

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC2  
<https://doi.org/10.5194/nhess-2021-384-AC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC2

Sigrid Jørgensen Bakke et al.

---

Author comment on "A data-driven model for Fennoscandian wildfire danger" by Sigrid Jørgensen Bakke et al., Nat. Hazards Earth Syst. Sci. Discuss.,  
<https://doi.org/10.5194/nhess-2021-384-AC2>, 2022

---

**AC: Many thanks for your time and efforts in evaluating our manuscript. We highly appreciate your positive and constructive feedback. In the following, we would like to respond to the comments.**

RC2: This manuscript uses machine learning methods to predict fire danger in Fennoscandia at approximately 0.25 degree spatial scale for 2001-2019. Here, the authors are using official statistics compared to MODIS burned area, with predicted fire danger probability models compared to the results from the Canadian Fire Weather Index. The method is novel and the comparison is rigorous, but the data and approach need to be explained more – and at times even cited better – to assess the efficacy of the model.

RC2: In general, this manuscript needs to be revised in order to understand why this method may be useful for predicting fire danger probabilities.

**AC: We will state our motivation clearer in the abstract. In the manuscript body, we believe the usefulness of the method is sufficiently justified. The reasoning is introduced, discussed and concluded; it links the background and the objectives (line 99-111) in the introduction, it is discussed in line 589-592 and Sect 4.5, and it is emphasised in the conclusion (line 637-640 and line 660-664).**

RC2: First, the authors should explain what fire danger is as opposed to fire occurrence.

**AC: We agree that we should clarify the difference between fire danger and fire occurrence. We will find a suitable place in the introduction to make this clarification.**

RC2: Second, why are burned area data used as 'fire occurrence' when satellite-based active fire detections are available?

**AC: The burned area data is used to get a binary fire/no-fire dataset based on the same resolution as found for many global climate models, to see if a data-driven model is able to make predictions of a dataset existing at this spatial scale. The active fire products detect burning at the time of overpass given relatively cloud-free conditions, which can be a problem for parts of Fennoscandia that are seldom cloud-free. The burned area product is considered less sensitive to cloud-cover and is also more reliable as it has a longer temporal**

**influence. Further, by detecting the structural consequences of fires, the burned area product have a more direct relevance to climate-relevant consequences, such as albedo and ecosystem functioning. We will make a comment on this in the revised manuscript. We acknowledge that more analyses comparing different target datasets would be an interesting continuation of our study. We made one such comparison of the target dataset by including a fire record of Norway. We chose this over a satellite-based active fire detection dataset because it clearly separates each fire occurrence from others and all known occurrences are registered regardless of the (e.g. heat) signal captured by the satellite.**

RC2: Finally, the manuscript does not describe fully many of the datasets used, including where to obtain them and what their uncertainty are.

**AC: We will carefully check the data section and add details, explanations and citations when lacking. We will include information of their uncertainties (as available), in particular for the Norwegian fire occurrence dataset. Most datasets are established datasets with known validation studies. We will add well-known uncertainties in the revised manuscript.**

RC2: Finally, the results seem to indicate that a single shallow soil moisture variable is driving the predictions (which is not usually considered in fire danger modeling like FWI). A major revision and resubmission is recommended.

**AC: The result indicate that a shallow soil moisture variable is the dominant predictor, however not sufficient alone to make a good prediction (emphasised e.g. in line 500-502). As you state, soil moisture is usually not considered in fire weather indices such as the FWI (this is commented on in general terms in line 637-638).**

Specific comments:

- RC2: The title is "A data-driven prediction model for Fennoscandian wildfires" but the thesis of the paper is to produce spatiotemporally resolved fire danger probability maps – which is not quite the same as predicting wildfires. Consider revising the title to be more specific.

**AC: The authors agree, and suggest the title: "A data-driven model for Fennoscandian wildfire danger"**

- RC2: Line 19: "which stores approx. 30% of the world's soil carbon pool" needs a citation

**AC: This is stated in the paper cited in the end of the sentence, i.e. Flannigan et al. (2009): "Boreal regions store about 30% of the world's soil carbon pool..."**

- RC2: Lines 26-27: "However, to the best of our knowledge, fire studies of the European boreal zone are limited." needs a citation.

**AC: We have not found a paper stating this specifically, and the statement here is therefore based on our literature search. This is why we emphasise that it is "to the best of our knowledge" in the beginning of the sentence.**

- RC2: Line 144: What is the spatial resolution of a European Space Agency Climate

Change Initiative (ESA145 CCI) product version 5.1.1cds? Please include that.

**AC: The spatial resolution is 0.25 deg longitude/latitude. We agree that we should state this earlier in the paragraph (it is mentioned in line 151-152), and will move the statement to the beginning of the paragraph.**

- RC2: Line 146-147: "and is based on Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Reflection information" is not correct way to right this. It should be "the reflectance product of the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on the Terra satellite". Can the authors please specify which reflectance information is used? Daily surface reflectance?

