Detecting and Assessing Pollution Events from Wildfires Using Remote Sensing and Meteorological Data

A Data Science Approach

Sofia Margarida Matias Rodrigues

Dissertação para a obtenção do Grau de Mestre em

Ciências dos Dados em Agricultura, Alimentação, Floresta e Ambiente

Orientador(es): Ana Russo

Célia Gouveia

Júri:

Presidente: [Grau, nome completo, categoria e local de trabalho]

Vogais:

(Versão provisória)

2025

Acknowledgments

Abstrato

Os incêndios florestais contribuem significativamente para a poluição do ar ao libertarem matéria particulada (PM) e gases tóxicos na atmosfera, com as alterações climáticas a preverem um aumento na atividade dos incêndios florestais e na propagação do fumo, agravando os riscos para a saúde. O trabalho a ser desenvolvido será apoiado por um quadro orientado por dados para monitorizar e avaliar a poluição do ar causada por incêndios florestais, uma questão urgente de saúde e ambiental que afeta a população global. O trabalho irá analisar os resultados dos modelos atmosféricos atuais, juntamente com indicadores de deteção remota, como o Fire Radiative Power (FRP) e o Fire Radiative Energy (FRE), combinados com dados meteorológicos e aprendizagem automática, para melhorar a deteção de eventos de poluição. O objetivo é analisar os impactos transfronteiriços das emissões de incêndios florestais, avaliar tecnologias de deteção remota (e.g., MODIS, SEVIRI, Sentinel) no monitoramento de incêndios florestais e examinar métodos de ciência de dados para monitorização ambiental. Na secção de dados e metodologia, será descrita a integração de dados meteorológicos e de deteção remota, com modelos de aprendizagem automática, como Random Forest, XGBoost e Redes Neuronais, usados para classificar eventos de poluição e mapear padrões espaço-temporais do fumo. A validação do modelo será realizada através da comparação dos resultados com eventos históricos extremos de incêndios florestais para verificar a sua precisão. O modelo será, então, avaliado pelo seu desempenho preditivo e fornecerá algumas perceções sobre os padrões de dispersão do fumo dos incêndios florestais, identificando fatores-chave que contribuem para os eventos de poluição. Para concluir o trabalho, serão apresentados os destaques do estudo, demonstrando como os dados de deteção remota e meteorológicos podem melhorar o monitoramento da qualidade do ar e apoiar o planeamento de políticas públicas. Serão propostos trabalhos futuros para melhorar as capacidades de monitoramento em tempo real, integrar fontes de dados adicionais e aplicar as descobertas em quadros ambientais e de saúde mais amplos. Esta pesquisa tem o potencial de informar intervenções estratégicas, reforçando ainda mais as ferramentas de tomada de decisão para gerir a poluição provocada por incêndios florestais.

Palavras-Chave: Fogos, Poluição do Ar, Dados Meteorológicos, Deteção Remota, Aprendizagem Automática

Abstract

Wildfires contribute significantly to air pollution by releasing particulate matter (PM) and toxic gases into the atmosphere, with climate change projected to increase wildfire activity and the spread of smoke, heightening health risks. The work that will be developed will be supported by a data-driven framework to monitor and assess air pollution from wildfires, a pressing health and environmental issue that affects the global population. The work will analyse current atmospheric models outputs coupled with remote sensing indicators, like Fire Radiative Power (FRP) and Fire Radiative Energy (FRE), combined with meteorological data and machine learning to improve pollution event detection. It aims to look upon transboundary impacts of wildfire emissions, evaluate remote sensing technologies (e.g. MODIS, SEVIRI, Sentinel) in wildfire monitoring, and examine data science methods for environmental monitoring. For the data and methodology section, it will describe the integration of meteorological and remote sensing data, with machine learning models, such as Random Forest, XGBoost, and Neural Networks, used to classify pollution events and track spatial-temporal patterns of smoke. Model validation will be performed by comparing results with historical extreme wildfire events to verify accuracy. Then the model will be evaluated by its predictive performance and have some insights into wildfire smoke dispersion patterns, identifying key factors contributing to pollution events. To conclude the work, the highlights of the study will be shown, demonstrating how remote sensing and meteorological data can improve air quality monitoring and support policy planning. Future work will be proposed to enhance real-time monitoring capabilities, integrate additional data sources, and apply findings within broader environmental and health frameworks. This research has the potential to inform strategic interventions, further strengthening decision-making tools for managing wildfire-driven pollution.

Key Words: Wildfires, Air Pollution, Meteorological Data, Remote Sensing, Machine Learning

Table of Contents

List of Figures

Não foi encontrada nenhuma entrada do índice de ilustrações.

