Response to the interactive comments to the manuscript **hess-2019-523**

"Flood trends in Europe: are changes in small and big floods different?"

by Miriam Bertola, Alberto Viglione, Julia Hall and Günter Blöschl

We reproduce and number in the following document all the comments of the Referee in *italic characters*, followed by our answers.

Referee #2: Duncan Faulkner

The paper makes an interesting and valuable contribution to the large volume of literature on trends in flood magnitude. The pan-European focus is particularly valuable, as is the separation by flood rarity and catchment size.

We thank the Referee Duncan Faulkner for the time he spent on our manuscript and for the useful and constructive comments that will help to improve the quality of the manuscript. We have carefully considered and addressed all his comments in the following.

Main comments

1. The paper acknowledges that no allowance is made for spatial correlation of floods, and that this may affect the estimation of uncertainty. It claims that the regional model is more robust than the at-site trend analysis. This raises the question of the extent to which the apparent increase in robustness is due to the same information being repeated several times over, if trends at nearby gauges are reflecting essentially the same flood events. I recommend that the authors consider ways of accounting for spatial correlation when quantifying uncertainty, such as a spatial nonparametric bootstrap or a likelihood correction (Sharkey and Winter, 2019). The authors quote a cross-correlation length of about 50km (section 4) which seems rather short in comparison with the spatial scale of some flood-producing weather systems.

The referee is right, spatial cross-correlation between flood timeseries at different sites is not accounted for in this study and it may affect the estimation of sample uncertainty. In particular, if the flood timeseries at different sites are strongly correlated, we expect the uncertainties to be larger than the uncertainties estimated in this paper. Therefore, at lines 299-301 we state that the estimated credible bounds should be intended as lower limits. However, we do not expect the trend estimates (i.e., in this case, the posterior median) to change, when cross-correlation is taken into account (Stedinger, 1983; Hosking and Wallis 1988 and 1997).

We thank the Referee for suggesting possible ways to account for spatial correlation. In particular, the likelihood correction approach in Sharkey and Winter (2019) seems to fit this case, as the parameter estimates derived from the corrected likelihood are unchanged and the independence likelihood is scaled down, resulting in inflated asymptotic variance of the posterior. Based on the results of Sharkey and Winter (2019), we expect the credible bounds to be up to 20% wider with the adjusted likelihood. We will apply the likelihood correction approach for the example region of section 3.1, using the likelihood correction factor as estimated in Sharkey and Winter (2019), in order to quantify the magnitude of the increase in the width of the credible bounds when spatial cross-correlation is taken into account. Based on the result, we will also introduce additional text in the discussion section 4.

The cross-correlation length of about 50 km has been calculated from the flood timeseries using distances between the catchment outlets. We will state more precisely how this has been calculated in the revised manuscript.

2. The discussion (Section 4) makes various statements that go some way towards attributing trends. These vary from confident assertions (flood trends in the Mediterranean are negative due to …) to more informal or speculative comments *using wording like "linked with", "suggest that", "could be found". I think many of us tend to use language like this when discussing trends, but in this context I would suggest the authors state more explicitly whether they are attempting a formal attribution of the trends or merely providing some hypotheses (or somewhere in between).*

Thank you for pointing this out; we understand that we need to clarify the nature of our statements in section 4. This work does not aim at formally attributing the observed flood trends to drivers and the statements in the discussion section, about the potential causes of the observed flood trends, are intended to be hypotheses or interpretation of the results, based on the literature and on the Authors' understanding of these processes. We will clarify the nature of our statements in section 4 and we will mitigate the statements that sound like confident assertations.

Minor comments

The paper makes frequent use of the return period terminology. This is conceptually awkward in non-stationary conditions. I would suggest that the authors at least acknowledge this, and perhaps refer to some of the literature on alternative ways of expressing flood rarity.

We understand that the use of the return period terminology in a non-stationary context may sound ambiguous to some readers. In the manuscript we refer to the return period (rather than to the annual exceedance probability) because, in the engineering practice, it is widely understood what a 100-year flood is. Examples of return period terminology used in a similar non-stationary context in the literature are Renard et al. (2006), Machado et al. (2015), Šraj et al. (2016). For these reasons we prefer to maintain the return period terminology in the manuscript. However, we will clarify the terminology used in the method section 2.1.