**AC: Thank you for pointing out the correct writing; this will be corrected in the revised manuscript. The main source of data are daily surface reflectance information in the red and Near Infrared bands. The algorithm theoretical basis is found under documentation at the reference given (specifically [http://datastore.copernicus-climate.eu/documents/satellite-fire-burned-area/D1.6.2-v1.0\\_ATBD\\_CDR\\_BA-FireCCI\\_MODIS\\_v5.1cds\\_PRODUCTS\\_v1.0.1.pdf](http://datastore.copernicus-climate.eu/documents/satellite-fire-burned-area/D1.6.2-v1.0_ATBD_CDR_BA-FireCCI_MODIS_v5.1cds_PRODUCTS_v1.0.1.pdf), which is based on [https://climate.esa.int/media/documents/Fire\\_cci\\_D2.1.3\\_ATBD-MODIS\\_v2.0.pdf](https://climate.esa.int/media/documents/Fire_cci_D2.1.3_ATBD-MODIS_v2.0.pdf)). We will include this reference in the revised manuscript.**

- RC2: Section 2.2 Norwegian fire occurrence dataset – the authors have not provided a citation to the dataset, where it can be accessed, and how it is collected. Are these truly wildfires or are these fires from all ignition sources (lightning plus human-caused)? Is there a burned area minimum that fires must meet to be included in this wildfire dataset? Please describe this dataset more.

**AC: We assume the reviewer is referring to Sect. 2.1.2 and not 2.2 here. We will provide more details and citation (<http://www.brannstatistikk.no>) to the dataset in the revised manuscript. We are unsure what you mean by "truly wildfires" (do you mean only the wildfires ignited by lightning?) as opposed to "fires from all ignition sources". The dataset comprise all fires in grass, cultivated land, forests and uncultivated land, regardless of ignition source. We do not define wildfires depending on the type of ignition source in our study. The data are based on the fire and rescue service reporting system in Norway (*brann- og redningstjenestens rapporteringssystem; BRIS*). There is no lower limit of burned area in this dataset, as it is based on fire responses of the fire department.**

- RC2: Line 164: Why were the months April – September selected?

**AC: The Norwegian fire occurrence dataset must cover the same months as the satellite based fire occurrence dataset, and the reason for omitting October to March in the satellite based fire occurrence dataset is given in line 152-153 (few fire occurrences). We will clarify this in the revised manuscript.**

- RC2: Figure 3: The authors are using burned area from the European Space Agency Climate Change Initiative (ESA145 CCI) product version 5.1.1cds but noting it as fire occurrence and number of fires. Can the authors describe how this was done with the burned area product?

**AC: The transition from burned area to fire occurrence is explained in Sect. 2.1.1 (line 147-154), and the transition from the national record to the Norwegian fire occurrence dataset is described in Sect. 2.1.2 (line 161-163).**

RC2: Is this the most appropriate comparison of burned area to number of fires in the official statistics? What is the original spatial resolution and what is lost when aggregated to 0.25 degrees?

**AC: None of the two datasets is directly comparable to the number of fires in official statistics because they are both aggregated in space and time. It is not an aim of the study to make the datasets directly comparable to official statistics, but rather see if a data-driven model is able to predict fire occurrences at the spatiotemporal resolution (0.25 deg regular grid and monthly time step) used in the study. The original spatial resolution of the burned area product is 250m. We have not evaluated what is lost when aggregated to 0.25 degrees, as the aggregated version is an established and verified dataset publically available. However, known uncertainties with the different datasets applied will be commented on when introduced.**

- RC2: Line 230: Can the authors explain how snow cover was used? Especially since the model was limited to monthly values from April to September over the period 2001–2019.

**AC: The (fractional) snow cover is a continuous variable describing the fraction of a given grid cell covered by snow at a given time step, and was used as a potential predictor. Our study region cover high latitudes and altitudes, and snow cover is present in some grid cells and months also in the period analysed.**

- RC2: Line 235: The land cover data and fraction of burnable area is not well described. Which land covers? Why were those chosen? Are all vegetation types are included?

**AC: Because the dataset is publically available, we do not elaborate in details choices made in their creation. However, we will consider commenting in general on their uncertainties along with key references. We have added a reference to the dataset for interested readers to look up (line 237).**

- RC2: Line 241-242: Can the authors provide citations for this statement (and for Norway and Sweden, specifically): "We chose FWI because it is developed for boreal forests and because it is used for fire danger forecasts in large parts of Fennoscandia (Norway and Sweden)."

**AC: We will provide citations for this statement in the revised manuscript. Norway: <https://skogbrannfare.met.no/>, Sweden: <https://www.smhi.se/brandrisk>. We will include Canadian; "...is developed for (Canadian) boreal forests...", to clarify that it was not originally developed for Fennoscandia.**

- RC2: Figure 6: Should readers interpret Figure 6 as the only important variable to be soil moisture anomalies in the layer 7-28 cm? It would be helpful for the authors to spend more time explaining why this figure is important for creating a data-driven model, i.e., variable selection.

