Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC1
https://doi.org/10.5194/nhess-2021-384-AC1, 2022

Reply on RC1

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-AC1, 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.

RC1: This paper uses Random Forests to estimate wildfire probability in the mostly boreal Fennoscandia region. Comparable studies using similar data and Random Forest models have been performed over various spatial domains but this study is the first one focusing on Fennoscandia in particular. The analyses are thorough and very well documented. There are a few issues I would like to see addressed before publication:

What was the motivation to perform the analysis at a 0.25° and not the native MODIS resolution, or at least at the finest meteorological resolution? You lose a lot of spatial detail in this way.Pixel product data are available at a 250 m resolution.

AC: We chose the 0.25 degree resolution to investigate if a data-driven model is applicable for use at the state of the art global climate models, rather than aiming for the highest spatial detail possible. Further, spatial dependency of fires (e.g. the same fire occurring in two or more cells) is reduced when using a coarser scale. We see that the reasoning behind the spatial scale chosen is not stated clearly in our manuscript, and we will do so in the revised version.

RC1: Not including dynamic vegetation predictors or specific land cover is a weakness. Recent work (e.g. Kuhn Regnier et al., 2021) has shown that adding vegetation dynamics has considerable impact on model skill. NDVI not being modelled by DGVMs is not a valid justification as several productivity-related indicators estimated by DGVMs are available from Earth observation. The same applies to (more static) land cover information, such as crop fraction or tree type (e.g. Forkel et al., 2019).

AC: Thank you for the references. As you state, several productivity-related indicators are estimated by DGVMs. Still, most climate model outputs are not based on runs for which the climate model has been coupled with a DGVM. For this reason, we wanted to limit the choice of predictors to those available from climate models without the need of DGVMs. We will clarify our reasoning, and acknowledge the possibility of productivity-based indicators estimated by DGVMs in the revised manuscript (e.g. in Sect. 4.2).

RC1: The same applies to socio-economic drivers such as population density. Fig.3 suggests that there is clear link between wildfire occurrence and population centres. Probably, including crop fraction as variable would already be a good proxy for this.

AC: We agree that socio-economic predictors would likely improve the model prediction. The main reason that we hypothesise this is that humans and human infrastructure are fire starters (line 547-551). Figure 3b of Norway suggests a link between wildfire occurrence and population centres. We suggest this is partly due to humans being a major ignition source as already mentioned, as well as the overlap between human settlement in Norway and burnable areas (a potential predictor included). However, as seen in Figure 2, which shows number of fires over Fennoscandia, the link between human settlement and fires is not clear. We chose to constrain our study to predictors available in global climate models. In a setting where no data restrictions are imposed, we agree that one should test the inclusion of socio-economic and vegetation based predictors.

Detailed comments

RC1: The title is a bit misleading: The model identifies the main hydrometeorological drivers of past wildfire occurrences. It estimates the probability of wildfire occurrence but it does not predict (i.e. forecast) wildfire occurrence itself; It this should be made clear in the title.

AC: We agree, and suggest to revise the title to "A data-driven model for Fennoscandian wildfire danger".

RC1: l68-69: It is misleading to state that fire-weather indices based on climate model and reanalysis data can be used for monitoring and forecasting.

AC: We agree and will rephrase these sentences.

RC1: l91: mention some of these limited studies using data-driven methods to predict intra- and inter-annual dynamics, e.g. Forkel et al. (2017, 2019) and Kuhn-Regnier et al. (2021), who predict monthly global patterns.

AC: Thank you for providing these references. We will include these (and potentially other relevant references) in the revised manuscript.

RC1: l92: How do you define a data-rich region? With recent satellite availability, practically all regions have become data rich and several studies ha

AC: We agree this is an unclear statement and will clarify it in the revised manuscript. The last part of your comment is unfortunately lacking, however, we trust your point has been made in this sentence part (please let us know otherwise).

