

Modelling fire frequency and area burned across
phytoclimatic regions in Spain using reanalysis data
and the Canadian Fire Weather Index System

Response to Anonymous Referee #1
RC C1596

J. Bedia*¹, S. Herrera¹ and J.M.Gutiérrez²

¹Meteorology Group. Dept. Matemática Aplicada y Ciencias de la Computación.
University of Cantabria (Spain)

²Meteorology Group. Institute of Physics of Cantabria. CSIC - University of
Cantabria (Spain)

November 12, 2013

Dear Referee,

First of all, we would like to thank you for the time devoted to the revision of our manuscript and the positive feedback provided, leading us to re-think some parts of our study and resulting in a new version with significant improvements. We are also grateful to the referee for pointing to some grammatical mistakes and writing inconsistencies. We have undertaken a thorough revision of the manuscript in order to correct them.

Following your comment/suggestions, and also considering the feedback received from the other two referees, we have undertaken substantial modifications to the original version. These are the most important new points addressed:

- The focus of our article is now more explicitly done on the daily predictability of wildfire occurrence from an operational point of view, although we also make some complementary analyses of burned area. This has led to a new title of the article: *Assessing the predictability of fire occurrence and area burned across phytoclimatic regions in Spain*.
- The three referees have coincided in pointing to the importance of anthropogenic factors in wildfire occurrence, not considered in the previous version. In the revised manuscript, we have included socio-economic and land use / land cover covariates in our analyses, in order to ascertain their contribution to the improvement of model performance.
- This has led to a new version in which the mechanisms behind the performance of the models at each phytoclimatic zone are more deeply analysed and discussed.

*bediaj@unican.es

- Finally, extended information has been included in the Supplementary Material, including Fig. 3 which was too complex and whose caption was too lengthy to be included in the main body of the article, but that in our opinion provides an extremely helpful visual overview of the modelling approaches tested.

In the following, we perform a point-by-point answer to the comments and questions posed by the referee. Note that the referee comments have been literally reproduced and indicated in boldface throughout the text.

General comments

The understanding of why models do not perform well, is limited and poorly described. The work would greatly benefit from exploring this key point more thoroughly, i.e. understanding why, where and when models do not perform well is, in my opinion, more important than just evaluating how they perform. This could shed some light on future improvements and important components that are missing from your modelling approach We agree with the referee on this point. With the inclusion of new socioeconomic and LULC covariates in the models we have made more emphasis on the implications of other factors apart from climatic ones in fire occurrence (and burned area), and included specific comments on possible ways of improving the models taking into account these factors.

I disagree that you demonstrate the usefulness of ERA-Interim reanalysis data. You demonstrate that using that data, along with the modeling approach, can provide interesting results. But to demonstrate the usefulness you would need to do some extra analysis, such as use multiple datasets and assess their impact on the capability of predicting fire frequency/burnt area. We agree with the referee, and we also think that this statement is misleading. What we meant here is that the output from ERA-Interim reanalysis can be successfully applied to fire modelling applications using real observed fire records, which does not mean that is the most adequate product to this aim. In the new version of the manuscript we have made an emphasis on the fact that ERA-Interim is used because it is representative of numerical model daily outputs, for instance as produced by the forecast models routinely used in operational weather prediction, while providing time series of sufficient length and homogeneity to enable modelling the historical fire-climate relationships. The added value of our study is to show how daily outputs from climate models can be successfully applied for the prediction of fire occurrence building upon the FWI System, so potential fire danger forecasts, as issued for instance by the EFFIS system in the Euro-Mediterranean countries of the EU (Camia et al., 2006; Camia and Amatulli, 2009), may be also translated to fire occurrence predictions over particular regions with a good level of confidence.

I have some doubts in the use of the word “skill” to describe model “accuracy” or “performance”. We agree with the referee in his/her doubts about its use in this context, provided that the term skill is used always when

comparing the results against a given benchmark. We have omitted this term from the text of the revised manuscript, and replaced it by *model performance*.

Specific comments

The specific comments given have been taken into account in the new revised version. We make a more detailed explanation on the final one in the following paragraph.

Please provide a more detail justification for the statement “bearing some sort of memory on the antecedent conditions” present in last paragraph of the Conclusions. This statement is done on the basis of the findings of several previous authors. For instance, Carvalho et al. (2011) indicate that DC estimated in summer still reflects spring time atmospheric conditions, justified by the slow reacting character of this indicators, that models variations on moisture of deep organic soil layers. Other studies conducted in Portugal show that the DC gives a good indication of wildfire behaviour and propagation and also of the relative hazardousness of a fire season due to its long-term response to daily weather variations (Viegas et al., 2004). This is explained in the text of the revised manuscript.

