

***Interactive comment on* “Identifying required model structures to predict global fire activity from satellite and climate data” by Matthias Forkel et al.**

Matthias Forkel et al.

matthias.forkel@geo.tuwien.ac.at

Received and published: 13 July 2017

Referee comments are cited in *italics* and author’s responses are written in normal font.

Dear Authors, First thank you for an extremely well-written manuscript about your exhaustive study of factors contributing to controls on burned area for climate models. I found the conclusions section especially well-written; it summarizes well the implications of your results and your conclusions are well-defended by the analysis you present. This paper is entirely suitable for publication in GMD, I have only a few comments that perhaps can suggest where your account can be clarified.

Printer-friendly version

Discussion paper



Dear Referee 1, we thank you for a very positive review and we are happy to respond to your questions.

As for the science itself, I have only a few relatively minor questions about your methodology. The strength of your conclusions is only moderate because the input datasets you have used, which are with a few exceptions the best available, are not very good, and specifically lack skill at capturing the truly relevant properties of vegetation. This is not a flaw of your study, but merely the state of the science.

We completely agree to this point. We mentioned and discussed this issue already in chapters 2 and 5.1. We hypothesize that the use of datasets from newer satellite sensors with possibly higher quality retrievals might result in higher model performances. However, this needs to be demonstrated in a potential follow-up study. Moreover, we think that fire modelling could advance from more fire-relevant satellite products (e.g. time series of biomass and fuel loads instead of FAPAR or VOD; or fuel moisture instead of surface soil moisture) than the ones that we used here.

I did wonder about your choice of variables to represent climate/weather effects. First, you use a dataset based on statistical interpolation of weather station data. A reanalysis would be a much more appropriate choice: statistical interpolation of weather station data will have obvious consequences for your analysis: for instance, interpolation of (dense) coastal weather data into (data-sparse) inland areas will produce erroneous results in near-coastal interiors. Reanalysis data would not completely solve this problem, but would surely better capture the weather in fire-prone areas.

We agree that re-analysis data might better resolve fire weather conditions in regions where interpolation-based datasets rely on remote information. However we had two reasons for using interpolation-based climate data (CRU and GPCC datasets). Firstly, the CRU and GPCC datasets are also commonly used as forcing within global vegetation/fire models (e.g. (Schaphoff et al., 2013; Thonicke et al., 2010)) or re-analysis

datasets are corrected by such datasets like in the CRU-NCEP dataset that is used in several vegetation model-inter-comparison projects like TRENDY or FireMIP (Rabin et al., 2017). As our study aimed to provide suggestions for the development of global vegetation/fire models, we here relied on comparable forcing datasets. Secondly, we nevertheless tested beforehand the influence of using alternative climate datasets but found no generally strong effect on model performance. As differences in climate forcing datasets are usually more associated to differences in precipitation than in temperature, we tested the capability of several precipitation datasets to predict burned area within the random forest machine learning approach (Figure 1). Specifically, we compared the predictive capabilities of the number of wet days from CRU (WET, pink in Fig. 1) with precipitation from GPCC (used in paper, brown in Fig. 1), CRU (yellow in Fig. 1), and GPCP (violet in Fig. 1) (Huffman et al., 2009). In GPCP, satellite information are additionally used to rain gauge data. Thus GPCP should potentially better account for the spatial-temporal variability in precipitation. We found marginally better performances in predicting burned area when using GPCP than GPCC or CRU in boreal, temperate and tropical forests but slightly worse performance in steppes and the Mediterranean. All precipitation datasets and the number of wet days resulted in very similar performances at the global scale. In summary, at local to regional scales the prediction of burned area is sensitive to the chosen meteorological forcing dataset. However, we could not identify a precipitation datasets that would result in clearly better fire model performances at biome- to global scales.

Second, the variables you use are “mean temperature, mean diurnal temperature range, mean number of wet days, and the total precipitation of the actual month and the 12 months before a fire.” 1) What is the role of “diurnal temperature range” with regards to wildfire? It seems like a very loosely related quantity. 2) Would you not get better results by using temperature and rainfall anomalies, rather than absolute values? Or perhaps this would make no difference in your analysis.

