Climate of the Past Discussions

Clim. Past Discuss., author comment AC2
https://doi.org/10.5194/cp-2021-160-AC2, 2022

Reply on RC2

Stefan Brönnimann et al.

Author comment on "Influence of warming and atmospheric circulation changes on multidecadal European flood variability" by Stefan Brönnimann et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2021-160-AC2, 2022

The authors identify a "long-lasting conundrum" in the literature such as past flood-rich periods occurred mostly during cold conditions, while more floods are expected with the ongoing global warming. They develop an approach aiming to explore changes in the atmospheric conditions (dynamic versus thermodynamic processes) that could explain these distinct patterns in flood activity in Europe. The scientific question is highly relevant as e.g. flood projections still encompass large uncertainties, partly because how the climate change may regionally modify flood hazard is still unclear. The study is rather well designed and the paper well written. However, i) the "long-lasting conundrum" as presented here is not fully consistent, ii) the objectives and the way of proceeding for some treatment / analyses need clarifications and, most importantly, iii) the findings partly rely on visual analyses, limiting the robustness of the conclusions. These points are detailed in the comments below.

Main comments

- *The "long-lasting conundrum" (section Abstract and Introduction)*

The authors introduce a "long-lasting conundrum" with the increased occurrence of floods during past cold periods and the expected increase of floods in the future with the climate change. This relies on the comparison between historical / paleodata and projections. However, the comparison is limited by differences in i) the time scales, ii) the studied catchments and iii) the return period considered. For instance, changes in flood activity from paleodata are mostly observed at longer timescales than those discussed here (centennial versus decadal). Paleodata also come from very small catchments (a few km²) compared to those studied here or those studied with historical data and projections (>1000 km²). Floods discussed in historical and paleodata are characterized by high return periods (>10-100 years), while the authors discussed here annual flood events. All these differences may result in a large range of flood-prone hydrometeorological processes and, thereby, in various responses in flood variability to the same climate change. This may easily explain this "long-lasting conundrum".

Thanks for the comments. This is an interesting point. We will reformulate the abstract, as our target are not these very long time scales. The author is right that time scales, catchments, and return periods are different. These three factors will be discussed better in the revised manuscript

A recent paper by Wilhelm et al. (2022) reiterates these points and shows that several warming periods during the last 9000 years induced a decrease in the frequency of large floods (return period >10 years). Our return period is shorter (it is actually slightly longer due to the seasonal focus, which selects slightly stronger events), which will be discussed.

The water vapour effects is expected to always play a role – no matter whether a warming or cooling is forced or due to internal variability, and not strongly dependent on the time scale. The circulation effect, however, may be different in the two cases, and it is certainly depends on the time scale analysed.

In addition, the authors only use 1 reference about flood projections (l.55), while many other studies have been published and show large differences even in the sign of the change (some of these references are used l.421-422). Thereby, selecting another one may also show a decrease in flood activity under warming conditions. The authors may more convincingly use the recent findings of Blöschl et al. (2020, Nature) to introduce this "conundrum" – indeed, they showed that past flood-rich periods occurred under cooler conditions, while the most recent one occurred under warm conditions with a more homogenous dataset.

Thanks, it may be a better introduction to use the historical period rather than the Holocene period. We will introduce this open question with Blöschl et al.

- *Treatment / analyses that need further explanations and/or quantifications*

Discharge data (l.72-74) - The authors apply many treatments to the discharge data without explaining the rationale behind. This needs clarification. I am also wondering why the authors do not use the raw data instead of this kind of index of "flood intensity", especially when this provides similar results as stated.

In our study we aggregate many different series, and for this they need to have similar properties. The normalization does that (at least to some extent). This will be explained in the revised paper. "Flood intensity" is simply our short name for "normalized peak streamflow". This can easily be changed.

Precipitation data (l.88-90) - The authors chose 2 precipitation indexes (Rx5d and Rx20d).

Again, there is neither a rationale nor reference to explain why the authors chose 5 and 20 days. The sizes of the studied catchment areas are very different and floods may be triggered by rainfall events of different durations. In addition, why a short and long duration? What do these two indexes represent here? This also needs clarification. At the end, only the Rx5d is used in the analyses.

Thanks for the question. The diagnostic for flood-propelling rainfall is Rx5d. We use Rx20d (more precisely, the seasonality of Rx5d) only for characterising the different streamflow series in the clustering process. Both will be better justified.

For typical river floods, shorter periods are relevant. We have checked this for one station (Basel) in a previous publication (Brönnimann et al., 2019) and found that the 2-3 days prior to the event are the most relevant (5 days prior to the event precipitation is already above the 75th percentile, but this is not extreme). There is a more systematic study on this by Froidevaux et al. (2015) concluding "that the consideration of a 3–4 days precipitation period should be sufficient to represent (understand, reconstruct, model, project) Swiss Alpine floods." Note that the size of catchments varies largely in our study; some are larger than those studied in Froidevaux et al., some are of similar size. Eventually we aggregate series regionally (both precipitation index and streamflow). Hence, Rx5day should be a good choice.

