Biogeosciences Discuss., 10, C5990–C6001, 2013
www.biogeosciences-discuss.net/10/C5990/2013/

Biogeosciences Discussions
Open Access

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.

S. Pueyo (Referee)

spueyo@ic3.cat

Received and published: 21 October 2013

1 Introduction

The manuscript *Fractal properties of forest fires in Amazonia as a basis for modelling pan-tropical area* (Fletcher 2013) addresses some concerns that I share completely. I agree with the authors in that Dynamic Global Vegetation Models (DGVMs) ignoring fire neglect a major biogeochemical and ecological process, in that DGVMs that model fire spread based on Rothermel equations are unnecessarily complicated and too difficult to parameterize, and in that an approach based on scale invariance is a promising alternative. I also welcome their focusing on Amazonia and other tropical forests, which I see as key regions of our planet for a number of reasons.

However, I have many technical disagreements with the ways in which the authors implemented this first step in their fractal approach. Considering these technical problems but also the importance of the topic that they are addressing, I encourage the authors to redo their calculations and rewrite their paper completely, along the lines that I indicate below. Since they have already done a large part of the work (at least, they have become familiar with the data and with the Pareto distribution), it should be feasible to redo the rest with a reasonable effort.

2 On the Pareto distribution

Unfortunately, much confusion surrounds the Pareto distribution in the literature. Fletcher et al.'s manuscript is no exception. Further progress will be difficult unless a few basic points are clarified.

Fletcher et al. distinguish the *power law distribution*, in which the probability of a fire of size A is proportional to some power of A, from the *Pareto distribution*, which satisfies

$$n_{X \geq A} = a_0 A^{-b} \tag{1}$$

In fact, the power law distribution and the Pareto distribution are exactly the same distribution. In the large-sample limit that is implicit in Eq. (1),

$$n_{X \geq A} = n_f P(X \geq A),$$

where n_f is the total number of fires in the sample. Equivalently,

$$n_{X \geq A} = n_f (1 - F(A)), \tag{2}$$

where F is the cumulative probability function. Following the authors in the convention of assigning the unit value $A = 1$ to the minimum area that is included in the distribution,

BGD
10, C5990–C6001, 2013

by definition $F(1) = 0$. Combining this with Eqs. (1,2), we obtain

$$F(A) = 1 - A^{-b} \tag{3}$$

for the Pareto distribution, as is well-known. Since, by definition, the probability density function (PDF) is $f(A) = dF(A)/dA$, it follows from Eq. (3) that the PDF of the *Pareto distribution* is

$$f(A) = bA^{-b-1}, \tag{4}$$

i.e. what the authors call *power law distribution*.

The authors state that they prefer the Pareto distribution rather than the power law because the first is continuous while the later would be discrete. We have seen that a power law is the same as a Pareto distribution, and it is apparent from Eq. (4) that it is continuous. They probably say that the power law is discrete because many studies use some plot similar to a histogram (on a logarithmic scale) for data that are claimed to display a power law, thus using a discrete number of bins. For each of the distributions treated in every textbook of statistics (e.g. a Gaussian), both the probability density function and the cumulative probability function are given, making clear that they are just two ways to express the same distribution. Frequently, histograms are also used for data following a Gaussian distribution (giving the well-known bell), but this does not mean that the Gaussian distribution is discrete. There is no reason not to apply all of these criteria also to the power law or Pareto distribution.

In empirical data, it is most often found that the power law relation breaks down above some given value A_{up}. This is what the authors call a *tapered Pareto distribution*. We do not have a general theory on the precise shape of the tapered part of the distribution. Often, many different functions are equally useful because of the few data in this region, and there is no reason a priori why the choice by the authors should not be valid. They assume

$$n_{X \geq A} = a_0 A^{-b} \exp(\frac{-A}{A_{up}}), \tag{5}$$

BGD
10, C5990–C6001, 2013

which, as can be easily shown, corresponds to

$$F(A) = 1 - A^{-b} \exp(\frac{1-A}{A_{up}}) \tag{6}$$

and

$$f(A) = (bA^{-b-1} + A_{up}^{-1} A^{-b}) \exp(\frac{1-A}{A_{up}}).$$

3 Data treatment

Fletcher et al. fit the Pareto and the tapered Pareto distribution to Amazonian fire size data using four different methods. They establish a distinction between *fitting* for the first method and *estimating* for the other three (which are small variants of one single approach), probably because the later are intended to use less data for practical applications. The first method is also used to test the goodness of fit. The four methods are problematic. I show this for each of them in turn.

