Biogeosciences Discuss., 11, C478–C480, 2014
www.biogeosciences-discuss.net/11/C478/2014/

Interactive comment on "Causal relationships vs. emergent patterns in the global controls of fire frequency" *by* I. Bistinas et al.

V. Lehsten

veiko.lehsten@nateko.lu.se

Received and published: 21 March 2014

The authors claim the novelty of their approach, but the same has been done (actually with even very related data) for the African Continent so the results should be similar (Lehsten et al 2009, this Journal).

However I still think it is valid doing it for the world.

I have some other major issues.

1) The Nesterov index sums up a drought unit (based on temperature and dewpoint temperature) and goes to zero once the precipitation is above 3 mm, at least this is the definition that is given in the stated reference. The dataset used by the authors does not allow to calculate this original index. So the authors call it the simplyfied

Nesterov which is fine. The use of diurnal temperature range as an estimate for the difference between dewpoint temperature and average temperature assumes that the air is saturated at night. This has been demonstrated to work in mountainous areas but is not the case in most fire prone regions at the time that the fire occurs. Just take any daily climate data set that contains these parameters and check yourself by using this proxy versus the real value. I also thought first it is a valud idea but it does not work at all, you get differences in orders of magnitude. Using monthly data to calculate an index which goes to zero in case of rain event of 3 mm per day is somewhat unsuitable I think how do you assess the temporal distribution of the rain? The monthly precipitation and the number of dry days does not allow this, except if you assume that all rain days in the month have the same precipitation.

2) Spatial autocorrelation and colinarity of input variables. The statement that the correlation between input data is not affecting the significance of the results is simply wrong. It is however true that the estimated parameters are only minor affected. If you estimate a GLM with highly correlated input data, the R-square will be strongly affected and you will overestimate the significance of the model.

3) The authors do not state how they found their final model. I assume that they used a forward step procedure, since they used the R-package. In this case it has been show that the order in which the variables enter the analysis strongly influences the result especially when the input variables are highly correlated. /Burnham & Anderson, Behavioral Ecology and Sociobiology 65:23–35; Whittingham et al. 2006. Journal of Animal Ecology 75:1182–1189). The R value as such is actually not usefull to state whether a model is significant. That has to be done using a p-value.

4) One of my main criticism is the final result which the authors claim to have found: "most notably, the widely assumed dependence of fire frequency on ignition rates – are evidently incorrect." With the data that they used, it is not surprising that they first found a negative relationship. The authors used the OTD lightning seasonal data. This data is highly interpolated in time. If you look at a single pixel you often see a sinus curve

with the maximum at the summer or rain season. This type of data is not representing the lightning that actually causes the fire. These flashes are so called dry lightning. They are only comrising a minor proportion of the total number of flashes, they will not be apparent in any seasonal product which is averaged over many years. While the independence of the fire on lightning in the tropics might be true since there most fires are caused by humans it is certainly not the case in the boreal region. Though even here I would argue that the statement as "the widely assumed dependence of fire frequency on ignition rates – are evidently incorrect." Since humans cause the ignitions there. If the authors would have found that fires in the world (and hence also in boreal regions) are independent of ignitions that would be very interesting, but neither the used data, nor the methods used allows such a statement.

However I think this is a valuable contribution and would strongly encorage a revision in which the authors:

Deal with the co-linearity of the input variables

Deal with the spatial autocorrelation (eg by using an ordination)

Describe how the variable selection was done

Stop using the temperature range (especially if it is a monthly variable) as a proxy for dewpoint.

Use a proper evaluation criteria for the statitistical significance of the model

Relate the quality of the input data to the possible relevance of the results. (the statement about the ignition independence can not be backed up by the data).

Interactive comment on Biogeosciences Discuss., 11, 3865, 2014.