Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2018-231-AC1, 2018

Natural Hazards and Earth System Sciences
Open Access
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.

Dirk Diederen et al.

dirkdiederen@hotmail.com

Received and published: 16 November 2018

Summary We would like to start with thanking mr Paprotny for his thorough and constructive review and apologise for the slight delay in reply.

First, we will address his points one by one. Second, we will provide a list with proposed improvements. We would be happy to hear from the reviewer, after reading the comments to his points, if we have missed/neglected/misinterpreted any major points.

Referee (D. Paprotny), main 1. How do the results look like? Ad. 1: it would very beneficial to presents and analyse the resulting dataset, as it is interesting to know how large

Printer-friendly version

Discussion paper

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?

Response (Diederen et al.): This paper expands the methodology of multi-site analyses to large-scale. As usual in multi-site analyses, the (tail-end) General Pareto Distributions (GPDs) connect magnitudes to probabilities (and therefore to return periods), which are captured locally (per location). These distributions are not the main focus of this study. We would like to point to p12,l9: "After the simulation, we transformed the margins of the synthetic 10 descriptor matrix to respect the fitted GPDs, thereby slightly distorting the dependence structure." This implies that, with this methodology, the discharge peaks computed with the dependence model are exactly the same as when the discharge peaks would be calculated independently (directly drawn from each individual distribution). In short, magnitudes and probabilities depend completely on the fits of the local GPDs, for which we applied a very simple methodology (Maximum likelihood, fitted to data above a fixed quantile), which we think is not too interesting to present. The main result in this paper is the capturing (Sect.3) and modelling/reproducing (Sect.4) of the spatial dependence structure. The captured and reproduced spatial dependence structure is displayed in Fig.8. (for the entire joint distributions) and figure 9 (specifically for joint tail-ends of the distributions).

Referee (D. Paprotny), main 2. What is the uncertainty in the results? 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?

Response (Diederen et al.): The notion of return period has no specific meaning within a multi-variate framework. In each individual set of peaks (each row that describes an event), each peak (at a different location) has a different exceedance probability (and therefore a different return period). Therefore, no return period can be assigned to a set of peaks. The uncertainty at individual locations could be investigated, which would imply investigating the GPD fits, which, as previously mentioned, is not the focus of this study.

However, we agree that uncertainty should be discussed and brought forward. We will provide better referencing to relevant recent studies that specifically focussed on uncertainty:

Hall, J. and Solomatine, D.: A framework for uncertainty analysis in flood risk management decisions Winter, B., Schneeberger, K., Huttenlau, M., and Stötter, J.: Sources of uncertainty in a probabilistic flood risk model)

Referee (D. Paprotny), main 3. What is the sensitivity to the choice of thresholds for the analysis? 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 panEuropean 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?

Response (Diederen et al.):

This is a very fair and important comment. Unfortunately, We have explored a couple of settings for each of the parameters mentioned, which classifies as trial and error. If necessary, we could explore a couple more settings and report on the difference between outcomes using different settings. However, because of the complex methodology applied to a large data set (long computational times), a full sensitivity analysis would not be feasible.

With regards to the settings of the statistical model, a recent study specifically addressed the settings using in the HT04 model: Winter, B., Schneeberger, K., Huttenlau, M., and Stötter, J.: Sources of uncertainty in a probabilistic flood risk model)

We suggest that the sensitivity of the methodology should be further tested in future research.

Referee (D. Paprotny), main 4. How the resulting dataset can be applied to a pan-European flood risk assessment? Ad. 4: The paper states in the title that the resulting dataset could be used for “largescale, 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?

Response (Diederen et al.): As part of the system-risk project (https://system-risk.eu), the task of the reconstruction of (spatially-coherent) hydrographs lies with the partner in Bristol, who do large-scale inundation simulations (Lisflood) and requested this set of synthetic discharge peaks (at these specific locations). To drive their inundation models, they will use the peaks to set up discharge boundary conditions (hydrographs), i.e. they do not require discharge peaks at each grid point. Alfieri et al. only address distributions per grid cell, which is something that can be done more easily on a high resolution. Addressing the dependence structure means that including each location gives an additional dimension (column) to the multivariate matrix. The number of locations was indeed for testing and could be easily expanded, with mainly the drawback of (again) longer computational times. However, it cannot be expanded to extremely large numbers of locations (high resolution grids) because of the dimensionality of the multivariate analysis.

Referee (D. Paprotny) What are the next research steps? 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?

Response (Diederen et al.): The main limitation of this method is discussed in 1.2.3 Handling dynamic events in a statistical event generator (p3, l1). This limitation is brought forward and discussed in the manuscript, because it is a limitation that does not specifically apply to this analysis only, but it applies to all large-scale, multi-site, event-based analyses. This type of analysis requires peaks (or other descriptors) at specific locations, whereas, within a large-scale framework, events do not produce peaks at all locations. This limitation implies that multi-site analyses will always require trade-offs (for the main trade-off of this study, see p8,l9). Climate change is not addressed in this study. It would require an expanded methodology (somehow incorporating trend analysis and dependent sampling within a multi-site framework) and would require (again) a lot of additional computational power and effort. Including trend patterns would reduce the quality of the spatial dependence structure, since such a methodology would lead to (statistical) trade-offs. Flood hazard zones and annual expected losses could be calculated after using these boundary conditions to drive a European-wide inundation model which then, subsequently, would be used to drive damage models. This study focusses on the generation of the boundary conditions, not on the input-output models belonging to the flood risk modelling chain.

Referee (D. Paprotny), minor 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

Printer-friendly version

Discussion paper

might be included as a supplement so that it won't slow down the pace of the paper.

Response (Diederen et al.): Although significant effort was put into finding the river network in the data set (using the mean flow per pixel and a river finding algorithm), we do not consider this part of the analysis important for publication, since the location of European rivers is well known, especially of the major ones. All that would theoretically be required is access to a database containing this information. The 298 locations were selected with the criterion of decent coverage of the river network. Other locations could be used, but this would not change the methodology, which we focus on in this paper.

Referee (D. Paprotny), minor 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.

Response (Diederen et al.): We agree and will omit the sentence.

Proposed improvements 1. We will improve the legends and scales in the figures (where required). 2. We will better reference to studies that consider uncertainty and sensitivity. We will add a discussion section as suggested, in which we will discuss uncertainty, sensitivity, applicability to European-scale FRA and future research outlook and we will see if we can move some discussion from the methodology section. In this section, we hope to address the reviewer's main points/objections [point 2 and 3] 3. We will see if we can provide better reference to the general HT04 fitting procedure. 4. We will completely revise the abstract as suggested [minor 1].

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