3.1 Method 1

The authors do not describe their first method, but their Fig. 2 strongly suggests that it consists of minimum square fitting on the cumulative plots that they obtained after rearranging Eqs. (1,5) as follows:

$$\log(n_{X \geq A}) = a - b \log(A) \tag{7}$$

for the (not tapered) Pareto distribution, and

$$\log(n_{X \geq A}) = a - b \log(A) - \frac{A}{A_{up}} \tag{8}$$

for the tapered distribution. In both cases, $a = \log(a_0)$.

They used this method both for fitting the parameters and for testing the goodness of fit. As a fitting method it is arguably suboptimal, but I consider it to be acceptable (certainly, it is widely used). As refers to the goodness of fit, it is clearly unacceptable. The main problem is that one of the assumptions of the minimum squares method is that the errors are independent for each data point. This assumption is violated by cumulative plots, in which no data point can take a value larger than its precedent or smaller than its follower. A cumulative plot of a completely wrong model can give a correlation that would look very large in the context of a correct application of, for example, ordinary linear regression, but this is just an artifact due to the forced alignment of the data points.

The authors give the squared correlations, but these are meaningless in this context. They also state *Both the Pareto and the tapered Pareto distribution explained a significant proportion (p values $< 2.210\text{-}16$) of the variance of the data (98 % and 99.8 %, respectively)*, but the attribution of this meaning to their correlations is incorrect for the above reason.

Instead of proving the suitability of their two models, a careful examination of their cumulative plots (Fig. 2) makes clear that neither model fits the data. For the Pareto distribution, the data points should follow, approximately, a straight line all over their Fig. 2a. For the tapered Pareto (Fig. 2b), the approximation to the straight line should increase as we move to the left of the figure. In both cases, the density of data points should increase to the left. However, in both cases we see that, an the leftmost end, the points separate from the straight line and become less dense. The decrease in density is very apparent and there can be no doubt that it is significant. This does not necessarily mean that Amazonian fires are not fractal: it could well be a problem of resolution of the data (Sect. 4), but the statistical fact is that these models do not fit these data.

BGD
10, C5990–C6001, 2013

The lack of fit is not negligible. Given the logarithmic scale in their Fig. 2, the total number of fires in the fitted distribution ($n_f = n_{X \geq 1}$) is several times larger than the real one.

Among the methods used by the authors, this is the only one that they do not apply to estimate the total burned areas. The bias mentioned above suggests that the results of doing so would be clearly unrealistic. However, this method is used to support the distribution hypothesized by the authors and to validate the estimates obtained with the other methods.

3.2 Method 2

From Eqs. (7,8), the authors obtain the estimators:

$$\hat{a}_{par} = \log(n_f),$$

$$\hat{a}_{tap} = \log(n_f) + \frac{1}{A_{up}}.$$

These are applied separately for each of several grid cells, but always using the value of A_{up} that was estimated by Method 1, so the only input is n_f. From the resulting a, they also estimate b:

$$\hat{b}_{par} = \frac{a_{par}}{\log(\max(A))}, \tag{9}$$

and

$$\hat{b}_{tap} = \frac{a_{tap} - \frac{\max(A)}{A_{up}}}{\log(\max(A))},$$

BGD
10, C5990–C6001, 2013

where $\max(A)$ is an approximation to the largest fire in the grid cell. That is to say, their estimators are

$$\hat{b}_{par} = \frac{\log(n_f)}{\log(\max(A))},$$

$$\hat{b}_{tap} = \frac{\log(n_f) - \frac{(\max(A)-1)}{A_{up}}}{\log(\max(A))}.$$

The same $\max(A)$ is used for all grid cells, and is obtained by averaging the maximum A in each of the cells.

When estimating the total area, this method might give better results than Method 1 in spite of the data not following the assumed distribution, because it is based on the actual number of fires. However, it has little point if the distribution is not Pareto or tapered Pareto (although this might be solved by addressing the resolution issue), and, if it is, I do not think that it is an optimum method.

