Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-396-RC1, 2020

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Do small and large floods have the same drivers of change? A regional attribution analysis in Europe" *by* Miriam Bertola et al.

Anonymous Referee #1

Received and published: 23 September 2020

The manuscript *Do small and large floods have the same drivers of change? A regional attribution analysis in Europe* by Bertola et al is the natural sequel of the previous HESS paper by some of the same authors (https://doi.org/10.5194/hess-24-1805-2020) taking the investigation from the detection to the attribution of changes in high flows of different frequencies. The manuscript is well organised and deals with a very interesting topic which I imagine will attract many readers. It is highly relevant for a European readership and presents an investigation of which physical variables appear to drive the magnitude of high flows in Europe differentiation between the common and the extreme high flows.

Printer-friendly version

Discussion paper

In the introduction the authors frame their study within the current literature giving a nice excursus of what the current state of modelling change is. I have some disagreement on some of the language they use, though. They mention several papers saying that most studies focus on the change in the mean annual flood, which they then contrast to their interesting new approach. On the other hand though most studies I have seen in the literature (including those cited) focus on explaining the change in the location parameter (or sometimes the scale parameter) - but typically the mean flood would be a combination of all distribution parameters. So modelling a change in location typically reflect on a change in the mean flood, but the model aims at modelling some slightly different quantity. More importantly, when location and scale are both allowed to change the mean flood would change as a function of both parameters, so the model for the mean flood would be rather complex.

In equation 4 it is not very clear to me how the model is regional and each station contribute information to the model. I understand that all station-years contribute to the likelihood and things are then corrected using the likelihood inflation? I mean this is not a multilevel model in which station-specific parameters are allowed, is that right?

Further, I understand that the model for the two quantities is estimated at the same time, so the q_2 is "hidden" in the x_{100} model: to make this maybe more obvious I would use a bracket before to "connect" equation 4a and 4b.

I am also not entirely sure why no ϵ_g was allowed in the growth factor model. For those who might want to code this up themselves it might be helpful to have the formulae translating parameters to quantile and even more, to be able to read the Stan code - I would recommend that the authors share their code either via GitHub or via some more academic-oriented repository such as Zenodo.

Printer-friendly version

Discussion paper

To summarise: I think the model could be described with more details, especially for those who have not read the first paper on which this builds.

Finally this is more of a curiosity, I was wandering what forms do the parameters functions take when one re-transforms the quantiles back to parameters. Can these shapes tell us something interesting about what types of functional relationship exist between the physical variables and the distribution parameters?

I find the modelling strategy of the authors quite interesting because they effectively model two quantiles which are indeed of interest rather than the parameters: should we then ditch the standard parametrisation of the Gumbel distribution or are the parameters still useful?

Regarding the choice of the priors: the authors choose to set a hard bound on the elasticity parameters: did this create any problem in the estimation? I mean: is the posterior distribution very concentrated on the lower bound or does it spread nicely?

I am somewhat dubious about the pooling of stations done by the authors and the use of averaged quantities across the rather large 200km x 200 km squares. To begin with the pooling will necessarily pool together information on small basins and large basins: this might not be problematic but I am more worried that with such large squares the pooling will put together very different types of basins (for example, alpine small basins and lowland larger basins): the response these basins have to drivers might be very different. Since from my understanding there aren't station specific parameters in the model, there might be some issues with the homogeneity of the groups and the ability of the model to identify the effect of the drivers on high flows. On the other hand, the average value of such large square might be not very useful to explain the variability of high flows for small basins and possibly inflate the variability of the results. I don't really see a way of out of this - I think the authors made some pragmatic decisions to be able to perform their study, but I wander whether we can fully trust their findings. In a similar vein: some areas are much more densely gauged than others, allowing possibly for a more precise estimation. This is not mentioned at all in the current manuscript.

Figure 8 is very interesting, but maybe I would complement it with two other visuals which would be relevant: the changes in the precipitation, soil moisture and snowmelt in each of the regions (to make more sense of how the curves morph from row to row in Figure 8) and final change in the different quantiles between the beginning and the end of the recording period (or any two moments in time).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-396, 2020.