REVIEW REPORT

Journal: HESS
Paper: hess-2024-334
Title: Return period of high-dimensional compound events. Part I: Conceptual framework
Author(s): Manuel Del Jesus, Diego Urrea Méndez, and Dina Vanessa Gomez Rave

GENERAL COMMENT.
For the reasons given below, I recommend a **REJECTION WITHOUT RESUBMISSION.**

This paper is useless, and introduces no novelties concerning the estimate of the RP of high-dimensional occurrences. The work does not contain any relevant advance: it simply suggests (in a non-technical and superficial way) to use well known algorithms to find the critical points of multidimensional functions.

Not to say about the "Conceptual Framework": it is already well known in hydrological/geophysical Literature since about 20 years. All the formulas presented (sometimes inexact from a mathematical point of view) do not show anything new, nor they represent any advance with respect to the present knowledge.

In addition, the fact that the occurrences considered are Compound Events does not emerge in any part of this work: only superficial comments and descriptions are presented, but the true core problem is nowhere investigated.

Furthermore, sort of indications for carrying out a multivariate analysis are sketched, but they are too generic, never discussed in details concerning the problems they are expected to solve, and most of all they are all already well known in Literature: this paper adds nothing to knowledge, all what is written has already been more precisely (and mathematically correctly) introduced in already published works, so what?

Finally, the mathematical notation is often wrong: e.g., the Authors confuse a function (say, F) with its value (e.g., $F(x, y)$), or even worse confuse $F(x)$ with $F(X)$, where the former is a real number, whereas the latter is a random variable.

Overall, the content of this paper can be summarized in a single paragraph, and recycled in the Introduction of the companion paper (Part II).

SPECIFIC COMMENTS.

Line(s) 55–56

AUTHOR(s). Vine copulas are a flexible class of dependence models consisting of bivariate building blocks.

REFEREE. The Authors do not mention the problems intrinsic to modeling via Vine copulas.

Line(s) 58–59

AUTHOR(s). For more theoretical details, please refer to (Sklar, 1959; Nelsen, 2006).

REFEREE. None of the references is pertinent to Vine copulas: more recent and relevant paper must be cited, starting from the basic one:

Aas, K., Czado, C., Frigessi, A., Bakken, H. (2009). Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics, 44(2), 182–198.

Line(s) 149–151

AUTHOR(s). For instance, Kendall's τ is more appropriate when the joint distribution is non-Gaussian (Serinaldi, 2008). Spearman's rank correlation is based on the rankings of variable values, whereas Kendall's rank correlation assesses the concordance and discordance between pairs of observations (Czado, 2019).

REFEREE. None of the references is pertinent. Much better references are the two books by Nelsen (2006)—a more theoretical one—and Salvadori et al. (2007)—a more practical one.

Line(s) 159–ff.

AUTHOR(s). Graphical tools for analyzing dependence...

REFEREE. For intellectual honesty, the Authors should warn the reader that the interpretation of graphical results always involves a degree of subjectivity, and should always be accompanied by objective formal tests.

Line(s) 175–176

AUTHOR(s). They also found that the strong bias and associated uncertainty raise doubts about the reliability of most empirical results reported in the hydrological literature.

REFEREE. See also Illustration 3.18 in Salvadori et al. (2007, p. 173), where numerical experiments were carried out both on empirical and simulated data.

Line(s) 242

AUTHOR(s). The literature proposes various alternatives for combining multivariate analysis and non-stationarity.

REFEREE. For intellectual honesty, the Authors should point out that, at present, non-stationarity is generally modeled by introducing a temporal dependence of the parameters of the marginals/copulas at play (e.g., by assuming linear and/or exponential changes of the parameters with time), but these remain mere mathematical exercises, not tested on empirical data.

Line(s) 292–294

AUTHOR(s). This test evaluates the null hypothesis that the empirical copula comes from a specific copula; if the null hypothesis is rejected, the empirical copula does not follow the distribution of the specified copula.

REFEREE. Statistically speaking, this sentence is not correct: Statistics can only offer guidance and suggestions, but never absolute truths. The words "copula does not" should be written as "copula may not".

Line(s) 422–ff. (3.6 Compound design events)

REFEREE. The Authors are quite confused about the difference between a density and a probability distribution function: they are not the same, and they have different meanings. For instance, in the cited paper by Salvadori et al. (HESS 2011), the two different strategies proposed were based either on a probabilistic base (the Component-wise Excess one) or on a likelihood/density base (the Most Likely one). The description and the explanation given by the Authors is a superficial and confusing one, especially for practitioners.

Here and elsewhere, use "most likely" instead of "most probable": in Probability Theory, a density induces a likelihood, not a probability (which, instead, is induced by a distribution function, i.e. the integral of a density). Line 427 refers to the Most Likely strategy outlined in the cited paper by Salvadori et al. (HESS 2011). Line 428 refers to the usage of "ensembles", as suggested in Salvadori et al. (HESS 2011).

Eq. (9) is a special case of Eq. (13) in Salvadori et al. (HESS 2011), using the density f as a weight function w. Note that $f_{XY\ldots W}$ is the joint density of the distribution $F_{XY\ldots W}$, not of the copula $C_{XY\ldots W}$.

For intellectual honesty, the Authors should make it clear, and warn the reader, that there is no guarantee that the maximum found by a numerical routine will be a global maximum, rather it is very likely that it will be a local maximum, and this will be more and more likely as the dimension of the problem increases.

Line(s) 489–ff. (Discussion, Conclusions)

REFEREE. This is not a Discussion, it is at most a replica of generic sentences already written in previous parts of the manuscript. Actually, in my opinion, in this paper there little to discuss about.

The Conclusions are a collection of statements that try to justify a paper with no content.