I see a alrge problem in assuming the same $\max(A)$ for all cells. The authors consider the hypothesis of a dependence between $\max(A)$ and n_f, which is the reason why they also develop Method 3. However, they seem to give the same weight *a priori* to this hypothesis and to the hypothesis of no relation between n_f and $\max(A)$ (which would justify Method 2), but I would give very little weight to the last hypothesis. It is an elementary fact of statistics that, for a given distribution, the larger the sample the larger $\max(A)$ is (see also Sect. 3.3). A constant $\max(A)$ would only be justified if there were some mechanism associating a larger n_f to a larger b in a precise way, but the authors did not propose any such mechanism and I do not think that it exists, at least in tropical rainforests.

Furthermore, the authors do not make clear if $\max(A)$ is assumed to be a kind of universal constant, as it seems that they assume for A_{up}. If they do not assume so, Method 1 losses the interest of needing only n_f as an input. Then, there is no reason not to use a method that makes a more exhaustive use of the data.

BGD
10, C5990–C6001, 2013

The problems of using a constant A_{up} are not so evident, but I do not either think that this is realistic (Sect. 4).

3.3 Method 3

Method 3 is similar to Method 2. However, it does assume that $\max(A)$ is a function of n_f. They define μ as an approximate estimator of $\max(A)$, which is calculated with the equation:

$$\log(\mu) = q \log(n_f), \tag{10}$$

where q is introduced as a constant that depends only on the resolution of the grid.

Let us check if this relation is realistic for a given sampling distribution. As mentioned by the authors elsewhere, we can use as an approximation that $n_{X \geq \max(A)} = 1$. From this and Eqs. (2,3),

$$\log(\mu) = b^{-1} \log(n_f).$$

Therefore, the relation is valid for the Pareto distribution. However, since $q = b^{-1}$, fixing q means fixing also b, so when they estimate b by replacing μ for $\max(A)$ in Eq. 9, they always obtain the same b. In other words, this method of estimation is circular. When they decide a given q they determine the b that they will obtain.

From Eqs. (2,6), the relation for the tapered Pareto distribution is

$$\log(\mu) + (bA_{up})^{-1}\mu = b^{-1} \log(n_f) + A_{up}^{-1}.$$

Therefore, Eq. (10) is approximately valid for a tapered Pareto, except for a small correction that is a function of A_{up}. This explains why they find some variability in b, but small. This variability does not reflect real differences in b among cells, but the fact that Eq. (10) is not fully accurate in this case.

BGD
10, C5990–C6001, 2013

3.4 Method 4

Since applying Method 3 they obtained almost no variability in b, they also considered fixing b. Even though the way they arrived to this conclusion is incorrect (Sect. 3.3), taking a constant b is very likely to be closer to reality than taking a constant $\max(A)$ as in Method 2. However, some problems remain: (1) that, while this might be an acceptable approximation, I do not think that it is entirely realistic (Sect. 4), (2) that I find less likely that it is acceptable taking a fixed A_{up} (Sect. 4), which is also a step in this method, (3) that some sound method has to be implemented for estimating b and A_{up} at least at the beginning (even though these estimates are later reused), and (4) that the data should really follow a Pareto or tapered Pareto (Sects. 3.1,4).

3.5 Method choice

The choice among methods 2-4 was based on their performance in reproducing the total burned areas. These areas were calculated by taking discrete pieces of the cumulative distribution in Eqs. (1,5). They can be calculated more accurately by integrating Eqs. (3,6), but I do not think that this is a big problem. However, the ways in which the calculated areas are used do pose problems.

Part of the comparison is based on the RMSEs in the plots in their Fig. 4. These plots do not satisfy the conditions for a linear regression, which implies that their RMSEs are not very informative. They all have a cloud of points at one end and a few other points scattered elsewhere. Taking logarithmic axes would probably give more information.

The reconstructed burned areas in Fig. 5 can be helpful, but cannot replace a more rigorous method.

BGD
10, C5990–C6001, 2013

4 Interpretation

As shown in Sect. 3.1, it is not correct to infer from the authors' application of Method 1 that the data follow a Pareto or tapered Pareto distribution. The correct conclusion is that they do not. However, this could well result from the problem with the resolution of the data that is pointed out by Referee #1. Apparently, the authors considered a minimum size of 25 ha. In Pueyo et al. (2010), using the areas of rainforest fire scars mapped from CBERS-2 images of Pando (Bolivian Amazonia) in 2005 (the same year analyzed by the authors), we did find evidence of scale invariance, but only above 30 ha, despite the fact that CBERS has a higher resolution than MODIS. The authors' Fig. 2 suggests that, if there is scale invariance in their data set, it does not start much below 200 ha, which seems consistent with the issue of resolution.

