

## ***Interactive comment on “Evaluating the efficacy of bivariate extreme modelling approaches for multi-hazard scenarios” by Aloïs Tilloy et al.***

**Aloïs Tilloy et al.**

alois.tilloy@kcl.ac.uk

Received and published: 10 June 2020

[R1 Introduction] The paper assesses the capability of six bivariate statistical approaches to model 60 distinct synthetic (bivariate) datasets, some of which possess asymptotic independence and others asymptotic independence. The results are used to develop a systematic framework for selecting among the competing statistical models. The framework is then demonstrated by way of two real world examples. The framework offers a novel approach for selecting among multivariate models. The manuscript is generally well written and relevant to the topic of “Advances in extreme value analysis and application to natural hazards”. In the opinion of this reviewer it is therefore worthy of publication in this special issue of Natural Hazards and Earth System Science. Nevertheless, the manuscript would benefit from a slight reorganization,

C1

additional explanation at certain points and a thorough review of grammar and spelling.

[AR to General Comments] We thank the Anonymous Reviewer 1 (R1) for this summary, and suggestion of reorganization, additional explanation and thorough review of grammar and spelling. Based on these comments, we have revised the manuscript, particularly restructuring parts of our manuscript. Below, we provide detailed replies to the reviewer 1 comments (Author Replies – AR).

[R1 Specific comments]

[R1 Comment 1] The abstract is missing a sentence describing the link between the synthetic datasets part and the application of the framework to environmental data (i.e., the work on the synthetic datasets aids in the creation of the framework). At present, the abstract states that the benefit of a systematic modeling framework are highlighted without any introduction/description of the framework.

[AR 1] We agree with the reviewer about the fact that a linking sentence is missing in the abstract. We therefore modified the abstract to define with more clarity the framework and its purpose (L10-11) and (L12-13). [R1 Comment 2] The manuscript would benefit from a more precise definition of what constitutes a hazard in relation to other recent literature on compound and cascading hazards. For instance, Zscheischler et al. (2018) defines compound events as “The combination of multiple drivers and/or hazards that contributes to societal or environmental impacts.”. Therefore, they may consider rain, lightening and hail as drivers and a landslide as a hazard, whereas here all four are considered as hazards. [AR 2] We thank the reviewer for this comment as we agree on the importance of defining what constitutes a hazard. We recognize that there are different definitions. Here, the term hazard will follow the definition by UNISDR (2009), which refers to a natural hazard (hereafter referred to as a ‘hazard’) as a natural process or phenomenon that may have negative impacts on society. As hail, lightning and extreme rainfall fall in the category of “hazards”. A sentence has therefore been added to the introduction.

C2

[R1 Comment 3] P2 L49 to P3 L83 in the introduction focuses on methodology and introduces specific subsections of the methodology section before the methodology section is introduced in the final (roadmap) paragraph at the end of the introduction. I recommend moving the text to the start of the methodology section. Perhaps, a paragraph giving a broad summary of the synthetic dataset work and the modelling framework including the link between them could be added.

[AR 3] We agree with the reviewer on this. Parts of the mentioned paragraph has therefore been moved. Paragraph P2 L52-59 has been moved to the start of the methodology section (P4 L101-111). The sentences P3 L 80-83 is now incorporated in the “roadmap” paragraph (P3 L88-190).

[R1 Comment 4] For the wildfire example, would  $\chi=0.7$  with marginals AC also be a relevant test case.

[AR 4] Thank you for this comment. It also raises the question of the selection of analogous datasets, which is currently quite subjective. To tackle this issue, we decided to add a small tool to help the selection. The tool is based on the empirical estimates of the 2 dependence measures ( $\chi$  and  $\eta$ ) and their uncertainty bounds. It therefore suggests analogous which have a dependence measure within the uncertainty bound. We believe this tool strengthen the whole methodology. Using this tool,  $\chi=0.7$  with AC marginal is not a relevant test case. We also reviewed our previous choices to be in accordance with the selection tool: we added the analogous AB and AC with  $\chi=0.3$  for wildfires and  $\chi=0.1$  AB and BB.

[R1 Comment 5] The (subjective) selection of the AND should also be discussed in the “Choices influencing the results of the simulation study” part of the conclusion.

[AR 5] Very good point. We now state the potential influence of the selection of the “AND” probability in the conclusion (P32 L780-782).

[Reviewer 1 Technical corrections] [AR to R1 Technical corrections] We thank the Re-

C3

viewer 1 for these detailed technical corrections. We have below replied to each one and taken on board the vast majority of them.

