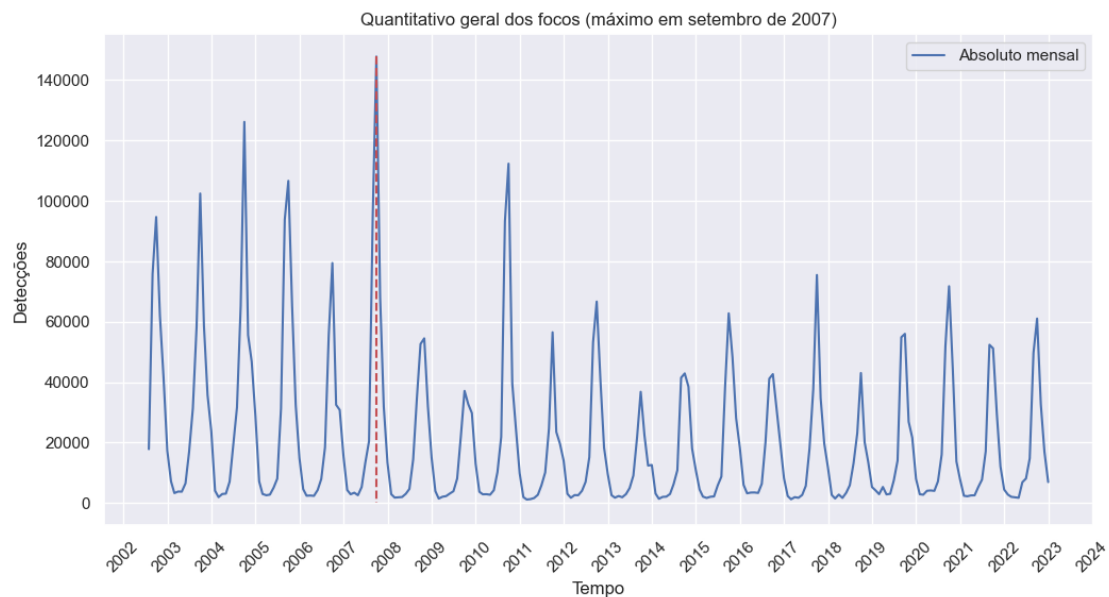
Figura 2.3 – Amostragem por tempo de cada satélite.



Fonte: O Autor

A partir de uma análise quantitativa dos dados com relação ao tempo de cada medida, exposto na Figura 2.4, podemos identificar uma grande sazonalidade, sempre tendo picos entre os meses de agosto e setembro. Desde o início da série, o mês que mais teve focos detectados foi setembro de 2007.

Figura 2.4



Fonte: O Autor

- P1. Fazer análise preliminar dos dados gerando alguns gráficos
- P2. Gráficos geral do Brasil com os focos de queimadas totais (REY; ARRIBAS-BEL; WOLF, 2020)

3 TRABALHOS RELACIONADOS

Muitos outros trabalhos antes deste versaram sobre os dados de queimadas do INPE e, em menor escala, especificamente sobre o cálculo das áreas queimadas. Nesse capítulo, serão apresentados alguns trabalhos que estão de alguma forma próximos deste e, no final, será discutido o que faz este de certa forma único.

O primeiro estudo refere-se ao trabalho que embasou o produto AQ1km (LIBO-NATI et al., 2015), desenvolvido pelo INPE. O objetivo do estudo é detectar áreas queimadas mensalmente por meio de sensoriamento remoto utilizando o sensor MODIS (satélites AQUA e TERRA). Para isso, o algoritmo identifica regiões que apresentaram fogo ativo usando diferentes satélites, extraídos da mesma base do INPE, e agrupa os dados a cada um mês. Em seguida, filtra essas regiões com base em um índice de vegetação sensível a queimadas no espectro do infravermelho próximo, em relação ao mesmo índice do mês anterior. Por fim, identifica regiões próximas às regiões classificadas como área queimada no passo anterior, mas que não apresentaram um indicador forte o suficiente, o que pode ser devido a uma queimada parcial ou baixa intensidade de fogo naquele local.

