Review Report
Author: Francesco Serinaldi

Title: About the return period of a catastrophe
Author(s): Mathias Raschke
MS No.: nhess-2021-86

General comments

I recognize that my report will look harsh, but reading this manuscript was a true nightmare. English is embarrassing, just to use an euphemism, and this makes many sentences/paragraphs completely incomprehensible. Sentences are generally disconnected one another, thus preventing the understanding of what the Author wants to communicate. The material is randomly spread throughout the paper without any logic like the paint in Pollock's artworks (... which would be good if NHESS journal were a Christie's auction); rationale, technical aspects, and results are mixed, and the Author moves back and forward among them without following any rational criterion. Technical aspects are introduced without clear terminology and specification of assumptions and mathematical derivations.

By the way, my frustration is exacerbated by the fact that (i) the material of the paper may be of interest if properly communicated, and (ii) the Author is not a young undergraduate student dealing with his first paper, but a researcher with some experience. Therefore, submitting a paper of such a low quality in terms of presentation is not a matter of lack of capability of writing a decent document. And this makes the poor presentation even less acceptable, as it denotes a sort of lack of attention and respect for editors, reviewers, and readers.

That said, I will try to provide some suggestions about how to reorganize the manuscript. However, let me clarify that a simple rearrangement of the material is not enough. Almost every line of the text requires rewording (to get a decent English) and restructuring (to make sentences and concepts understandable).

Specific comments

Abstract: The structure should be: motivation/problem, proposed technique, and results. Here, the terms of the problem are quite simple and the abstract should look like this:

"Natural catastrophes are spatial process affecting a given area; however, natural hazards and their impacts/effects are generally monitored/measured locally. In order to quantify the degree of rarity (probability of occurrence/exceedance or return period (RP)) of a spatial event (hazard/loss), we need suitable metrics enabling to assess areal risk from the local one. In this study, we propose a metric called "combined return period" (CRP), which is the (weighted) average of local return periods, and can be shown to be a proper return period itself. CRP is characterized by some properties that allow the calculation of the areal/spatial expected loss for a given areal RP or the expected RP for a specified areal loss starting from local RPs, hazard values, and exposure. The paper also discusses the effect of considering max-stable and non max-stable spatial dependence, and introduces bias correction methods for local RP estimates, etc. As a case study, the proposed framework is used to quantify RP and losses of winter windstorms over Germany recorded from 1999 to 2019. Results are compared with those reported in the literature and show that...bla, bla."

Of course this is only a suggestion, but it gives an idea of how to reword the abstract in plain language, following a logical structure (a story, if you want), without mentioning things such as "*testable reproductivity*" (whatever it means) or "*pseudo-polar coordinates*", which require a technical introduction to be understandable.

Introduction: This should slightly expand the abstract. References are OK. However, the message is not conveyed. If I understand, the Authors wants to say that the general approach in the existing literature is to classify hazard events evolving in space and time via simple indices, such as the Richter magnitude scale for earthquakes, and then assigning an RP or probability to the observed values of these indices. Conversely, the Author suggests assigning an RP to a spatial event by combining the RPs of the original hazard/loss variables recorded locally. And this approach has the properties/advantages mentioned (very confusedly) by the Author throughout the text. If my interpretation is right, the Author should make it clear.

By the way, please avoid expressions like "*the destruction's extent of the destruction*", or "*In sum, previous approaches are not very fruitful.*"... perhaps the Authors means "previous approaches are not very effective". Please also avoid sentences like ''*Also, their statistical models include assumptions and pitfalls*": all methods/models are based on assumptions, while pitfalls should be specified or properly referred via suitable references.

