REVIEW REPORT

Journal: HESS
Paper: HESS-2019-358
Title: Time-varying copula and design life level-based nonstationary risk analysis of extreme rainfall events
Author(s): Pengcheng Xu, Dong Wang, Vijay P. Singh, Yuankun Wang, Jichun Wu, Huayu Lu, Lachun Wang, Jiufu Liu, Jianyun Zhang

GENERAL COMMENTS.

In my opinion, essentially this paper only adds "noise" to the existing Literature: the techniques used have already been published in other works, the only novelty (clearly, not a methodological one) could be the case study, but any new case study must represent a newness over previous ones (otherwise it would be a replica). Most importantly, the work is in general statistically weak, and affected and flawed by fatal errors: the conclusions of the Authors may not be supported by the analyses they carried out. Apparently, the Authors (incorrectly) interpret the results according to their convenience, in order to prove what they want to prove, as shown below. In addition, referencing is often imprecise and/or improper and/or missing: always give credits to whom deserve credits. My recommendation is: REJECTION.

SPECIFIC COMMENTS.

Line(s) 49–54.

Authors. Copulas, a useful tool for modelling the structure of dependence between hydrological variables regardless of the types of marginal distributions, have been widely used for multivariate frequency analysis + references…

Referee. Historically, the paper by Salvadori and De Michele (2004) was the first one to deal with (copula) multivariate frequency analysis—later works are copies or small variants: this paper is not cited. Please, always give credits to whom deserve credits.

Line(s) 75–ff.

Authors. There are three kinds of joint return period methods…

Referee. NO. In Literature there are, at least, four kinds of joint return periods. The references given are incorrect. In Salvadori and De Michele (2004) the OR, AND and Kendall cases were first introduced. In Salvadori et al. (2013) a further survival-Kendall approach (not mentioned by the Authors) was outlined. Referencing is often imprecise, almost random: for instance, why citing Jiang et al. (2015) here? It has nothing to do with the original formalization of the four return periods mentioned above. Incidentally, the reference "Salvadori and Michele, 2010" is "Salvadori and De Michele, 2010" (it seems that the Authors wrote the references by hand, instead of using some suitable software…)

Line(s) 95–97.

Authors. Note that following the idea of Rootzén and Katz (2013) we regard the term hydrological risk as the possibility of a certain extreme event occurring and not as a quantification of expected losses.

Referee. Then, probabilistically and statistically speaking (and hydrologically as well!), you should better use the term "hazard" instead of "risk".

Line(s) 126.

Authors. Detailed information about copulas can be found in Nelson (2007).

Referee. NO. It is Nelsen (2006), not Nelson. For an engineering approach, you may also cite Salvadori et al. (2007). As a strong suggestion, the Authors should carefully check the correctness of all the references (it is easy to do it on the Internet), and add the missing ones.

Line(s) 138–139.

Authors. ... θ_C^t is the dynamic copula parameter which is a linear function of time.

Referee. The Authors must justify this choice. Please do not reply that "the model was taken from this or that paper": it is not a scientific reason, for a model must be validated on the available data. Also, the results of suitable Goodness-of-Fit statistical tests must be shown.

Line(s) 143–144.

Authors. It is however possible that the nonstationary behavior may exist in both the marginal and joint distribution function.

Referee. Such an issue was already clearly pointed out and discussed in Salvadori et al. (2018), where a similar case study was investigated, and a thorough statistical analysis was carried out. The Authors must mention this fact, and follow the (proper statistical) guidelines outlined in that paper.

Line(s) 158-Figure 1.

The flow-chart shown in Figure 1 provides wrong indications (see also later comments). In fact, the Authors confuse GoF tests with selection criteria. The flow-chart must be rewritten.

Line(s) 161–164.

Authors. In this part, the Generalized Extreme Value (GEV) distribution was used to... (Cheng and AghaKouchak, 2014).

