Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2019-83-RC3, 2019

Natural Hazards and Earth System Sciences
Discussions

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

Anonymous Referee #3

Received and published: 21 May 2019

In this paper, the authors present a method/case study to reconstruct a continuous times series of annual maximum discharges in order to estimate return times for flood discharges for the Rhine at Lobith. The study uses modern data from 1901 onwards, discharges reconstructed from water level measurements back to 1772 and information from historical flood events back to the 1300. Extending a time series with this information leads to a reduction of uncertainty and to more stable return times. The paper is well structured and written, and the topic is of relevance for flood risk estimation. However, there are some major issues which need to be addressed before this paper can be published from my point of view:

One general problem I have with this manuscript is that the authors refer to and use data from many other studies, especially the one from Toonen (2015). It is difficult to follow the article for reader if one is not familiar with these studies because it requires reading many secondary sources to gain insight on how all the different data(-sets) were collected and obtained, e.g. how was the regression analysis by Toonen (2015) performed, how were the historical floods in Cologne by Herget and Meurs (2010) reconstructed, etc.

This paper includes a lot of different data sets (systematic, historical, plus various bootstrapped time series), it would be beneficial for readers to include a table with a short description and overview of the properties of these data sets and to name them consistently throughout the paper.

The term "normalize" is used in different contexts (e.g. for historical floods, for the 1900-2008 data set, for the data set of Toonen (2015) which is not normalized but used as normalized data). I find this confusing since it does not become clear what is actually meant by this and what has been done to "normalize" each of these data sets. A more thorough explanation on this matter would be useful.

In section 2.2 the authors describe the Toonen (2015) data set which uses a linear regression to compute water levels at Lobith. This method leads to a reduced variance of this data set (c.f. table 1). How would this affect the bootstrapping later on, if samples from the so called "systematic time period" with different variances (1772-1900, 1901-2018) are drawn?

From my point of view, the section 2.3.2 presenting the normalization of historical flood events leaves some open questions which need to be addressed. Using a coupled 1D/2D model to route the discharges from Cologne to Lobith seems a reasonable approach given the circumstances of the data, but the dike breach model and the underlying assumptions need more explanation. Is it valid to assume dike breach parameters from today's river geometry for historical times? Is there any historical evidence that there were dike breaches in the past, especially the 1374 event? Especially the reduction of the 1374 flood peak from Cologne to Lobith needs some sound justification/explanation. Why is this reduction only occurring for this specific event? Were there also dike breeches for the other historical events? What exactly is meant by "the upstream discharge shape is varied" (p.6, line 12)? There is a lot of uncertainty in this, which somehow contradicts the aim of the paper to reduce uncertainty. Furthermore, it would be interesting to know if any of 12 historical flood events where winter events, where ice draft/ice jams could/did play a role. Furthermore, assuming a normal distribution of uncertainties is valid for discharge measurements, but is this also the case for the estimation of historical extreme floods? Or is any discharge values in the uncertainty range equally possible? The reconstruction of the events in Cologne is based on the Manning equation and the uncertainty range results from different roughness coefficients. But do all of these follow a normal distribution?

Section 3: The bootstrap method to create continuous times series is a reasonable approach, however it would also be possible, to use the maximum likelihood method and incorporate the uncertainty range of the historical discharges as well as the discharges lower the perception threshold in the parameter estimation. From my point of view this approach is straight forward and should yield similar results. Could the authors explain/discuss the benefit of the bootstrapping approach?

In Section 4, the authors state that there are many distributions and fitting methods for flood frequency analysis and that the only use the GEV with maximum likelihood method. It seems justified, that only one combination is used to quantify the reduction of the uncertainty, but in practice there are many different distributions and parameter estimation methods - which again cause higher uncertainties in the estimation of return times, especially for the upper tail extremes. The authors should include a comment and if possible a quantification of this effect on this in the discussion. In Section 5.2., the authors argue that the reconstruction of historical flood events is complicated and time consuming and that this can be overcome by bootstrapping. However, the information

from rare and large historical flood events is still required as is stated at the end of the section. This sounds like an inconsistency in the line of argumentation. Furthermore, this whole section is somewhere between results and discussion. I suggest that the authors try to separate more clearly between results and discussion.

In the discussion, the effect of a hypothetical future extreme flood on the robustness of return times is addressed, which is somehow obvious from my point of view. This aspect does not add much value to the paper and can either be omitted or be moved to the results section.

Some specific comments:

Page 3, line 3f.: Why are uncertainties not symmetrical due to missing continuous data? Don't these result from the non-linearity of the rating curves?

Page 4, line 7f.: ACDP-measurements are in general not free of uncertainties, this assumption is not correct.

Page 11, line 2: Where does this confidence interval of 7400m^3/s come from?

Page 15, line 1: Same as above, modern discharge measurements are not free of measurement errors!

Page 15, line 5f.: See above, this is not a novel results and can more or less be expected. Furthermore, the statement that "flood managers can be less nervous" sounds awkward and is not really correct, since the uncertainty caused by different distributions/parameter estimation methods is not addressed.

Figure 2: Should be replaced by a "conventional" map, including national boundaries, a scale bar etc. Readers from outside of Europe might not be familiar with this region.

Table 1: The results of Toonen (2015) can be omitted in this table from my point of view.

Figure 6 and 7: The colours/line styles of the different curves are difficult to distinguish and should be changed to make these figures better to read.

References: To my knowledge, Meurs 2006 is a diploma thesis, not a PhD thesis.

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