L17-20: These lines are an example of what I mean when I say that many sentences are disconnected. For better reading and understanding, it should read as "Natural catastrophe (NatCat), such as large windstorms or earthquakes, are natural hazards evolving in space and time. This means that the definition/identification of a NatCat event is not unique, and generally relies on both hazard magnitude indices and socio-economic aspects, such as the interest in short-term or long-term effects on the affected areas. Irrespective of a specific definition of NatCat event, this study deals with the assessment of RPs of complex hazard events and corresponding losses affecting multiple locations/areas and spanning a given time interval of interest..."

L44-55: This part is an example of the general lack of clarity characterizing the discussion of technical concepts. These lines introduce the key concepts to understand the rest of the paper, and this is done superficially, without the required premises. The Authors merges methodological concepts and empirical aspects, and uses meaningless terms such as "*A Poisson point process... is also a random element*": why "*also*"? is a Poisson process anything else? what is a "*random element*"? Which paper or book does refer to a Poisson point process as a random element?

As an example, this part could read as follows:

"To put our discussion in the context, let us assume that an environmental process of interest, such as river flow or wind, is monitored at a given location by gauging devices that measure for instance river stage/discharge or wind speed and direction. A NatCat event occurs when the measured variable X assumes a value equaling or exceeding a critical value x, i.e. $X \geq x$, thus resulting in possible damages. The occurrence process at a given location can be described by stochastic process, which is a collection of random variables. In particular, a Poisson point process or briefly Poisson process is a convenient model to describe the occurrence of independent events such as rare NatCat events. In Poisson point processes, the number of events, K, over specified time intervals $\mathcal{I}$ (e.g., a season or a year) follows a Poisson distribution with expected value

$$\mathrm{E}[K(x)] = \Lambda(x), \tag{1}$$

where $\Lambda(x)$ is the exceedance frequency function, EF, and $K(x) = \sum_{i \in \mathcal{I}} \mathbf{1}(X_i > x)$. The reciprocal of the local EF yields the local RP

$$T(x) = \frac{1}{\Lambda(x)}. \tag{2}$$

Since local EF Λ is uniformly distributed, and the relationship between Λ and T is monotonic, according to the rule giving the distribution of functions of random variables (e.g. Kottegoda and

Rosso 2008; pp. 133-142), the EF function of T has Pareto form

$$\Lambda(T) = \frac{1}{T}. \tag{3}$$

etc. ”
This should give an idea on how to present such a kind of things. Analogously, Eq. 4 should be written as

$$\begin{cases} R = T_1 + T_2 \\ V = \dfrac{T_1}{T_1 + T_2} \end{cases} \iff \begin{cases} T_1 = RV \\ T_2 = R(1-V) = T_1 \dfrac{1-V}{V} \end{cases} . \tag{4}$$

Again, when referring to books such as Coles (2001), Beirlant et al. (2004), and Falk et al. (2011), please indicate the exact pages, as going through a whole book to double check what an author writes is almost impossible. For Falk et al. (2011), please provide reference details (publisher and address).

Section 2.2: Please reword “*Opportunities and implications*”... opportunities? perhaps consider something like “CRP properties: from local to spatial RPs and losses”. Again, this section is not clear at all. Please avoid digressions on e.g. goodness-of-fit tests, or other things that are not relevant. This section should explain how CRP can be used to assess spatial RPs and losses by aggregating local RPs, quantiles and exposures. It should be merged with sections 4.2, which attempts to explain (with the usual lack of clarity characterizing this manuscript) the rationale of the method, and section 4.1, which provides a summary of the ‘scaling’ algorithms. Figures 4, 5, and 6 should be described in detail, as they can help clarify methods and scope. Please, spend time on this, as this material is the actual methodological body of this manuscript.

The material in section 3 should also be included in the methodological section after removing parts referring to the German case study. Indeed, this section presents e.g. loss ratios and other concepts that enter the ‘scaling’ algorithms.

Finally, the empirical results in section 3 should be merged with sections 2.3 and 2.4 to create a unique section (with possible subsections, of course) describing all empirical results, and the various applications of the concept reported in the methodological part.