Diurnal temperature range (DTR) has been long used as predictor for fire weather con-

[Printer-friendly version](#)[Discussion paper](#)

ditions. DTR is sensitive to stable weather conditions, i.e. DTR is usually high under high pressure systems with low cloud cover that allow high maximum temperature during day time but low temperatures at the morning because of the strong long-wave radiation loss during night-time (Lewis and Karoly, 2013). Such weather conditions are usually associated to low humidity and thus are supportive for fire activity. Therefore, DTR has been used as predictor variable in fire weather indices such as the Nesterov index (Nesterov, 1949; Venevsky et al., 2002) or is also used in fire models such as SPITFIRE (Thonicke et al., 2010). Also newer analyses based on satellite and climate data have shown that DTR shows a strong sensitivity to burned area (Bistinas et al., 2014). We also found in our preparatory analyses for this study that DTR is one of the most important (rank 5) predictor variables for global spatial-temporal dynamics of burned area (Figure 2).

We also initially thought that using anomalies of precipitation, temperature, soil moisture, or vegetation variables would be more relevant than using absolute values. To test the importance of several variables for the prediction of burned area, we computed several statistical properties of each variable and used random forest to quantify the importance of variables (Figure 2). Statistical properties were for example monthly anomalies relative to the mean seasonal cycle or averaged absolute values and anomalies over several pre-fire months (in total 132 variables were included in this analysis). Surprisingly, we found generally a higher importance of the absolute variables than of the anomalies. For example, short-term anomalies of precipitation (GPCC.P.anom) or soil moisture (CCI.SM.anom) had very low importance (below rank 53, not included in Fig. 2). We think that the lower importance of anomaly-based variables is caused by the fact that measurement noise is more prominently included in anomaly time series than in absolute-value time series. However, anomaly-based variables might have more predictive capabilities than absolute variables at regional scales which will strongly depend on the regional data quality. The most important anomaly-based variable was the average of the anomaly of wet days in the actual month and the 12 months before a fire (CRU.WET.anom.filter13, rank 3). However such a variable has only a limited phys-

[Printer-friendly version](#)[Discussion paper](#)

ically meaning for fire activity but it likely represent an indirect effect of precipitation on vegetation productivity and thus fuel production. For the development of SOFIA models, we finally selected the set of predictor variables based on their importance (Figure 2), their interpretability, and based on how closely they are related to fire activity (by avoiding variables that account for indirect effects).

Besides that question, I have only two other minor comments: Line 350: “(e.g. quantiles 0.01 to 0.02)” this is not clear to me; generally when I hear “quantiles” I think “bottom 20%” or “top 25%” or things like that.

The quantiles 0.01 and 0.02 are the percentiles 1% and 2%, respectively. The terms quantiles and percentiles are often not precisely used. Here, we mean (for example) that we computed the quantiles 0.01 (i.e. 1% of values are below this value) and 0.02 (i.e. 2% of values are below this value) of annual burned in each region. Consequently, some regional grid cells have annual burned areas that fall between the quantiles 0.01 and 0.02. From these grid cells, we took a random sample to be used in the training dataset. We repeated this procedure for all 0.01 quantile ranges between minimum (quantile 0) and maximum (quantile 1) to include in the training dataset grid cells that represent the entire regional statistical distribution of burned area.

Line 419: “explained reasonable” -> “explained reasonably” This was the only typo I encountered in the entire manuscript!

We will change this typo (and a few others) in the revised manuscript.

Figure captions

Figure 1: Effect of different climate and vegetation datasets on the regional index of agreement between observed (GFED dataset) and predicted (by different random forest models) burned area. Shown is the index of agreement for the evaluation data subset.

Printer-friendly version

Discussion paper



Figure 2: Importance of several variables to predict monthly burned area using random forest. Importance is expressed as the percentage increment in mean squared error if a certain variable is not included in random forest. Thus, the most important variables cause the largest increment in MSE. Variables that include “orig” or “anom” indicates original absolute values and anomalies (relative to the mean seasonal cycle), respectively. “filterX” indicates mean values over the X months before the actual month for which burned area should be predicted. In total 132 variables were included in this analysis but variables below ran 53 are not shown in this figure).

