

We would to thank both reviewers for the helpful suggestions and comments. Our replies are included below in blue

- 1) L130: "Namely ..." This seems to be an incomplete sentence; suggest to revise. [done](#)
- 2) L807-819: There are a couple of explanatory references to Fig A3 (context of 'nival alpin' and 'nival meridional'), but most statements are not backed up by references to the illustration or wider literature, so the authors should insert respective references. [This section in the Appendix describes the results presented in Figure A3. We have added additional references to Figure A3](#)
- 3) Fig 5 (schematic) would benefit from some sort of axes labeling. [We have added axes labels to Figure 5](#)
- 4) I think the addition of Table 3 and the related text in response to Reviewer 2's comments is a very good idea and very helpful. I have a couple of thoughts on this:
  - 4.1) There is a nomenclature for satellite data that sounds very similar (see for example here: <https://www.earthdata.nasa.gov/learn/earth-observation-data-basics/data-processing-levels>). Users who regularly use EO data and are hence accustomed to this system might be confused by the one used by the authors. The authors could consider changing their nomenclature or to stress the difference, but maybe I am just overthinking this aspect. [We added a comment that the classes differ from the satellite data classification](#)

4.2) The levels stand for increasing processing effort. In both the caption and the main text, the authors relate this processing with a decrease in reliability by labeling L1 as 'most reliable'. I understand their intention, but I think that they are being unnecessarily harsh on at least their L2 output by this. After all, the tables in the appendix and the description of L2 and L3 in Table 3 mention the presence of strong assumptions where applicable. And in many cases of L2 (pure spatio/temporal aggregation) the processing does not constitute a decrease in reliability in my opinion as long as readers/users understand that for example a catchment average depends on spatiotemporal sampling in the respective station network. I do not have a clear recommendation but nevertheless encourage the authors to explore ways to weaken the link between data level and reliability for these cases. Maybe replace 'reliable' by 'objective'? [Changed the wording from "reliability" to "processing" level](#)

4.3) Table B4: I do not understand how a climatology or an anomaly can be L1 according to the authors' definition, unless these instances are directly contained in the input data used by the authors (in which case this could be made clear in section 4.5). Suggest to revise to minimum L2 or explain. [We adjusted the L1 - L2 levels to L2. The initial thought behind this was to keep a discrimination between climatologies derived from spatially interpolated fields via catchment aggregation and the direct measurement based climatologies derived from streamflow measurement stations \(Q\). But we agree that L2 is the only meaningful level also for these types of climatologies.](#)

Reviewer 2:

In the previous comments I suggested to replace "included stations" with "stations included". That created some confusion as to what the intent of the comment was and apologies for not being clear. It was simply on the word order. I think both are correct, but personally preferred the second as it reads easier. [Thank you for the clarification, done](#)

L38: I suggest to drop "drought" before propagation, as it is already mentioned [done](#)

L47: I suggest to drop "Novel" here as you later introduce HYD-RESPONSES as novel - and here

refer primarily to existing datasets if I understand correctly [done](#)

L55: Here I would replace "novel" with "unique" [done](#)

L59: Check the references - I do not think these should be in-line references here. [done](#)

L91: "hydrological Switzerland" is a bit confusing. Initially I wanted to suggest "for the hydrology of Switzerland", but the term is again used later in the manuscript. If it is a proper noun then it should be a capital H. I leave it to the authors what to do, but bear in mind it may be a term that not every reader immediately understands what it is (I assume it is the area that includes full catchment areas that drain into CH, even if for some these partially lie outside Switzerland. [done](#)

L103: There seem to be some catchments missing.  $94+56+9=159$ , which is less than 184. Also the percentages do not add up! Maybe you omitted the 8 very large ones - but that still does not add up to 184. Please clarify!

Thank you for pointing out this inconsistency. Indeed there are roughly two thirds of catchments falling into the small-to-midsize catchment size category. The other categories stay as they were. We adjusted the sentence accordingly:

**Roughly two thirds** of the catchments **64.7 % (n=119)** are small to mid-size with an area of between  $10 \text{ km}^2$  and  $500 \text{ km}^2$ . Only 9 (4 %) catchments are smaller than  $10 \text{ km}^2$  and 56 (30.4 %) catchments are larger than  $500 \text{ km}^2$ . The dataset contains eight very large catchments with areas between  $10000 \text{ km}^2$  and  $50000 \text{ km}^2$  (max. area =  $35'878 \text{ km}^2$ ), associated with the three largest rivers in Switzerland: Aare, Rhine and Rhone.

L130: I would suggest to replace "Namely stations which provide" with "These are stations that provide" [done](#)

L131: Stations are therefore excluded if they ... [done](#)

L248: The placement of Figure 6 and Figure 7 is a bit unclear to me. Are these now part of the main text or the appendix? if it is the latter then perhaps the numbering and referencing should be adjusted. Otherwise perhaps the position of the figure(s) should be adjusted. [Figures 6 and 7 are in the main section of the paper. Unfortunately the Latex rendering moves some of the figures to the end of the paper. We are positive that this can be resolved in the final editing phase.](#)

L267: the time series for the most recent gap-free time period (I think that reads easier) [done](#)

L357: Assume that restrict means the values are constrained to +3 and -3 and not replaced with missing values [done](#)

L458: Previously swallow holes were mentioned in relation to Karst areas. Should these also be mentioned here? [Thank you for point this out, we added the swallow holes](#)

L521: Suggest "allow for comparison across catchments" [done](#)

L616: Should FOEN not be mentioned here instead of MeteoSwiss twice [Mentioning MeteoSwiss twice is correct. We want distinguish between data sources from MeteoSwiss only and from MeteoSwiss and the SLF jointly](#)