Referee #2

The authors have conducted a study over 93 Swiss catchments that cover a range of elevations. They analyzed moderate annual and seasonal low and high flows under the influence of climate change (RCP 8.5), using a modeling chain comprising of 20 RCMs and one hydrological model (PREVAH). The high and low flows were evaluated based on (1) magnitude, (2) emergence, (3) changes in seasonality, and (4) frequency.

With careful revisions, I believe the article will be ready for publication. My comments below are organized by the respective section of the paper they are relevant for.

Reply: *Thank you very much for your thorough review. We very much appreciate your inputs and suggestions.*

Abstract:

From the text, it can be gleaned that the annual and seasonal moderate high and low flows were analyzed according to: (1) magnitude, (2) emergence, (3) changes in seasonality, and (4) frequency. However, this is not explicitly stated in the abstract – it is not until lines 79-81 within the introduction that the analysis is clearly laid out. I would suggest that the authors spend some time reworking the abstract so that the reader can very quickly understand what was analyzed.

Reply: *The abstract will be rewritten according to your suggestions and the suggestions raised by the other reviewer. Also, we will make clear that we analyze low and high flows in terms of magnitude, emergence, seasonality, and frequency.*

Line 13 – suggest removing the term 'downscaled' before regional climate models

Reply: *Will be removed.*

It would be nice to highlight the changing behavior of high flows per season (see lines 16 and 28/29, where high flows are referred to), as was done for low flows (see lines 15 and 26/27). If the climate model agreement is too poor to draw conclusions on changes to high flows per season, I think this needs to be explicitly stated in the abstract.

Reply: *We will discuss changes in high flows in more detail in the abstract and mention the difficulties to draw conclusions because of the weak model agreement.*

Introduction:

Lines 58-59 read, "high flows may also cause severe damages and significant costs. Hence, potential changes in high flows have to be integrated in water management and infrastructure planning, as well". It seems much more common that infrastructure planning is made resilient against very extreme events. Also, severe damages from flooding are more so associated with more extreme events rather than moderate extremes as defined here. I was hoping that the authors can either support their language more clearly and state how these examples are relevant for moderate extremes, otherwise I would suggest that the language be toned down.

Reply: *Thank you for pointing this out. We will rephrase this sentence toning down the language and add that moderate extremes are also important for water use and ecology.*

Data:

On line 89, the authors mention that 22 glaciated catchments were analyzed, but the number of catchments representing the other regime types considered are not provided. Suggest adding these numbers to make the text clearer.

Reply: *Numbers will be added and Figure 1 will be changed to make a clear distinction between high-elevation and low-elevation catchments.*

At least one sentence should be dedicated to why RCP8.5 was selected as the sole scenario, as opposed to the others available, and an explanation of what that pathway represents.

Reply: *The reasons for using RCP8.5 will be discussed in more detail. Several reasons led to the decision to use RCP8.5. For example, RCP8.5 is the worst-case scenario and changes in low and high flows are stronger than under other scenarios which makes the interpretation easier. Also, the number of simulations available in the Hydro-CH2018-Runoff ensemble differs between the RCPs which may affect some of the results. And last, a large number of simulations within an emission scenario increases the robustness of our results, particularly for the analysis of the time of emergence.*

Figure 1- please indicate blue shading is for water bodies. Also, see my first comment within the Results section below.

Reply: *Figure 1 will be changed according to your suggestions and the suggestions raised by the other reviewer.*

Methods:

The authors state that the time of emergence 'may not be stable' (line 141). Could the authors expand upon what they mean by 'stable'? Did the authors also find when the KS test is rejected repetitively? This indicator could seemingly be made more robust by requiring more than just one rejected KS test.

