Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/hess-2022-207-RC1, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review of hess-2022-207

Anonymous Referee #1

Referee comment on "Future changes in flash flood frequency and magnitude over the European Alps" by Mar J. Zander et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-207-RC1, 2022

This very well written manuscript proposes a European scale analysis of flash flood frequency combining downscaled ERA5interim data with the distributed hydrological model wflow. The method and results are overall well described; there is however a considerable lack of reference to actual hydrological processes involved in flash flood generation and what this means in terms of requirements for the modelling chain. I cannot recommend publication of this manuscript. It has the merit to attempt a regional-scale analysis of an important phenomenon. But the modelling exercise is detached from flood generating mechanisms. I anticipate that revising the work would require a profound modification of the modeling chain.

As becomes clear below, I do not see evidence that the modelling framework is fit to model flash flood frequency. It rather seems like the model applies some filtering to the precipitation events but we do not know how realistically this generates flash floods. Accordingly, the added value of the hydrological model as compared to frequency analysis on the precipitation directly is not convincingly demonstrated or discussed. This problem is enhanced by the fact that according to the text (methods section), only 9 flash floods are contained in the used data base for the entire region and study period. For these floods, the model validation is summarized as: "For all recorded flash floods, the modelled peak daily discharge was heightened compared to the preceding and following days." (Section 3.3). This seems a weak argument. The regional approach to flash flood validation (an observed event is validated if something happened within the three days in one of the 5 large subbasins, Fig. 1) is also not convincing: in as far does this demonstrate that the hydrological model has any added value compared to the precipitation input?

Furthermore I do not understand how the flash flood producing threshold is validated, i.e. I do not understand what is actually validated (lines 215-220). This threshold is however the key to judge if flash flood frequencies increases or decreases.

Finally: the study does not work with downscaling or bias correction (see discussion: "it was explicitly chosen not to apply a bias correction, downscaling, or a delta change approach to the climate model data as these techniques can disturb the change signal”). I have a hard time to understand how the output of the hydrological model is of any use in this case, in particular because it aggregates model outputs from catchments for which the bias of precip., temperature or their variability might be important but spatially different This problem is not reduced by comparing similar simulations for the reference and the future period since flash flood analysis corresponds to the analysis of extreme events that might be crucially influenced by biases, and differently under current climate compared to future climate.

Detailed comments

- Introduction: I would suggest to add references / arguments for “The intensity of flash floods and thereby their impacts may increase”. Is precipitation intensity the only driver? What is the role of soil moisture, infiltration capacity? Is looking at precipitation intensity sufficient or would you need to consider compound events of high precipitation intensity and low antecedent moisture? I do not mean to suggest that you need to address all this but it would be good to give the bigger picture and to not oversimplify the case of flash floods; we also would need to have information (in the methods part) if the chosen model can reliably reproduce Hortonian and saturation-excess floods and who this is validated for the selected catchments
- Introduction: the sentence “Although Kay et al. (2015) showed that finer resolution CP-RCMs (..) for large-scale river flooding,” is surprising, you just said before that using CP-RCMs for hydrological impact modelling is novel
- Line 56 following: unclear if you talk about your own study or about the study in Norway?
- Introduction: “However, they conclude to finding no clear added value of the CP-RCM simulations due to lacking realism in the temporal distribution of rainfall intensities at a sub-daily scale and/or total precipitation amount per rainfall event (Reszler et al., 2018).” Did you check in your work that the used precip series (without downscaling) have a realistic temporal distribution?
- Study area: a map with mean precip properties (annual totals, intensity of e.g. 1-day precip events) would complete the picture; with the presented information, we have no idea how spatially variable the meteorological drivers are
- Study area: “Most flash floods in the European Alps occur in late summer and autumn”; do you have a reference? later on, “in France xxxx” : these sentences are misleading since this analysis is only about a specific part of France, Italy etc. Do you have detailed information on flash floods in the catchments that you actually include in the study? What about flash floods in the Northern parts (Aare / Rhine catchments); are there many?
- Hydro model: could you perhaps mention the number of model parameters that are estimated / assigned following “same a priori parameter estimation methodology as Imhoff et al.”? which parameters potentially influence flash floods the most? Does the model have Hortonian and saturation-excess surface runoff?
- Can you elaborate on why the vertical hydraulic conductivity parameter is the most sensitive one? And is this result from previous work transferable to here? I guess the sensitivity of wflow parameters depends on what hydrological time scale is dominant for

the considered catchment / scale: either processes leading to water mobilization at the hillslope scale or processes of lateral transfer (surface and subsurface) to the stream network or routing within the stream network; if you only assess hydraulic conductivity you implicitly assume that water mobilization is dominant? But what about lateral flow to the stream network?

- What is a "constant multiplication factor of 100"? how does it influence hydraulic conductivity?
- Why did you use ERA5 for the sensitivity study and not downscaled ERAinterim? Should you not test the model sensitivity with the type of input data that you use thereafter? What is the added value of the validation with ERA5? If the model performs better with ERA5, how is this useful for the final aim, generating flash floods under future climates?
- What was the time step of ERA5 combined with wflow? Hourly and daily, this is unclear? What is the value of daily model performance assessment for a model that is later on used to derive flash floods? I miss some convincing arguments that the analysis framework is actually able to reproduce extreme streamflow events of the flash flood type
- Is the quality of the modelling chain with CP-RCM driven by a Global Climate Model assessed in any way? E.g. in terms of the flash flood frequency for present day? Otherwise, how can you justify that this model chain gives valuable results?
- How are lakes treated in the model? Does flash flood analysis below large lakes make any sense?
- The paper would highly benefit from a sketch that summarizes the work scheme (what data went into what model and how the output was assessed / used)? Instead of Table 1, which does not mention to what the model output is compared to
- The focus of the paper is on the European mountain range where snowmelt processes will necessarily play a key role at elevations roughly above 1600 m asl., even in summer and especially in autumn; in summer, flash flood frequency will crucially depend on the saturation level of the catchment, which in turn is conditioned by the snowmelt season (duration, intensity); in autumn in exchange, potential early snowfall with or without subsequent rain-on-snow events can strongly modulate how intense precipitation events translate into flash floods or not
- 2.6: daily maximal specific discharge for flash flood identification: why would the maximum discharge of a daily time step be relevant for small catchments?
- Flash flood definition: I see that only rainfall induced flash floods are considered; are rain-on-snow events never assimilated to flash floods? Given that the analysis expands on the alpine area, this should be made very clear; besides: are you sure that flash floods cannot occur early in spring (the analysis is on summer (JJA) and autumn (SON) period)? Especially in the future?
- Results: the fact that the he annual cycle of discharge with low flows in winter and snowmelt in May to June does not validate the hydrological model; any model will do so as long as there is any form of freezing and snow model in it; given the high annual cycle, we have no idea what the reported KGE values mean (any model that has some annual cycle will have KGE values beyond 0.6 I would guess); how could you validate the flood generation frequency of the modelling chain otherwise?
- Line 211: incomplete sentence (and update all figure references)
- Discussion: I do not understand how you can judge if the frequency of events increases but at the same time admit that it is impossible to estimate return periods ("Just as in Rudd et al. (2020) we do not attempt to estimate return periods of extreme events as the simulation periods are too short to warrant such an approach."). I guess that answer is in Rudd et al. but it would be good to understand this here also