Climate of the Past Discussions

Clim. Past Discuss., author comment AC1
https://doi.org/10.5194/cp-2021-160-AC1, 2022

Reply on RC1

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-AC1, 2022

This study uses streamflow, daily weather data, reanalyses, and reconstructions to explore variations in European flood frequency during the past 2 centuries and reconcile indications of European flood frequency to be higher under warm as well as cold background climate conditions. The study illustrates how the relative role of atmospheric circulation and moisture content for moisture flux convergence, hence precipitation, changed historically. Based on the moisture content contribution becoming predominant in recent, warmer times, the manuscript discusses implications of such findings for projected floods.

I found the study overall well-conceived and the manuscript well organized and well written. I especially appreciated the efforts to combine different sources of information including observations, model output and reconstructions. I found the analysis overall sound and the conclusions well supported by the results. However, I have a few comments on the study that I ask the Authors to consider in a revised version of the manuscript.

On a general note, I would appreciate a stronger focus on the statistical analyses supporting the existence of linkages between the considered processes, for instance in terms of significance of co-variability between time series. I provide a few specific comments below to illustrate the occasions when I felt the interpretation of results requires further support. Similarly, the comparison between information from different sources appears occasionally to be only qualitative. This left me wondering about the purpose of some of the comparative analyses provided in the manuscript: central or just ancillary to show uncertainty? I think the manuscript would benefit from a bit more guidance by the authors about the purpose (and expected outcome) of some of the analyses. Again, I highlight the few occasions when this occurred in the specific comments below.

Concerning the adopted methodology, the only main question I have regards the normalization: If I understand the method correctly, normalization is over the whole length of a time series. The relative representation of trends is affected by the fact that series of different length are compared (those including the trend period and those extending further back, during period of little or no trend). If this is true, maybe a word of caution on this approach when comparing trends in figures 2 and 3 or when different discharge time series are averaged out in Figure 1b can be appropriate. An alternative approach could be to normalize over periods as similar as possible across the different

series (for instance using 1900-2000). A few more specific methodological questions follows in the specific comments.

Thanks for this comment. Indeed, the normalization was done over the maximum possible length. In the revised manuscript we will use a normalisation on all data after 1900. Results are very similar

Specific comments

Section 2.1: The authors use annual maximum streamflow as a reference for their analysis of floods. What if more floods occur within one year? In my understanding this possibility is not accounted for in the analysis, but might be relevant for the overall assessment of flooding statistics. In section 2.4, it comes clear that daily streamflow series are available for only two stations, so I guess this aspect is difficult be assessed. Nonetheless I feel some discussion in section 2.1 would be worth it.

The reviewer is correct that we have only two series with daily data, which allow addressing that question. We will discuss the question of flood definition more prominently in Sect. 2.1 and add, for these two series, lines of flood frequency (number of events exceeding the 98th percentile, declustered) to the Supplement. The general results are similar, a sentence will be added.

Line 89-90: "From the precipitation series we calculated Rx5d and Rx20d, i.e., the annual maxima of precipitation sum over periods of 5 and 20 days, respectively." Is the temporal connection with the flood event checked? As far as I see it, especially for "flood intensities" that are about average this may not be reflective of a true connection between precipitation and discharge.

In our paper, Rx20d is used to characterise catchments hydrologically mainly with respect to seasonality. 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 3 days prior to the event are the most relevant; 5 days prior to the event precipitation is already increased (above the 75th percentile) but not extreme. There is a more systematiy study on this by Froidevaux et al. (205) 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. Hence, Rx5day should be a good choice.

We will discuss this paper in the revised manuscript and will be more clear that Rx20d is not used to characterise floods, but to characterise catchments according to their seasonal behaviour.

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., and Martius, O.: Flood triggering in Switzerland: the role of daily to monthly preceding precipitation, Hydrol. Earth Syst. Sci., 19, 3903–3924, https://doi.org/10.5194/hess-19-3903-2015, 2015.

Line 52: typo in controversy

Thanks

Figure 2 and 3: "All series are smoothed with a 30-yr moving average". It looks like a backward smoothing, not centered (the data reach well into the 2000s). Maybe it should be explicated. Is the type of moving window considered when identifying the three multidecadal periods of flood variability analysed later on?

Thanks for this comment. No, it is centered, but we had calculated the filter to the end of the series (when only 15 data points were available). In the revised paper we will set a minimum limit to 20 available data points. A special note will be added to explain our conditions (e.g., at least 50% of series must be available to aggregate regions, etc.).

Line 143-144: "selected from the 1x1° grid such as to best represent atmospheric processes relevant for the region)" is this based on some skill metric like correlation on some target? Some more words would help here, especially if in contrast with an alternative approach such as to spatially average the reanalysis data over several gridpoints.