**AC: No, Figure 6 should not be interpreted this way. The figure shows the importances of the subset of predictors used in the final data-driven model, and is therefore rather showing the opposite; multiple predictors are important, and a model of the soil moisture anomaly alone would not perform well. This is further emphasised by Figure S1, which shows that model performance reduces when reducing the number of predictors, and by Figure 7, which illustrates that swvl2\_anomaly alone is not a sufficient predictor. See e.g. line 500-503. This figure is not important for creating a data-driven model, rather it is a result of the final data-driven model.**

- RC2: Table 1: Should NDVI be included in this as a potential predictor?  
**AC: We considered including NDVI in this table, but concluded not to because the NDVI experiments were performed separately from the main analysis.**
  
- RC2: Figure 8: The red-blue scheme is not colorblind safe. Can the authors change these figures to make them colorblind safe? Tools like colorbrewer can help.  
**AC: We tested the figures for colour blindness using <https://www.color-blindness.com/coblis-color-blindness-simulator/> and the app "Color Blind Pal". We did not find any difficulty for the different colour blind views with this figure. Given your comment, we wonder if we have overlooked a colour blind view. If so, please let us know for which colour blind view this figure is a problem for, so we can correct it? It is of high priority to us to make the figures interpretable for all colour views.**
  
- RC2: Figure 8: At first look, a reader may think that the fire danger probability maps did not perform well, especially compared to the satellite-based fire occurrence (which is really burned area dataset). Using the active fire products from MODIS or VIIRS may provide a better match than the burned area. Further, consider changing the title and better explaining fire danger in the Introduction so that interpretation of the Results is more straightforward.  
**AC: The satellite-based fire occurrence dataset is used to construct the model, which is why we use this dataset in Figure 8. We are unsure if the active fire products would provide a better match, as the main reason for mapping regions of fire danger probability and not to predict fire occurrences as such. Rather, the lack of no fire occurrences is likely related to the lack of ignition source. However, regions with fire occurrences are often mapped with high probability, indicating a good model prediction. We agree that the title is unclear, and suggest to revise it as stated in our answer to your comment 1. We will make sure to carefully go through the description of fire danger and clarify (better explain) the text to ensure it is well understood.**
  
- RC2: Figure 9: Same comment as for Figure 8. Is this colorblind safe? The colors chosen are hard to interpret, particularly in Figure 9c.  
**AC: We tested the figure for colour blindness using <https://www.color-blindness.com/coblis-color-blindness-simulator/> and could not find an issue with the colours for the different colour visions. The same colour scale is used for all three maps to ease the comparison. Figure 9c shows high correlations (above 0.8 for the whole study domain) and thus, is only represented by two of the colours from the scale.**
  
- RC2: Line 500: Most of the figures and results in the manuscript highlight the importance of swvl2\_anomaly only. The manuscript needs to better describe the input and importance of other variables.  
**AC: We disagree that most figures and results highlight the importance of swvl2\_anomaly only. There is only one figure (Fig. 7) in which swvl2\_anomaly is the only predictor shown. This is justified by the relatively high importance of this predictor as compared to other predictor as shown in Fig. 6. All other figures relating to the predictors, show either all potential predictors (Fig. 4), or all selected predictors of a given model (Fig. 6, S3b, S4b, S5b and S13b). All input variables are described in Sect. 2.3, and the selected variables other than swvl2\_anomaly are discussed in the lines 503-535; following the line of your comment. We will carefully go through the text and add information on**

**predictors as seen needed for the interpretation of the overall results. However, commenting on the role of all predictors in more details, also those of less importance, we believe would lengthen an already long text and divert the attention from the key findings.**

- RC2: Lines 535: The authors need to better evidence to say that reanalysis products are helpful when what was used in this study is mainly reanalysis.  
**AC: Stating that “the use of reanalysis products is useful” does not imply that it is useful compared to another alternative, but simply that reanalysis products can be used to construct a well-performing model. We will consider making it clear that other types of data can be useful as alternatives and that this was not tested in our study.**
  
- RC2: Conclusions: Since the subsurface soil layers are the best predictors, can the authors provide some description of this dataset and the uncertainties / validation of the product? This is not described in section 2.3.3.  
**AC: We agree that this can be a valuable information, and will search the literature for description, uncertainties and validation of the dataset and include the relevant information we find.**
  
- RC2: The authors have not shared the data or code and these should be provided. How was this study conducted? In R? In MATLAB? Please provide these details.  
**AC: The data is openly available online, except for the details concerning the Norwegian fire dataset, for which we have provided the source. We will add a ‘code and data availability’ section in the end following the Copernicus template where we will repeat the information given in the data section and acknowledgements related to data availability. We support the general efforts to make code used in publications available to make analyses reproducible. Unfortunately, the code is not in a state appropriate for sharing. However, we will add “Code is available upon reasonable request from the corresponding author” under the ‘code and data availability section’. Here, we will also state that we used R for the SPI and SPEI calculations, and Python otherwise.**