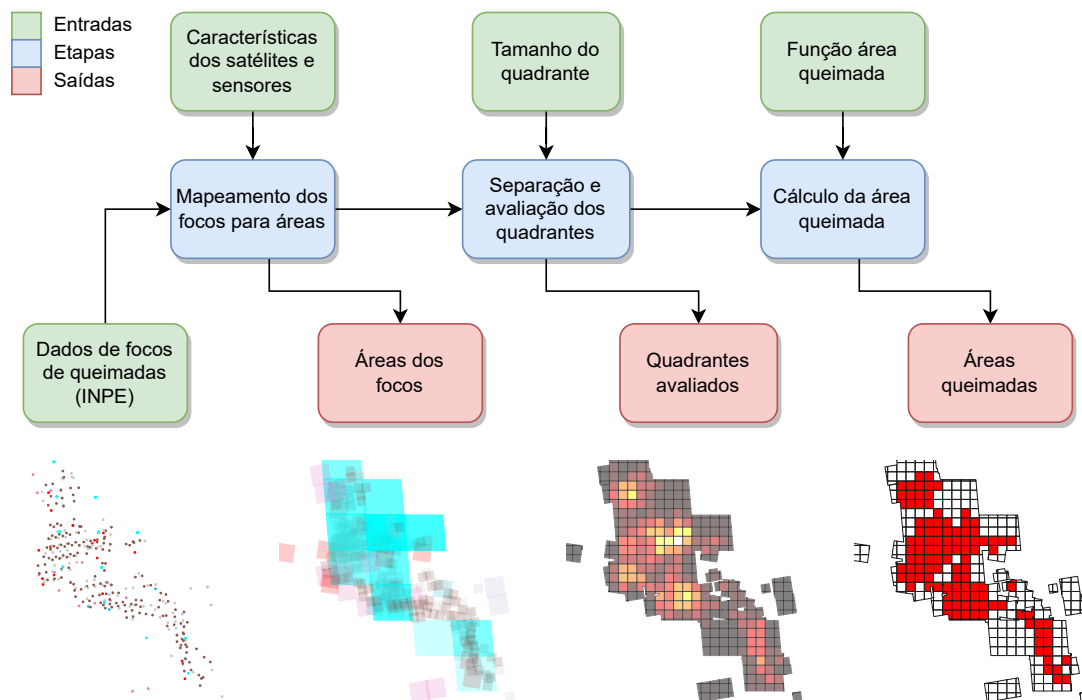
4 METODOLOGIA

Neste capítulo será abordado como método utilizado para calcular as áreas queimadas foi desenvolvido. Começa com uma visão geral do método, depois entra em cada etapa de forma separada e mais detalhada.

4.1 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. 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.

Figura 4.1 – Diagrama do método.



Fonte: O Autor

O método completo, conforme apresentado na Figura 4.1, é composto por três

principais etapas (cartões em azul): mapeamento dos focos para áreas, separação e avaliação dos quadrantes, e cálculo da área queimada. Cada etapa recebe dados de entrada e produz uma saída que, com exceção da última etapa, é usada como entrada para a etapa subsequente.

A primeira etapa transforma os pontos dos focos em áreas que representam a abrangência do foco. Essa etapa recebe dados de focos de queimadas do INPE, previamente filtrados para o intervalo de tempo de interesse, e produz como saída as áreas de cada foco. Essas áreas estão diretamente relacionadas às características do satélite e sensor que captaram a imagem original, como abordado em 2.3, e, portanto, também precisam ser uma entrada.

Na segunda etapa, o espaço é dividido em quadrantes e avaliado. Os quadrantes recebem as áreas dos focos calculados na etapa anterior e o tamanho do quadrante, que pode ser definido pelo pesquisador. A saída são quadrantes com valores que representam a intensidade das queimadas ocorridas dentro de cada quadrante específico. A função de avaliação é proporcional à diversidade e quantidade das áreas de foco que interceptam os quadrantes.

