

About the return period of a catastrophe

By M. Raschke

Submitted to Natural Hazards and Earth System Sciences

MS-NR: nhess-2021-86_R1

REFEREE REPORT

Firstly, I thank the Author for his reply to my comments on 10.5194/nhess-2021-86. He reorganized the structure of the text, which is now more accessible and readable. However, language is not yet good enough. Really, text is still full of grammar and syntax errors. Below I try to track errors and suggest changes up to line 105 or so. Eventually, I gave up because almost every line requires corrections or rewording.

Honestly, it is a pity that such an interesting content is not supported by decent language and presentation, thus preventing effective communication.

As far as I see, the Author intends to use a proofreading service after revision due to the cost of these services; I can understand, but the problem is that the poor language can also prevent an effective review. In these cases, reviewers have to interpret not only the technical content, but also figure out the actual meaning of each sentence. Some of my previous remarks were related to lack of clarity, and some of these issues are now a bit clearer in the revised version.

I repeat my previous conclusion: the material is interesting and I would like to see it published, but publishing something that is unreadable makes little sense.

I leave to the Editor the final choice. For sure, at least grammar errors should be fixed.

Reply to Author's general Response

(Author's responses are reported in [blue](#))

Please, do not get me wrong.

- Like you, I'm not a mother tongue, and I know very well the difficulty of writing in a foreign language. However, if we decide to play the publishing game in English journals, we have to produce texts in decent English. It's part of the job, we like it or not. Otherwise, we can publish in Italian or German... If we play a game, we have to play with its rules.
- Mathematical notation: like English, I do not care if the notation was revised by a mathematician. If you use the same symbol to denote multiple variables in multiple contexts,

this generates confusion, period. If your mathematician colleague thinks different, (s)he is wrong, I'm sorry.

- "trivial for a student in mathematics": I can guarantee (as a reviewer an associate editor) that the mathematical skills of the average audience of hydro-climate journals is far below those of mathematical students. Many risk "experts" are geographers or environmental scientists who never attended a calculus class, an often neglect also the difference between population means and sample means, leaving aside the concepts you discuss in your work.
- Style: In my daily job, I deal with papers published in maths journals, physics journals, medicine, biometrics, environmetrics, etc. Therefore, I know that there are different styles. However, in the past, when I approached for instance physics journals, I wrote those papers trying to follow the style of those journals, as it is more comfortable for editors, reviewers, and readers. Let us say, using a mathematical style (e.g. axiom, theorem, proof) on a hydrological journal makes little sense if you want to communicate to the audience of that journal. Let me better explain: sometimes, statisticians publish their methods/models on journals of theoretical statistics with the typical mathematical style (lots of symbols, few lines of text, theorem-proof organization, etc.); however, they change style when they apply these methods, and propose an environmental application in Journal of Hydrology, Water Resources Research, etc.. In these papers, you can recognize the stats' background of the authors, but also their attempt to fit the journal style for a different audience. I suggest being balanced in this respect and not too strict.
- Again, do not confuse some freedom of style with lack of clarity: Although different journals in different disciplines have different styles, there is always a rationale behind them, aiming at guaranteeing readability.

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.

Firstly, in my comment, I never said that you should not use the term "return period"; secondly, please read more carefully what I wrote: "*In order to quantify the degree of rarity (probability of occurrence/exceedance or return period (RP)) of a spatial event*": I talk about degree of rarity and I suggest some options in parentheses. I never wrote that exceedance probability is identical to return period.

Please, do not focus on irrelevant points that do not even exist in my comments, overlooking the general meaning of my statements. This strategy does not work in review processes.

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.

You contradict yourself: look at L121-122 of the revised paper.

If you have two random variables X and Y , you should write $F_X(x) = P[X \leq x]$ and $F_Y(y) = P[Y \leq y]$ to avoid confusion. $F_Y(x) = P[Y \leq x]$ makes no sense if e.g. $X = \text{flow}$, and $Y = \text{rainfall}$.

If you use “ x ” to denote the value assumed by whatever random variable (X, Y, T , etc.), the final notation becomes incomprehensible, despite the opinion of your mathematician friend.

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

This reply makes no sense. x is not equal to $T(x)$. Your notation is just confusing.

“My notation has been validated by a mathematician.”, also my notation was validated by a mathematician! Is my mathematician better/worse than yours?... please, we are not 10-year old kids.

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.”.

No, confusing notation generates misunderstanding. Please, use consistent notation.

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.

Basically, you analyse annual maxima (i.e. maxima over two seasons) with size n , let's say, and the series of the maxima over two half seasons of each year, i.e. a sample with size $2n$. OK! I'm sorry, but the text was not clear in this respect. The new version is clearer.

