Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2018-231-RC1, 2018
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Stochastic generation of spatially coherent river discharge peaks for large-scale, event-based flood risk assessment" *by* Dirk Diederen et al.

D. Paprotny (Referee)

d.paprotny@tudelft.nl

Received and published: 28 August 2018

The study which I had the opportunity to review aims at producing a synthetic series of discharge peaks that preserves spatial dependence across various locations in Europe. The manuscript is an important contribution to solving the problem of generating pan-European events using statistical methods. The methodology is explained well and the paper is pleasant to read. However, the paper has one major drawback and a number of smaller points related to how the research is presented.

The main issue is that the study abruptly ends just when it gets most interesting. Most of the paper is covered by methodology (which is understandable given the scope of

the paper), then followed by a section on validation and suddenly ends with a very short conclusion. Therefore, the reader is left with a number of questions, that would be normally addressed in a proper results and discussion sections:

1. How do the results look like?
2. What is the uncertainty in the results?
3. What is the sensitivity to the choice of thresholds for the analysis?
4. How the resulting dataset can be applied to a pan-European flood risk assessment?
5. What are the next research steps?

Ad. 1: it would very beneficial to presents and analyse the resulting dataset, as it is interesting to know how large are the pan-European events? What is the total peak discharge of an average event, annual maxima, or a 1 in 100 years event? How different is the discharge computed with the authors' methodology from discharge computed independently for each station? Is there spatial variation within Europe how much the discharge differs between the assumptions of dependency between locations and of complete independence?

Ad. 2: using copulas has the benefit that the uncertainty can be easily obtained. Can this information be used to show what is the uncertainty in the discharge during a synthetic pan-European flood with a large return period, e.g. 1 in 100 years? Or at least the uncertainty in discharge for individual locations?

Ad. 3: The paper mentions selecting parameters through "trial and error" no less than four times. That includes parameters of the noise reduction, time window of the pan-European events, threshold parameters of GPD fitting and multivariate analysis. How sensitive are the results to the choice of parameters? How much would the discharge at a given location change if one of the parameters is modified, or the total discharge of a pan-European event with a given return period?

Ad. 4: The paper states in the title that the resulting dataset could be used for "large-scale, event-based flood risk assessment". However, it is rather unclear how would it be applicable to such studies. The authors only calculate the discharge for 298 locations, which is around 0.1% of all EFAS grid points, or less than 1% of flood zone calculation subdomains used by Alfieri et al. for a pan-European flood hazard assessment. Therefore, what is the utility of such a small number of locations of Europe-wide assessments? Is it possible to scale it up? I guess that might be a problem, as more locations means much more computational burden for the multivariate analysis. Is the 298 locations the maximum feasible number, or was chosen for quick testing of the method and can be easily increased?

Ad. 5: related to previous points, it is unclear from the papers what are the limitations of the method (if any) and what should be done to improve it. Also, how can future climate change be added to analysis? How can the result be applied to e.g. computation of flood hazard zones, annual expected losses in Europe or pan-European losses with a given return period?

I really don't expect the authors to tackle fully all my questions, but a proper discussion section in the paper is needed to make the reader familiar with the method's limitations, uncertainty, sensitivity, applicability to European-scale FRA and future research outlook. The aspect of application to FRA and overall the paper's contribution should also be an important point in the conclusions, which at present read more like a part of an abstract. Also, the questions posed by Niall Quinn in his short comment are good points that should be taken into consideration in a discussion section. Where necessary, the authors also might want to move some paragraphs from the methodology to the discussion (e.g. P8L8-15) to avoid overlaps and make the methodological sections more on point.

Minor points:

1. The abstract needs to be rewritten, as presently it consists almost entirely of back-

ground and seems to serve as a introduction to the whole paper. It should be a balanced representation of the rationale, methods, results and conclusions. It also shouldn't contain any references. I would suggest then to move the two first paragraphs of the abstract into the paper's introduction (as the literature review in the abstract is relevant but not repeated in the main text), and use the text from the conclusions to make a new abstract.

2. In section 2, the data selection is not explained. It is written that "the network was reduced to the major streams and tributaries", but no information what threshold was applied here. Next, it is written that "we selected 298 representative locations within the network of major European rivers", but no selection criteria is provided, making it particularly unclear why those locations are "representative" and do not include some regions of Europe (as per Fig. 1). The extra information might be included as a supplement so that it won't slow down the pace of the paper.

3. Figures in the manuscript need various improvements. Fig. 1 misses both a legend and a scale, including explanation what variable is shown in the background and what is the data source. Fig. 2 has a rather uninformative description "Framework". Fig. 3 need indication of panels (a,b,c), as they are mentioned in the text, and a legend. The Y axis label should be "discharge" rather than "flow" and the units should conform to the journal's guidelines. In Fig. 4 and 5 full axis labels should be used. In Fig. 6, the data source is not stated (which station is it?), and I have doubts whether the axis labels are correct – please check. Also, it might make a better visual impression if the font sizes were uniform.

4. The authors use "large-scale" many times, including the title, in the sense of analysis covering a large area. However, in cartography "large-scale" has opposite meaning to the common one, and refers to maps covering a small area. Therefore, I would suggest to change "large-scale" to "continental-scale" to avoid the confusion.

5. P2L7: the authors mention a rather loosely-worded definition of a flood, which is

neither strict nor of much relevance as the paper deals with river discharge peaks regardless whether they cause a flood or not. I think it's best to omit the first sentence of the paragraph.

6. P3L31: citation missing.

7. P6L9: "River discharge waves generally propagate through the network in downstream direction". I guess it should be "always" instead of "generally".

8. In the description of HT04, it is unclear which cited paper gives detail of the fitting. The authors also might consider providing a supplement with details of the technical aspects of the multivariate analysis.

I am looking forward to the authors' revision of this important paper.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2018-231, 2018.