#RC2

Reviewer Comments

The authors in this paper present a comprehensive and methodologically robust assessment of human-caused wildfire ignition probability across diverse European landscapes. By combining machine learning techniques, specifically Random Forest models, with high-resolution geospatial and socio-environmental data, they deliver both localized and regionally integrated ignition probability models. The overall quality of writing is good, with a clear structure, appropriate referencing, and a sound methodological framework.

The topic is highly relevant and timely, particularly given the increasing wildfire risk under changing climate and land use dynamics in Europe. Importantly, the authors' focus on the interplay between local ignition drivers and their generalization into a full model provides valuable insight into the complexity and variability of fire ignition processes. This shift from local to pan-European modelling is of great significance for the development of integrated fire management strategies at the EU scale.

Overall, the manuscript is a solid contribution to the scientific understanding of ignition patterns and offers operationally meaningful outcomes for fire risk management and prevention across Europe.

Thank you for your thorough and constructive comments. We have carefully addressed all your suggestions and concerns. All our responses are highlighted in **red** for clarity. Furthermore, we have added a Supplementary Material section that provides additional details on data, methods, and overfitting analysis.

Below I provide a series of detailed comments and questions that may help the authors strengthen the manuscript even further:

Specific Comments and Questions

- **Section 2.3.3**: How did you identify the so-called "mixing areas" algorithmically? Some additional details about the method used would be appreciated.

  We did not explicitly delineate "mixing areas" as unique classes. Instead, for each cell we calculated the percentage of forested, agricultural, and urban land cover. To better capture the effect of mixed-use zones and the associated increase in human activity, we also included in the model the interfaces between wildland and urban, agricultural, and grassland areas. To avoid reader confusion, in the manuscript we substitute the term by *intermix of* by *presence of urban, agricultural*…(L173)

- **Section 2.3.4**: What are the FirEUrisk fuel classes used in the study? Can these be differentiated enough to capture important distinctions in land cover such as eucalyptus in Portugal, which, despite being a broadleaf, behaves quite differently due to its high flammability? Also, how did you project a 10-meter resolution (CLC+ Backbone) raster to 100 meters, considering the categorical and delicate nature of land use data?

We thank the reviewer for this observation. Regarding the fuel classes, we relied on the classification developed within the **FirEUrisk project**, designed for the pilot sites and able to differentiate broad vegetation types relevant to fire risk (e.g., coniferous forests, broadleaved forests, shrublands), although it does not reach the level of detail of specific species such as eucalyptus stands. For further details on the categories, we refer to Aragoneses et al. (2023) (https://essd.copernicus.org/articles/15/1287/2023/). The assignment table was provided in the manuscript as Supplementary Table S2. Concerning the projection from 10 m to 100 m, since this is a categorical variable, we applied a majority vote criterion, assigning each 100 m cell to the dominant class, which preserves spatial consistency. We have updated the supplementary material further details have been added in section 2.2.4 and the following table in supplementary material.

FirEUrisk fuel type		FirEUrisk fuel type	
Code	Description	Code	Description
1111	Open broadleaf evergreen forest	23	High shrubland [≥ 1.5 m]
1112	Closed broadleaf evergreen forest	31	Low grassland [0–0.3 m]
1121	Open broadleaf deciduous forest	32	Medium grassland [0.3–0.7 m]
1122	Closed broadleaf deciduous forest	33	High grassland [≥ 0.7 m]
1211	Open needleleaf evergreen forest	41	Herbaceous cropland
1212	Closed needleleaf evergreen forest	42	Woody cropland
1221	Open needleleaf deciduous forest	51	Wet and peat/semi-peat land – tree
1222	Closed needleleaf deciduous forest	52	Wet and peat/semi-peat land – shrubland
1301	Open mixed forest	53	Wet and peat/semi-peat land – grassland
1302	Closed mixed forest	61	Urban continuous fabric
21	Low shrubland [0–0.5 m]	62	Urban discontinuous fabric
22	Medium shrubland [0.5–1.5 m]	7	Nonfuel

- **Line 208**: What exactly is meant by "null model"?

We thank the reviewer for this comment. By *"null model"* we were referring to an *initial model*, in which we optimized the hyperparameters of the Random Forest algorithm for each Pilot Site before running the final experiments. To avoid confusion, we have replaced the term *null model* with *initial model* throughout the manuscript (L214).

- **Line 210**: The term *number of predictors* should be highlighted, perhaps using italics or quotation marks, for clarity.

Done.

- **Line 214**: Reference to "Section 0" is likely a formatting or numbering error and should be corrected.

(Removed from text.

- **Line 214 (continued)**: Was the Autocorrelation Control (AC) also used in the full model with all the regions? If so, what was the bounding box adopted?

Yes, the Autocorrelation Control (AC) term was also used in the full model. The bounding box was defined by the extreme coordinates of the pilot sites: the

northernmost point of PS1 (Northern Europe), the easternmost point of PS5 (East Attica), and the southernmost and westernmost points of PS3 (Central Portugal). Please see section 2.4.1 (L223-L225).

- **Line 217**: The AC strategy deserves more clarification. It seems to include:

  - Distance from the center of each PS

  - Distance from the corners of each PS bounding box

  - x and y coordinates in the adopted CRS

Yes, the AC strategy includes these variables, which correspond to the Euclidean Distance Features (EDFs) described by Milà et al. 2024 (https://gmd.copernicus.org/articles/17/6007/2024/). These variables were incorporated to account for spatial autocorrelation effects. Please see 2.4.1 section (L220-223).