Reply: *The concept of time of emergence is based on a statistical test between two distributions (reference vs future). However, changes in runoff may not be linear over time. Even though changes in low/high flows are tested as significant in one period, they may not be significant in all periods afterwards (e.g., due to non-linear effects of enhanced snow melt, decreasing snow cover, increasing glacier melt and decreasing glaciation). We will include two figures in the Appendix showing the rejections of the p-value over time for low and high flows (see figure 1 below). Further explanation on this point will be made in the methods section.*

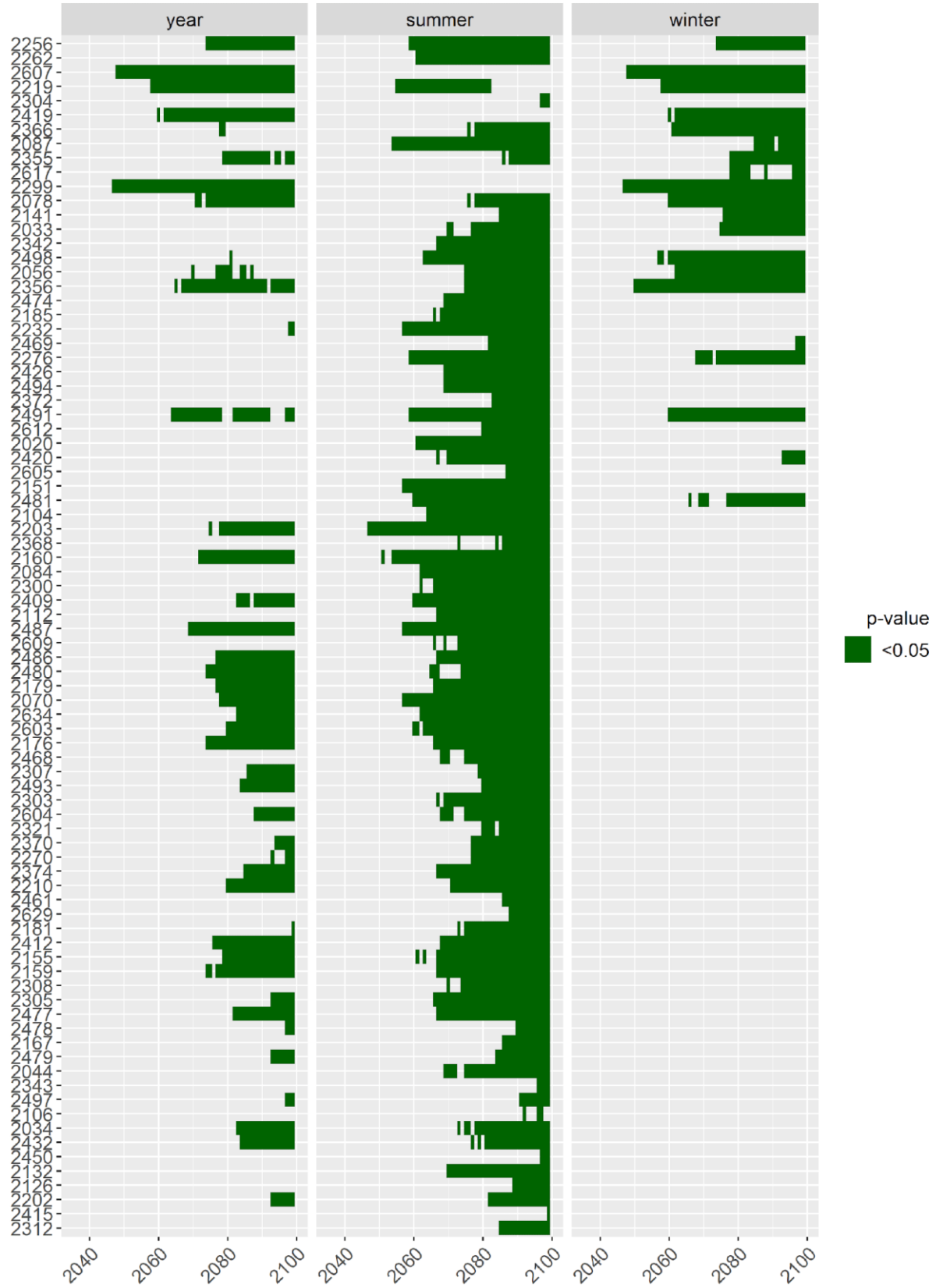

Figure 1: Temporal evolution of rejections of p-value for low flows.