Na terceira 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 4.1 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 e 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 $33,32km^2$.

4.2 Mapeamento dos focos para áreas

Quando um foco de queimada é detectado pelo INPE, ele representa uma área quadrada aproximadamente do tamanho da resolução do sensor que capturou a imagem original. Ou seja, um foco detectado a partir de imagens do satélite AQUA, que utiliza o

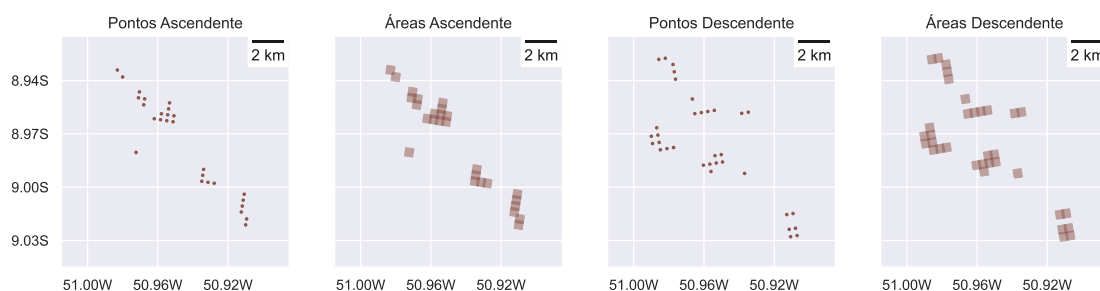
sensor MODIS, representa uma área de $\approx 1Km^2$. Para um satélite com o sensor menos preciso, como o GOES-13, que utiliza o sensor GOES I-M com resolução espacial de $4Km$, a área representada seria 16 vezes maior, indicando uma menor precisão.

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 $36Km$ (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 em um sentido. Por outro lado, se a trajetória for descendente, do norte para o sul, a rotação deve ser no outro sentido.

A Figura 4.2 apresenta dados coletados pelo satélite Suomi NPP no dia dois de setembro de 2022. O primeiro par de imagens 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 ganham a área do sensor ($500m$) e encaixam perfeitamente entre seus vizinhos. A partir desse ponto até o final do documento, a área coberta pelo foco será chamada de medição.

Figura 4.2 – Comparação entre pontos e áreas dos focos detectados pelo satélite Suomi NPP.



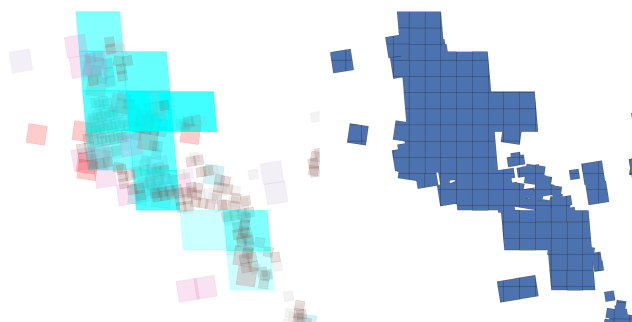
Fonte: O Autor

4.3 Separação e avaliação de quadrantes

A separação entre quadrantes 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 4.3, temos uma grade regular que é usado para as operações de avaliação de serão descritas a seguir.

O tamanho do quadrante é uma entrada dessa etapa e pode ser definida pelo pesquisador. O valor recomendado desta entrada fica em torno de 0,004 a 0,005 graus, o que coincide com o tamanho da menor resolução de sensor presente nos dados, que é o VIIRS de 500m.

Figura 4.3 – Separação entre quadrantes.



