Development and Optimization of a Machine Learning Fire Risk Prediction Model Using Multimodal Geospatial and Climate Data in Cyprus

Maria Prodromou[1,2], Stella Girtsou[3], Panagiotis Vasiliou[1], Georgios Leventis[1], Alexis Apostolakis[3], Marios Tzouvaras[1,2] , Christodoulos Mettas[1,2], Giorgos Giannopoulos[3], Haris Kontoes[3],Diofantos Hadjimitsis[1,2]

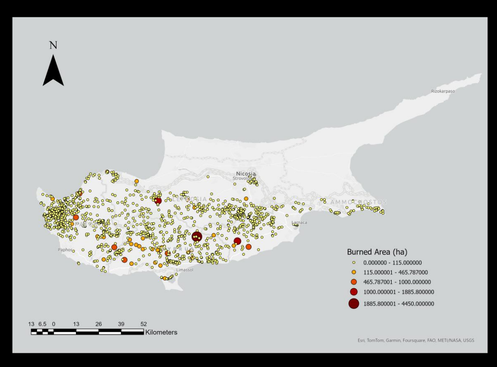
[1]ERATOSTHENES Centre of Excellence, Limassol, 3012, Cyprus
[2] Department of Civil Engineering and Geomatics, Cyprus University of Technology, Limassol, 3036, Cyprus
[3] National Observatory of Athens, Operational Unit BEYOND Centre for Earth Observation Research and Satellite Remote Sensing IAASARS/NOA, GR-152 36 Athens, Greece

Problem Definition

- In recent years, extreme fire events have been on the rise, becoming more destructive and deadly in many areas worldwide, and the recent incidents in Europe have highlighted the severity of the situation
- Wildfires occur in all ecosystems but are particularly characteristic in the Mediterranean region
- Mediterranean countries have recorded maximum burned areas during the last decades due to extreme weather events, especially heatwaves
- Evaluating the integrated spatiotemporal patterns of natural hazardous events is crucial for effective disaster risk management.

This study aims to leverage machine learning for wildfire risk assessment using multimodal geospatial and climate data in Cyprus.

Study Area



Cyprus is a member of the European Union and is located at the north-eastern end of the Mediterranean basin covering an area of 9251 Km^2. The island features a pre-dominantly mountainous terrain and a distinctly Mediterranean climate, characterized by dry and hot summers that typically extend from May to October.

Based on the records provided by the Department of Forests in Cyprus from 2000 to 2023, forest fires in Cyprus destroyed **over 55240 ha of burnt areas**, including state forests and surrounding areas.

Also, **37.2%** of fire events in Cyprus were **deliberate**, followed by agricultural activities that correspond to 19.7%.

Methodology

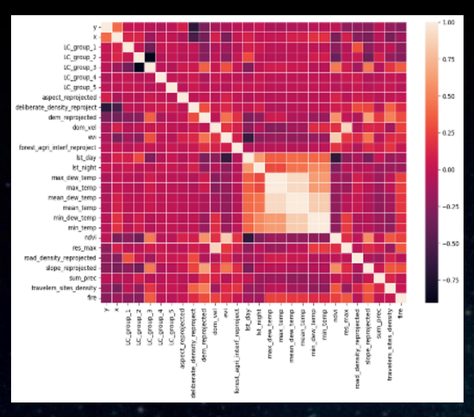
•This study presents the actions that are currently being conducted through a demonstration project in the framework of the EXCELSIOR project, entitled "Development of the National Next-Day Wildfire Risk Prediction AI Model for Cyprus" between the ERATOSTHENES Centre of Excellence and the National Observatory of Athens.

•Features: Anthropogenic, Environmental, Meteorological, Topographic, Fire-related.

• We consider grid cells 500 m wide to cover the whole of Cyprus. Considering that each of these cells "generates" daily instances for a 7-month fire period and a 24-year interval

•Preprocessing: Handling NaN values and normalization.

