Scalable & explainable ML for wildfire risk modeling in Southern Europe: A case-study in Portugal

Ophélie Meuriot[1], Jorge Soto Martin[1], Beichen Zhang[2], Francisco Camara Pereira[1], Martin Drews[1]

[1]DTU Management – Climate and Energy Policy, Division for Climate Economics and Risk Management, Technical University of Denmark
[2]Lawrence Berkeley National Laboratory, Climate and Ecosystem Sciences Division, Berkeley, United States.

1 Motivation

Wildfires in Europe

- Wildfires are among Europe's most destructive natural hazards, affecting ecosystems, economies, and communities (EEA, 2024).

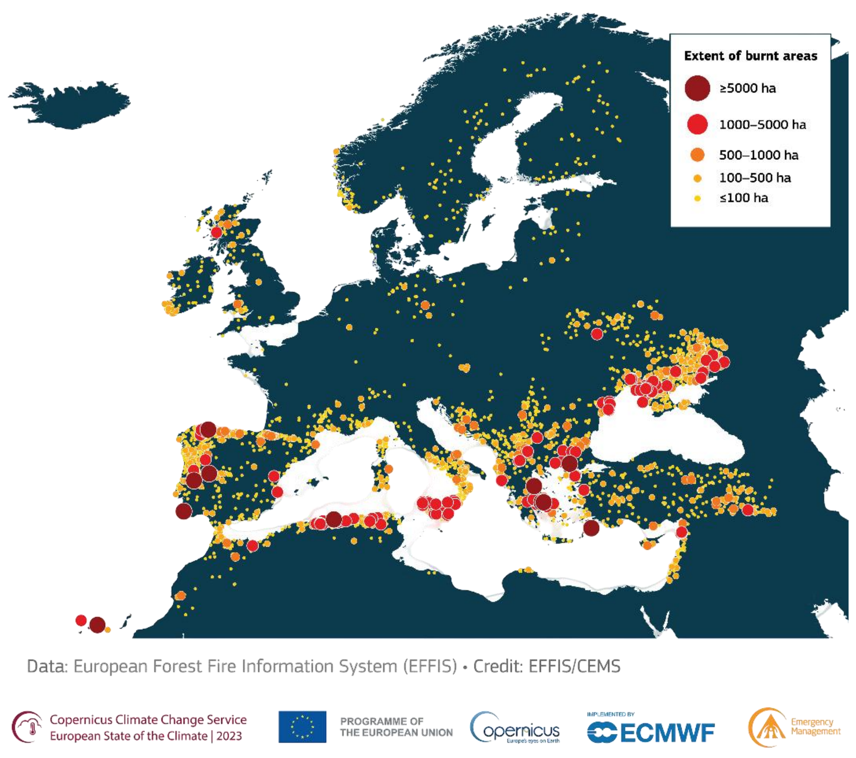
Wildfires in Europe

- Traditional fire indices rely mainly on weather, missing key human and landscape drivers

Wildfires in Europe

- Remote sensing data enable ML-based wildfire risk assessment through data integration (Bountzouklis et al., 2023; Oliveira et al., 2012; Dorph et al., 2022).

Burnt areas across Europe and the Mediterranean in 2023



Distribution and extent of burnt areas across Europe and the Mediterranean in 2023.
Data source: European Forest Fire Information System (EFFIS). Credit: EFFIS/CEMS.