Therefore, we cannot infer from the authors' observations that Amazonian fires are scale-invariant but this is plausible considering other previous findings. However, I do not agree in that this should imply self-organized criticality (SOC). SOC is just one among many possible explanations for the Pareto distribution, however popular it might be. The SOC forest fire model is based on the assumptions that, locally, there is a negative feedback between biomass and fire, and that, globally, the system has reached a steady state pattern of biomass distribution resulting from this interplay. This might be true for boreal forests, but rainforest fires very often result from anthropogenic disturbance and extreme drought in areas that might not have burned for thousands of years, rather than affecting stands that have become flammable because they have just accumulated the right amount of biomass in a process that started after the last cyclical fire. In fact, there is evidence of a positive feedback in rainforest fire (Cochrane1999), rather than a negative feedback as needed for SOC. Fletcher et al. give no hint of how there could be SOC in a system like this. In Pueyo et al. (2010) we suggest that the tapered Pareto distribution in rainforest fire might result from scale invariance in patterns of anthropogenic disturbance, and perhaps also in the distribution of the few

BGD
10, C5990–C6001, 2013

types of rainforest that are naturally susceptible to fire, such as igapó.

One of the stated intentions of the authors is testing the hypothesis that *vegetation and climate variations do not affect the distribution parameters directly, but instead influence the number of fires or fire fronts that occur. These, in turn, determine the amount of forest area burnt.* I do not think that this is realistic. For example, the large fire activity in Acre in 2005, which can be appreciated in the authors' Fig. 5, was much larger than in previous years. Clearly, this did not result from an abrupt behavioral change with Acreans igniting many more fires than previously, but from the extreme drought that affected SW Amazonia in 2005. Probably, the spatial patterns in the figure reflect in part the distribution of ignitions, but also of the previous forest degradation and of the drought, which was concentrated in the SW.

This does not mean that the authors' intention to infer the complete fire distribution from the number of fires should be a bad idea. All the fires analyzed by the authors had already reached a considerable size, which means that they were able to propagate. Therefore, in this context, n_f is largely an indicator of fire propagation, and we can expect it to covary with other indicators of fire propagation, such as b and A_{up}. In a different context (boreal forests) I found major changes in the parameters of the fire size distribution as a function of weather, and also an effect of weather on the number of fires, even though I was working with much more resolution (Pueyo 2007, Climatic Change 82: 131-167). I found a clear effect of weather on both b and A_{up} (in a type of tapered Pareto distribution slightly different from Eq. 5), and the later had the strongest effect on the burned area. Considering this quantitative precedent and my qualitative observations elsewhere, I am not convinced that the assumption that A_{up} is fixed can be useful. A fixed b (as in Methods 3 and 4) is not either realistic, but it is probably more realistic than the assumption that all of the changes in the fire size distribution are due to changes in b (as in Method 2).

BGD
10, C5990–C6001, 2013

5 Suggestions

In view of all of these points, I suggest that the authors redo their calculations and rewrite their manuscript completely. I suggest that they:

1. Carefully select the minimum fire size that is really within the scaling range, and do not deal with smaller fires or treat them separately (for issues of interpretation, it will be interesting to know if this lower limit is also a limit to the reliability of Modis' fire detection).

2. Use some standard test of goodness of fit to test the distribution (of the same types that they would apply to other distributions).

3. Apply a sound method to fit the parameters. If the data are really well fitted by the theoretical distribution, maximum likelihood can be a good choice. If this is not completely clear, they can take advantage of the more robust methods that I used in Pueyo (2007) (their Method 1 might also be an acceptable approximation, but they should apply it only to the scaling range, and should not use it to test the goodness of fit).

4. Study how b and A_{up} covary with n_f. Avoid aprioristic assumptions such as those in Methods 2-4, unless they result from some theory.

5. Remember that scale invariance does not imply SOC, and that SOC is unrealistic in tropical rainforest. Either study the causes of scale invariance in this context, or treat scale invariance just at a phenomenological level.

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