List of Tables

List of Abbreviations

IPMA – Instituto Português do Mar e da Atmosfera

CAMS – Copernicus Atmosphere Monitoring Service

CDS – Climate Data Store

PM – Particulate Matter

FRP – Fire Radiative Power

FRE – Fire Radiative Energy

MODIS – Moderate Resolution Imaging Spectroradiometer

WRF-Chem – Weather Research and Forecasting model coupled with Chemistry

WRF – Weather Research and Forecasting

SEVIRI – Spinning Enhanced Visible and Infrared Imager

MSG – Meteosat Second Generation

Introduction

Background:

- Air pollution, responsible for an estimated 7 million preventable deaths annually, affects over 90% of the global population with polluted air. Pollution from both natural and anthropogenic sources frequently transgresses national boundaries, impacting distant regions.
- Wildfires contribute significantly to air pollution by releasing particulate matter (PM) and toxic gases into the atmosphere. With climate change, wildfire activity is expected to increase, leading to greater smoke dispersal and heightened health risks across affected populations.

Problem Statement:

- Although atmospheric dispersion models are commonly used for tracking wildfire smoke, they have limitations, such as heavy computational demands and discrepancies between predicted and actual pollution concentrations. Alternatively, remote sensing offers near-real-time data with products like the Fire Radiative Power (FRP) and Fire Radiative Energy (FRE), which allow for effective monitoring of wildfire emissions.

Research Aim:

- To detect and assess wildfire pollution events using a combination of remote sensing, atmospheric monitoring data, and machine learning approaches, focusing on FRP and FRE outputs as indicators of wildfire-induced pollution.

Objectives:

- Identify key indicators from remote sensing and meteorological data that correlate with pollution events due to wildfires.
- Develop a machine learning-based model to detect pollution events using data from sources like CAMS, WRF-Chem, FRP/FRE.
- Analyze the spatial and temporal impacts of wildfire smoke on air quality in affected areas.
- Analyse the impact of compound extreme events on wildfire-related pollution events

Literature Review

Air Pollution:

- Discuss the impacts of particulate matter and toxic gases emitted from wildfires, with a focus on the transboundary nature of pollution and its exacerbation due to climate change.

Remote Sensing in Wildfire Monitoring:

- Review remote sensing tools, including MODIS, SEVIRI, and Sentinel products, used in monitoring wildfire activity and emissions. Discuss the utility of FRE and FRP as measures of biomass combustion rate and their effectiveness in tracking pollution.

Data Science in Environmental Monitoring:

- Explore data science methods in atmospheric and environmental sciences, including machine learning models used in pollution detection and forecasting.

Data

Data Collection:

- Meteorological Data: Gather atmospheric data, such as temperature, humidity, wind speed, and direction, from sources like CAMS and monitoring stations.
- Remote Sensing Data: Use FRE and FRP data from MODIS, SEVIRI, and Sentinel-3 for monitoring fire radiative power and energy release.

Data Processing:

- Integrate meteorological and remote sensing data with pollution monitoring station data to create a comprehensive dataset.
- Preprocess data by cleaning and transforming it to address missing values, noise, and scaling for model input.

Methodology

Model Development:

- Supervised Machine Learning Models: Implement and compare models (e.g., Random Forests, XGBoost, and Neural Networks) to classify pollution events and identify anomalies related to wildfire smoke.
- Spatial and Temporal Analysis: Use geospatial tools to map the spatial reach of pollution plumes, identifying the temporal dynamics of smoke dispersal patterns.

Evaluation and Validation:

- Validate the model using metrics such as accuracy, F1-score, and area under the curve (AUC) for classification tasks.
- Cross-reference predicted pollution events with historical extreme events (e.g., Portugal's 2017 megafires) to validate spatial and temporal accuracy. Analyse the impact of compound events (e.g. droughts and heatwaves) to the magnitude of fire-driven pollution events.

Results

Model Performance:

- Present the results of model performance, discussing the predictive accuracy in detecting pollution events and identifying contributing factors from FRE, FRP, and meteorological data.

Discussion

Spatial and Temporal Impact Analysis:

- Discuss the spatial and temporal patterns of wildfire smoke distribution. Evaluate how FRP and FRE data correlate with pollutant concentrations across affected regions.

Limitations and Future Improvements:

- Discuss potential limitations of the model, such as sensitivity to specific atmospheric conditions or data quality, and suggest directions for future improvement.

Conclusion

Summary of Findings:

- Summarize the effectiveness of remote sensing and meteorological data in detecting and mapping wildfire-induced pollution events.

Implications:

- Discuss the potential of this approach to support decision-making interventions and policy planning.

Future Work:

- Suggest advancements in real-time monitoring systems, the inclusion of additional data sources, and potential integration with other environmental and health monitoring frameworks.

References

Não existem fontes no documento atual.

Detailed list of all references cited, including recent studies on air pollution, wildfire activity, and data science applications in environmental monitoring.

Appendices