

## ***Interactive comment on “Uncertainty in flood frequency analysis of hydrodynamic model simulations” by Xudong Zhou et al.***

**Anonymous Referee #2**

Received and published: 14 September 2020

### General comments

This paper is based on large scale hydrologic-hydrodynamic simulations to investigate different sources of uncertainty in flood risk estimation, with the use of flood frequency analysis tools. The chosen topic deserves some interest, though the analysis is based on a specific configuration of a set of available hydrological model output (from the Earth2Observe project) and an in-house hydrodynamic model (CaMa). However, the focus on the global domain makes it of larger interest for a wider community.

- Among the main limitations of the manuscript is the sub-optimal use of the english language, including both terminology, grammar, typos and structure of the sentences, which makes it hard to read and at times hampers the understanding of the content. I strongly suggest to work and improve it with the help of a native speaker.

C1

- Another important comment is related to the general framing of the analysis. In the current version a number of analyses are performed, focusing on different aspects, though in my opinion it lacks a consistent storyline and some reasoning behind why they were made and clear statements about what we learn following their results.

- The manuscript is too long compared to the information content it brings. I suggest shortening following the comments below. A number of figures should be removed, improved or put in the supplement material, for the reasons I explain below in the specific comments. In particular, I'm speaking about Figures 4 and 5 wrt the issues with fitting analytical functions with different degrees of freedom (comment #10), Fig. 6 (comment #18), Fig. 10, 12, and 14 (comments #24, #27, #29)

### Specific comments

1. p2, l8-9: acronyms should be defined with “full name (acronym)”, e.g., Global Runoff Data Centre (GRDC). Same for p3, l5 and l26. 2. p2, l14: Pearson type III 3. p3, l1: suggested “connected” → “analyzed the relation between ...” 4. p3, l3-5: Sentence not clear. Please rephrase. 5. P4, l3: please define the acronym SAR 6. p4, l10: “various runoff inputs” is too general. Please add details here or a reference to the details included in Sect. 2.2 wrt the inputs used. 7. P4, l13: I suggest adding an introductory sentence here to give more details about the experiment itself, before jumping to the uncertainties to investigate. 8. P4, l14-16: please improve this part. Also, I find the variable names V1\_(rivdph) and V2\_(sto2dph) not very intuitive. Why not simply calling them depth and storage? Especially sto2dph creates confusion on whether it is a storage or a depth. 9. Table 1: I suggest removing “Various” in the caption. 10. P5, l12: Note that the Gumbel and the Gamma distributions have 2 parameters. In fact, results in Figure 5 seems to me the natural consequence of fitting a series of points with mathematical functions with different degrees of freedom, where the 5 parameter distribution is able to fit the data more skillfully (though it doesn't mean it will be more skillful in predictive mode for future floods), Then the 3 parameter distributions and the 2-parameter Gamma and Gumbel as the least skillful. One would obtain similar results

C2

when fitting the series of data with polynomials of grade 5,3 and 2, because higher grade polynomials can fit better the input data. 11. P5, l13: I suggest renaming this section (e.g., "Fitting performance" or similar) 12. p5, l15: calculated 13. p5, l19-20: This should be expressed more clearly. E.g.: "Smaller aic denote higher fitting performance" or similar, which is actually better written in p6, l23-24 14. p6, l24-26: Use active rather than passive form (e.g., "we compare") 15. p7, l6-7: Is the normalization the real reason? Also, I suggest giving more details on how to weigh the aic values. What is the optimum? What are normally considered good or bad values? It is not intuitive for those who have never used it. 16. P8, l8: "The later peak" -> "the latter" 17. Figure 3: Interesting to see how the pdfs of gamma and gumbel have similar peaks to the other distributions only for the storage, but not for the river depth. Indeed it is clearly visible also in Fig. 3c. Would be interesting to investigate and motivate the reasons. Now it is only mentioned but no justification is given. 18. Figure 6: How does this analysis relate to the FFA and to the rest of the paper in general? I'm not sure of the value of these maps, given the little information the readers have on the 7 runoff inputs, and also because there is no clear pattern identified. Perhaps the main information one can obtain is that anu and univu tends to be on the lower side, while cnrs and univk on the higher side. Yet, this doesn't say anything about the skills of these estimates, which would imply validation with gauge data at a number of stations. 19. P13, l2: after -> downstream 20. Note that Figure 8 is referenced before Figure 7 21. Sect 4.1 refers to return periods in Fig.7, hence in Fig.7 I advise to show return periods in place of frequencies. In any case, to be correct you should refer to those as annual frequencies of occurrence, to avoid confusion. Also, in Figure 7c, why not all distributions are shown? 22. P14, l2: please give some details and possibly a reference on the downscaling procedure. 23. P14, l3-6: To aid the assessment of water depths I suggest showing in Fig.8 a map or contour of the permanent water bodies. Clearly it is normal to have higher water depths in rivers and lakes, compared to areas normally dry. Also, I cannot find information about the terrain model, in particular whether it represents the river bed or some reference water level. This is important for this analysis.

C3

24. Figure 10: results shown in this figure are rather obvious. I suggest removing this figure as it brings little information. Over large inundation depths it is normal to have good agreement on whether there's inundation or not, as having poor agreement would mean huge differences in the results of the model used (hence very poor skills for some models). 25. P17, l11: return periods should not be expressed as percentage 26. p18, l10 and Figure 12: Is this the mean inundation of the 7 models? Clarify 27. I find the analysis in Figure 12 of limited use, being a qualitative visual comparison with two other publicly available maps, but also resulting from modeling exercise with limited calibration. Similarly, the comments in p19, l14-18 are partly speculative. More rigorous validation with observed flooded areas would give much more strength to the paper. 28. P 21,l6: for flood impact assessment it is more interesting to know (even smaller) inundation depths in areas where people live or where economic assets are, rather than the inundation in the main channels, which has fewer fields of application. 29. Figure 14 is unreadable and of limited use in the present form. It is impossible to get enough spatial details of a global inundation map at such small scales. Furthermore, the left and right column are almost indistinguishable. I suggest removing this figure and rather put it in the supplement, together with a number of inset panels zooming into some areas, especially those where the authors want to comment the results. 30. Figure 15: What do you mean by the third (and fourth) row and the second row, in the caption? Is it related to the rows of Figure 14? If so it should be clearly stated. 31. P23, l14-15: To be improved 32. p24, l16: this is a model result for just one point in the entire world, hence it is completely irrelevant. Even more when looking at figure 6. Also (see lines 20-22), being in the middle of the 7 outputs doesn't mean it is more skillful. Validation with observed data is recommended.

---

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2020-272>, 2020.

C4