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-Universität Erlangen-Nürnberg

1 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, therefore supporting the development of more effective wildfire management and prevention strategies.

2 Research Questions

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

3 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 an overview of the Data Pipeline.

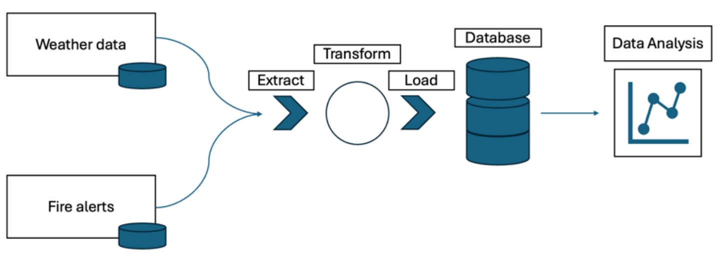

Figure 1: Data Pipeline

4 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 & week:** The year and week in which the data was recorded.
- **avg_temp:** The average temperature for each week, measured in degrees Celsius.
- **min_temp:** The minimum temperature for each week, measured in degrees Celsius.
- **wind_speed:** The average wind speed for each week, measured in meters per second.
- **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.

Wildfire Occurrences: This section contains weekly counts of wildfire alerts, indicating the frequency and intensity of wildfires over the specified period.

The merged dataset spans from week 41 of 2018 to week 41 of 2022.

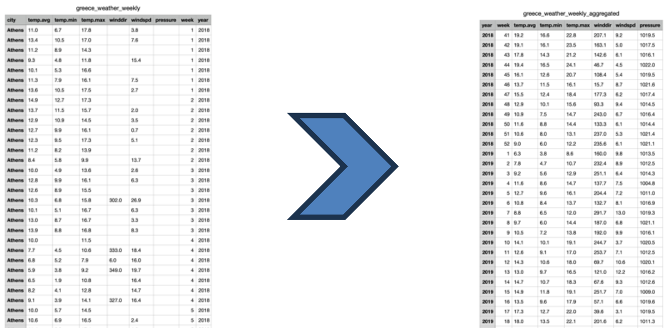

Figure 2: Data Preparation

Data Preparation Process
To prepare the dataset for analysis (Figure 2), the following steps were undertaken:

a. **Filtered for greece:** Included only relevant data.
b. **Date conversion:** Standardized date columns to datetime format.
c. **Weekly sggregation:** Aggregated weather data to weekly averages.
d. **Dropped unnecessary columns:** Removed latitude, longitude, and city columns.
e. **Renamed columns:** Standardized column names for clarity and consistency.

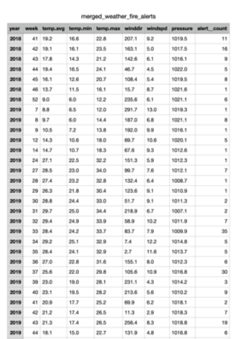

Figure 3: Aggregated data

These data preparation steps were necessary in transforming the raw data into a structured and clean format.

5 Analysis

Method Used

Addressing the research questions regarding the correlation between weather and wildfire:

1. **Data Preparation:**
   - A scatter plot was created to visualize the relationship between average temperature and the number of wildfire alerts. Similarly, another scatter plot was generated for wind speed and wildfire alerts.
   - Pearson Correlation Coefficients (PCC, r-values) were calculated to assess the strength of the correlation between the variables.

2. **Visualization of Relationships:**
   - A 3D scatter plot was created to visualize the combined effect of temperature and wind speed on wildfire occurrences. This plot includes data points for all weeks in the dataset and helps in understanding the multi-dimensional relationship between these variables.

3. **Interaction Effects Analysis:**
   - Interaction terms between temperature and wind speed were included in a regression model to explore how their combination affects wildfire occurrences. This model helps understanding whether the combined effect of highs temperatures, wind speeds leads to more significant risks.

Results

1. **Temperature and Wildfire Alerts:**

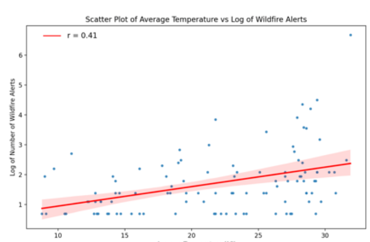

Figure 4: Correlation of average Temp & number of wildfire alerts

   - The scatter plot by the regression line shows a moderate positive correlation between average temperature and the log of wildfire alerts (r = 0.41), indicating that higher temperatures are associated with an increase in wildfire occurrences.

2. **Wind Speed and Wildfire Alerts:**

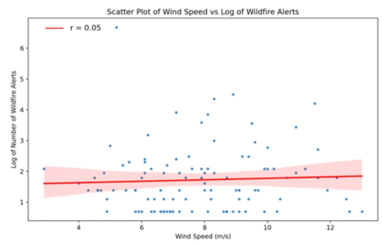

Figure 5: Correlation of wind speed & number of wildfire alerts

- The scatter plot for wind speed and log of wildfire alerts reveals a very weak positive correlation (r = 0.05), suggesting that wind speed has a small impact on frequency of wildfires.

3. **Combined Effect of Temperature and Wind Speed:**

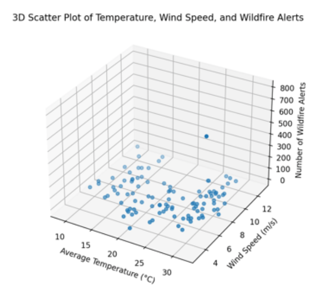

Figure 6: Interaction effects of temp & wind speed

- The 3D scatter plot indicates that regions with both higher temperatures and higher wind speeds tend to have more wildfire alerts. However, the effect of temperature is more dominant compared to wind speed.
- The regression model including the interaction term supports this observation, showing that the combined effect of temperature and wind speed significantly influences wildfire occurrences.

6 Conclusion

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

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.

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

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.

7 Limitations

a. **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.
b. **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.