RC1: l97: "In addition, a bottom-up approach is typically less straightforward in its data requirements and methodology as compared to the process-based approaches" -> explain

AC: We will clarify this in the revised manuscript.

RC1: l123: unclear whether this dataset is used for training or as independent validation reference. If used as target in model development, this doesn't come out clearly in Fig.1 (as it should also be split up into training and testing)

AC: The local (Norwegian) fire occurrence dataset is used as an independent validation reference for research question 2, and used for training (a target in model development) for research question 3. For research question 3, we split up the Norwegian fire occurrence dataset into training and test datasets. This is described in line 375-380. Figure 1 shows the data-driven approach for the Fennoscandian domain (i.e. the one developed using satellite-based fire data as target), and not that of the Norway alone. It is stated in the figure caption, but we will clarify it in line 132 as well. Further, we will make a line break after punctuation in line 375, to separate the two applications of the Norwegian dataset. We considered including it is the figure, but concluded it would make the figure more messy than clarifying.

RC1: l127: How can the machine learning algorithm both be simpler and more sophisticated?

AC: The simpler and more sophisticated machine learning algorithm are two separate algorithms (Decision tree is the simpler and AdaBoost is the more sophisticated one). We will clarify by rephrasing the sentence.

RC1: l181: Why are dynamic vegetation predictors not included? Recent work (e.g. Kuhn Regnier et al., 2021) has shown that adding vegetation dynamics ha considerable impact on model skill.

AC: With reference to our earlier comment on the subject (based on your question regarding DGVM), we will clarify our reasoning in this section in the revised manuscript.

RC1: l220: why is wind speed included as predictor? More a predictor of fire spread than of occurrence

AC: For a fire to occur in the burned area dataset, it must have been of a size recognisable for the satellite. Thus, the fire must have spread to some degree (due to wind or not). Another effect of the wind is drying of the ground and vegetation by increasing evapotranspiration prior to the fire. Regardless of the reason, wind was found to be a selected predictor, indicating its importance in predicting the fire occurrence dataset.

RC1: l314: Which threshold was used beyond which no more predictors were removed?

AC: We used no threshold; a predictor subset was made for all (each) number of predictors (Np), as stated in line 314. This can also be seen in Fig. S1 that shows the average cross-validation score for each combination of max depth and number of predictors from one to all (30) predictors. The Np selected for the final model was selected as described in Sect. 2.6.1.

RC1: l394: Why did you not assess the impact of a predictor that is

AC: Unfortunately, the last part of your comment is missing.

RC1: Section 3.1/l415: The final set of predictors, which mostly excludes anomaly-based indicators, seems to suggest that the model is tuned to predict fire occurrence climatology rather than typical fire weather situations. Is this correct?

AC: We do not fully agree that the model predict fire occurrence climatology rather than typical fire weather situations. First; even though most of the anomaly-based potential predictor are not included in the final set of predictors,

the shallow soil water anomaly stands out as a clear dominant predictor as compared to the other selected predictors. Secondly, the predictors have a high annual variability in monthly values. Notable differences from year to year for the same month can be seen in the fire danger probability maps produced by the model (in Fig. 8 and S6-S9), for example July 2017 (Fig. S9d) versus July 2018 (Fig. 8d).

RC1: L419-421: is the minor difference between the RF model and the FWI predictors really significant?

AC: We did not test for significance, but we agree that this difference is likely not significant. We will change to a more precise language (e.g. at which digit they differ) in the revised manuscript.

RC1: l442-447: To me it's not very surprising that simply including NDVI does not improve model skill as it's climatology closely follows that of soil moisture and meteorological variables. Did you also test the inclusion of NDVI anomalies?

AC: We did not include NDVI anomaly. NDVI can be viewed as a potential estimate of burnable biomass (in particular in the Nordic landscape that has a high variability in burnable biomass) and it is therefore preferred to include the absolute NDVI value instead of the NDVI anomaly. The close relationship between NDVI and hydrometeorological variables, such as temperature and snow cover, further supports the potential of developing models without NDVI. As acknowledged earlier, many more variables could have been included in our study (NDVI anomaly being one of them), however, some constrains in predictors included had to be made at the start of our study.