Fonte: O Autor

Com a distribuição dos quadrantes definida, é necessário atribuir a cada um deles um valor que represente a probabilidade de que a área contida tenha sido queimada. A avaliação dos quadrantes deve valorizar aqueles que apresentam 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 (veja a Figura 2.3), o que indica uma queimada mais prolongada. Em ambos os casos, queimadas mais intensas e duradouras sinalizam um maior potencial de o fogo se espalhar para outras áreas da vegetação.

Sendo q o quadrante (polígono) a ser avaliado; M um conjunto de todas as medições (área coberta pelo foco); operação $area(p)$ retorna a área de um polígono p qualquer; operação $unique_satellite(Q)$ retorna todos os satélites diferentes do conjunto Q ; min_area é a área mínima de uma medição; $threshold_satellite$ 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 intersecção com cada medição

(m) contida nele, resultando em um conjunto Qm :

$$Qm = \{m \in M \mid m \cap q\} \quad (4.1)$$

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

$$Qm' = \{m \in Qm \mid \text{area}(m) \geq \text{min_area}\} \quad (4.2)$$

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

$$Us = \{m \in \text{unique_satelite}(Qm')\} \quad (4.3)$$

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

$$ia = \left(\sum_m^{Qm'} \text{area}(m) \right) \div \text{area}(q) \quad (4.4)$$

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

$$c = 1 - \min \left(1, \frac{|Us|}{\text{threshold_satellite}} \right) \quad (4.5)$$

Finalmente, a avaliação final do quadrante (aq) é obtida pela expressão:

$$aq = |Us|^2 + ia - ai \cdot c \quad (4.6)$$

A fórmula final 4.6 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 cabe 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.

Desta forma,

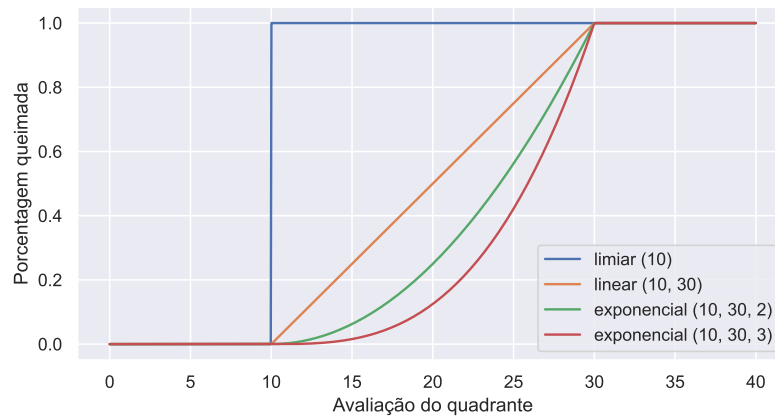
4.4 Cálculo da área queimada

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 ($eval(v)$) que receba o valor do quadrante, calculado na etapa anterior, e retorne um número real entre zero e um. Essa função também pode receber o maior e o menor valor presente na avaliação dos quadrantes, que podem ser usados para normalizar as áreas.

Para facilitar o trabalho do pesquisador, a implementação pode fornecer funções embutidas comuns que definem a função $eval(v)$. Na Figura 4.4, são apresentadas algumas possibilidades de definições para essa função. A função mais simples é chamada de limiar, que estabelece que toda a área do quadrante deve ser considerada queimada se a avaliação for maior que um determinado valor e nenhuma área deve ser considerada queimada se não alcançar esse valor. Outra função simples é a linear, que faz o valor da área queimada crescer de forma linear dentro de um valor máximo e mínimo. A última função é a exponencial, semelhante à linear, mas com uma exponenciação que faz o valor crescer mais lentamente no início e de forma mais acentuada no final do intervalo.