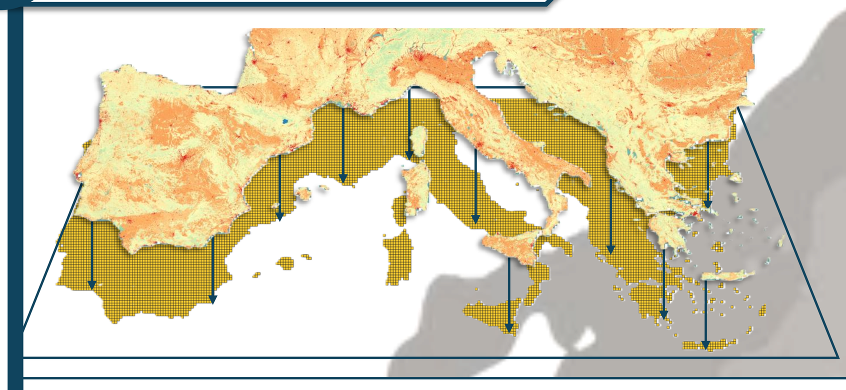
2 Objectives

I Develop a data-driven model to estimate daily wildfire probability

II Use explainable AI to quantify the impact of key predictors

III Assess model scalability and robustness in an unseen region

3 Methodology



Input: The data is aggregated to 12 x 12 km grid

Predictand

- Fire season (Jun – Oct)
- 50 ha 2018 – 2023
- 50% fire – 50% no fire

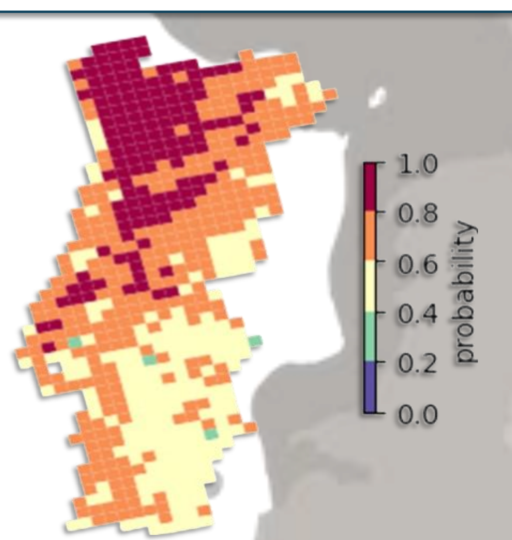
Predictors

- Weather data (temp, pr, ws, rh)
- Land cover
- Topography
- Human activity

Model Selection and Explainability

- 4 classification models
- Hyperparameter tuning on training data using a 5-fold CV & final model selection on testing data
- SHAP used for **model explainability**

Training 2008 – 2021; Testing 2021 – 2023



Model scalability

- Best performing model is run in Portugal (**unseen region**)
- Focus on the Oct. 2017 fires during which 51 causalities were recorded
- Key predictors of the event identified using SHAP

4 Model Selection & Performance

Models trained

- Support Vector Machine (SVM), Random Forest (RF), K-nearest neighbors (kNN) and Logistic Regression (LR)

Model performance

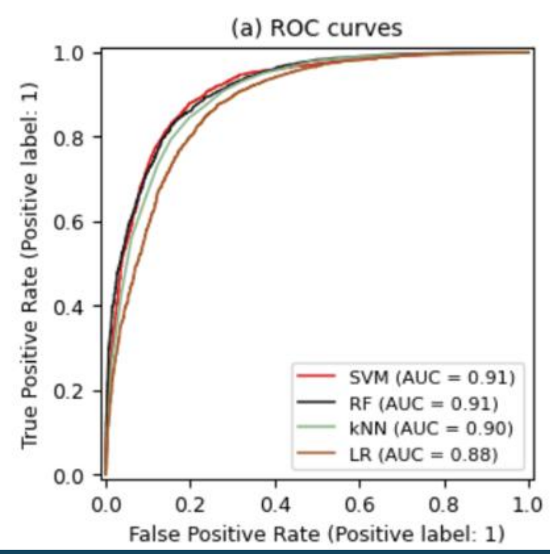
Model	F1 score	Log loss
SVM	0.91	0.38
RF	0.91	0.38
kNN	0.90	0.72
LR	0.88	0.44

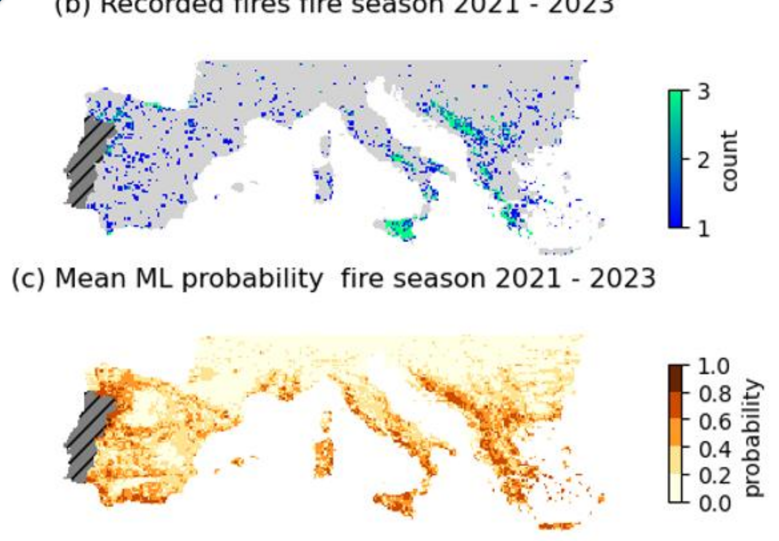
RF model selected for scalability and computational cost