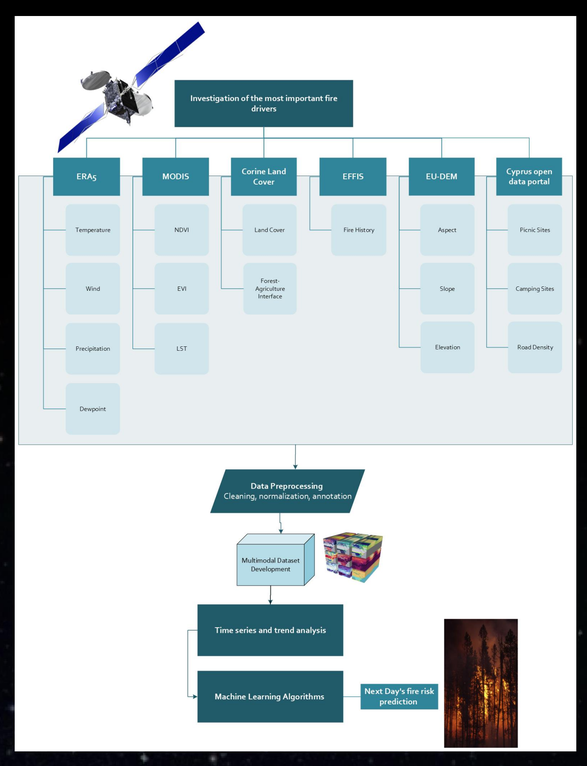
•Machine Learning: Random Forest Classifier, Hyperparameter tuning using RandomizedSearchCV.



Cross Validation

- Dataset : 2001-2024 multimodal data.
  - ❑ 2001-2022: Train-Validation dataset (Fire period)
  - ❑ 2023-2024: Test Dataset
- GroupKFold cross-validation for model robustness
- Performance Evaluation: Precision, Recall, F1-score

$$Recall = \frac{TP\ (True\ Positives)}{TP(True\ Positives) + FN(False\ Negatives)}$$



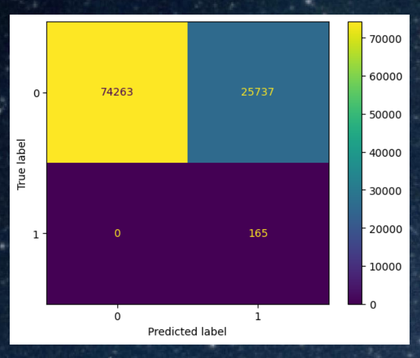
Preliminary Results

Model Accuracy & Recall Performance

- **Cross-Validation Approach:**

To manage large training data and computational costs, we under sample no-fire instances. Using k=10k=10, we prevent data leakage and overfitting by ensuring data from the same day stays within a single fold.

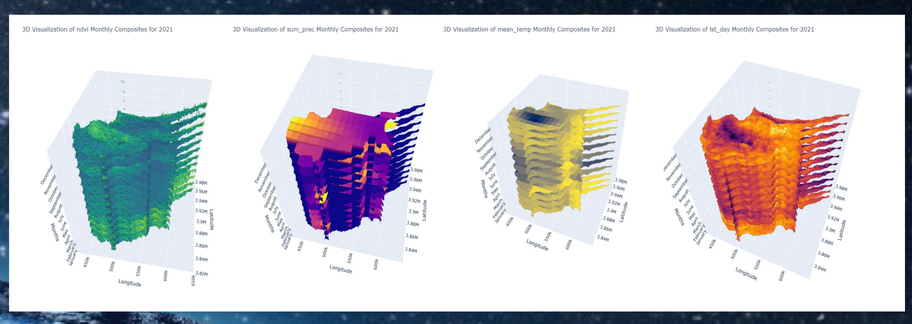
- The model evaluated based on the validation set



- **Top-performing model achieved on the test set:**
  - ❑ Recall =100% for fire and no-fire classes
  - ❑ Recall =74% for no-fire classes
  - ❑ Balanced recall rates ensure effective fire prediction

Practical Implications

- Model shows high potential for continuous monitoring and real-time fire risk assessment
- Can be used as a commercial wildfire prediction service
- Supports authorities in early warning and risk mitigation



Future Work

- **Comparison with other models: Decision Trees, Support Vector Machines (SVM), and Artificial Neural Networks (ANNs)**
- **Goal: To further improve wildfire prediction reliability and robustness in early warning and risk mitigation**

Acknowledgments

The authors acknowledge the 'EXCELSIOR': ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming project (www.excelsior2020.eu). The 'EXCELSIOR' project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510, from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development and the Cyprus University of Technology.

CONSORTIUM
AFFILIATED ENTITIES
ACKNOWLEDGEMENTS

This project has received funding from the Government of the Republic of Cyprus through the "Directorate General for European Programmes, Coordination and and Development".

This project has received funding from the European Union's "Horizon 2020 Research and Innovation Programme" under Grant Agreement No 857510".

This project is co-funded by the Cyprus University of Technology.