L63: ‘*random element*’??? If T_1 and T_2 are random variables, every quantity resulting from their combination is also a random variable.

L70: Please consider something like “Exploiting the properties of Poisson processes, the univariate CDF of maximum RP values occurring in k unit periods can be expressed in terms of the EF $\Lambda(x)$ in Eq. (1) (see e.g. Stedinger et al. 1993; Ch. 18, pp. 37-38)

$$G_K(x) = \exp(-k\Lambda(x)) = \exp(-k/T(x)). \tag{5}$$

etc.” I am not sure that Eq. (3) is needed here.

L73-75: “*The independence gives this max-stability of the dependence structure between pseudo angle V and pseudo radius R in (4) (Coles, 2001).*”??? Please clarify, and report pages of Coles (2001) discussing this property.

“*The occurrence of the pseudo radius is once again a point process with EF* $\Lambda(x) = 2/x$ *- the double of (3).*” Please provide a reference (with pages, if it is a book). By the way, if this sentence refers to pseudo radius, it should be $\Lambda(r) = 2/r$. My understanding is that the Author uses x as a generic variable when he presents an EF of some quantity (e.g. T, R, etc.). However, this introduces

lot of confusion, and makes reading and understanding very difficult, leaving aside possible errors when handling and combining equations. Please use a consistent notation.

L98: Please, clarify.

L106: "*The CRP T_C represents the expectation (or its estimate).*"... of what?

L117: "*The reason is explained in Section 3.1 and the appropriateness of the Gumbel distribution for the block maxima of local event intensities and corresponding computation of RP per event with bias correction.*"??? What about using subject, verb, and object? Just to write a sentences with some meaning.

L141: "*The scatter range of the half seasons is smaller than for two seasons due to different sample sizes.*" In my opinion, the difference depends on the fact that the 'two-season' sample actually merges data from two seasons that are expected to be less correlated, as they are likely non-homogeneous (seasonality effects, etc.).

L151: '*this confirms the non-max-stable behavior of Kendalls τ.*' How can rank correlation coefficients be max-stable?

L183: "*is more minor than*" → 'less than'

L241: Please provide reference or derivation for Eq. (13)

L290: "*Our estimation variants are formulated by (11).*" Variants of what? SARS-CoV-2? What about making things readable? For example, "In this section, we show how to use Eq. (11) to derive alternative estimates of this and that... bla, bla"

L297: "*The estimation is based on following stochastic relations and assumptions (or proxies)*" Estimation of what? Proxies? Is it so difficult to start a section trying to explain what is gonna be presented?

L299: "*The origin is (5); the well-known delta method (Coles, 2011) for computation of propagation of errors is also a base. A more illustrative explanation is provided for the loss scaling by Figure 6 a.*" Origin of what? Why talking about delta method without any justification? Is it a base for what? I hope the Author will recognize that these sentences are presented without any logic and explanation. The scope of a paper is the communication of ideas; this manuscript is more similar to a collection of personal notes reporting only some keywords for Author's record, and neglecting the fact that a reader is not clairvoyant, and cannot read the Author's mind to shed light on those short notes.

L359: "*Further arguments...*" → 'There are *further arguments...*': subject, verb, object... it is not so difficult, I think.

To conclude, as mentioned above, this manuscript is one of the most badly written documents I handled in the last months as reviewer and editor. Nonetheless, the topic may be of interest; so, I think it deserves a chance to become a readable paper. However, I want to be clear: cosmetic changes are not enough. Every sentence, paragraph, section, and the overall structure require to be carefully revised. The Author can consider involving colleagues that can help in this respect. Concerning the language, the Author can consider the use of proofreading services.

Sincerely

Francesco Serinaldi

References

Kottegoda, N.T., Rosso, R. (2008). Applied statistics for civil and environmental engineers. Malden, MA: Blackwell.
Maidment, D.R. (1993). Handbook of hydrology. New York: McGraw-Hill.