Questions

1. **In the abstract and conclusions you mention that fire frequency predictions are more suitable for past fire history reconstruction and pose several advantages over burned area. What are those advantages? This should be in the manuscript.**

Inter-annual correlations between observed and modelled fire occurrence are much higher than those obtained with burned areas, and therefore fire occurrence is better modelled from climate data alone than burned area. In the framework of future climate impact assessment, the projections of future fire danger scenarios are most often based on the simulation output of GCMs (either downscaled or not) run in transient mode. This implies that model predictions do not have a day-to-day correspondence with real climate, and their value lies in the ability of the models to represent the mean state of climate, its variability, trends . . . throughout relatively large climatological periods. This suggests that the estimation of inter-annual fire frequencies from simulated model outputs for sufficiently long time slices (typically 30-year periods) is able to provide a more robust estimation of future impacts than area burned, the latter often yielding too inflated, unrealistic future estimations, as shown in several previous studies (Amatulli et al., 2013; Balshi et al., 2009; Carvalho et al., 2010; Flannigan et al., 2005). This issue has been explained in more detail in the text.

2. **The definition of phytoclimatic regions was based on Spanish Meteo Agency data and not in ERA-INTERIM. Did you make any comparison between both meteorological datasets? If they**

have significant discrepancies what do you think will be the impact on your results and major findings?

Yes, this issue is specifically addressed in a previous comparative study by the authors (Bedia et al., 2012). According to our results, data derived from reanalysis in general, and from ERA-Interim in particular, tend to underestimate the magnitude of the index (i.e., yields a negative bias) with regard to the observations. However, this effect is not by itself deleterious for the models. Nevertheless, other problems arise from the fact that ERA-Interim data have a relatively coarse horizontal resolution, and they present some deficiencies in the representation of the tail of the distribution (we analysed the 90th percentile in particular), and therefore it is not able to capture all extreme events. This effect was shown to be of different magnitude depending on the location. Reanalysis data are likely to provide a less accurate representation of the observed climate in coastal locations and zones with complex topography, like northern Spain.

However, given the coarse scale at which the phytoclimatic regions are delineated in the map by Allué (1:1.000.000), we believe that the mismatch between reanalysis data and phytoclimatic regions has a negligible effect on the models.

3. Why do you think that in most phytoclimatic regions, the time series of fire frequency and burnt area follow so closely? Some works have shown that, for instance in agricultural regions, the fire frequency can be high but with very low burnt area.

We guess that this comment refers to data displayed in Fig. 2 of the manuscript. In this case, please note that we are not displaying time series, but time-averaged data. The close relationship depicted is partly due to the aggregation of the data in time (period 1990-2008), and space (gridboxes from each phytoclimatic zone have been averaged). Variability is larger when considering disaggregated data, for instance single pixels or monthly time series. In Table 1 we show the low cross-correlation values between burned areas and number of fires when considering the monthly time series for the period 1990-2008 ($N = 228$ months), as opposite to the close correspondence found when considering the monthly-averaged data for the whole period ($N = 12$ months, as depicted in Fig. 2).

	2-3	4	5	6	7-8	9	10-11-12	13-14-15
(i) monthly	0.12	0.11	0.21	0.10	0.21	0.07	0.19	0.10
(ii) aggregated	0.59	0.66	0.66	0.41	0.78	0.46	0.45	0.55

Table 1: R-squared values of the monthly time series of burned area *vs.* number of fires at each phytoclimatic region (i), and considering the data aggregated for the whole 19-year period (ii) (as in Fig 2 of the manuscript.)

As the referee points out, the relationship is not that close actually when considering the time series instead of the aggregated means. This is because usually, many small fires occur throughout the year with little influence on the total burned area records, whereas a few large fires contribute most to this value.

4. **Did you try to make any temporal aggregation of fire frequency data? Weekly, biweekly or monthly, for instance. Predicting daily fire frequency is challenging, do you think that, both in terms of research and fire management applications, a larger time step would be as useful (or even more) and that results would be better? If not, please justify.**

We fully agree with the referee in his/her appreciation on the challenging nature of modelling daily data, and also in the likely improvement of the results with the temporal aggregation of data. The reduction in variability derived from temporal aggregation would surely help to improve model fits (see e.g. Bedia et al., 2013, using annual time aggregation for FWI validation purposes). However, the main value of using daily data lies in the greater advantages that it provides from an operational point of view, as we have indicated previously. This aspect of our study is emphasized in the new version of the manuscript.

5. **In section 2.6.1, for grid-box model training you sampled fire absences based on information from the entire phytoclimatic zone. Why? Moreover, don't you think that this procedure is bringing the grid-box and areal approaches closer together? As I mentioned previously, I think that difference between both is one of the most important results of this work.**

The gridbox approach does not mean that models are independently developed for each gridbox, otherwise it wouldn't be necessary to use the phytoclimatic regions in order to stratify the territory, as fuel-climate relationships would be particular for each gridbox. We assume that within each phytoclimatic region, the sampling space is homogeneous in this sense.

