GlobalBlaze: Analyzing Global Trends in Wildfire Burned Areas and Emissions(2002-2023)

1.Introduction

Wildfires are recognized as a most important environmental challenge, it impacts on different landscapes, biodiversity and human beings. Wildfires will impact directly or indirectly on the climate change and human health. This report explains the methodologies used to analyse the comprehensive datasets on wildfire burned areas and emission gases. This project mainly aims to answer these four questions:

- What are the long-term trends in wildfire burned areas and emissions globally, and are there significant changes over time?
- How do seasonal variations affect wildfire activities and emissions, and are there specific months or quarters with higher prevalence or levels?
- Which countries or regions experience the most wildfires, and what factors contribute to their prevalence?
- Is there a correlation between wildfire area burned and emissions gases?

By addressing these questions, we can understand which land types are more likely to get wildfires, determine which regions or countries experience more wildfires and to find a correlation between burned areas and emission gases.

---

2. Methods

2.1. Data Sources

In this project we are taking datasets from the Global Wildfire Information System(GWIS), they are stored in the CSV format, with no missing values in both the datasets and are relevance to analyse our main questions. These datasets include the historical wildfire burned area and emission gases from 2002 to 2023.

- **Global Monthly Burned Area Dataset(GMBAD)** [2002-2023] contains the details with monthly burned area records, years, several country names and codes, and different land cover types measured in hectares. This dataset contains 0 values which indicates that there is no wildfire occurred during that time. **GMBAD**

- **Global Monthly Emission Dataset(GMED)** [2002-2023] contains the details with monthly burning emissions by pollutant, years, it also includes many country details, and different pollutant gases measured in tons. This dataset contains 0 values which indicates there are no emission gases during that period. **GMED**

2.2. Data Pipeline

The data pipeline is created using Python, utilizing libraries such as pandas, numpy, requests, zipfile, io, and sqlalchemy for data handling, extraction, transforming , loading, and database integration.

- **Extraction:** The data pipeline uses Python’s request library to fetch ZIP files with burned area and emissions data from specific URLs, then extracts their content into memory for further processing and analysis.

- **Transformation:** During transformation, both burned area and emissions data are cleaned by removing irrelevant columns, filtering out rows with zero values, renaming columns for consistency, and adding a quarter column for seasonal analysis.

- **Loading:** After transformation, data is loaded into pandas Data Frames, merged based on common attributes, and stored in an SQLite database for storage and analysis.

2.3. Data License

The datasets GMBAD and GMED are licensed under the Creative Commons Attribution 4.0 International **(CC BY 4.0)** license link. Under this license, it allows to freely use the data by copying, sharing, and modifying. So, I am using the datasets from Global Wildfire Information System (GWIS) and transforming those datasets for our project. Giving proper credits by acknowledging GWIS as the source of the data and providing a link to the license. This ensures that others know where the data originated from and can access the original source for further information or verification. By complying with these terms, we contribute to open access and encourage the continued sharing and use of valuable data for research, analysis, and innovation. I will make sure that GWIS receives proper credit in all the reports and documents.

---

3. Results

To answer the objective questions, I filtered the dataset to focus on the top 10 countries based on total burned area and emissions, from the original dataset containing data for 195 countries. I conducted exploratory data analysis on the datasets including wildfire burned area and emissions.

3.1. Long-Term Trends

Over the years, forests have maintained stable wildfire patterns due to effective fire management practices. The extreme high burned areas are shrubs grasslands and savannas suggest that these ecosystems are highly fire prone. However, there has been a slight decline over time, likely due to improved land care and the effects of climate change. The CO2 emissions chart shows fluctuations with notable peaks and troughs, particularly in recent years, suggesting specific large-scale events or activities have influenced these trends as shown in Figure 1. This underscores the importance of focusing on reducing CO2 emissions to effectively combat climate change. The low levels of other pollutants suggest effective maintenance, but ongoing monitoring and regulations are essential to further enhance air quality.