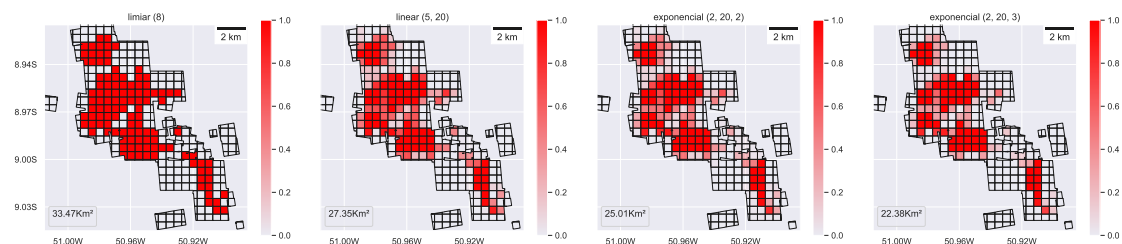
Figura 4.4 – Funções embutidas para cálculo de área.



Fonte: O Autor

A Figura 4.5 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 parecem mais próximas da realidade, já que consideram apenas uma porcentagem da área dos quadrantes próximos aos agrupamentos de alta intensidade como queimados.

Figura 4.5 – Exemplo da aplicação das funções embutidas.



Fonte: O Autor

5 IMPLEMENTAÇÃO DO MÉTODO

Após a formalização do método de forma abstrata, esse capítulo descreve sobre a implementação de fato do método. Primeiro está discutido como os dados do DBQueimadas foram coletados da base de forma automatizada. Em seguida, descreve-se sobre a implementação do método

5.1 Coleta 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, com intervalo de tempo máximo um ano para cada pedido. 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 o site tenha boa usabilidade, seria impraticável baixar todos os dados do Brasil manualmente sem sobrecarregar os servidores no INPE. Por isso, foi necessário entender como a solicitação dos dados era 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 da URL, incluindo os filtros, e-mail e formato de arquivo desejado. Um exemplo de uso dessa API por meio de uma chamada CURL pode ser encontrado no ??.

Para automatizar o processo, desenvolvemos um script em Python que solicitava os dados de cada 30 dias, totalizando 300 requisições desde 1998 até 2022. Para não sobrecarregar os servidores do INPE, adicionamos uma espera de um minuto entre as requisições. Além disso, implementamos um sistema de registro de requisições em arquivo, que armazenou o estado de cada uma delas, evitando requisições 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 do arquivo por

meio do link enviado por e-mail. Utilizamos 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 para 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. Caso uma requisição permanecesse por mais de 30 minutos no estado pendente, o script fazia uma nova solicitação 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 gigabytes de armazenamento em disco e, juntos, somam exatamente 43.782.758 linhas. Ao final, eles foram recompatados em um único arquivo zip (450 megabytes) que está disponível para download em <<https://bit.ly/3IgHIXH>>, independentemente dos servidores do INPE.

Também utilizamos 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. Todos os arquivos baixados estão no formato Shapefile, que é o formato responsável por armazenar dados vetoriais geográficos.

5.2 Implementação 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 reprodutibilidade das execuções e a organização do código, foi adotado o Jupyter Notebook. Todos o código 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 do Pandas que adiciona uma coluna especial chamada 'geometry'. Essa coluna permite

armazenar as coordenadas e o formato dos dados em um sistema de coordenadas pré-definido. Além disso, o GeoPandas integra-se à biblioteca Shapely, que é usada para realizar cálculos e operações em estruturas geométricas espaciais. Com o uso dessas bibliotecas, é possível executar operações avançadas da 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.

Após a coleta de dados mencionada anteriormente, os dados foram carregados no Python e algumas otimizações nas estruturas foram realizadas. Os 300 arquivos das quais foram lidos pelo Pandas e concatenados em um único DataFrame. 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 que muitos valores repetidos ocorriam nessas colunas.

Para otimizar o processamento e a leitura dos dados, também foi utilizada a biblioteca Dask para realizar o processamento paralelo. Essa biblioteca oferece a capacidade de criar clusters locais ou se conectar a clusters remotos, 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 programa, ao invés de carregar novamente os arquivos CSV, o programa fará 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 ocorreu 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. Em comparação

com a coluna original dos satélites, que tinha 32 valores possíveis, a nova coluna contém apenas 22 valores possíveis. Para cada satélite, foi escolhida uma cor única, conforme apresentado na Figura 5.1, que é utilizada em todos os gráficos apresentados neste documento.