Flood seasonality (Fig. S1 and l.93 and following) - The authors perform a selection to get a set of cold- and warm-season floods, assuming they mirror distinct, regional hydrometeorological processes. The considered seasons are here long of 6 months. Why 6 rather than 1 or 3 months needs to be explained.

From an atmospheric point of view, heavy rainfall is associated with specific weather patterns such as elongated troughs or cut-off lows (see Stucki et al. 2012). Winter events tend to be related to different pattern (e.g., zonal flow) than summer events. Moreover, the role of convection is stronger in summer. It therefore makes sense to discriminate a cold and a warm season, but a finer partitioning would probably not result in more different weather regimes. Conversely, it would strongly decrease the sample size. In the revised manuscript we will explain that better.

The authors also discard 5 flood series because their triggers may include e.g. snow processes (l116-117). However, much more series do not show a good correspondence between the highest values in the precipitation indexes and the highest occurrence of the annual peak streamflow (Fig. S1), suggesting that almost half of the series represent mostly floods triggered by a mix of processes in which precipitation is not a dominant driver. Or that the chosen precipitation indexes are not the most relevant. The selection process of the series is thereby questionable.

The strongest precipitation event in a year does not necessarily cause the highest streamflow. As explained in the paper, other factors contribute. We would like to avoid selecting flood events for which the atmospheric disposition was not relevant; this would dilute our sample. We also do not want to be select only a handful out of 47 series. We have good reasons for removing the five series, which will be further discussed in the text. It is correct that the remaining series are not purely atmospherically driven, and some comments on that will be added in the context of correlations between Rx5d and peak streamflow.

Correlation test (l.248 and following, Fig. 4) – A correlation test (but which one is not indicated) has been performed between peak streamflow and Rx5day (why not also with Rx20d?). Among the results shown (8 rivers among 43?), the values are rather low for most of them (< 0.35). First, results for the other results should also be shown in e.g. a table in Supplementary Material so that the reader can have an overview of its relevance.

Thanks. The correlation text will be better explained. It is a t-test and was performed at the level of the regions for Rx5d (as mentioned above, Rx20d was only used for discriminating series), not at the level of the individual rivers. But in the revised manuscript we now add a column to Table S1 with this information on the level of individual series. In the submitted manuscript, we performed the correlation for 4-yr average, to be consistent with Blöschl. However, this introduces a new time scale, which may be confusing. Therefore, in the revised manuscript, we use only the unfiltered data for the correlation (and then use cross-wavelets to address the relation between two series as a function of time scales).

Correlations (on the unfiltered data) are in the range of 0.35-0.4. Whether this is low or high is another question. As mentioned above, the strongest precipitation event in a year does not necessarily cause the highest streamflow. Some series will have their strongest Rx5d event always in summer and the highest peak streamflow always in winter. So, we essentially correlate a summer series with a winter series. Do we expect a higher correlation than 0.3 between the two? This is another reason for the seasonal stratification. We now give more explanation of this in the revised manuscript.

Second, the general low values suggest that precipitations explain only a small part of the variability, limiting the relevance of the following analyses to explain changes in flood variability. This point is not discussed. About the correlation between CONV5d and peak streamflow, it is only assessed visually, while it is a key link for the following analyses.

As mentioned above, the low correlation on unfiltered (or 4-yr filtered) data does not mean that precipitation is unimportant. Annual maxima of Rx5d and annual maxima of peak streamflow often capture different events, so their time series may not be strongly correlated (we do expect some correlation). We ask whether there is a multidecadal variation that is common to both, which would then lead to stronger correlations on a low-frequency scale. In addition to the visual judgement, we now also perform cross-wavelet analyses (see below).

Similarly, the respective contributions of the circulation change, water vapour change and interactions on changes in the annual peak streamflow is also based on visual "correlation" (l. 294 and following; Fig. 5). Therefore, the findings mainly rely on visual comparisons, strongly limiting their robustness. Instead, we expect that a correlation test as well as a significance test to be applied systematically to each correlation discussed and supporting the findings.

All correlations are tested. For the case of Fig. 5. we perform cross-wavlet analyses of the unfiltered data. These analyses confirm significant relations on multidecadal scales between Rx5day and streamflow, between CONV5d and Rx5d, and between CONV5d and streamflow.

- *The relative contribution of atmospheric processes to changes in peak streamflow*

The authors stated that periods with higher flood intensity prior to 1950 are mainly due to circulation changes, while the period with higher flood intensity after 1950 is more related to changes in water vapour changes. However, looking at Fig. 5b, the contribution of circulation changes is also increasing after 1950, better mirroring the increase of CONV5d than water vapour changes. Therefore, a quantification of the respective contributions of atmospheric processes to changes in CONV5d is really needed to objectively assess them.