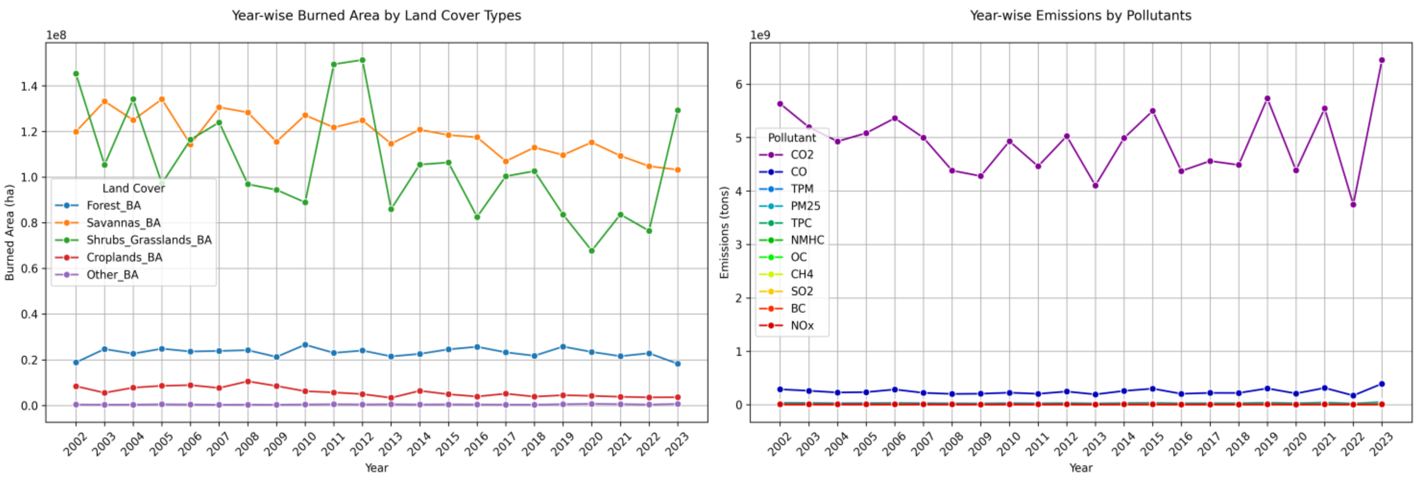

Figure 1: Year - Wise Burned Area and Emissions from 2002 to 2023

3.2. Seasonal Variations

Both burned area and emissions increase significantly from Q1 to Q3, peaking in Q3 (Figure 2), indicating that wildfires are more likely to occur in summer months, specifically from July to September. Summer months often bring high temperatures and low humidity, creating dry conditions that are highly favourable for wildfires to start and spread quickly. Savannas and shrub grasslands show particularly high burning during Q3, as seen in the graph. Also, human activities can also contribute to these wildfires during summer. There is a noticeable decrease in Q4 for all land types and pollutants. During this period, which corresponds to late autumn to early winter in many regions, temperatures drop, and humidity levels increases. To combat the increased risk of wildfires during summer, effective fire management practices are crucial.