Given this, I would expect that a local PS model relying too heavily on the AC variables (as visible from importance rankings) could be overfitting the ignition patterns of its training set rather than capturing true statistical drivers of ignition. The dummy variable for the PS code used in the full model seems to act similarly to the AC, but at a larger scale. Could you please clarify what is meant by "The AC control successfully alleviated spatial autocorrelation…" (line 217) and confirm whether "disregarding AC control" means simply removing all AC variables from the feature set?

In our study, Euclidean Distance Features (EDF) were included to control for spatial autocorrelation. It is true that EDF can potentially induce overfitting (Milà et al., 2024), particularly when these features rank among the top predictors in variable-importance measures. Consequently, importance ranks in models that include such predictors should be interpreted with great caution, and emphasis should be placed instead on explanatory variables (Meyer et al., 2019; Wadoux et al., 2020). Please check Reviewer 1 author's reply, where we show the results of a model's overfitting performance. Otherwise, the same plot is included in supplementary material (Figure S1)

Regarding the clarification on line 217, it should be noted that incorporating the spatial autocorrelation (AC) term as a predictor substantially reduced the number of realizations showing residual spatial autocorrelation, with an overall reduction of about 66% across pilot sites.

Meyer, H., Reudenbach, C., Wöllauer, S., and Nauss, T.: Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction, Ecol. Model., 411, 108815, https://doi.org/10.1016/j.ecolmodel.2019.108815, 2019.

Milà, C., Ludwig, M., Pebesma, E., Tonne, C., and Meyer, H.: Random forests with spatial proxies for environmental modelling: opportunities and pitfalls, Geosci. Model Dev., 17, 6007–6033, https://doi.org/10.5194/gmd-17-6007-2024, 2024 .

Wadoux, A. M. J.-C., Samuel-Rosa, A., Poggio, L., and Mulder, V. L.: A note on knowledge discovery and machine learning in digital soil mapping, Eur. J. Soil Sci., 71, 133–136, https://doi.org/10.1111/ejss.12909, 2020

- **Figure 2 (Line 250)**: This figure highlights an extremely important aspect. The choice of presence and pseudo-absence points can shift AUC from 0.4 to 0.9, as in the case of East Attica. This issue is critical and often neglected in wildfire susceptibility literature.

  This variability evidences the poor performance of the model in this pilot site due to data scarcity and uneven ignition distribution. Expanding the ignition sample—if governmental datasets become open—or complementing the analysis with similar regions could substantially improve model robustness in eastern mediterranean basin. To amend this limitation, we employed repeated subsampling and reported the distribution of AUC values rather than relying on a single realization. For cartographical representation we select the model that has an AUC closer to median AUC of the different repetitions. We now stated the East Attica overfitting problem in 3.1 (L 253-254) and 4.4 sections (L421-L423).

- **Line 260**: Please specify that "dry and warm season" refers to the Swedish climate, which may not be intuitively understood by all readers.

  We added this clarification: "*which in the context of the Northern Europe climate correspond to its relatively warm and dry summer months*" in lines 270-271

- **Line 266 and elsewhere**: The phrase "chances of ignition" may be misleading. Human sources of ignition (e.g., arson, negligence) are usually orders of magnitude higher than the fires actually recorded. What determines whether a fire is recorded is its success in developing beyond a minimal threshold. It would be more precise to refer to "chances of successful ignition" throughout the manuscript.

  Changes done.

- **Caption of Table 2**: The sentence should be revised to: "The top 5 variables **for each column** are highlighted in grey."

  Done,

- **Figure 4 (Line 310)**: While the figure is clear and well done, the discussion could be enriched by acknowledging a critical issue in model interpretation: susceptibility values from RF (ranging 0 to 1) cannot be meaningfully compared across pilot sites. A 0.999 value in Sweden does not equate to a 0.999 in Attica. This is where the full model provides value by smoothing across regions. In some of my previous work, I have addressed this using quantile ranking—i.e., describing a pixel as "top 5% susceptibility" within its region, rather than relying on the raw RF voting score. Consider discussing this approach or acknowledging the issue.

  We thank the reviewer for highlighting this important point. As noted, raw Random Forest (RF) probability values are calibrated from the local training data of each pilot site and thus cannot be directly compared across regions—a 0.999 value in Sweden does not represent the same absolute risk as a 0.999 in Attica. The full continental

model partly mitigates this issue by pooling data across sites, which smooths differences in scale. Importantly, our focus is not on comparing absolute probability values between sites, but rather on analyzing the drivers and spatial distribution of ignitions. We now explicitly refer to probability values in terms of quintiles in section 3.3 and updated the figure 4. in which the results are represented by quintiles. Following your suggestion, this approach improves interpretability and highlights the relative distribution of probabilities more clearly.

- **Line 325**: Since temporal variability is removed by taking average values, I assume your model highlights "spatial areas where extreme dry events tend to occur" rather than correlating specific years with ignition. Is that correct?

  That is correct in part. The mean values capture the general spatial tendency of where extreme dry events are more likely to occur, as in Mediterranean basin. However, by also incorporating Z-Scores, we account for interannual variability, allowing the model to identify whether a given year is characterized as extreme or not.

- **Line 339**: "Modest climate conditions" is ambiguous. I believe what you mean is that Attica's fire season is, under climatic/weather conditions, uniformly and persistently extreme. It would be more accurate to use the phrase "fire-prone climate" here instead of "favourable climate".

  Changed

Additional Literature Suggestions

- **On RF versus other techniques, and wildfire susceptibility modelling performance** (Section 2.4, line 195): Trucchia, A., Izadgoshasb, H., Isnardi, S., Fiorucci, P., Tonini, M. (2022). *Machine-Learning Applications in Geosciences: Comparison of Different Algorithms and Vegetation Classes' Importance Ranking in Wildfire Susceptibility*. Geosciences, 12(11), 424. https://doi.org/10.3390/geosciences12110424

  Incorporated in line 205