Comments from Referee #2

We thank the reviewer #1 for his/her thorough comments on our manuscript. Please find below a detailed feedback to individual comments and questions. All reported typos will be accounted for in the revised manuscript.

Major points

Methodology in general is lacking a lot of key information:

- Did you take the snowfall variable directly from the RCMs or how?
- What exactly was ADAMONT used for and how?
- What did you use S2M for: as past reference, or only in the bias adjustment? It seems like you concatenated S2M with a GCM-RCM, but please correct me if I'm wrong. And if yes, why did you not use the historical data of the GCM-RCM instead? I would expect some kind of break when merging a reanalysis and a climate model.

These three questions were answered in the second paragraph of our Data section. In the revised manuscript, we will add the following explanation:

"Annual maxima of daily snowfall are computed in three steps. First, daily precipitation from the EURO-CORDEX project (Jacob et al, 2014) are downscaled and corrected at the massif-level and every 300 m of elevation using an advanced quantile-mapping method (so called ADAMONT, see Verfaillie et al. 2017). This ADAMONT method relies on the S2M reanalysis as reference and is applied separately on four seasons and four weather regimes. Then, daily precipitation is disaggregated at the hourly time step using analogues and partitioned between rain and snow with the threshold 1°C and an additional quantile-mapping correction is applied (see Verfaillie et al., 2017, for details). Finally, daily snowfall are computed by aggregating hourly snowfall and the annual maxima of the daily snowfall values are computed (annual values taken from August to July, so as to cover the hydrological year)"

Therefore to answer the initial questions of the referee: We did not concatenate S2M with GCM-RCM. For each GCM-RCM we consider both historical data and the data for the scenario RCP8.5. We do not consider the snowfall variable from the RCMs, but the precipitation from the RCM. First, this precipitation is downscaled at the massif-level and every 300 m with the ADAMONT method with the precipitation from S2M as reference. Then, hourly precipitation data from ADAMONT is partitioned into rainfall and snowfall using a threshold at 1 degree celsius for the transition from rain to snow. Finally, annual maxima of daily snowfall are computed. But in the historical period these annual maxima differed with annual maxima of the S2M reanalysis, because the chronology of meteorological events differs between climate model simulations (RCM) and the S2M reanalysis. This is the reason why in our non-stationary GEV model we rely on annual maxima of snowfall from S2M as reference.

GEV approach:

- I would have expected a few more references on GEV in your description. For readers it would be useful to have an overview paper for general non-stationary GEV, which is what most people use. Then an adaptation to GEV with covariates (what you call non-stationary, but in principle this works with any covariate) and has been developed years ago.

  Yes we agree, and do not claim to have developed the adaptation of GEV to covariates. For clarity, and for readers of the Cryosphere, our objective was to present our non-stationary GEV model as concisely as possible, and point interested readers to our previous article for a more thorough understanding of the non-stationary GEV model. In the methodology, we will add a citation to Coles 2001 (and reference therein) which is a reference book for non-stationary GEV. In climatology non-stationary GEV have been used extensively, e.g. (Maraun et al. 2009). The originality of our methodology is to consider a non-stationary GEV model fitted both on GCM-RCM pairs and reference data.

- Your application to GCM-RCM pairs seems like a prime example of using a random effects specification to reduce the number of parameters while accounting for model variability and interdependence. I'm not up-to-date on recent GEV literature and developments, but I guess there should be some random effects GEV models (not necessarily in your specific field). Would be nice to see this as a discussion point at least.

  Yes, that's exactly the point of our GEV model. In the introduction and the discussion of Le Roux et al. 2022, we reported articles in climatology that are using similar techniques (additional and/or multiplicative random effects) for GCM-RCM ensemble. For instance, Brown et al 2014 estimated a non-stationary GEV distribution with observations and a single GCM-RCM and introduced additive coefficients for each GEV parameter. Other authors considered multiplicative coefficients (Hosseinzadehtalaei et al., 2021). To the best of our knowledge, we do not know any comparable GEV models outside of climatology. The idea to rely on random effects stems from the ANOVA framework, which have been applied to partition the uncertainty of GCM-RCM projections (Hawkins and Sutton, 2009; Evin et al., 2019).

- Since almost no natural physical process follows a piecewise linear function with abrupt breakpoints, I wonder why you chose this method? What is the benefit to more flexible approaches like a spline basis, which has a similar number of parameters but does not depend on choosing breakpoints and number of breakpoints (which you note as limitation in the discussion)? Then, even if you did some validation in your snow load paper, I would expect a goodness-of-fit test also here, since snowfall might behave differently than snow load. So as to justify why a piecewise linear function is fitting better than a simple linear model (or one without covariates, by-the-way).