The catchment sizes used in our study vary in size, so the averaging regions would be very different and shapes could be complex. Therefore, we opted for "point information", as for Rx5day. Furthermore, using 3 hourly atmospheric data on pressure levels from 80 members over a 210 year period requires a substantial amount of downloading, extracting and processing the calculation of entire fields. We did this only for selected grid points.

Line 171: check typo "for only for"

Thanks

Line 214: PMIP maybe worth to be explicated

We now write "Paleoclimate Modelling Intercomparison Project (PMIP)"

Line 232-233: Maybe this statement requires a bit more support. In my understanding, non-climatic anthropogenic influences on river runoff processes (e.g., river network changes, dams, etc.) may enhance/dampen multidecadal runoff variability or at least affect the autocorrelation of the discharge time series hence the detectability of multidecadal fluctuations above the red-noise background.

Thanks. This is correct. River works or new dams can induce step changes. One series (Rhine, Basel) was corrected for two such changes (we will ad another not at this instance). Step changes do affect annual peak streamflow, an example is the Upper Rhone, which was excluded from our study (Hingray et al. 2010). The paper however also shows that large floods remained unaffected. We will cite and discuss that paper in the revised manuscript.

Hingray, B., Schaefli, B., Mezghani, A. & Hamdi, Y. (2010) Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments. *Hydrol. Sci. J.* **55**(6), 1002–1016.

Line 253-254: for me it was somehow difficult to check this statement by comparing the figures, especially given the premise provided in the preceding statements. I see that the documentary versus observational evidence is not central, but in its current form this aspect of the study appears to be missing some elaboration, either in the text or as additional analyses (for instance, I was just thinking that some bivariate wavelet analysis may work here).

Thanks, that is a relevant comment. Yes, the agreement is hard to spot for individual regions, which may be due to the fact that the Blöschl data are an interpolation. Interpreting them locally may be misleading. As mentioned in the text, we focus on some salient features: the flood-rich decades in the middle and late 19th century in Central Europe, in the early 20th century in Northwestern Europe, the Europe-wide flood-poor period after 1950, and the recent increase in flood intensity. Therefore, in the revised manuscript we will show the comparison between peak streamflow series and the Blöschl

series in a spatial aggregation: Northwestern Europe (UK and Southern Norway) and Central Europe (all other regions). This will be shown in a Supplementary figure. These salient features do appear in both series; at this aggregation level the comparison is much better.

Figure S4: the red line looks more like brown?

We will check the colours.

Line 269-270: correlation of 0.21 appears rather low to me in terms of shared normalized variance, especially for smoothed/temporally aggregated time series, as I understand is the case here, which may contain a significant trend. Significance levels appear to be missing and should be provided, possibly accounting for autocorrelation of the series.

Yes, it is not very high (yes, the aggregation here was 4-yrs). We had used the 4-yr aggregation for reasons of consistency with the Blöschl comparison. However, we will replace this with correlation in the unsmoothed data (they are actually higher) and add more details on the correlation test (it is a t-test, $p<0.05$, since always at least one record has no significant autocorrelation, the test is valid and autocorrelation does not need to be accounted for). In addition, the significant correlations are now highlighted in italics in Fig. 4. The relation between Rx5d, peak streamflow, and CONV5d is further evaluated using cross-wavelet analysis. This reveals an association mainly on the multidecadal scale.

Figure 5a: how is the regional mean calculated? As I understand the calculation, as we move further back in the past, less time series contribute to the average, so this could lead to an inhomogeneity that can explain the discrepancy in the early period between time series. Possibly some illustration of standard error can reveal this uncertainty…

Thanks. We used all available series. Plotting the standard error is difficult as then three shadings would overlap. In the revised manuscript, we will use a consistent rule (50% of regions must have data). As commented above, we will additionally use bivariate wavelet analysis to corroborate the association.

Figure 6b: can regions with non-significant regressions be indicated (for instance through shading)? Both map show a clear imprint of topography, which might be related as well to different variances in seasonal temperatures, significance would help to illustrate such effect for the T-FPI connection.

Yes, we will do that in the revised manuscript. In the summer case, almost the entire domain is significant. In the winter case, the cooling in the Alps is significant as well as the warming in the East.

Line 322: cyclonic weather type centered where?

Over Switzerland (is now added)

Line 352-355: this may resemble a negative Eastern Atlantic rather than a negative NAO…

Thanks, this is a valid point. We will add this in the revised manuscript.

Line 376: where is significance shown? Hatching is mentioned in the caption but I do not see it in the figure (rather I see red contours, that may encompass regions of significance?)

Yes, the orange line show significance – the caption was wrong and is now.

Paragraph 3.4: This is another occasion when there is mostly a qualitative presentation of the comparison between different sources. Can this be improved?

Thanks. We will add the ensemble spread for each line to Fig. 6 to show that the circulation effects are outside the ensemble spread. We will also change the text.

Line 385: external forcing comes out a bit out of the blue here. Maybe some further elaboration would help.

The focus of the paper are not external forcings, but they are in the model boundary conditions, so we need to state that. We will rephrase this.