Nat. Hazards Earth Syst. Sci. Discuss., 2, C3888–C3891, 2016
www.nat-hazards-earth-syst-sci-discuss.net/2/C3888/2016/

Natural Hazards and Earth System Sciences
Open Access
Discussions

Interactive comment on "Signatures of the self-affinity of fracture and faulting in pre-seismic electromagnetic emissions" *by* S. M. Potirakis et al.

Anonymous Referee #2

Received and published: 7 August 2016

Review of "Signatures of the self-affinity of fracture and faulting in pre-seismic electromagnetic emissions".

My background: I have worked with rescaled range, detrended fluctuation analysis, and power-spectral analysis for time series analysis of environmental time series.

I have read through this paper several times and remain unconvinced by the argument given. The data used is clear 'what' it is, but limitations and uncertainty in the data not discussed. I also believe that many more days (i.e., many months) of data need to be used to be convincing. The methods used are 'standard' techniques in the statistical physics/complexity community but have many limitations in them, which are not ac-

knowledged. The benchmarks also have no benchmark testing performed which limits their utility given the data (i.e., using time series with known statistical properties on the techniques themselves). The argument would also be much more convincing if statistical hypothesis testing is used, something I feel is very important any time 'precursors' are discussed. Although the authors do not use the word 'prediction' of earthquakes, it is implied in the precursory phase information. Finally, the argument is one that is laden with jargon, not always defined (and where I am an expert, some misunderstandings of the current literature). In short, this adds up to me just not sure if what has been found is statistically significant, clear, and reproducible on larger data sets. Below are specific comments in more depth.

(A) General. Major. Statistical testing and need for much longer time series. To be convincing, I would expect to see some statistical hypothesis tests done, that would include much longer time series. I find that the precursor activity that is 'visually' shown in the time periods shown are not convincing. What happens when one looks at 200 days of activity (vs. 11 days). Would one find that the precursor (call it A) it always followed by 'activity' (call it B). If we look at the 'strength' of A (how much the signal goes up and down) vs. the strength of 'B' what happens? Is A always followed by B, is B always preceded by A? I would be much more convinced if both a much longer time series were examined and if statistical hypothesis testing were done.

(B) General. Major. Who is the audience? The paper is an incredibly jargon rich, often without explaining terms, and difficult to read through. The way the introduction (and indeed much of the manuscript) is currently written, a scientist would need to be fairly familiar with the techniques that are being used to follow them, and even with that expertise at times it is not easy to follow. If your audience is an intelligent outsider, then please add to the text so that others can follow the argument that by its nature is very jargon rich and mathematical. It is fine to have some of the analyses as 'see ***' so long as these actually allow one to work from data to methods to results.

(C) Description of data. The data itself is described, but would benefit by a stronger description as to limitations and uncertainty in the values, and the recording technique. I've never found a data set that is 'perfectly' measured, so would appreciate this background.

(D) Rescaled range (R/S) [Major]. Many existing studies have however shown that R/S analysis is an exceptionally non-robust technique for synthetic time series that are (a) small number of values, (b) strongly non-Gaussian. It has also been shown that R/S is only appropriate over a given range of long-range persistence, unlike DFA and power spectral analysis. These are all limitations which should be mentioned and put into context, or R/S just not used, because of many inherent limitations of this method. There are a number of studies, dating back to the 90s, which have taken fractional noises of different strengths of persistence and different underlying frequency size distributions, and compared the strengths and weaknesses of R/S, DFA and power-spectral analysis. These should be mentioned, so the reader understands the strengths and weaknesses of the techniques to be applied.

(E) Detrended fluctuation analysis (DFA). [Major]. DFA also has many limitations due to the one-point probability distribution of the data examined, and the length of the time series. Some of these limitations can be removed by using higher orders of DFA (DFA3, DFA4). Somehow, the time series that you are examining (for R/S, DFA, PS analysis) need to have benchmarks input to them with properties similar to your time series, so that we can better understand the limitations of each technique in terms of the resultant applications to your 'real data'. This will give us uncertainty bounds.

(F) Fig. 2 and other figures where the results of 'fits' are shown (estimators). Need to see some of the fits as examples. What is window being used for DFA, R/S, etc., and to what lags (this is described in the text as samples, but can we see this in the figure in timeâĂŤit is hard going between the text and the figure for time/samples)? This is unclear and might impact any potential predictive power.

Other:

Abstract: This reads more like an introduction to a paper, not a factual data (and specifics), methods and conclusions. Please rewrite more in the form of a quantitative summary of the paper.

1.0 Introduction. Spatial vs. temporal. Although much of the introduction is about time series, in some places spatial patterns are brought in. This gets a bit confusing going between spatial patterns, frequency-size of the attributes of the spatial patterns, and the time series. How are these related, or should they be related?

Rescaled range [minor] Probably best not to say the scaling is as 'time' increases' but as the period considered increases. You might also want to note he found this for certain environmental phenomena.

Spectral analysis. I would not measure the fit using r2, but rather, using Maximum Likelihood Estimation (or some other robust technique for measuring the power spectrum).

When you discuss items like "1024 sample long windows" tell us how much this is in real time.

Fig. 2b and 2c: I'm not clear if the resultant H is plotted in middle of window or to right of window boundary.

"and running time average of four windows with 25% overlapping". This was unclear.

"Only beta exponent values. . .with rˆ2 > 0.5 were considered here". How many were rejected?

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 2981, 2014.