

Reply to reviewer comments (Francesco Serinaldi/RC1) regarding MS "About the return period of a catastrophe" - nhess-2021-86

Dear Dr Francesco Serinaldi,

Thank you very much for all your effort regarding my manuscript (MS) and the helpful comments, notes and advice. I will consider some of these directly in a revision. However, I will also reject some of your concerns by arguments. Before, I would like to give following general reflections.

I am aware that my English is poor and awkward. Therefore, I used already the help (proofreading) by two colleagues. The outcome of this procedure is unfortunately not good. Besides, I am not in the privileged position to be paid for research and writing scientific papers in English. I rarely publish also because I do science with my personal limited resources. Nonetheless, I will use a more professional service after a revision. The mathematical notation was already validated by a mathematician for the current submission.

Furthermore, there is no uniformity in scientific writing. The explanation style is extremely short in Mathematics compared (e.g.) with social sciences. I am more oriented to the first and prefer sparsity.

Besides, I was not sure about the level of mathematical/stochastic expertise of the different (and fragmented) science communities which deal with natural catastrophes. According to your and the other reviewers' comments, more mathematical explanations are needed. I will add a subsection. However, I don't want to explain too many details which would be trivial for a student in mathematics (maybe 3rd semester). A higher level of statistical expertise of the reader is assumed what I will mention in a revised introduction.

In addition, I will change the structure of the MS in a revision. Nevertheless, I underline that there are very different accepted or even prescribed structures for scientific papers (e.g., [Nature Scientific Reports](#)). And I keep the distinction between the new approach (CRP as main result) and technical details for the demonstration example.

My detailed replies to your comments are below.

Once more, thank you very much for your advice, notes and comments.

Sincerely, Mathias Raschke

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 NHES 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).

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.

Replay: I will improve the English in a revision. However, I am also convinced that it is acceptable to mention the theoretical basis of the new concept in the abstract. Therefore, I will also mention the *pseudo polar coordinates of extreme value statistics/theory* in the abstract of a revision. Besides, the term *return period* (RP) is already common in the NHESS and is frequently used in abstracts. A recent example is the paper by [Letson et al. 2021](#). Furthermore, an actual return period is not the same as an occurrence probability even though there is relation between both.

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.

Replay: I will consider your suggestions and concerns in a revision. Probably, I will shift a part of the current section conclusion in the introduction following a suggestion of the other referee.

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...”

Replay: I will consider your concerns in a revision. However, your suggestion is very long.

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?

Reply: Please aware, that the term *random variable* is mentioned in the previous sentence. Obviously, *random element* is a generalisation and include *random variables, Poisson point processes* and so on. I will not provide a special reference since this is a trivial and self-explaining term and is even described in [Wikipedia](#). For more examples, please use google search (catchwords [stochastic “random element”](#) - more than half a million results).

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

$$E[K(x)] = \Lambda(x), \quad (1)$$

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

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

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

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

etc. ”

Reply: Thank you for the suggestion, I will consider parts in my revision. However, your notation does not convince (e.g., equation (3)). In stochastic, one notation is used for the scale of real numbers. For example, random variables X and Y have cumulative distribution functions (CDFs) $F_x(x)$ and $F_y(x)$. The domain scale is symbolized by x in both CDFs. My current notation was validated by a mathematician with special expertise in extreme value statistics and theory.

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

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

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).

Reply: I will use your suggested notation (4) in a revision. I will also mention the pages of references in the text (not in the reference list since I do not want to list the same literature twice).

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

Reply: As explained before, *random element* is a general term in stochastic. And T_1 and T_2 are not random variables. “ T is also a point event” as written in line 54.

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)

$$G_K(x) = \exp(-k\Lambda(x)) = \exp(-k/T(x)). \quad (5)$$

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

Reply: I will modify my expression and will not use your suggestion since it is longer. My notation has been validated by a mathematician.

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.

Reply: Regarding the notation, please see my replies before. May be your misunderstanding of the notation the reason for your confusion. In the reversion, I will cite Coles (2001) “*In other words, the angular spread of points of N [the point process] is determined by H , and is independent of radial distance.*”.

L98: Please, clarify.

Reply: I will add a new sub section about Schlather’s theorem/approach.

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

Reply: I will clarify.

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.

Reply: I apploigize and will correct.

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.).

Reply: No, the reason is the difference between the sample sizes. For every estimate applies that the corresponding standard error decreases with increasing sample size. The samples of the half seasons are concerned by the same seasonality as the sample of a complete of two seasons (two years). The season is divided according to reasonable physical criterion. The first half is from September to December, the second half from January to April. The sub sample of the first half do not differ significantly from the sub sample of the second half according to t-test and F-test.

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

Reply: If the dependence structure/copula is max stable (the same for one or ten seasons or years), the corresponding dependence measure (which also parametrise the copula) must be also stable. As aforementioned, I will add a further sub section which explains this amore in detail.

L183: “*is more minor than*” → ‘less than’

Reply: I will consider it in a revision.

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

Reply: As written in the MS short before equation (13), its parameters correspond with parameters in equation (12). In addition, the relation between extreme value distribution and frequency function was already presented in section 2.1. I will consider your concerns it in a revision.

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”

Reply: Thank you once again for the helpful notes. I will consider it in a revision.

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?

Reply: I will modify the sentence it in a revision.

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.

Reply: When an equation is presented immediately before a sentence (no new paragraph) then it is very likely, that the latter is related to the equation. Why is a justification of the delta method needed? The delta method used in many statistical analyses and is well explained, for example, by Coles (2001). There are further course books for statistics which explain this approach. However, I will reformulate the sentences in a revision.

L359: “*Further arguments...*” → ‘There are *further arguments...*’: subject, verb, object... it is not so difficult, I think.

Reply: Even though, we could discuss about the style, I will consider your concerns in a revision.

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.

Reply: Thank you very much for all your advice, notes and suggestions. I will consider these in a revision.

References (which are not listed in the MS)

Letson, F. W., Barthelmie, R. J., Hodges, K. I., and Pryor, S. C.: Intense windstorms in the northeastern United States, *Nat. Hazards Earth Syst. Sci.*, 21, 2001–2020, <https://doi.org/10.5194/nhess-21-2001-2021>, 2021.

Wikipedia https://en.wikipedia.org/wiki/Random_element