Referee. This reference makes little sense: the features of the GEV have already been stated and described since decades in other (seminal) works. Please use proper references.

Line(s) 166–ff.

Authors. The GEV distribution consists of three control parameters...

Referee. The GEV distribution is well known to hydrologists, there is no need to tell again a story that everybody knows.

Line(s) 176–179.

Authors. In this study, two kinds of nonstationary GEV models (GEVns-1 and GEVns-2) are developed with the shape parameter being constant. It should be emphasized that modelling the time variance in shape parameter needs long-term observations, which are often not available in practice (Cheng et al., 2014).

Referee. I recently rejected a paper very similar to the present one, where the GEV shape parameter was kept constant. The shape parameter is the most important one, for it rules the generation of extremes. The assumption adopted is definitely questionable: what (extreme) climate change could you really hope to model with a constant shape parameter? Practically, you are trying to model climate changes where the statistics of the extremes do not change with time: it makes little sense.

In addition (see also later comments), some estimates of the GEV shape parameter are positive and other negative (Table 3). This entails that, in some cases, the corresponding GEV law is upper-bounded, i.e. unable to model an extreme behavior: this is a well known feature of the GEV. I agree that the GEV is the right distribution to be used in your analysis (Block Maxima), but the question is: how can you claim that the phenomenon you are modeling is an extreme one when upper-bounded GEV's are involved? The statistical results seem to tell another story…

Line(s) 184–186, Eq.s (3)–(4).

You must justify the assumptions/relations implicit in these equations. Why should the position and scale parameters change according to Eq.s (3)–(4)? Did you carry out any valuable/reliable fit? What are the p-Values? And, again, why should the shape parameter be constant instead? Incidentally, these are the same equations used in the paper I recently rejected…

Line(s) 191–194.

Authors. Simultaneously, the Deviance Information Criterion (DIC) and Bayes factors (BF) for different stationary and nonstationary models were calculated to select the best fitted marginal model. The minimum DIC value yielded the best performance, while BF smaller than 1 indicated the best fitting.

Referee. This is a typical fatal error of practitioners. These are only selection criteria, not Goodness-of-Fit tests. You must first use (non-stationary) GoF tests to check whether a model is admissible! Otherwise, without first checking the models via suitable GoF tests, you may end up choosing non-admissible ones. This work has no statistical bases.

Line(s) 191–194.

Authors. In multivariate hydrological frequency analysis, two kinds of copulas, named elliptical and Archimedean copulas are widely used in hydrological applications.

Referee. So what? The fact that these copulas were used in other works is not, and cannot be, a scientific justification. This is the usual approach of practitioners that use the copulas provided by Matlab. Given my experience, I do not really think that Nature (especially considering the generation of Extremes) gets stick to just these dependence structures—see also later comments. And, worst of all, you did not even check these copula models via suitable multivariate GoF tests (which are available in Literature, and some certified software is even for free—see below): this work has no statistical bases.

Line(s) 203–205.

Authors. The Gaussian copula was not used in this study because of its deficiency in describing dependencies of extremes (Renard and Lang, 2007).

Referee. The Authors are clearly considering the concept of Tail Dependence. Well, also the Frank family has no tail dependence, while the Clayton family only has lower tail dependence (possibly, of no interest here), the Gumbel family only has upper tail dependence, and the Student family has both lower and upper tail dependence (but they must be equal, and, most of all, they both must exist at the same time!). There are more suitable families of copulas for modeling extremes: again, the ones used by the Authors are simply those provided by Matlab, as (unfortunately, too) many practitioners do, preventing a reliable/valuable investigation and modeling of the phenomenon of interest.

Line(s) 208, Eq. (5).

Again, as above, you must justify the assumptions/relations shown in this equation. Why should the copula parameter change according to Eq. (5)? Did you carry out any investigation? What are the p-Values?

Line(s) 212–214.