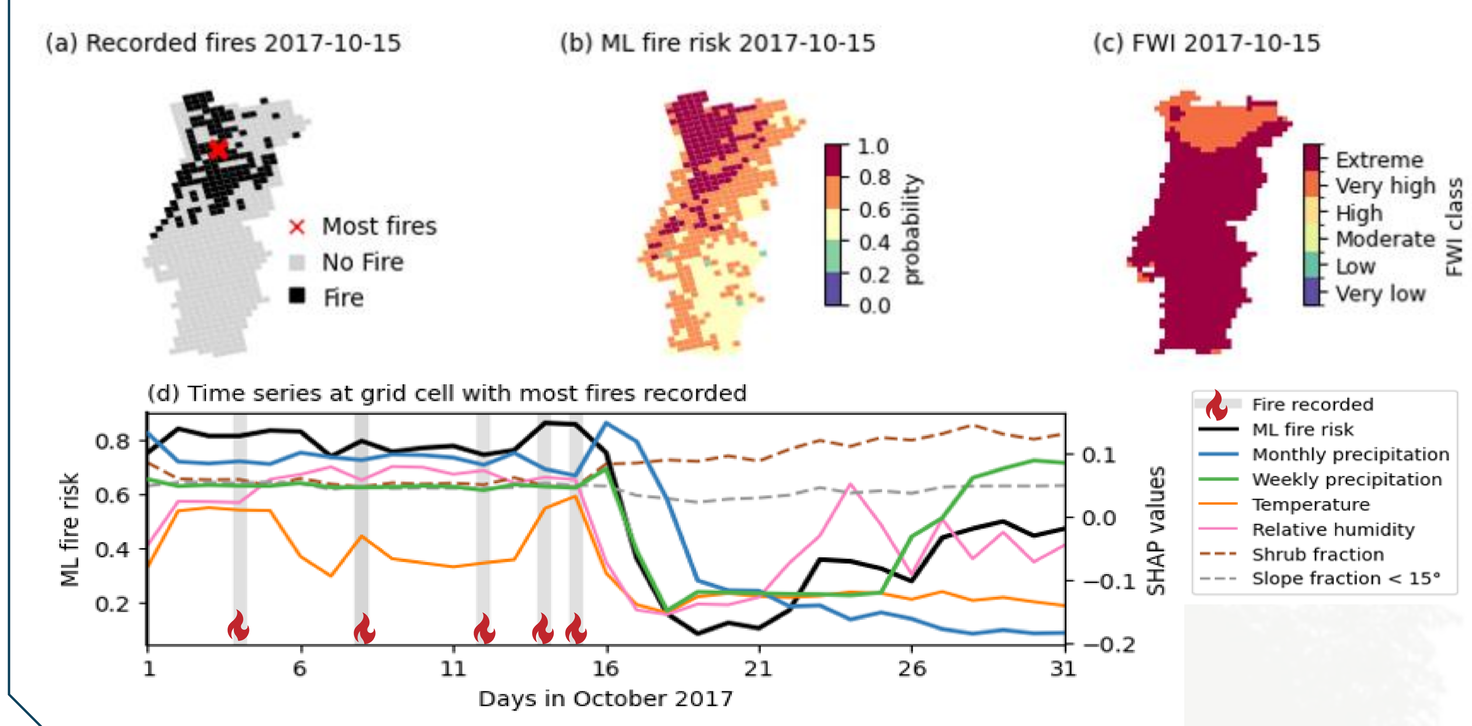
RF model evaluation over 2021 - 2023

- Model run daily with the classification probability used to assess wildfire risk.
- Good agreement between historical record of fires and mean wildfire probability.