In the revised manuscript we quantify this by addressing the trend since 1963 (when streamflow reached a minimum) in CONV5d and in its contributions. This is based on unfiltered data (as all analyses - filtering is only used for plotting). This shows that CONV5d increases, the contribution of circulation changes has no trend, wheres the contribution form water vapour changes increseases even more strongly than CONV5d and is highly significant. This is added to the manuscript.

In addition, the authors state that this explains why flood-rich periods have been mostly observed during cold periods in the paleodata. However, large changes in temperature have also been reconstructed over the last millennia. So, we may also expect that changes in water vapour played a role further back in time?

Exactly! Any change in temperature, irrespective of its underlying cause, would be expected to cause higher water vapour concentrations and therefore higher CONV5d. This is what we show in Fig. 5b. However, if that temperature change is caused by a change in circulation, then this circulation change (in our study regions and seasons) operates in the opposite direction. Because circulation dominates at the multidecadal times scale we focus on, we do not see the temperature effect in the past (only our decomposition uncovers this). In contrast, the water vapour contribution becomes obvious in recent years because the warming was not (or not strongly) driven by circulation changes and therefore circulation does not counteract the water vaopur signal in this case. This leads back to the initial comment of this reviewer: It is indeed an open question whether the same holds for centennial variations. It may turn out that on these scales, water vapour changes always dominate.

We will explain this better in the revise manuscript.

Minor comments

L.117. The authors removed five series. They should name the series they removed.

They will be mentioned.

- *Typo: "in order to"*
  Thanks

l.180. Why 5 days? What this duration correspond to?

This window length and weighting was taken from a previous study (Brönnimann et al., 2019) and was based on analyses of daily discharge, precipitation, and water flux convergence on the preceding days. For the current study we had also tested using varying window lengths l depending on catchment size, using an equation l = sqrt(A)/47.3 + 2. Results were very similar. We also tested using different windows for water flux convergence. As the methods are already very complex, we chose not to elaborate on all the myriads of tests we performed. In the revised manuscript we add the sentence: "This window length and weighting was taken from a previous study (Brönnimann et al., 2019) and was based on analyses of daily discharge, precipitation, and water flux convergence on the preceding days. "

l.221. The authors may refer to Fig. S1 instead of S2?

Yes, correct, thanks.

l.227. Typo "such less pronounced peaks"

Thanks

- *"1919-.. exhibit low values)". This is unclear.*
  Changed to: local maximum at a time when warm-season series exhibit low values
- *"at 4-yr aggregation" again, why 4?*
  This is the resolution (voxel size) in Blöschl et al. (2020). This is now mentioned in the revised manuscript.

Fig. 4 and S4 are very similar. So, Fig S4 could replace Fig. 4.
The agreement demonstrates the robustness of the results. We prefer to show the seasonal analysis in the paper (Fig. 4) and the annual analysis in the Supplement for the reasons stated above (we are then more sure to have captured the relevant atmospheric processes, even though the result is the same). We think that it is important not to mix winter-dominated flood series with summer-dominated Rx5d series.

Fig. 7. Why the 80 members are not shown here? What is the curve, the mean of the 80 members?

Yes, it is the mean. This will be made clear. As the figure shows 32 time series, it is challenging to add the ensemble standard deviation to each one. In the revised manuscript we add bars (centered around zero) at the beginning and at the end of the series with their length corresponding to 2 ensemble standard deviations (+/-1). As the standard deviation decreases approximately linearly from the beginning to the end, it is sufficient to show it for the first and last year. This will be added in the revised manuscript. At the same time, the figure confirms that the multidecadal variations of the circulation contribution varies by more than 2 standard deviations.

Fig. S1. The way the seasonality is identified is sometimes misleading since there is not always a correspondence between the six months of high precipition and high discharge. Thereby, for some series, months with the most frequent annual peak flow are not considered.

There are two streamflow series where the month with the maximum peak streamflow is not in the selected flood window. The clustering put these two series in the same clusters as other series with a different seasonal cycle. Note that we perform all analyses also for annual series and show the results in the supplement.

Fig. S3. Why these series are excluded from the analyses?

This was explained in the paragraph l. 105-127 and Figure 2: Five of them concern high-altitude catchments that are either affected by snow melt or by hydraulic installations, but in any case do not show the same flood regime as on the same further downstream. The sixth turned out to be a one-series cluster and is reportedly affected by snow melt and rain-on-snow events. We nevertheless did not want to hide the results from these, and so put in a supplementary figure. In the revised manuscript we add to the Caption of this supplementary figure: (see Section 2.2 and Fig. 2). Furthermore, we add another reference supporting the effect of hydraulic installations in one case.

Fig. 1a. Why the numbering of the discharge series starts by #2 (instead of #1)? SI Table 1: why the cluster column is empty? If so, it can be removed.

Thanks should be #1, then #3. The numbers of the cluster are added.

Thanks for this very careful review and for the thoughts about the difference of our setting and that of palaeostudies.