Authors. The Corrected Akaike Information Criterion (AICc; Hurvich and Tsai, 1989) was employed to make a goodness-of-fit…

Referee. NO. This a typical fatal error of practitioners. The AIC (corrected or not) is only a selection criterion, not a GoF procedure. You must first show that a copula is statistically admissible, e.g. via suitable Monte Carlo Cramer-von Mises or Kolmogorov-Smirnov tests, as in the R package "copula". Then, and only then, you may compare (only) the admissible copulas (if any) and select the "best" one according to some suitable criterion (e.g., the AICc, the BIC, the NLL, etc…).

Line(s) 214–216.

Authors. Obviously, the presence of nonstationarity in the copula parameter was determined by comparison of the AICc value.

Referee. This sentence is obscure. Are you saying that, since the non-stationary model performs better, then the phenomenon is non-stationary? If so, this makes no statistical and philosophical sense. It looks like you are using your models to "decide" how the real world should work: this is contrary to every scientific principle. This work is also bugged from an epistemological perspective.

Line(s) 217–ff., Sec. 2.3.

Authors. "2.3. Joint return period and risk analysis based on KEN's and AND's methods"

Referee. Multivariate failure probabilities have been well mathematically formalized in Salvadori et al. (2016), by originally defining and exploiting suitable Hazard Scenarios and copulas' relations. The Authors must take this work into serious account, and mention it.

Line(s) 241, Eq. (7).

See the more general approach and discussion in (Salvadori et al., 2016, Eq.s (33)-(35)).

Line(s) 262, Eq. (11).

Why in Eq. (11) the parameters of the marginals F_X, F_Y, used as arguments in the copula C, do not vary with time?

Line(s) 268–269.

Authors. The most likely event at the T_0-year level can be calculated as (Graler et al., 2013)…

Referee. NO. The Most Likely technique was first introduced in Salvadori et al. (2011): always give credits to whom deserve credits. In addition, it is not the only possible one, as shown in the same paper (viz., the Component-wise Excess method). Moreover, further approaches are outlined in Corbella and Stretch (2012) and Salvadori et al. (2014). Why was the Most Likely approach chosen in this work?

Line(s) 289, Eq. (17).

In Eq. (17), why is the modulus used? Obviously ΔR will always be positive. And even in this latter case, there is no quantification of any "scale" on which ΔR should be evaluated (when is it large? when is it small?). Such a number tells nothing to me.

Line(s) 330–331.

Authors. As shown in Figure 3, concurrences of univariate and bivariate trends, the nonstationarities in rainfall extremes can be detected at several stations...

Referee. This is simply because you use a 10% critical α-level, entailing a large probability of rejecting the Null Hypothesis of non-stationarity. For instance, at a standard 5% level, no one of the Univariate and Multivariate MK tests would fail, only two (at most three) out of 12 of the Univariate Pettitt tests would fail, and only one out of 6 of the Multivariate Pettitt tests would fail. In turn, the conclusions of the Authors are definitely questionable: in my opinion, in general, there is no clear statistical evidence of non-stationarity (not to say if the standard 1% level were used, for in this case stationarity would be fully supported). Apparently, the Authors manipulate statistics according to their convenience, in order to show what they want to show.

Line(s) 353–354.

Authors. The location parameter (μ) and scale parameter (σ) are regarded as time variant, while the shape parameter κ is time invariant...

Referee. As above, it is a dream to try and model time-variation of extremes using a constant shape parameter: it is the only one that matters in these kind of analyses. In addition, why should the other parameters vary according to Eq.s (3)–(5)? Simply because the same relations were used in other papers (again, without justification)? This paper has no scientific objective grounds.

Line(s) 356–358.

Authors. Despite the exception of Im for station 4, the shape parameter κ for most fitted models was in the interval of [-0.3,0.3]...

