Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2019-83-RC1, 2019

Natural Hazards and Earth System Sciences
Discussions
Open Access

Interactive comment on "Decreasing uncertainty in flood frequency analyses by including historic flood events in an efficient bootstrap approach" *by* Anouk Bomers et al.

Elena Volpi (Referee)

elena.volpi@uniroma3.it

Received and published: 17 April 2019

General comment

The manuscript investigates how the estimation uncertainty affecting flood frequency curves can be reduced by including historical flood events. The Authors analyses the case of the Dutch Rhine river delta, by using systematic information recorded at the site of interest, systematic information interpolated from close river sections and historical events reconstructed by using hydraulic models. The model is used to "normalize" the historical discharges for anthropogenic and natural changes occurred in the river

Printer-friendly version

Discussion paper

system. The study case is o interest for the large amount of hydrologic/hydraulic information and for sure offers the possibility of interesting discussions.

The manuscript is well written and organized; the methodology is almost well described, even if additional details could be included to help for reader understanding. Summarizing, the topic is of interest for the scientific community and the manuscript could be eventually considered for publication in this Journal after some issue are addressed by the Authors. My comments about the work are listed in the following paragraph; I hope that they will be helpful for manuscript improvement.

With appreciation,

Elena Volpi

Specific comments

1. The Authors introduce the bootstrap approach (l. 5-9 p. 3) as a solution to overcome the problem of isolated historical events for which confidence intervals are typically not symmetrical. It is not clear what the Authors mean by symmetrical confidence intervals; this issue should be explained since it is the motivation (together with the easy application of FFA) for reconstructing a continuous data set. Further, bootstrap is not necessary for confidence interval estimation (l. 9-10 p. 3) yet still necessary for continuous data set reconstruction.

2. The hydraulic model is used to propagate the discharge for the historic flood events reconstructed by Meurs (2006) from Cologne to Lobith; to this aim the Authors state that they use the current geometry of the riverbed and floodplain in order to correct the historic floods for anthropogenic interventions and natural changes of the river system, which is referred as “normalization” in the manuscript (l. 10-14 p.3). This approach is unusual based on my experience

(Calenda et al., 2005); historical flood events should be simulated by reconstructing the historical conditions (the river geometry as in the period the flood occur), that is what Authors would have available if measures would have started in the ancient past. In essence, I am not convinced that propagating the ancient floods in the current riverbed is the correct approach to solve the "homogenization" problem; conversely, this "gives insight in the consequences of an event with the same characteristics of a historic flood event translated to present times" (as stated by the Authors themselves at l. 17-18, p. 3).

3. Based on my opinion the Authors should "naturalize" the estimated discharge, by computing the discharge that they would have observed in absence of some anthropogenic change in the riverbed or in the catchment (l. 14-16 p.3). This means that are the recent events that should be reported to pre-dike conditions and not the opposite (as done in Section 2.3.2). The presence of the dike artificially alters the natural regime of the extreme flood events; the anthropogenic alteration of flood regime should be of deterministic nature, even if its estimation is characterized by a certain degree of uncertainty.

4. Why do the normalized events almost always lead to a higher discharge than the historic event (l. 16-17, p. 3)?

5. Section 2. For the sake of clarity, a table summarizing the type of information and the related uncertainty for the different time periods should be included.

6. L. 14-15, p. 4. The Authors should clarify the distance and the characteristics of the nearby gauging locations.

7. The procedure discussed in Section 3 is based on a non-parametric approach; alternatively a parametric method, based on the same assumption that ancient flood events follow the same statistical behavior of those systematically recorded, could have been considered. See Stedinger and Cohn (1986) and Francés

(1998); for an application to a case study similar to that discussed in this paper see also Calenda et al. (2009). Which is the advantage of the approach used in this work?

8. L. 2-7, p. 9. The Authors states that “the available goodness-of-fit tests for selecting an appropriate distribution function are often inconclusive. Those tests are more appropriate for the central part of the distribution than for the tail (Chbab et al., 2006), where we are interested in since the tail determines the investments required for future flood protection measures.” I agree with the Authors that goodness-of-fit tests might be inconclusive, as discussed deeply in Serinaldi et al. (2018); on the other hand they provide a first indication on which models, among several competing ones, could be excluded due to the poor performance (see, e.g., Laio, 2004). In such a sense, I suggest the Authors at least to rephrase the sentence, also because there are different goodness-of-fit test which focus on the statistical behavior of the tails, such as the Anderson-Darling test and the Modified Anderson-darling test (Laio, 2004).

9. Following the argument of previous comment, I do not believe that restricting the analysis to a single probability distribution model (although it is the Generalized Extreme Value distribution commonly used in literature to perform an FFA) is a good choice. Since the interest is in evaluating how the confidence bounds of extreme quantile estimates reduce when adding the historical information (l. 18-21 p. 9), it should be considered that confidence bounds depend not only on the length and information content of the dataset but also on the probability model itself. Hence, results could be different if a different model is taken into account.

10. L. 10-12 p. 9. Do you the Authors mean that they assume an upper bounded distribution? This issue should be clarified.

11. Figure 5 is unnecessary, It could be removed.

12. Figure 6. The largest extreme events are not included in the uncertainty bounds. The corresponding sample bounds could be included as well to text the model performance (see comment 9).

13. Section 5.2. I am not sure I fully understood the rationale and the approach behind the analysis performed here. The historical events are some of the highest events observed in the whole observation period. If a sample is reconstructed by simply resampling the events observed in 1901-2018 (without including the largest historical events but with the same length of that used in previous sections), the largest events might only be those observed in the more recent period; as a consequence, the fitted model is expected to be characterized by, e.g., a smaller variance, which implies narrower uncertainty bounds. I do not see this behavior in figure 7 (upper panel). What I see in figure 7 is that the fitted model in the two cases is almost the same, while the uncertainty bounds are significantly different. I can explain this only if the reconstructed samples have a very different length. Please provide a deeper explanation.

14. L. 20-22 p. 14. It is not clear how the extended data set with normalized reconstructed discharges can capture the long-term climatic variability (see also previous comments).

15. L. 35, p. 14. Isn't it the 1374 event?

16. Fig. 8. Adding one event equal to the largest one over a record is expected to affect somewhat the estimated model if the record is 100 years while non changes in the model are expected if the record is about 700 years. Hence, which is the lesson learned from this analysis?

17. Within the Conclusion Section a detailed list of the limitations of the approach proposed here should be provided.

References

- Calenda, G., Mancini, C. P., Volpi, E. (2005). Distribution of the extreme peak floods of the Tiber River from the XV century. Advances in Water Resources, 28(6), 615-625.
- Calenda, G., Mancini, C. P., Volpi, E. (2009). Selection of the probabilistic model of extreme floods: The case of the River Tiber in Rome. Journal of Hydrology, 371(1-4), 1-11.
- Francés, F., 1998. Using the TCEV distribution function with systematic and non-systematic data in a regional flood frequency analysis. Stochastic Hydrology and Hydraulics 12, 267–283.
- Laio, F., 2004. Cramer-von Mises and Anderson–Darling goodness of fit tests for extreme value distributions with unknown parameters. Water Resources Research 40, WR09308.
- Serinaldi, F., Kilsby, C. G., Lombardo, F. (2018). Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Advances in Water Resources, 111, 132-155.
- Stedinger, J.R., Cohn, T.A., 1986. Flood frequency analysis with historical and paleoflood information. Water Resources Research 22 (5), 785–793.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-83, 2019.