The model agreement that you are highlighting for magnitude (Figure 3) and frequency (Figure 9) is 90%, whereas the model agreement highlighted for emergence (Figure 6) is 66%. Can you please explain your reasoning for highlighting different levels of model agreement?

Reply: *We chose these thresholds following the IPCC terminology where 90% refers to "very likely" and 66% to "likely". The threshold of 66% is used for the determination of time of emergence. The definition of time of emergence in this study relies on the definition used by Mahlstein et al. (2011). We will extend the description of time of emergence and add a statement on the threshold definitions.*

Mahlstein, I., Knutti, R., Solomon, S., and Portmann, R. W.: Early onset of significant local warming in low latitude countries, Environ. Res. Lett., 6, 3, 034009, 10.1088/1748-9326/6/3/034009, 2011.

Results:

The results are described in terms of low and high elevation regions, however the reader is not provided a clear picture or threshold of how the authors have separated these grouped catchments nor are these regions indicated within Figure 1 (please see for instance Figure 2 from Brunner et al., 2019 – Science of the Total Environment: https://www.sciencedirect.com/science/article/pii/S0048969719306576).

Reply: *Figure 1 will be changed according to your suggestions and suggestions raised by the other reviewer.*

For Figures 3, 4, 5, 6, 7, 8, 9 – would suggest replacing the word 'YEAR' with 'ANNUAL' since this corresponds to the terminology used in the main body text.

Reply: *Thank you for highlighting this inconsistency. "YEAR" will be replaced with "ANNUAL".*

Figures 3 and 9 indicate where models agree 90% of the time, whereas Figure 6 indicates when 66% of the models agree. Also, Figure 6 uses grey circles to show non-agreement, which is different from Figures 3 and 9, which use muted tones. What is the reason for designing these figures so differently? Where possible, suggest making the figure design cohesive.¨

Reply: *The content of figures 3 and 9 requires a different visualization than figure 6. Figures 3 and 9 indicate changes over time. There is a value (%) for the multimodel median for every catchment. However, not all changes are robust among the climate models. Therefore, we highlight catchments with a strong model agreement (>=90% of models) on the sign (positive or negative) of change with a black circle. We will also add a black circle indicating good model agreement in figure 6 (will be the same as in figure 3). Figure 6 shows the time of emergence and such a time of emergence cannot be determined for all catchments. For such cases we just use a grey color. This will be added to the legend.*

Discussion:

On lines 287-290 and 315-317 you offer some discussion of the significance of your work, but this comes across as very brief. In general, the discussion section is a good opportunity to substantiate the overall implications of this research. Indeed, this work is relevant for the agricultural and hydropower industries as you mention – I encourage you to make stronger statements about the relevance of your work and look for connections in literature. As an example, I highlight the relevance of your work to Switzerland's hydropower concessions, which are strongly influenced by

projected low and high flow behaviour – described here: https://hess.copernicus.org/articles/24/3815/2020/ and especially by the following authors (just an example of suggested authors):

Dr. Ludovic Guadard

Prof. Dr. Fanco Romerio

Prof. Dr. Hannes Weigt

Reply: *Thank you very much for this comment. The discussion section will be extended by statements on the significance of our study on other areas. Thank you also for the reference and the list of authors. We will incorporate some implications of our study on future research.*

On line 292: '…will decrease in the projections' should be changed to '…has been shown to decrease within the projections' or '…is likely to decrease in the future'.

Reply: *The sentence will be changed according to your suggestion.*

Conclusion:

Any mention of the greater implications of your work is generally absent from your conclusion section (please see similar comment in the discussion section above).

Reply: *See answer above. A statement on the implications of our study will be made.*