Abstract: “match” implies that synthetic data was based on specific examples of environment data “be representative of” maybe a more appropriate expression. We totally agree, thanks, we have now made the change. P2 L46: “Copulas” refers to a specific type of model whereas “multivariate model” is more general, perhaps replace the latter with a more specific description of the models. We replace “multivariate models” by “Multivariate extreme models”. We did not find a more specific appellation.

P4 L94-95: The sentence “A theoretical background on extreme value theory is given in Supplement S1.1.” should appear earlier in the paragraph, after the sentence which starts “Extreme Value Theory : : :”. Change done

P4 L106: Typo. Remove “Then”. Change done

P4 L116: Typo. Remove “.”. Change done

P6 L166: Typo. Replace “F\_X2 (x\_2)” with “F\_X2 (x\_2)”. Change done

P6 L170: Use “e.g.,” here and elsewhere before citations where there are other relevant papers omitted due to the need for brevity. Change done.

P7 L175: “Formally, the application of a copula model can be summarized in four main steps”. It should be made clear here and elsewhere that “the application” refers to the application in this study and may vary elsewhere for instance in terms of marginal distributions etc. Change done

P7 L181: “joint distribution” may refer to the full multivariate distribution including marginals. Consider changing to “dependence structure” or similar. Change done

P7 L210: “extrapolate the conditional model to simulate new extreme data.” Do you do this? Not exactly, that sentence was misleading and not well written, it has been rephrased to improve clarity.

C4

P8 L218: Subjective term. Remove “very”. Change done

P11 L186: Perhaps change “amount” to “number” to be more specific. Change done

P13 L323: “Confront” seems like an unusual term to use here. Consider replacing it with “compare and contrast” or simply remove “confront or”. Change done

Table 2: Spelling. Change “thorough” to “throughout”. P13 L334: Typo. “Random variables x”. Consider capitalizing X and Y as they denote random variables. Change done

P14 L338: Why characterize log-normal distributors by the coefficient of variation? This has been the topic of a long debate between authors. The idea behind using the coefficient of variation is to (i) relate to previous literature working with log-normal distribution. (ii) characterise the log normal distributions used with one single parameter instead of two. We clarified this by adding a sentence composed of the two aforementioned points (P14 L370-371). Table 3: Typo. Replace “dataset” with “datasets”. Change done

P16 L379 Typo. Add space “test(”. Change done

P16 L378 & L379 & L392: Typo. Remove first names “(Arnold, Taylor and Emerson, John, 2011)”. Change done

P16 L389: Grammar. “issue is” or “issues are”. Change done

P16 L391-393: “The measures mentioned above are not suitable as they imply parametric distributions to be compared against observations (Stephens, 1970; Arnold, Taylor and Emerson, John, 2011). It is then not possible to compare the goodness-of-fit on the whole range of the data.” I am not clear as to the exact limitation being discussed here. The main limitation being discussed is the need for models to be parametric (e.g., copula) to be compared with the discussed method. As we worked in this study with parametric and non(semi) parametric models, it is not possible to use the previously discussed Goodness of Fit methods. We changed the text to make it clearer.

C5

P16 L399: A description of the reference level curve is required that it comes from the “underlying bivariate ( $X_1, X_2$ ) distribution of the data” as stated in Figure 6 should be added to the main body of text. Thank you, Change done

P19 L461: Consider adding “in general” or similar before “outperform all the other: : :”. Change done

P19 L468-469: Grammar. Rephrase “Gumbel and Galambos copulas show very similar behaviours with respectively 68% [53–93%] and 68% [52–93%] cases with  $w_d < 0.1$ ”. Change done

P20 L473 Grammar. Change “less” to “least”. Change done

P20 L494 I believe “Fig. 6” should be “Fig. 5”. Change done

P20 L497 Grammar. Change “abilities of each models for the selected datasets in step (ii)” to “abilities of each model for the datasets selected in step (ii)”. Change done

P20 L510 & P21 L519: “London Heathrow airport, UK” and “Heathrow airport (LondonUK)” be consistent with names, only need to specify it is in the UK on first mention. Change done

P21 L515: Typo. Add spaces: “Season(Hawkes”. Change done

P22 L544-545: Repetition: Remove “Daily rainfall data from E-OBS (Cornes et al2018) and wind gust data (maximum 3 s wind 545 velocity in a day) from the Met Office (2019).” as the information is already given in the text. Change done

P23 L555-556: Repetition. Remove “Daily rainfall data from E-OBS (Cornes et al., 2018) and wind gust data (the 555 maximum 3 s wind velocity in a day) from the Met Office (2019).” as the information is already given in the text. Change done

P23 L560: I believe “Figs. 9 and 6” should read “Figs. 9 and 5”. Change done

P23 L565 - 568: Consider replacing circle bullet points with numbers 1-4 so the cases

C6

correspond to the relevant rows of Table 4. Good idea! Change done

P23 L573: Add “average” before “confidence score”. We are unclear here if it makes sense to describe the “confidence score” as an average and would require further clarification from the reviewer.

