Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2021-345-RC2, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

# Comment on hess-2021-345 Brunner and Slater "Extreme floods in Europe: going beyond observations using reforecast ensemble pooling"

Peter Salamon (Referee)

---

Referee comment on "Extreme floods in Europe: going beyond observations using reforecast ensemble pooling" by Manuela I. Brunner and Louise J. Slater, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-345-RC2, 2021

---

*General comment:*

The manuscript presents an approach to increase the sample size for the estimation of the frequency of flood events. The approach is based on pooling of reforecast ensemble members and has not been previously assessed for flood frequency analysis.

The paper is overall well written and structured. The approach presented in this manuscript is of high interest as estimating flood frequency in practice is often hampered by short observational records. The discussion section outlines the possible limitations of the approach. My main concern is related to the data used for the study. The study uses EFAS v3.0 historical simulations to assess whether the selected stations have a good performance when comparing simulations and observations. However, the EFAS reforecast data set used for the ensemble pooling is based on EFAS v4.0 which includes a completely new model calibration, upgrades to static fields for the hydrological model LISFLOOD and a change from a daily timestep to a 6 hourly timestep. Overall, EFAS model performance from v3.0 to v4.0 has increased significantly and therefore it is not recommended to select stations based on v3.0 and perform an analysis using reforecasts that are based on EFAS 4.0. As this has an impact on all results and analysis in the manuscript a major revision is required.

*Main concern:*

As described in the general comment EFAS reforecasts are based on EFAS v4.0 as is also indicated in the metadata on the Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/dataset/efas-reforecast?tab=overview ). However, EFAS historical simulations v3.0 were used to pre-select stations with a good fit

between simulated and observed discharge. Given that EFAS 4.0 contains a completely new model calibration with more calibration stations (1137 for v4.0 instead of 717 stations for previous EFAS versions), and upgrades to static fields for the hydrological model LISFLOOD and a change from a daily timestep to a 6 hourly timestep as is described in detail in the EFAS wiki (see here: https://confluence.ecmwf.int/display/COPSRV/EFAS+v4.0 ) it is not recommendable to use EFAS v3.0 model performance to pre-select stations and then use those pre-selected stations with EFAS re-forecasts from EFAS v4.0.

Furthermore, the authors do not describe in detail how the simulated data was extracted from the EFAS simulations. EFAS output has a spatial resolution of 5km x 5km. The coarse spatial resolution of the hydrological model LISFLOOD used in EFAS requires an upscaling of the river drainage network from a high resolution datasest to the 5km x 5km grid scale. This means that coordinates of gauging stations cannot be used directly to extract simulated timeseries of discharge from the EFAS simulations as original gauging station coordinates may be located just next to the main river channel on the coarse grid scale. Instead, before extracting simulated time series it has to be checked whether the drainage area of the EFAS grid pixel corresponds to the drainage area as provided by the data provider (here GRDC). While smaller differences in the drainage area are expected due to the different spatial scales, if there is a large difference, it means that coordinates have to be shifted to ensure an adequate match. For this purpose the drainage area of the LISFLOOD/EFAS network is available on the C3S CDS (https://cds.climate.copernicus.eu/cdsapp#!/dataset/efas-historical?tab=overview ). This is especially important for gauging stations with very small drainage areas which seem to have been used predominantly in this study (Fig. 9).

Furthermore, LISFLOOD simulates lakes and reservoirs as points on the channel network. It is not recommended to extract simulated time series at the same pixel where the reservoir or lake is located but to either extract the time series on the upstream or downstream pixels of lakes and reservoirs (depending on the location of the gauging stations for observations). More info can be found on the model documentation of LISFLOOD (https://ec-jrc.github.io/lisflood/ ). The location of lakes and reservoirs on the EFAS grid can be found also on the EFAS map viewer (https://www.efas.eu/efas_frontend/#/home ) .

Finally, we have found several data quality issues with the observed discharge data in GRDC in the past. We recommend strongly to have at least a visual check of the observed data that is selected for the analysis.

*Minor comments:*

Chapter 2.1, page 3: EFAS 4.0 as well as EFAS3.0 have been calibrated using Kling Gupta efficiency and not NSE. For more details on the EFAS versioning and what changes are included in each EFAS version please see here for a detailed description: https://confluence.ecmwf.int/display/COPSRV/EFAS+versioning+system The reference to Smith et al. refers to previous and outdated EFAS model versions that is not available on

the C3S CDS.

Chapter 2.2, page 5, line 109: "…model stability…": Please describe what is meant here with model stability? In terms of what?

Chapter 2.2, page 5: It is stated that Spearmans rank correlation can only be computed for AM and not directly for POT (lines 114-116). However, in the following sentence you write that you calculated Spearmans correlation for POT. Please clarify!

Chapter 2.3, page 8, line 180: … the the…. Please correct.

Chapter 3.1, page 8: This is a repetition of Chapter 2.2. Please remove Chapter 3.1!

Chapter 3.3, page 13, lines 269-271: Please describe the evidence for claiming that relative differences between simulated and observed best estimates and uncertainty bounds are independent of model performance.

Chapter 3.3, Fig. 11: I disagree with your statement that flood quantiles are positively related to mean precipitation. According to Fig. 11 there is only a very weak positive relation.

Chapter 3.3, Fig. 11: Please explain why there is such a strong positive relation to Latitude according to Fig. 11? This is not mentioned at all in the text.