References

Bistinas, I., Harrison, S. P., Prentice, I. C. and Pereira, J. M. C.: Causal relationships versus emergent patterns in the global controls of fire frequency, *Biogeosciences*, 11(18), 5087–5101, doi:10.5194/bg-11-5087-2014, 2014.

Huffman, G. J., Adler, R. F., Bolvin, D. T. and Gu, G.: Improving the global precipitation record: GPCP Version 2.1, *Geophys. Res. Lett.*, 36(17), L17808, doi:10.1029/2009GL040000, 2009.

Lewis, S. C. and Karoly, D. J.: Evaluation of Historical Diurnal Temperature Range Trends in CMIP5 Models, *J. Clim.*, 26(22), 9077–9089, doi:10.1175/JCLI-D-13-00032.1, 2013.

Nesterov, V. G.: *Gorimost' lesa i metody eio opredelenia*, Goslesbumaga, Moscow., 1949.

Rabin, S. S., Melton, J. R., Lasslop, G., Bachelet, D., Forrest, M., Hantson, S., Kaplan, J. O., Li, F., Mangeon, S., Ward, D. S., Yue, C., Arora, V. K., Hickler, T., Kloster, S., Knorr, W., Nieradzik, L., Spessa, A., Folberth, G. A., Sheehan, T., Voulgarakis, A., Kelley, D. I., Prentice, I. C., Sitch, S., Harrison, S. and Arneeth, A.: The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions, *Geosci Model Dev*, 10(3), 1175–1197, doi:10.5194/gmd-

[Printer-friendly version](#)[Discussion paper](#)

10-1175-2017, 2017.

Schaphoff, S., Heyder, U., Ostberg, S., Gerten, D., Heinke, J. and Lucht, W.: Contribution of permafrost soils to the global carbon budget, *Environ. Res. Lett.*, 8(1), 014026, doi:10.1088/1748-9326/8/1/014026, 2013.

Thonicke, K., Spessa, A., Prentice, I. C., Harrison, S. P., Dong, L. and Carmona-Moreno, C.: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model, *Biogeosciences*, 7(6), 1991–2011, 2010.

Venevsky, S., Thonicke, K., Sitch, S. and Cramer, W.: Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study, *Glob. Change Biol.*, 8(10), 984–998, 2002.

Interactive comment on *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2016-301>, 2016.