Because the phytoclimatic zone is homogeneous, the idea of the gridbox experiment is to consider individual gridboxes to determine fire absence/occurrence, but always within the region of interest. In contrast, in the case of the areal models, the fire occurrence/absence is considered globally at the phytoclimatic zone (i.e. considering whether fire occurs at *any* of the gridboxes within the area or not).

Please note that Fig. 3 has now been moved to the supplementary material. In this figure, gridbox homogeneity is illustrated. We have also included here further comments to better clarify the methodology for data analysis.

6. **The way you performed the fire occurrence model training involved bootstrap techniques (page 6,sec 2.6.1)?**

The procedure for sampling no-fire days is just a simple random resampling of the non-fire series with the same sample size as the fire occurrences (in order to have balanced datasets). This procedure is repeated 100 times in order to estimate the sensitivity of the results to the resampling procedure. Note that no replacement is followed in this process and, therefore, the method cannot be considered a bootstrap approach.

7. **Can you provide a more detailed information of the main causes behind the bimodality in region 10-15? How well does your**

model fit to the annual cycle in these regions? (see the residuals analysis suggestion above and I suggest reading Le Page et al. 2010).

With the inclusion of the new socio-economic/LULC variables we can now attribute much of the bimodal behaviour to anthropogenic factors, provided the importance attained by socioeconomic and LULC covariates in the models for this phytoclimatic zone (actually the greatest of all zones).

8. **What about region 9, don't you think this exhibits some sort of bimodality also?**

Yes, it also exhibits some bimodality. This has been indicated in the text.

9. **Monthly-varying threshold – So is this monthly or seasonal? In page 9 sec 3.2 you mention seasonal. From what I understand this is done at a monthly basis without “looking” to the other months, i.e. you don't do a “moving temporal window” approach, so if I understood correctly, the term “seasonal” cannot be used or should be used in another way.**

Yes, the referee is right in his/her interpretation. We mentioned seasonal by error, we meant monthly. The text of the revised version has been changed accordingly.

10. **I think that it was expected that areal approach would provide better results than the grid-box approach. For management purposes working with broad and large phytoclimatic regions can be a limiting factor. Do you have any suggestion on how to “decompose” or “break” this broad regions into smaller contiguous areas that would enhance the usefulness of your work for fire management purposes?**

One possibility is using political/administrative units (e.g. Spanish Provinces, NUTS2), or even smaller administrative units (NUTS3) provided a climatic information source of and adequate resolution (ERA-Interim data is too coarse for such an approach). This approach would be not as adequate in order to obtain homogeneous areas in terms of fuel/climate relationships, but on the other hand it would allow to ensure a better homogeneity in terms of fire alert and suppression means, whose action is strongly determined by administrative boundaries.

11. **In section 3.2, how did the sample size (N) vary with the burnt area threshold? Could the results shown in the first paragraph of this section be significantly affected by the sample size and condition your findings? I suggest you put the N in tables 3 and 4.**

Surely the results are affected by the sample size. In fact, it can be seen in Table 3 how the variability in the ROC areas consistently increase with larger area thresholds. Following the suggestion of the referee, we provide the N in the corresponding table in the revised manuscript, and make a comment with this regard in the results.

12. **What do you think are the main reasons for the correlation decrease from smaller to larger area thresholds? It seems that this result is contradictory with figure 3, please confirm.**

The ability of the models to detect large fires decreases since those events are largely controlled by other variables not taken into account in the models, related to the physical conditions needed for triggering and controlling the spread of this kind of fires, like the velocity of reaction of fire suppression means, or the spatial continuity of fuels across landscape (100 ha are still below the spatial resolution of our analysis). However, the favourable conditions with high fire danger values are likely to give raise to numerous small fires that only under particular circumstances can become large fires. As a result, the predictability of small fires is easier because it depends on less factors and can be more accurately modelled with the climate/socio-economic/LULC covariates used in the analysis.

13. **In Table 3, the bimodal phytoclimatic areas do not have significant models' performance. What do you think are the main reasons for this, how does this limit your work's findings and how can this be overcome in the future? Is the earlier fire peak captured by your model predictions? (again the residuals suggestion, above).**

The reasons behind this fact are the strong anthropogenic influence of fire regimes in this area of Spain (and specially in the north-western corner), a phenomenon well described by previous authors (see e.g. Martínez et al., 2009). We make special reference to this issue in the new revised version of the manuscript.

