Final Report

*Assessing the impact of weather conditions on wildfires in Greece*

Methods of Advanced Data Engineering

2024/07/03

**Arvanitidis Alexis - 22558160**



Faculty of Engineering

Methods of Advanced Data Engineering

Friedrich-Alexander University of Erlangen-Nuremberg

# Introduction

This project aims to assess the impact of weather conditions on wildfire occurrences in Greece. By analyzing historical data, we seek to identify trends in weather patterns and their correlation with the frequency and severity of wildfires. Our study focuses on key weather variables, such as temperature and wind speed, to determine their influence on wildfire events. This investigation is crucial for enhancing our understanding of how climatic factors contribute to wildfire risk, thus supporting the development of more effective wildfire management and prevention strategies. The results of this study will also provide valuable insights for refining climate models and improving disaster preparedness.

# Research Questions

1. “What are the trends in weather conditions in Greece and how do they correlate with wildfire occurrences?”
2. “How do specific weather variables, such as temperature and wind speed, influence the frequency and severity of wildfires?“

# Data Pipeline

To address our research question, we constructed an ETL (Extract, Transform, Load) data pipeline that generates the final dataset for analysis. This pipeline automates the extraction, processing, and storage of data from diverse sources, ensuring the data is thoroughly cleaned, properly structured, and ready for analysis. The following figure presents a conceptual overview of the Data Pipeline.

A diagram of a data processing process

Description automatically generated

Figure 1: Data Pipeline

# Used data

This report utilizes a merged dataset produced by an ETL data pipeline, derived from two primary sources: Kaggle **and Global Forest Watch. The resulting dataset offers a comprehensive overview of weather conditions, including temperature** and wind speed, alongside wildfire occurrences in Greece.

**Description of Dataset**

The dataset is structured as a merged CSV file and includes the following columns:

* **year:** The year in which the data was recorded.
* **week:** The week number within the year.
* **avg\_temp:** The average temperature for each week, measured in degrees Celsius.
* **min\_temp:** The minimum temperature for each week, measured in degrees Celsius.
* **max\_temp:** The maximum temperature for each week, measured in degrees Celsius.
* **wind\_direction:** The predominant wind direction for each week.
* **wind\_speed:** The average wind speed for each week, measured in meters per second.
* **pressure:** The average atmospheric pressure for each week, measured in hPa.
* **fire\_alerts:** The number of wildfire alerts recorded each week.

**Weather Data:** This portion of the dataset includes weekly averages of various weather conditions, such as temperature and wind speed, collected for Greece. It provides a detailed view of the climatic factors that could influence wildfire occurrences.

**Wildfire Occurrences:** This section contains weekly counts of wildfire alerts, indicating the frequency and intensity of wildfires over the specified period. The data is collected from Global Forest Watch and focuses on high-confidence alerts to ensure accuracy.

The merged dataset spans from week 41 of 2018 to week 41 of 2022, allowing for a comprehensive analysis of the correlation between weather conditions and wildfire occurrences in Greece. This integrated data facilitates the investigation into how specific weather variables impact the frequency and severity of wildfires.

A table of numbers and a number of people

Description automatically generated with medium confidence A table of numbers with numbers

Description automatically generated

Figure 2: Data Preparation

**Data Preparation Process**To prepare the dataset for analysis as shown in figure 2, the following steps were undertaken to ensure data quality and relevance:

## **Filtered to Include Only Records from Greece:**

* + The initial datasets were filtered to include only data pertinent to Greece, focusing the analysis on this specific geographic area.

## **Converted Date Column to Datetime Format:**

* + The date columns in the datasets were converted to a standard datetime format to facilitate accurate time-based operations and analyses.

## **Extracted Week Number and Year from the Date:**

* + The datetime data was further processed to extract the week number and year, enabling the aggregation of data on a weekly basis for consistent comparison and analysis.

## **Aggregated Data to Provide Weekly Averages:**

* + Weather data, including temperature and wind speed, was aggregated to provide weekly averages. This step ensured that the data was suitable for identifying trends and correlations over time.

## **Dropped Unnecessary Columns:**

* + Columns such as latitude, longitude, and city were deemed unnecessary for the analysis and were removed to streamline the dataset and focus on the relevant variables.

## **Renamed Columns for Clarity:**

* + Column names were standardized and renamed to ensure clarity and consistency, making the dataset easier to understand and work with during the analysis phase.

A table of numbers with numbers

Description automatically generated

Figure 3: Aggregated data

These data preparation steps were critical in transforming the raw data into a structured and clean format as shown in figure 3, ready for detailed analysis to investigate the impact of weather conditions on wildfire occurrences in Greece.

# Analysis

### Interpretation

For weather conditions, we obtained significant p-values (*p <* 0*.*05), indicating strong evidence against the null hypothesis.

# Conclusion

**Q1.: “What are the trends in weather conditions in Greece and how do they correlate with wildfire occurrences?”:**

By analyzing the weekly aggregated data of temperature and wind speed alongside wildfire occurrences, one could identify potential correlations. For instance, higher temperatures and increased wind speeds might correspond to higher frequencies of wildfire alerts.

**Q.2: “How do specific weather variables, such as temperature and wind speed, influence the frequency and severity of wildfires?“**

Increased average, minimum, and maximum temperatures likely contribute to drier conditions, which can enhance the likelihood of wildfires. Higher wind speeds can exacerbate wildfire conditions by spreading flames more rapidly and over larger areas. Wind can carry embers to new locations, igniting spot fires ahead of the main fire front. Consistent and strong winds in a particular direction can also influence the spread and severity of wildfires, as they drive the fire front and increase the difficulty of containment effort.

# Limitations

* 1. **Geographical scope:** The weather data pertains to the capital city of Greece, while the wildfire data covers the entire country. This discrepancy can limit the comparability of the datasets.
  2. **Sample size**: The number of observations (N) is relatively low, especially after filtering and merging the datasets, which could impact the robustness of the findings.

Therefore, the research question cannot be fully answered with the current dataset. To achieve a more comprehensive and accurate answer, additional and better-suited datasets are needed.