Referee. Tables 3 provide little statistical information, for no suitable confidence intervals are shown: this may have considerable consequences regarding the conclusions drawn by the Authors in later sections. In fact, they did not carry out any Monte Carlo analysis, and hence their results do not take into account the estimates' uncertainties (as if the Authors were stating the absolute Truth). To be clear, no confidence bands are plotted in later figures. This is not a scientific way of proceeding: the Authors must provide plots such as the ones shown in Salvadori et al. (2018), which may give an idea of the uncertainties at play (which may be huge, especially when a GEV is used, and may completely change the interpretation of the results, as I suspect).

In addition, as above, some of the fitted values of the shape parameter would imply that the corresponding GEV is Upper Bounded, entailing that the corresponding variable cannot be an Extreme one. Furthermore, the fact that the range of the shape parameter is "in accordance with previous studies" is not significant and relevant at all (also given the fact that the range is quite large).

Line(s) 359–360.

Authors. The best fitted model was selected by performing the minimum DIC criterion combined with the Bayes factor (BF) test.

Referee. Again, you did not show that it is an admissible one! This work has no statistical bases.

Line(s) 380–382.

Authors. Table 4(a)-(b) illustrates the results of best fitted copula, based on the minimum AICc and maximum logllikelihood value (LL).

Referee. Again, AIC and LL are not GoF criteria: the chosen models can be non-admissible! This work has no statistical bases.

Line(s) 433–435.

Authors. Although the copula model for station 5 was stationary, it was regarded as a nonstationary model because of the marginal nonstationary GEVns-2 model for Ps or Im, which existed at other stations.

Referee. This makes no sense. The Authors do not understand the basic fact that the dependence structure is independent of the marginals (as stated by Sklar's representation Theorem): even if the marginals are non-stationary, the copula may be stationary. The introduction of non-stationary copulas is arbitrary, without any justification: you cannot manipulate the results in this way!

Line(s) 440–ff.

Authors. Figure 5 shows isolines of Kendall return period and AND-based return period. . .

Referee. Given the uncertainties mentioned above (not considered by the Authors), I strongly suspect that the interpretation of the results shown in Figure 5 could be quite different if suitable confidence bands were plotted. This work lacks of elementary statistical bases.

Line(s) 537–ff., Sec. 4.6.

In the light of the objections given above, the "Further discussion" section (4.6) makes no sense.

References

Corbella, S., Stretch, D. D., 2012. Multivariate return periods of sea storms for coastal erosion risk assessment. Nat. Hazards Earth Syst. Sci. 12, 2699–2708.

Salvadori, G., De Michele, C., 2004. Frequency analysis via copulas: theoretical aspects and applications to hydrological events. Water Resour. Res. 40, W12511, doi: 10.1029/2004WR003133.

Salvadori, G., De Michele, C., Durante, F., 2011. On the return period and design in a multivariate framework. Hydrol. Earth Syst. Sci. 15, 3293—3305.

Salvadori, G., De Michele, C., Kottegoda, N., Rosso, R., 2007. Extremes in Nature. An approach using Copulas. Vol. 56 of Water Science and Technology Library Series. Springer, Dordrecht, ISBN: 978-1-4020-4415-1.

Salvadori, G., Durante, F., De Michele, C., 2013. Multivariate return period calculation via survival functions. Water Resour. Res. 49, 2308–2311, doi: 10.1002/wrcr.20204.

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., Petrella, L., 2016. A multivariate Copula-based framework for dealing with Hazard Scenarios and Failure Probabilities. Water Resources Research 52 (5), 3701–3721, doi: 10.1002/2015WR017225.

Salvadori, G., Durante, F., Michele, C. D., Bernardi, M., 2018. Hazard assessment under multivariate distributional change-points: Guidelines and a flood case study. Water 10 (6), 751–765.

Salvadori, G., Tomasicchio, G. R., D'Alessandro, F., 2014. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. Coastal Engineering 88, 1–14, doi: 10.1016/j.coastaleng.2014.01.011.