GMDD

Interactive
comment

Printer-friendly version

Discussion paper



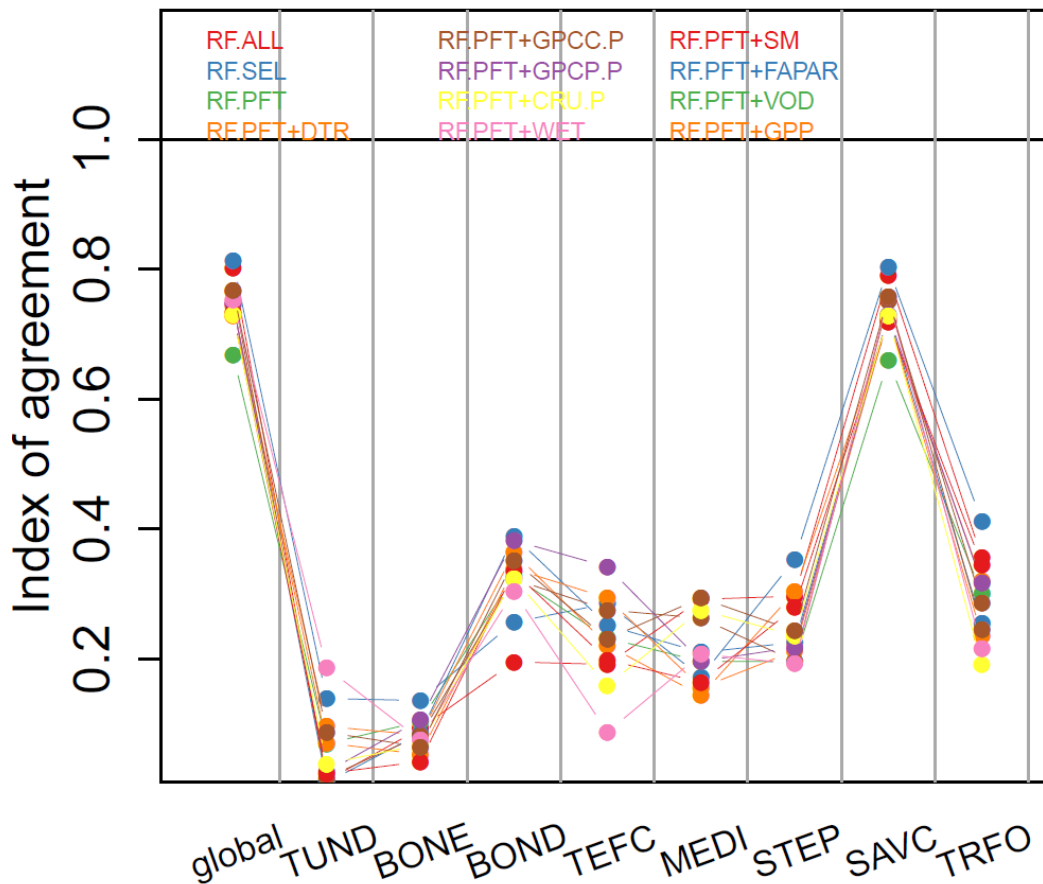


Fig. 1. Effect of different climate and vegetation datasets on the regional index of agreement ...

Printer-friendly version

Discussion paper



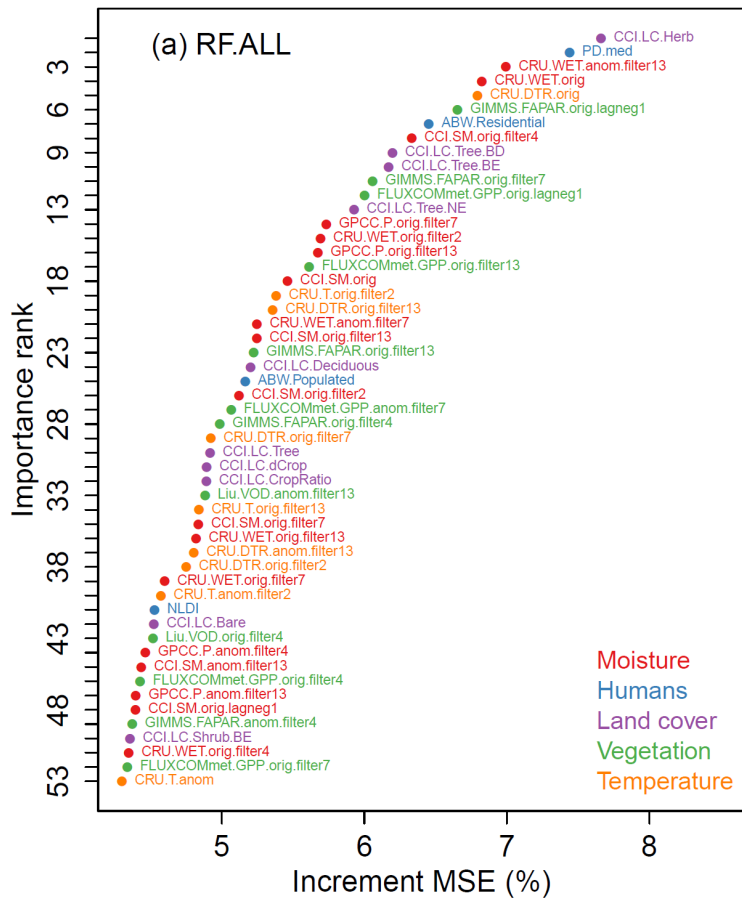


Fig. 2. Importance of several variables to predict monthly burned area using random forest ...

Printer-friendly version

Discussion paper