In the literature, most non-stationary GEV models rely on linear functions of the covariate. Thus, such models cannot model a more complex trajectory, e.g. increase of mean maximum followed by a decrease in mean maximum. In order to allow a modeling of more complex behavior, while having as few parameters as possible, we decided to rely on piecewise linear functions, which is the same as degree 1 splines. In other words, we choose to rely on the most simple function that is capable of approximating more complex behavior. We will add goodness-of-fit tests with the selected models, simple linear models and constant models.

- Have you assessed the shape parameter of your GEV functions? The shape parameter decides on Gumbel, Weibull, or Frechet type, and each of these have a very different behaviour on tails and expectance of extremes, and your use of adjustment parameters might change the type of GEV in the future – so I was wondering whether you have analyzed this also?

  No, we did not analyze this parameter. We did not display this parameter because our objective was to provide an article accessible to readers that do not necessarily have a strong background in statistics, and because this parameter can be difficult to interpret. In the revised manuscript, we will add in the Result section sentences on the shape parameters based on Figures (maps and/or trajectory of the shape parameter with GMST) that will be located in the Supplementary Materials.

Annual maxima from daily values (? I guess, this is not specified in the manuscript) are notoriously unstable. This deserves some discussion on the challenges in using annual maxima of snowfall, which already suffers from high spatial and temporal variability.

In the Discussion section on the data, we will add the following sentence “Snowfall amounts at high elevations are thus more uncertain, and an underestimation of precipitation above 3000 m can be suspected (Vionnet et al, 2019). Annual maxima of daily snowfall are likely affected by these uncertainties and by the high spatial and temporal variability of annual maxima from daily values.”

Finally, from a user and climatological point of view, I would say that besides maximum 1-day snowfall, 3-day and 5-day snowfall are of particular interest, even more than the 1-day one. In fact, many assessments of future climate change focus not only on 1day annual maxima but also on 3, and 5 day precipitation extremes (cf. Rx*day indices from ETCCDI). This is beneficial in multiple ways: First it does not introduce an arbitrary break at time=0 for the accumulation of daily values. Second, it also considers changes in circulation regimes, such as increased persistence. Third, problematic recent snowfall extremes in the Alps were often multi-day events. This study could benefit strongly from also including 3 and 5 day extremes in the analysis, making the results both scientifically more robust and with more societal relevance.

We agree with this remark. We will include in the revised manuscript a brief analysis of maximum 3-day snowfall which is, e.g. the most relevant variable for avalanche risk assessment.

Minor points:

- L58: I suggest not calling a reanalysis an “observation”, not even for simplicity. This can be very misleading for a reader who does not carefully read every detail.

  In the revised manuscript, we will remove this sentence. In the rest of this article, we will replace the term “past observations” with “reference data”.

- Fig1c: Smoothed, how exactly? Please indicate.

  In the revised manuscript, we will add in the Data Section of the revised manuscript that smoothing was applied using cubic splines.

- L127: missing sentence?

  In the revised manuscript, we will reformulate this sentence as “For the mean annual maxima, relative changes are often larger (increasing trends are greater, decreasing trends are less marked) in the Northern French Alps than in the Southern French Alps.”

- L129: -40% is not a substantial change?

  We will remove this sentence in the revised manuscript. This sentence was only meant to explain that changes were not as high as we expected.

- L205: “huge step forward” is not a justified conclusion. Just using a more complicated model does not automatically imply better results. To justify this sentence, I’d expect some tests of goodness-of-fit (see also related comment above).

  In the revised manuscript, we will remove the sentence mentioning a “huge step forward” from the discussion. Yes, as explained above, we will add some goodness-of-fit.

References:

Maraun et al. 2009, The annual cycle of heavy precipitation across the United Kingdom: a model based on extreme value statistics, https://doi.org/10.1002/joc.1811

Hawkins, E. and Sutton, R. 2009, The potential to narrow uncertainty in regional climate predictions, https://doi.org/10.1175/2009BAMS2607.1

Brown, S. et al. 2014 Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases, https://doi.org/10.1007/s00382-014-2080-1

Evin, G et al. 2019, Partitioning Uncertainty Components of an Incomplete Ensemble of Climate Projections Using Data Augmentation, https://doi.org/10.1175/JCLI-D-18-0606.1

Hosseinzadehtalaei, P et al. 2021, Climate change impact assessment on pluvial flooding using a distribution-based bias correction of regional climate model simulations, https://doi.org/10.1016/j.jhydrol.2021.126239