Biogeosciences Discuss., 10, C7140–C7146, 2013
www.biogeosciences-discuss.net/10/C7140/2013/

Biogeosciences
Discussions

Interactive comment on "Fractal properties of forest fires in Amazonia as a basis for modelling pan-tropical burned area" *by* I. N. Fletcher et al.

I. N. Fletcher et al.

inf201@exeter.ac.uk

Received and published: 4 December 2013

The authors would like to thank the reviewer for the time and effort they put into reviewing our manuscript. We appreciate all of the comments, and are pleased that you think the material has the potential to be important in the literature. Based on your suggestions, and those of the other reviewers, we have made considerable changes to the manuscript, and hope that they satisfactorily address any concerns you have about the validity of the methods or reliability of the results. Below, we first of all outline the main changes we have made, and then address each of your suggestions/comments individually. We have included in each case the exact comment that we are addressing.

Main changes to the methods

- A new distribution is used, which allows for a tail at both ends of the distribution: $n_{X \geq A} = aA^{-b}\exp(-\frac{1}{A} - \frac{A}{\theta})$ $= n_f A^{-b}\exp(1 - \frac{1}{A} + \frac{1-A}{\theta})$
- Only one parameterisation is presented, with a comprehensive explanation of its physical interpretation.

Response to comments

"A great deal of the paper is taken up by considerations of fitting the distributions to the processed data, and I am not convinced that all that space is necessary to describe the process."

This comment has made it clear to us that we did not adequately explain the distinction between what we have called "fitting" and "estimating" the parameters. The first step, fitting, involves estimating the distribution parameters for each grid cell, using least-squares regression on the log cumulative frequencies, as provided by the data. This is the most accurate indication we have of the "true" values of the parameters. The second step (section 2.3), which consists of the simple approximations for each parameter as derived from the distribution functions, is the "estimation" stage - this is the development of methods with which to estimate the parameters, when the full set of data is *not* available, i.e. when the only available information is the total number of fires that have occurred, with no indication of their respective sizes. Since the purpose of the study is to show whether it is possible to predict the fire size distribution and hence burnt area of a region using only fire counts as an input, we feel that it is important to describe this parameter estimation process thoroughly.

BGD

10, C7140–C7146, 2013

Interactive Comment

“The comparison between the Pareto and tapered Pareto is obvious and probably spurious. One should not be surprised if the tapered Pareto, with an additional parameter, gives a better fit."

Although we have now replaced these two distributions with a single, alternative distribution, we feel it is worth pointing out that the purpose of including both distributions in the initial manuscript was to determine whether, ultimately, one gave better burnt area estimates than the other. We agree that it is unsurprising that the tapered Pareto works better when fitted to the data, due to the additional parameter, but the parameters of this distribution are harder to estimate that those of the standard Pareto, hence could have resulted in BA estimates further from the observed values.

“Specialized methods of fitting are available from the literature and these would generally be preferred." Your concerns about the choice of fitting methods for the initial stage of the analysis (“fitting" rather than “estimating" the parameters) are completely valid, and we have put a lot of time into considering the various options. Although we have now taken out the comparison between the Pareto and tapered Pareto distributions, we have used the same method to test whether the data fit the new distribution, namely least-squares regression.

We have included more information in the manuscript about our reasons for choosing least-squares regression in section 2.2 of the revised manuscript. This is pasted below:

> We check that this distribution fits the data by estimating parameters b and θ using least-squares regression on Eq. (4), and comparing the resulting fitted cumulative frequencies to the data points. This is not an optimal fitting method, since a condition of least-squares optimization is that the errors be independent of one another. This is obviously not the case when cumulative frequencies are used. However, alternative methods such as maximum likelihood regression or the method of moments are not suitable in this case. These methods are commonly used for similar problems in

BGD

10, C7140–C7146, 2013

the literature, using binned data (e.g. Pueyo, (2007), Pueyo et al., (2010), Moreno et al. (2011)). Binning the data results in the loss of information about extreme fire sizes, hence our reluctance to use this technique in this instance. If the data is used unbinned, we encounter the problem of trying to fit a continuous, monotonically decreasing probability density function to a set of data in which many sizes can take the same frequency, and some intermediate fire sizes do not occur at all (this pattern can be seen in the top-right plot of Fig. (2)). Ultimately, this results in a clear underestimation of fire frequencies. Least-squares regression, although not a perfect option, provides decent approximations of the parameters.

To reinforce this, we have included a figure below, which shows the results of fitting the distribution to the data using maximum-likelihood regression. It is clear to see that this is considerably worse than using least-squares regression, if compared to Fig. 2.

"The R-squared value has little meaning in this context." We agree that the R^2-value is meaningless when fitted to a cumulative distribution, and hence have omitted its inclusion in the revised manuscript.

"It is very difficult to tell by eye whether the tapered Pareto is really better in Figure 2, or whether the way the data is plotted makes the data look better" As explained above, we have not been able to use a different method to fit the distribution to the data, but for the new distribution we have replaced Figure 2 with

(a) a log-log plot of the observed and fitted cumulative frequencies against fire sizes,

(b) a log-log plot of the observed and fitted frequencies against fire sizes,

(c) a log-log plot of fitted against observed cumulative frequencies with a 1:1 line, and

(d) a log-log plot of fitted against observed frequencies with a 1:1 line.

We feel that this makes it easier to assess the fit of the distribution to the data.

“If a figure such as Figure 2 is to be included, then all the parameters ought to be written on the plot" We have added the estimated parameter values and their uncertainties to the amended Fig. 2, as suggested.