14. **In conclusions, if the good model performance in terms of RSA does not directly translate into a good reproducibility of fire frequencies then what does this say about the capability of RSA to provide reliable model performance indications? What do you suggest as alternatives? (I suggest you take a look at the Model Efficiency Index developed by Nash and Sutcliffe, 1970)**

RSA is an adequate metric for model performance assessment. There are two different issues jointly affecting the reproducibility of inter-annual fire frequencies: On the one hand, this is affected by model performance (RSA). Obviously, those models with low RSAs have less ability to adequately reproduce the inter-annual fire frequencies. This can be seen in the examples of zones attaining low RSA values (e.g. 13-14-15), specially with the area threshold of 0.1 ha (Table 3). However, it is also important to take into account the inter-annual variability of fire occurrence, which in turn is related with the prevalence of the phenomenon. For instance, most models attain relatively high RSA values for the 100 ha area threshold, but the inter-annual variability of these large fires is lower, because there are few fires larger than 100 ha every year (perhaps zero some years at some regions), and definitely much less than small fires above 0.1 ha. As a result, the estimated inter-annual frequencies are more drastically affected by false alarms / false positives in the case of large fires than in the case of small ones, even though in both cases RSA values are good. In Fig. 4 it can be seen how the best inter-annual correlations correspond to those

zones where the inter-annual variability is higher. This variability tends to reduce as the fire area threshold increases, leading to worse results. We agree with the referee on the relevance of this fact and specifically address it in the discussion of the new revised version of the manuscript.

References

- Allué, J.: Atlas Fitoclimático de España. Taxonomías, Tech. rep., Instituto Nacional de Investigaciones Agrarias. Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain, 221 pp., 1990.
- Amatulli, G., Camia, A., and San-Miguel-Ayanz, J.: Estimating future burned areas under changing climate in the EU-Mediterranean countries, *Science of the Total Environment*, 450-451, 209–222, 2013.
- Balshi, M., McGuire, A., Duffy, P., Flannigan, M., Walsh, J., and Melillo, J.: Assessing the response of area burned to changing climate in western boreal North America using a Multivariate Adaptive Regression Splines (MARS) approach, *Global Change Biology*, 15, 578–600, 2009.
- Bedia, J., Herrera, S., Gutiérrez, J., Zavala, G., Urbieto, I., and Moreno, J.: Sensitivity of Fire Weather Index to different reanalysis products in the Iberian Peninsula, *Nat. Hazards Earth Syst. Sci.*, 12, 699–708, 2012.
- Bedia, J., Herrera, S., Camia, A., Moreno, J. M., and Gutierrez, J. M.: Forest Fire Danger Projections in the Mediterranean using ENSEMBLES Regional Climate Change Scenarios, *Climatic Change*, In press, doi:10.1007/s10584-013-1005-z, 2013.
- Camia, A. and Amatulli, G.: Weather Factors and Fire Danger in the Mediterranean, in: *Earth Observation of Wildland Fires in Mediterranean Ecosystems*, edited by Chuvieco, E., pp. 71–82, Springer Berlin Heidelberg, Berlin, Heidelberg, URL http://www.springerlink.com/index/10.1007/978-3-642-01754-4_6, 2009.
- Camia, A., Barbosa, P., Amatulli, G., , and San-Miguel-Ayanz, J.: Fire danger rating in the European Forest Fire Information System (EFFIS): current developments, in: *Vth International Conference on Forest Fire Research*, edited by Viegas, D., Figueira da Foz, Portugal, 2006.
- Carvalho, A., Flannigan, M. D., Logan, K. A., Gowman, L., Miranda, A. I., and Borrego, C.: The impact of spatial resolution on area burned and fire occurrence projections in Portugal under climate change, *Clim. Change*, 98, 177–197, 2010.
- Carvalho, A., Carvalho, A., Martins, H., Marques, C., Rocha, A., Borrego, C., Viegas, D., and Miranda, A.: Fire weather risk assessment under climate change using a dynamical downscaling approach, *Environ. Modell. Software*, 26, 1123–1133, 2011.
- Flannigan, M., Logan, K., Amiro, B., Skinner, W., and Stocks, B.: Future area burned in Canada, *Clim Change*, 72, 1–16, doi:10.1007/s10584-005-5935-y, 2005.

- Martínez, J., Vega-García, C., and Chuvieco, E.: Human-caused wildfire risk rating for prevention planning in Spain, *Journal of Environmental Management*, 90, 1241–1252, doi:10.1016/j.jenvman.2008.07.005, URL <http://linkinghub.elsevier.com/retrieve/pii/S0301479708001758>, 2009.
- Viegas, D. X., Reis, R. M., Cruz, M. G., and Viegas, M. T.: Calibração do sistema canadiano de perigo de incêndio para aplicação em Portugal, *Silva Lusitana*, 12, 77–93, URL http://www.scielo.oces.mctes.pt/scielo.php?pid=S0870-63522004000200007&script=sci_arttext, 2004.