Reviewer #1 comments

Major comments

Comment 1: Structure
Please consider to move 'Study area' to the beginning of Chapter 2. Maybe make a new chapter called 'Study area and data' and a new separate chapter 3 'Methods'.

The difficulty is that the "Study area" section includes information from the datasets presented under "Data" (Figures 2 and 3). It would make sense, however, to divide into two chapters as suggested: "Data and study area" and "Methods".

Comment 2: Map Ripley's K
I am not sure why you do not (also) make the analysis on the grid-cell level of RhiresD. Such a map could be quite informative and provide more detailed spatial information. Such a figure would be better resolved as maps on sub-basin level.

RhiresD is given on a 2x2km grid scale, but its effective resolution is on the order of 15x15km (15km being the typical inter-station distance). Therefore, an analysis at the 2x2km grid scale might give a false sense of confidence. In addition, the clustering is more relevant for flood hazard at the catchment scale (100-1000km^2) rather than the grid-point scale (1-10km^2).

Comment 3: Figure header
I realized that it was a bit annoying to have to search in the figure captions what seasons the individual figure panels represent. Is it possible to add a header such as a) DJF? Or does this go against journal figure guidelines?

We agree and it doesn't seem to be in contradiction with figure guidelines. All corresponding figures were updated accordingly.

Comment 4: Scheme to explain methods
You use a lot of data and quite some different analytical tools. It was not easy to keep an overview on the individual numbers, windows, variables, watersheds, regions,... you use/calculate. Maybe consider making a scheme that illustrates your workflow. It would really help, I think. I do not think that such a scheme it is a requirement, but it would facilitate the readability of the manuscript.

Please see the updated Figure A1. As we already have many figures, we added it to the supplemental material.

Specific comments

L68-75 Maybe also say here that you also use different rainfall products.

Good suggestion, we expanded the second sentence: "*First, we aim to quantify the sub-seasonal clustering of precipitation extremes in time across Switzerland using several gridded station- and satellite-based datasets.*"

L34: flood hazard

Corrected, thanks.

L218: I do not really see this 'significant temporal clustering of precipitation extremes is generally found along the Alpine ridge' in Fig 4a or A1 b. Only a fraction of the Alps (as defined in Fig. 1b) shows significant signals. You also write this in the conclusion (L365). Please consider to be more specific here.

The clustering over the Alps in winter is less obvious in RhiresD than it is in the other gridded datasets, which indicate a wide region of clustering significance stretching along the Alpine ridge from southwestern to eastern Switzerland (Figure 5-a). Still, you are right to point out that the signal is not as clear in RhiresD. Thus, we suggest reformulating the sentence you refer to as follows: "*In winter, significant temporal clustering of precipitation extremes is mainly found in central Switzerland, along the Alpine ridge, at the 15-25 and 25-35 day timescales*". We also suggest reformulating the sentence comparing clustering in winter with the other precipitation datasets as follows: "*Clustering significance over the Alps in winter is also present in the coarser-resolution data, but with a wider extent than in RhiresD.*" For the conclusion, we suggest the following statement: "*Various station- and satellite-based datasets point to generally significant clustering over the Alps in winter, particularly their central part, and over Southern Switzerland during fall.*"

L134 and L184: Better do not refer to your result figures in the method section already. This was a bit confusing for me.

At line 184 we refer to the figures of Kopp et al. (2021), not to ours. At line 160 we refer to figures 2 and 3 which are not part of our results.

L212: Why 10 days? Can you add one sentence here why you choose 10 days? You use a lot of different windows (5, 10, 21,…), so it is important to give a bit more justification, I think.

Froidevaux et al. (2015) showed that while flood timing was mainly associated with extreme precipitation accumulations in the 2-3 days before the flood event, many floods were nonetheless associated with wet conditions (not "extreme wet") in the 4-14 days before the event. These longer-lasting wet conditions could possibly play a role in the persistence of high discharge, hence why we choose to start calculations 10 days before the beginning of

persistent high-discharge periods. We added the following sentence to the text: "*We choose to begin 10 days before because Froidevaux et al. (2015) showed that moderate wet conditions occurred in the week preceding many flood events in Switzerland, hence the need to look beyond the few days preceding persistent high-discharge periods.*"

L223: You actually do not compare the different rainfall products, right?

We limit our comparison of the different datasets to superposing their clustering significance maps and identifying areas of agreement.