The Gumbel parameters are modelled as varying with time according to a log-linear relationship. Perhaps the authors could comment on any alternative ways they considered of modelling trend, such as other mathematical forms of the relationship with time, or inclusion of physical covariates in an attempt to improve the identification of the time trends.

Thank you for pointing this out. We will introduce additional text about alternative ways of modelling trends in section 4. The use of physical covariates in order to attribute the trends in flood quantiles is actually planned in the next phase of our research.

The meaning of gamma and S in the equations around lines 103-4 was not clear to me.

Thank you for spotting this lack of clarity in these lines. The gammas are parameters of the model that control the scaling with catchment area and S is catchment area. We will specify it in lines 103-107.

The assumption of homogeneity of the windows (section 2.3) seemed to me to need some justification.

Thank you for rising this point; we understand that we need to clarify it in section 2.3. As described in the manuscript in section 4 (lines 314-321), we have not formally tested this assumption (i.e. the statistical homogeneity of the 600x600 km regions in terms of the flood change model used here), because formal procedures to assess the regional homogeneity, such as for example those used in regional flood frequency analysis (e.g., Hosking and Wallis, 1993; Viglione et al., 2007), are not available at the moment. Furthermore, while deviation from regional homogeneity would probably invalidate estimates of local flood change statistics from the regional information (e.g., as in the prediction in ungauged basins, see Blöschl et al., 2013), we expect its effect on the average regional behavior to be less relevant. However, this assumption seems to have a small influence on the results, as we have not observed significant differences in the regime changes when changing the size of the moving windows (this was done in preliminary tests and is not shown in the results). We will introduce additional text in section 2.3 to clarify this point.

I was impressed with the design of Figs 2 and 3, which pack in a great deal of information. I would suggest that the authors either remove or justify the extrapolation of the model to catchment areas ten times smaller and ten times larger than those included in the dataset.

We agree with the Referee's suggestion of removing the extrapolation of the model to very small and very large catchment areas. We will modify Figures 2 and 3 accordingly.

The description of Fig 5 mentions larger positive trends in NW France for big floods than for small floods. I could not see that effect from comparing the pairs of maps.

Thank you for noticing it; we will correct the description of Figure 5.

Reference

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H. (2013) 'Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales', Cambridge University Press. doi: 10.1017/CBO9781139235761.

Hosking JRM, Wallis JR. The effect of intersite dependence on regional flood frequency analysis. Water Resources Research 1988; 24:588-600.

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, Water Resources Research, 29, 271–281, https://doi.org/10.1029/92WR01980, 1993.

Hosking JRM, Wallis JR. Regional Frequency Analysis: an approach based on L- Moments. Cambridge, UK: Cambridge University Press, 1997

Machado, M. J. *et al.* (2015) 'Flood frequency analysis of historical flood data under stationary and non-stationary modelling', *Hydrology and Earth System Sciences*, 19(6), pp. 2561–2576. doi: 10.5194/hess-19-2561-2015.

Renard, B., Lang, M. and Bois, P. (2006) 'Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data', Stochastic Environmental Research and Risk Assessment, 21(2), pp. 97–112. doi: 10.1007/s00477-006-0047-4.

Sharkey, P., & Winter, H. C. (2019). A Bayesian spatial hierarchical model for extreme precipitation in Great Britain. Environmetrics, 30(1), e2529.

Stedinger JR. Estimating a regional flood frequency distribution. Water Resources Research 1983; 19:503-510.

Šraj, M. *et al.* (2016) 'The influence of non-stationarity in extreme hydrological events on flood frequency estimation', *Journal of Hydrology and Hydromechanics*, 64(4), pp. 426–437. doi: 10.1515/johh-2016-0032.

Viglione, A., Laio, F., and Claps, P.: A comparison of homogeneity tests for regional frequency analysis, Water Resources Research, 43, 545 1–10, https://doi.org/10.1029/2006WR005095, 2007.