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.

OK, it is clear that we are speaking a different language, which prevent communication. I do not say that you should introduce the rationale/description/equations of Delta method; I say that a document should state something like “I will use Delta method for this reason... to obtain this... to quantify that... etc.” in order to guide the reader throughout the text.

I understand you like “axiom-theorem-corollary-proof” style and “sparsity” (whatever it means), but this is not a mathematical paper to be published in Annals of Mathematics or like. Most of my remarks do not concern theory, but presentation, which is important to make the paper more accessible and successful. Anyway...

Remarks to R1 version

Despite Author's effort, the level of language, grammar, and syntax is still insufficient. I tried to suggest changes (almost line-by-line) up to L105, but eventually I gave up for two reasons:

- 1) Rewriting parts of a manuscript goes beyond reviewers' task, and
- 2) The Authors already declined my previous suggestions, without recognizing that I tried to make unreadable sentences just readable.

L6: "How often does such an event occur~~s~~? What is ~~their~~ its return period (RP)?"

L10: "event and the~~r~~ spatial"

L12 "For this purpose, we extend previous construction for max-stable random fields from extreme value theory and consider a recent concept from NatCat research"

"a recent concept"? Which concept? If you refer to something mentioned above, please use "the above mentioned concept of..." otherwise "a recent concept from NatCat research, i.e. CONCEPT X".

L20-23: please consider this version:

"Moreover, we propose a new statistical index highlighting that the spatial dependence decreases with increasing block size. We compare our risk estimates with existing risk models that do not allow for such a detailed control of spatial characteristics and dependence."

Please do not discard my suggestions for text changes outright. It is not a matter of style, it is a matter of syntax.

I spent time to suggest alternative versions of some parts of the text because yours is difficult to read.

L25: "... the question arises, how does"? are you sure? I prefer the version "... a question arises: how often...", alternatively "... a question arises: what is the frequency of occurrence...?"

L27-29: please consider "definition of event. This definition may vary by peril, but it is not the topic of this study even though it influences our research object – the RP of hazards and NatCat events."

L30: "index is ~~so far~~ frequently used"

L31: "magnitudes scales... ~~But~~ However, their RP may not correspond ~~well~~ with"

L34: "measure; ~~but~~ however,"

L36: "The different definitions of these indices result"

L38: "In sum, as previous approaches are not satisfactory regarding the stochastic quantification of a hazard or NatCat event, we develop a new approach."

L40-43: "Building on results of extreme value theory and statistics, we mathematically derive the concept of combined return period (CRP), which is the average of RPs of local event intensities. As we will show by a combination of existing and new approaches from stochastic and NatCat research, the concept of CRP is strongly related to the spatial association/dependence between the local event intensities, their RPs, and corresponding block maxima, such as annual maxima."

L44: "Spatial dependence is not suitably considered in previous research about NatCat. This issue..."

L45: "... by Mitchell-Wallace et al. (2017, Section 5.4.2.5) about NatCat modelling for insurance industry. Jongman et al. (2014)..."

L48: "... assumption; examples are the snow model..."

L52: "Also Raschke et al. (2011) proposed a winter storm risk model for a power transmission grid in Switzerland without validation of stability assumption."

L56: "Youngman and Stephenson (2016) suggested a statistical model to simulation hazard events."

L55-61: No, your interpretation is wrong. Neither Youngman nor Papalexioiu simulate "local annual loss from catastrophes".

Youngman generates events with local POT, and then they extract spatial "catastrophes": they explicitly state that some of their events are not extreme in terms of financial losses.

Papalexioiu perform continuous simulation of the whole process (e.g. daily random fields/time series): extreme events in terms of financial losses or whatever else can be extracted from the whole process.

Criticizing papers is legitimate and necessary, but be sure of what you criticize before doing that.

L66-67: "*The corresponding assignment of Asadi et al. (2015) of one local flood peak to peaks at other sites does not convince us completely.*"

I'm sorry but this kind of statement is uninformative and not fair. As mentioned above, criticism is legitimate, but we cannot say "I do not like this or that, period" without explaining why.

Such personal opinions without supporting arguments are good for newspapers, not for scientific journals.

My suggestion is to focus on your own work rather than expressing vague opinions on others' work. If you want, you can write a discussion paper focusing on those criticisms.

L68-70: please, reword in a readable language.

L75: "opportunities for" → "approaches to"

L78: "opportunities"... again; please, use appropriate terms.