L224: closing bracket missing

Corrected, thanks.

Fig A2: K displayed here is calculated as the average of of all K form the different data sets? Why? It would be interesting the see a comparison between the rainfall products? Does this make sense? Did you try this? Is it out of scope. Please justify.

The point of figure A2 is to give a sense of the average number of expected extreme precipitation events in the neighbourhood of any random extreme event (hence why we show the average Ripley's K value). But one shouldn't draw too much from it. Because the different datasets have different lengths, the same Ripley's K value will not be associated with the same level of significance in each dataset. Additionally, we limit our comparison of the various datasets to identifying regions where they agree on the presence/absence of temporal clustering significance. We do not seek to explain why a given region is significant in one dataset and not in another. This would go a bit far for the present study. Each dataset also has its own biases and could tend to over- or underestimate clustering significance in different regions. Looking at significance across the datasets therefore allows to highlight regions where they tend to agree, regardless of potential biases and data sources.

L225: Fig. 5 b is not white. Hence at least one, sometimes two or three datasets, suggest significant clustering. Please consider to rephrase this sentences as 'there a no signs' seems a bit strong.

You are correct, we need to qualify this statement. We suggest to reformulate as follows: "*Temporal clustering during spring is generally less significant across Switzerland (Figure 6-b). Two datasets indicate significant clustering locally in northwestern Switzerland, but none do along the northern and southern borders where significance was found in RhiresD.*"

Fig. 9 caption: It should be 'within a 10-day (respectively 20-day, 30-day) window […]'

Thanks, corrected.

Fig. 11: Can you explain a bit more what is in this figure. There only is a short description in the method section, but I think it is an important figure and would deserve another one or two sentences of explanation and description.

The figure shows the odds ratio of extreme discharge occurrence in the period during and up to 5 days after cluster events. The equation for the odds ratio is given section 3.4, but to make it clearer we suggest expanding this section by adding the following: "*The odds ratio compares the likelihood of extreme discharge occurrence in the presence of a precipitation cluster to its likelihood in the absence of a precipitation cluster. The higher it is, the stronger the relationship between the occurrence of extreme discharge and precipitation clusters.*"

L279: SST is sea surface temperature? Please also add a reference here.

You are right, we removed the acronym. Tuel and Martius (2021) discuss the relevance of sea-surface temperatures for temporal clustering at various timescales.

L314: Change 'strongly impact' to 'increase'?

Good suggestion, thanks.

L335: As you investigate on a seasonal level, lower runoff (compared to other seasons) is no explanation for low impact of precipitation clusters on discharge, I think. The main reason is snowfall, as you also mention. Rainfall just is solid and stored in snowpacks.

Precipitation extremes are investigated on a seasonal level, but not discharge extremes which are defined based on fixed percentiles. Thus, discharge extremes are by construction much less frequent in winter than in summer in snow- and glacier-dominated catchments.

L351: 'heavily biased': As far as I can see, you only investigate the hazard component. As you never had the ambition to assess risk in the first place, I would try not spend too much time in the discussion on this.

Yes, the point of this short paragraph is to underline that we only look at the hazard component of risk, not exposure or vulnerability. We may shorten it as follows: "*Our results only focus on the hazard component of flood risk. We do not take into account exposure and vulnerability, which may differ substantially between catchments due to variability in population density, infrastructure, flood management capacities, etc.*"

Conclusion: Maybe you can add a sentence on the other rainfall products used? Are they also suitable for such analysis...

The third sentence can be expanded as follows: "*Various station- and satellite-based datasets point to generally significant clustering over the Alps in winter, particularly their central part, and over Southern Switzerland during fall.*"

Reviewer #2 comments

Minor comments

l. 79 : do you have a reference for RhiresD dataset ?

The main reference for RhiresD is Frei and Schär (1998) (interpolation algorithm) which we cite already. We suggest expanding the last sentence of the first paragraph in the "Precipitation" section as follows: "A detailed description of this dataset can be found at https://www.meteoswiss.admin.ch/home/climate/swiss-climate-in-detail/raeumliche-klimaanalysen.html and the interpolation algorithm is described in Frei and Schär (1998)."

l. 84 : please specify the resolution of ERA5

The spatial resolution of ERA5 is 0.25°.

l. 96 : please specify the year of Muelchi et al.

We updated the reference with the year (2021).