RC1: l445: High fire danger (luckily) most of the times does not lead to actual wildfire activity as an ignition source is required.

AC: We agree. The relation to line 445 is unclear to us, and we suspect the reviewer intended to refer to line 455.

RC1: Fig.9: it seems that the correlation patterns closely follow the border between Finland and Russia (ans to lesser degree Sweden). How can this be explained?

AC: This is an interesting observation, and we are left with speculations when trying to explain the pattern. Figure 2b also shows a Finland-Russia divide in the number of fires, where more fires are found in Russia. As a consequence, the data-driven model may have been better tuned to Russian conditions as compared to Finnish conditions, whereas the FWI performance is independent of the fire occurrence density. This may be one reason for the higher correlations between the two approaches in Russia compared to eastern Finland. We will briefly comment on this in the revise manuscript.

RC1: Can it be that the superior skill of FWI over the RF model is because FWI describes anomalous conditions whereas your model more relates to describing fire weather climatology and spatial patterns?

AC: In our understanding, FWI does not describe anomalous conditions, but rather estimates moisture (in surface, intermediate and deep organic layers) and spread conditions regardless of what is "normal". Since soil moisture anomaly is a dominant predictor in the data-driven model, the emphasis on anomalous conditions is a notable feature of the data-driven model rather than FWI.

RC1: l496: In this context, reference should me made to Forkel et al., 2012, who showed that antecedent moisture conditions are better predictors of fire occurrence in a Boreal environment than FWI and precipitation anomalies.

AC: Thank you for this good suggestion, we will include it in the revised manuscript.

RC1: l498: to what extent is soil moisture an indicator of litter fuel conditions? This is usually where fires start, not in the tree crowns.

AC: We expect a strong relation between shallow soil moisture and litter fuel conditions (favourable fuel conditions for low soil moisture). We agree with your statement and will consider adding a remark to the manuscript.

RC1: l531: This statement underestimates the role observations play in reanalysis.

AC: We agree and will make this clear in the revised manuscript.

RC1: l533-534: Could it be that wind is not directly but indirectly related, i.e. by the dominant weather patterns? High-pressure conditions, which are favourable to fire weather, are typically associated with low wind speeds. Vice-versa, westerlies bring high wind speeds and precipitation.

AC: Yes. We will consider mentioning this in the revised manuscript.

RC1: l539: are latitude and months of the year not already implicitly included in the other predictors?

AC: In some ways, yes, but they could have guided the model in cases such as the example presented commenting on the different effect of SPEI3 during the growing season as compared to the snow accumulation period (line 540-541).

RC1: l545: vegetation variables like fAPAR and LAI would be more obvious candidates than NDVI as these are simulated by DGVMs (which is an argument you brought up earlier).

AC: As commented on earlier, we excluded vegetation variables represented by DGVMs and limited the choice of variables to what is available form more common climate models (not including dynamic vegetation).

RC1: l546: Vegetation Optical Depth from microwave satellites has been proposed as fuel moisture indicators (e.g Forkel et al., 2017, 2019).

AC: We will check the suggested references (thanks for providing these) and adapt the text accordingly.

RC1: l582: Several studies have done this before as proved by the references below. Please rephrase.

AC: We will rephrase this sentence in the revised manuscript according to the analyses done in the references provided.

RC1: l611: I'd be careful with the word easily here as in other regions others drivers can be dominant, some of which may not even have been originally tested here. Besides, high-quality datsets such as the EOBS and observation-heavy reanalysis data may be unavailable or have reduced skill, respectively, in other regions and hence lead to a different model. Also fire management is different in many parts of the globe (e.g. rangeland burning management in Africa or deforestation).

AC: Yes, we agree and will remove the word easily from this sentence.