L88: "2nd Section and illuminate..." → "Section 2, where we recall the concept of max-stability for single random variables, bivariate dependence structures (copulas), and random fields"

L90: "*The more recent approaches of **hazard event related area functions** (Raschke, 2013) and survival functions (Jung and Schindler, 2019) of local event intensities within a region are implemented therein to characterize spatiality.*" ???

Do you mean

"The more recent approaches to area functions (Raschke, 2013) and survival functions (Jung and Schindler, 2019) of local event intensities within a region are implemented therein to characterize spatial properties of areal hazard events."?

L92: "In the 3rd Section" → "In Section 3"

L93: “scaling opportunity” → “scaling property”... please, stop using “opportunity” in every sentence and out of context. Please, use terms that are more appropriate.

L97: “to remain clarity of the main paper and limit its extent” → “to keep the paper clear and focused on main results”

L98-99: OK. However, please note that you use “x” to denote different things in the same section and same equation. This is not simplicity but only confusion. Please, see my example reported above about $X = \text{flow}$, and $Y = \text{rainfall}$.

You should establish a clear notation. If “X” denotes a random variable and “x” the values assumed by “X”, if you introduce another random variable “Y” its values cannot be denoted by “x”. Please think about that... and avoid answers like “my mathematician friend told me... bla, bla”, as this does not justify bad notation.

L99: “We also expect that the reader is more familiar with statistics and stochastic than only with basics about random variables. Statistical significance, goodness-of-fit tests, random fields, or a Poisson (point) process (Upton and 100 Cook, 2008) should be familiar terms.” →

“We also expect that the reader is familiar with statistical and stochastic concepts such as statistical significance, goodness-of-fit tests, random fields, and Poisson (point) processes (Upton and 100 Cook, 2008).”

L104: “Before we formulate the CRP and discuss their opportunities, we must present, discuss, and extend a corresponding topic – max-stability in extreme value statistics especial of random process and fields.” → “Before introducing CRP and its properties, we discuss and extend the concept of max-stability in extreme value statistics with focus on random process and fields.”

L121: “can replace by a copula” ... grammar...

L123: “a universal distinction”??

L124: “principals”??? → principles!

L157: “verities of standard deviation”... verities ???

L227: “exception” → “expectation”

L248: “it’s” → “its”

L309: “3.2 The scaling opportunity” → “3.2 Scaling property of CRP”

L325: “We introduce an alternative method. According to (18), the expectation of an unknown ELRP \mathcal{TE} is the CRP $\mathcal{T}c$ of the local event intensities; the CRP is an estimate of the ELRP (max-stability between ELRP and CRP provided). To get a good estimate of ELRP, we must average the $\mathcal{T}c$ of many events with the same event loss. We cannot observe such, but we can stochastically scale historical events respectively their local intensity observations.”

This concept is the fundamental message of this study!
Please, emphasize it and put it clearer!

For example

“We introduce an alternative method. Under the assumption of max-stability between ELRP and CRP, according to (18) with $T_1 = \text{CRP}$ and $T_2 = \text{ELRP}$, the expectation of an unknown ELRP TE corresponding to (conditional on) the CRP Tc , is the CRP Tc of the local event intensities. This means that the CRP is an estimate of ELRP under max-stability assumption, that is, we can use widely available hazard data to make inference on the risk of less available loss records. In fact, (18) implies averaging the Tc of many events with the same event loss to get a good estimate of ELRP. However, observations of events with same loss are not available. Nonetheless, we can exploit the stochastic scaling property of CRP to rescale the local intensity observations of historical events to get the required information”

L335: “be adjusted in an iteration until the defined event loss is the result of (26).” → “be adjusted iteratively until the result of (26) converges to the desired event loss.”

L343: “Delta method, well explained by Coles (2001, Section 2.6.4),” → “Delta method (Coles, 2001; Section 2.6.4),”

L345: “derivates” → “derivatives”, “derivatives”... English!!!

“In its meaning, we can also average the event loss for a fixed/determined CRP respectively its scaled variant.”... reword in understandable English!!!!

L346-359: please, reword in decent English!

L413: “accepts” → “does not reject”: statistical tests cannot accept anything; they only reject or not reject. The latter outcome only means that data/info is not enough to discard H_0 .

L423: “half-season maxima and two season maxima”... “two season” is confusing, as you actually have season/annual maxima and two half-season maxima per year. I suggest being consistent throughout the text, and use “season maxima” or “whole-season maxima” referring to the period from September to April, and “half-season maxima” for the data corresponding to Sept-Dec and Jan-Apr.

L426-429: Please, reword.

L435-436: English, please!

Please double check the language in the rest of the paper as well.

Sincerely,

Francesco Serinaldi