“The justification for the use of a specific fitting procedure finally used is by way of results from Figure 5. I am not sure if these images tell us much, except that the specific fitting method chosen gives something reasonably close to the observed burnt areas.” Your comments made it clear that we were not explicit enough about the reason for our choice of a final fitting method. The decision was made not only based on Figure 5, but also from the results shown in Figure 4 and Table 1 (in the original manuscript), i.e. we looked not only at the spatial distribution of burnt area, but also the mean errors of the estimates from the observed burnt areas, and the total burnt area estimate for the whole study region. In the revised manuscript, we have not tested multiple fitting options for the sake of clarity. However, we have still tested the same aspects of the burnt area estimates (Figures 5 and 6, Table 1).

“There is no need to have to justify the model as an example of SOC for it to be acceptable, provided it is a good model." We agree that the discussion about whether the data show self-organized criticality is unnecessary in the context of our study, and have removed this from the manuscript entirely.

“Unfortunately, I do not think the model is particularly well-justified, and on balance it is probably not a good model." We hope that the revised manuscript and our new results can convince the reviewer otherwise. We are always happy to answer any other queries.

“The use of a 500m cell size is possibly over-optimistic as a basis for detecting fire events, although it is an arguable issue" We agree with the reviewer that this is not the best resolution for detecting fire scars. However, the use of moderate resolution data has an advantage over higher resolutions, specially for the tropics because of their

BGD

10, C7140–C7146, 2013

high temporal resolution. This minimizes cloud obstruction problems. Moderate resolution is likely to underestimate small fire scars due to the spatial resolution of the data. High resolution data can underestimate burn scar area because of its low temporal resolution, and the higher probability of cloud cover. Moreover, methods for producing regional to global scale burned area estimates generated using change detection algorithms applied to time series of surface reflectance from moderate resolution spectrometers, typically at 500 m to 1 km resolution, are well consolidated (Randerson et al. 2012, and references in this publication).

Another important point is that despite our final product being at a 500 m grid-cell size, the mapping is done based on 250m pixels from spectral bands 1 (620-670 nm) and 2 (841-876 nm) where a large part of the information about photosynthetic capacity and structure of the vegetation is concentrated.

Interactive comment on Biogeosciences Discuss., 10, 14141, 2013.

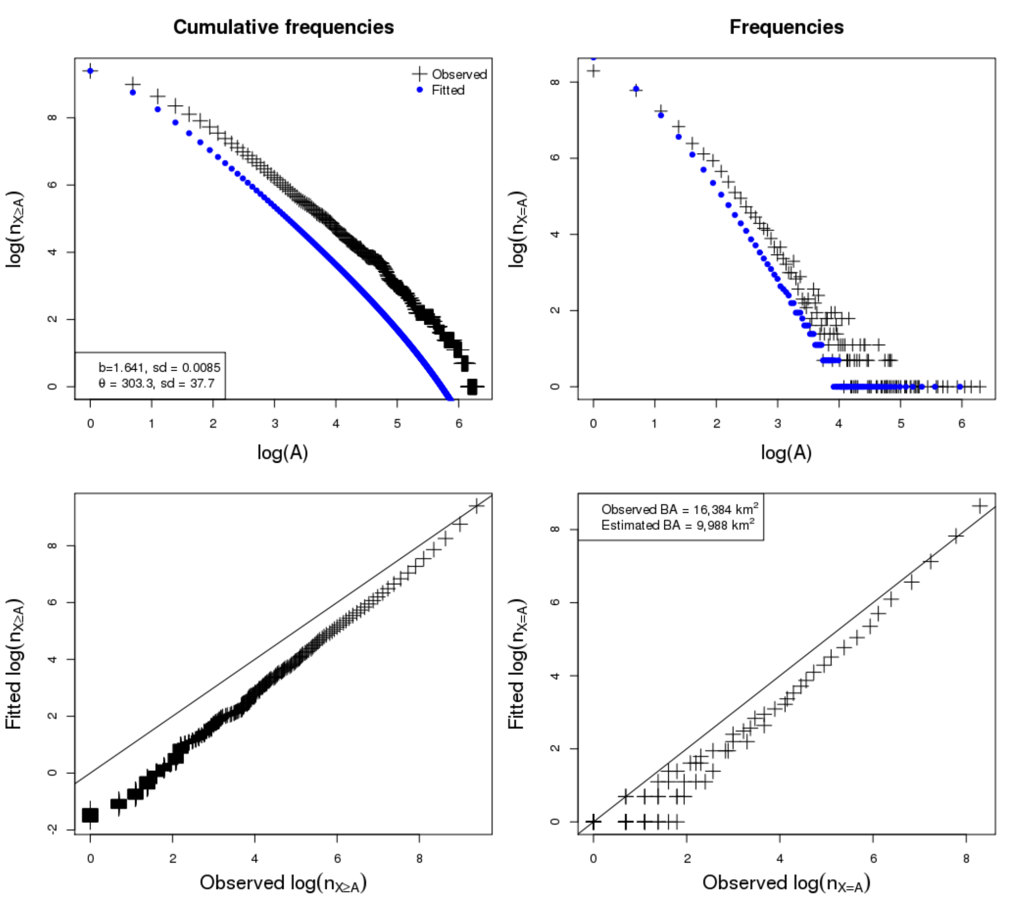

Fig. 1. Results of fitting the distribution to the data using maximum-likelihood estimation