General comments

The manuscript assesses the influence of temporal clustering of precipitation on discharge response. This is assessed with a forward and backward approach, for catchments in Switzerland and Europe. Temporal clustering of precipitation is an important factor in the generation of several natural hazards and a proper understanding of its influence on flood hazard, and natural hazards in general, is still lacking. In this regard, the manuscript provides novel and interesting results.

My main comment regards the identification of precipitation clusters:

- l.110 You identified clusters of precipitation events with a pairwise approach, categorizing each event depending on the distance in time from the previous closest event. The whole precipitation cluster is not identified, and consequently some events belonging to different categories may in reality be dependent, being part of the same precipitation cluster. I imagine that some events in a specific category may be preceded by other precipitation events and the whole story may influence the final discharge characteristics, rather than only the last two events of the cluster. Does this occur in your datasets? How much do you think this may affect the results?

- l.99 and l.114 Have you tried also lower precipitation quantiles or different pairs of precipitation and discharge quantiles? Are the results sensitive to this choice? Also (l.117) I do not think to be meaningful to chose discharge quantile starting from the chosen quantile of single precipitation events. Mainly considering that you show that extreme discharge is more often caused by precipitation events close in time than isolated.

Regarding the structure of the manuscript, I find it in general well-structured, with some exceptions reported in the technical comments below.

Specific comments

- l.77 Why did you not operate the same selection of catchments used for the Switzerland dataset (human influence, lakes, stationarity of the series . . .)?

- l.145-147 From this section it seems that all combinations of periods here reported are considered as persistent high discharge periods, included the pair (10, 1). However, also reading l.235, this is a non persistent period, right?

- Fig.11 I did not understand what is reported with cluster frequency.

- l.250 Why did you not use a measure, normalized for L, for example, that is comparable between the different pairs of (N, L)?

- l.293 Have you considered also the possibility of using rainfall rather than total precipitation?

- Fig. 13 Why did you choose an absolute precipitation magnitude rather than a magnitude relative to the quantile in each specific catchment (i.e. the excess over the threshold)?

Technical corrections

- l.69 *these catchment* → these catchments

- l.70 *The data is* → The data are

- l. 71 *We selected catchments among all available ones based on several criteria* → I would rewrite it like this: "Among all available catchments, we selected a subset of them based on several criteria:"

- l.76 *Daily discharge data for Europe comes from the Global Runoff Data Center dataset (GRDC).* → The second dataset consists of daily discharge data for Europe and it comes from the Global Runoff Data Center dataset (GRDC).

- l.92 *yielding two precipitation datasets* → I would remove this. You had two precipitation datasets also before.

- l.109 *events which occurred between n-1 and n weeks after another event are put into the "n-week" category, where n ∈ {1, 2, 3, 4, 5, 6, 7, 8}.* → I would substitute *another event* with *the previous extreme event*, this to make it clear that you are selecting the smallest window, or otherwise rewrite it more similarly to the explanation in the caption of Fig. 2.

- l.122 I think this part not to be reported clearly. In particular, I find misleading that you write that the probability is averaged across all extreme precipitation events belonging to a clustering category. Also, I would explicitly specify that the days (60 and 30 days) are the ones after each precipitation event. This is what I would probably write, but feel free to write it differently:

  For each catchment, clustering category, and for each of the 30 (Switzerland data) or 60 (GRDC data) days following extreme precipitation events, we calculate

  1. *daily discharge percentiles, averaged across all extreme precipitation events in each clustering category;*
  2. *daily high discharge probabilities;*
  3. *daily high discharge odds ratios.*

- Sec. 3.1.1 I believe that it is not always clear if the comments reported refer to all the catchments or only low/high elevation ones. In ll 172-187, for example, you are referring only to Fig. 3, that collects results of low elevation catchments but at the same time you are comparing it with Fig. 5, that, if I understood correctly, reports results for both subsets.