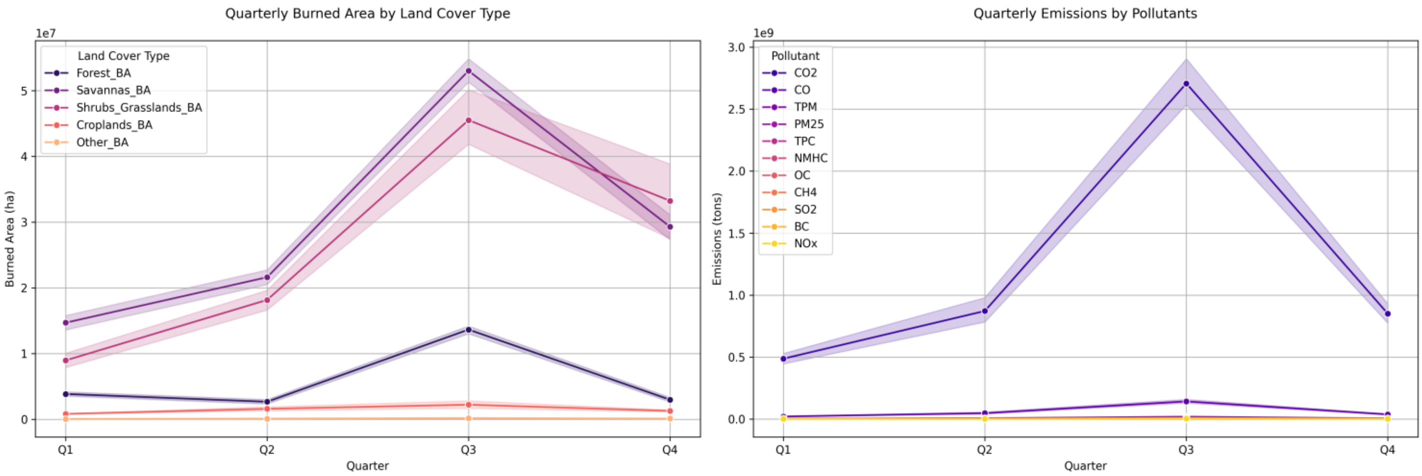

Figure 2: Quarterly Burned Area and Emissions from 2002 to 2023

3.3. Top Countries Affected by Wildfires

Australia has the highest burned area approximately 1 billion hectares, indicating it has extensive wildfires due to the dry climate and frequent droughts. Angola, Democratic Republic of the Congo, South Sudan and Zambia belongs to the African Countries where they contain extensive savannas and grasslands. Democratic Republic of the Congo and Angola have contributed significantly to the highest emissions. In contrast, Brazil and Russia are having highest emissions and do not report the highest burned area(Figure 3). Brazil has Amazon Rainforest it can experience fires that are not always extensive in terms of burned area but are extremely intense and produce large number of emissions. Russia's vast Siberian and boreal forests contain large amounts of carbon-rich peat and organic soils. Fires in these regions can burn deeply into the ground, releasing significant amounts of carbon dioxide and other pollutants, even if the surface burned area is not the largest.

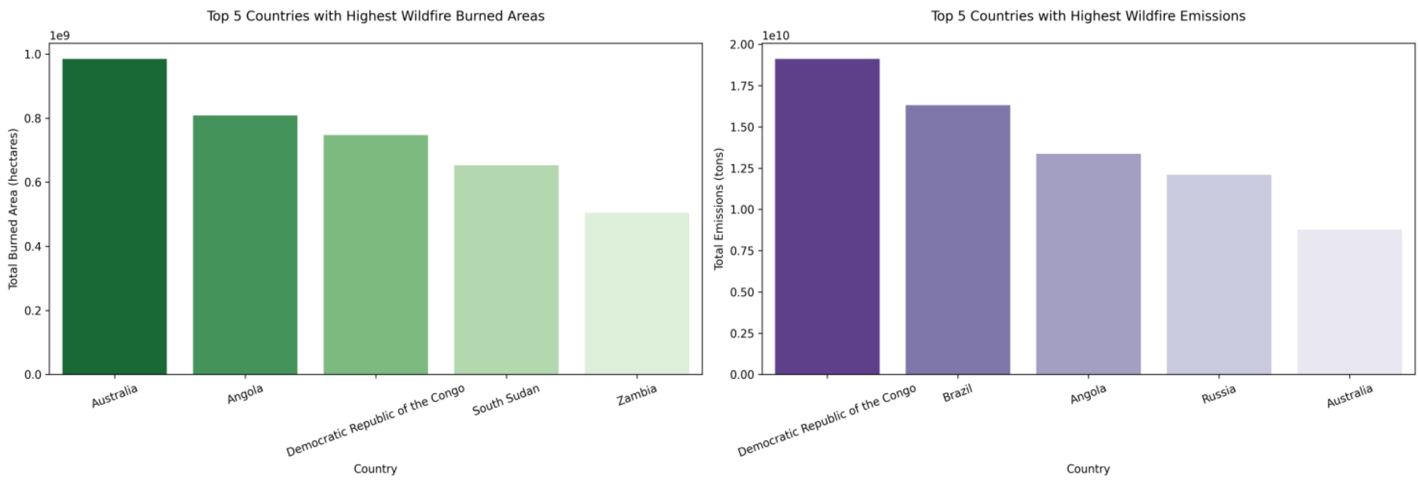

Figure 3: Top Countries Affected by Wildfires from 2002 to 2023

3.4. Correlation Between Burned Areas and Emissions

From the Heatmap we can conclude that emissions are highly correlated with each other, indicating that these emissions tend to increase or decrease together and likely to share the common sources. Forests and savannas are moderate to strong contributors to these emissions , particularly CO2, CO, TPM and NOx, while the shrubs and grasslands contributes less when compared to other areas. Croplands have least impact on emissions by showing weak or negligible correlations. NOx has strong correlations with forests (0.64) and savannas (0.70), indicating that these two types of burned areas are major contributors to NOx emissions(Figure 4). So, reducing one type of emission could help to decrease others due to their strong correlations. Focusing on forest and savannas areas could significantly reduce multiple emissions including NOx.