P24 L578: Grammar. Replace “model” with “models”. Change done

P24 L591: Grammar. Replace “curve” with “curves”. Change done

P25 L612 & P25 L612: “2395 km<sup>2</sup>” and “1,1778,146” be consistent with use of commas to separate numbers here and throughout the manuscript. Change done, Thank you

P27 L643 – 644: Repetition. Remove “Daily mean temperature data from E-OBS(Cornes et al., 2018) and wildfire data from Pereira et al. (2011).” as the information is already given in the text. Change done

P28 L653-654: Repetition. Remove “Daily mean temperature data from E-OBS (Cornes et al., 2018) and wildfire data from Pereira et al. (2011). as the information is already given in the text. Change done

P28 L660: I believe “Figs. 13 and 6” should read “Figs. 13 and 5”. Change done

P28 L666 - 669: Consider replacing circle bullet points with numbers 1-4 so the cases correspond to the relevant rows of Table 6. Change done

P28 L666 - 669: Replace “B-C” with “BC”. Change done

P30 L713: Typo. Replace “four following” with “following three”. Change done

P31 L742-744: Grammar: Change “One of these disadvantages is the that the parametric nature of copulas lead to a lack of flexibility while going to higher dimensionality.” to “One of these disadvantages is that the parametric nature of copula leads to a lack of flexibility when going to higher dimensionality.” or similar. Change done

P31 L747: Pair Copula Construction and vine copula are not synonymous rather vine

C7

copulas are a type of Pair Copula Construction. Thank you for this comment, we changed the sentence accordingly.

P32 L787: Typo. “\_Results” add “.” Change done

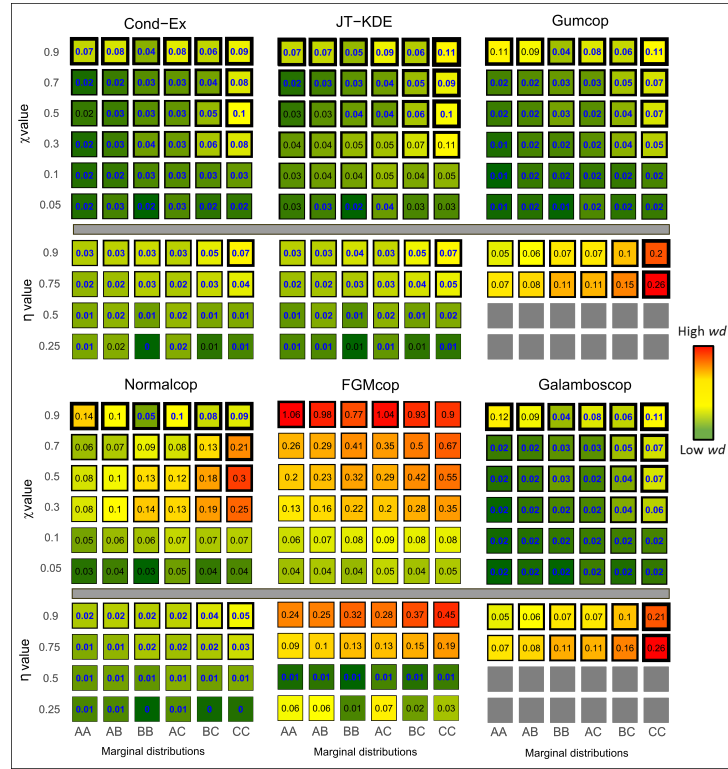
P32 L783: Perhaps add a sentence to make it clear that the RMSE comes from the calculation of the dependence measures for the 100 realizations of the 60 datasets. Change done

P36 L887: Remove “pp.” Change done

---

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2020-28>, 2020.

C8



**Fig. 1.** Weighted Normalized Euclidean Distance (wd) to the reference curve for all 60 different synthetic datasets. Fitting capacities of each model are represented. Values in cells and colours represent the