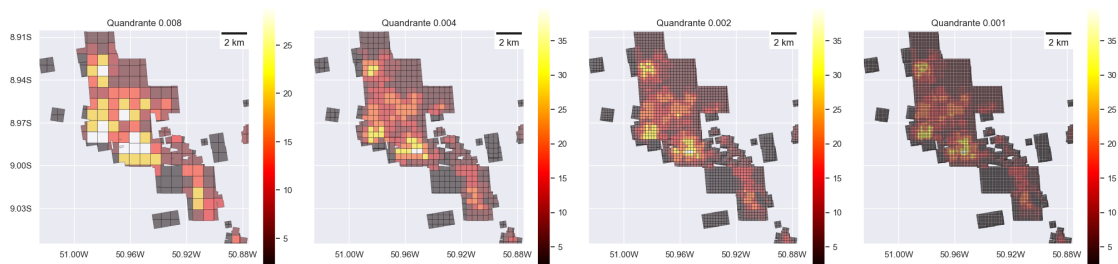
Figura 5.1 – Cores escolhidas para cada satélite.



Fonte: O Autor

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, com a utilização de índices espaciais, usados nos cálculos de intersecção, o processamento dos quadrantes é ainda mais otimizado. Quanto maior o quadrante usado, menos avaliações são necessárias, porém a precisão da área queimada também diminui. Como se trata de um espaço bidimensional, a complexidade computacional da avaliação em relação à quantidade de quadrantes é $O(n^2)$.

Figura 5.2 – Diferença da avaliação para diferentes tamanhos de quadrantes.



Fonte: O Autor

Com base em experimentos, foi constatado que o uso de quadrantes muito pequenos (com menos de 0,002 graus quadrados) não aumentam significativamente a precisão dos resultados e tornam a execução muito mais demorada.

O GeoPandas também oferece suporte à operação de junção espacial (spatial join) de forma otimizada. O spatial join 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, o spatial join é utilizado para obter todas as medições que intersectam cada quadrante (interseção entre dois polígonos), que é a principal informação usada para avaliar um quadrante.

[falar sobre suporte a Spatial index <https://autogis-site.readthedocs.io/en/2019/notebooks/L3/spatial-join.html>]

6 RESULTADOS E DISCUSSÃO

6.1 Aplicações do método

O método em si é geral para qualquer entrada de focos de satélites, e retorna no final polígonos representando áreas geográficas associadas a um valor que indica a área queimada dentro desse polígono. Variando a entrada do método e o agrupamento dos resultados preliminares, pode-se obter diferentes resultados finais.

7 CONSIDERAÇÕES FINAIS

REFERÊNCIAS

- ÁREA queimada. 2023. Disponível em: <<https://queimadas.dgi.inpe.br/queimadas/aq1km/>>. Acesso em: 24 de abr. 2023.
- Divisão de Geração de Imagens. **Geração de imagens**. 2023. Disponível em: <<https://www.dgi.inpe.br/>>. Acesso em: 16 de fev. 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. 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>>.
- INPE. **Programa Queimadas Perguntas frequentes**. 2023. Disponível em: <<http://www.inpe.br/queimadas/portal/informacoes/perguntas-frequentes>>. Acesso em: 23 de jan. 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. An algorithm for burned area detection in the brazilian cerrado using 4 μ m modis imagery. **Remote sensing**, v. 7, n. 11, p. 15782–15803, 2015.
- REY, S. J.; ARRIBAS-BEL, D.; WOLF, L. J. **Geographic Data Science with Python**. [S.l.: s.n.], 2020. Disponível em: <<http://geographicdata.science/book/notebooks/>>. Acesso em: 5 de jan. 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>. Acesso em: 1 de fev. 2023.