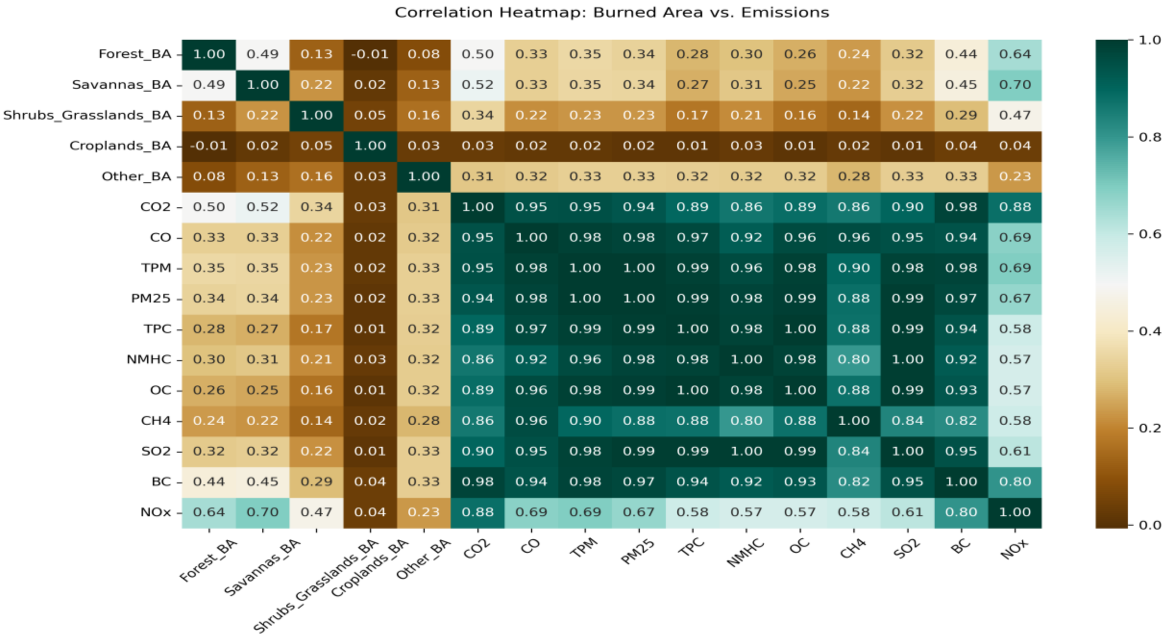

Figure 4: Correlation Heatmap between Wildfire Burned Area and Emission Gases from 2002 to 2023

4. Conclusions

The analysis of wildfire burned area and emissions from 2002 to 2023 shows that savannas and shrubs grasslands are most prone to the wildfires. These wildfires peak in the summer months, particularly from July to September. Australia, Angola, and the Democratic Republic of the Congo are the most affected regions due to their dry climates. There is a strong correlation between the burned area and the emission gases particularly CO2 and NOx. Effective fire management practices and targeted emission reduction strategies, especially in forests and savannas are crucial for mitigating the impact of wildfires on the environment and human health.

4.1. Limitations

- The analysis does not fully explore all the factors contributing to wildfire prevalence such as human activities, land use changes, and specific climatic conditions .
- The data is aggregated monthly which may not overlook in detailed temporal variations such as daily or weekly fluctuations in wildfire activity and specific events such as sudden spikes in wildfire incidents due to extreme weather events or human activities, might not be captured adequately in monthly averages.

4.2. Future Work

- Evaluating the broader ecological and socio-economic impacts of wildfires including biodiversity loss, health effects and economic costs.
- Utilizing advanced machine learning algorithms for predicting wildfire occurrence, behaviour, and emissions. Exploring novel algorithms, ensemble methods, and deep learning architectures to improve accuracy and reliability in each detection systems.