RF model can successfully represent high wildfire risk areas







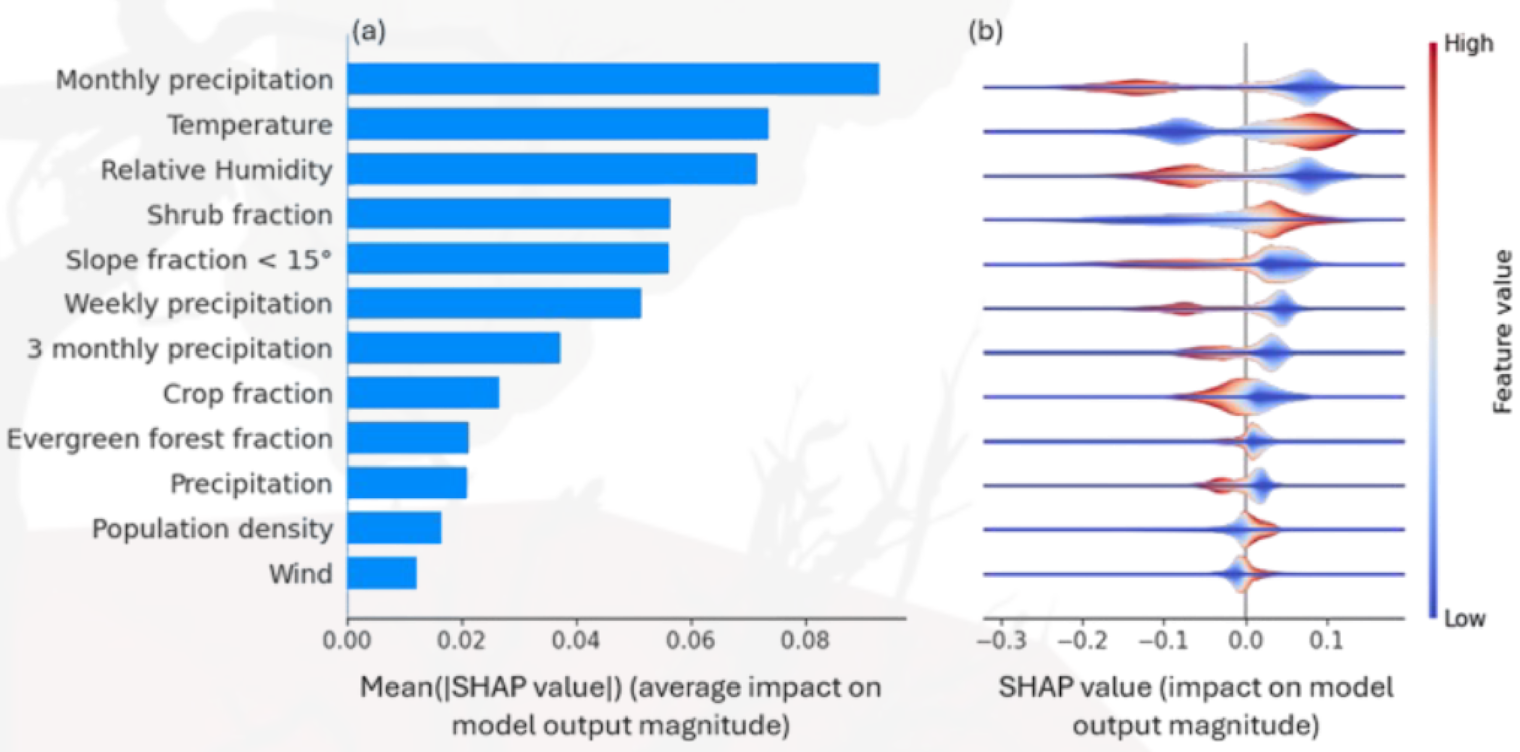
Spatial assessment:

- Close alignment between the RF model **high probability** (Fig a.) and the **recorded fires** on the same day (Fig b.)
- The **FWI** (Fig c.) is **not able to capture** the areas where wildfires where recorded.

Temporal assessment (Fig d.):

- ML fire risk **high when fires were recorded** and drops in the second half of the month
- Drop in fire risk associated with drop in the SHAP values of the **weather variables** but

5 Results



High model output

Low

- Precipitation
- Relative humidity
- Slope < 15°

High

- Temperature
- Shrub fraction

6 Conclusion & next steps

Suitability of data driven models to represent wildfire risk in Southern Europe

Explainability of ML model used to identify influential predictors for wildfire risk

Scalability of model demonstrated by running model in unseen region

Pathways to Impact

- Light-weight and adaptable ML model
- Identification of key wildfires drivers
- Support for policy & land management decisions

Next steps

- Extend analysis to other regions
- Future projections of wildfire risk

7 References

C. Bountzouklis, D. M. Fox, and E. Di Bernardino. Predicting wildfire ignition causes in southern France using explainable artificial intelligence (xai) methods. Environmental Research Letters, 18,4 2023. ISSN 17489326. doi: 10.1088/1748-9326/acc8ee.

A. Dorph, E. Marshall, K. A. Parkins, and T. D. Penman. Modelling ignition probability for human-and lightning-caused wildfires in Victoria, Australia. Natural Hazards and Earth System Sciences,22:3487–3499, 10 2022. ISSN 16849981. doi: 10.5194/nhess-22-3487-2022.

S. Oliveira, F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J. M. Pereira. Modeling spatial patterns of fire occurrence in mediterranean Europe using multiple regression and random forest. Forest Ecology and Management, 275:117–129, 7 2012. ISSN 03781127. doi: 10.1016/j.foreco.2012.03.003.