

This manuscript presents a thorough and novel analysis of rain-on-snow (ROS) events in the French Alps. More importantly, it identifies direct links between ROS and torrential floods, floods, and landslides, improving the current understanding of ROS-associated impacts. The analysis presents sufficiently high accuracy with the RTM database, and overall, I think that the authors have presented an important contribution.

Below I have included several comments, per section, most of which are requests for minor changes in figures and rephrasing in the text. However, I also have two major concerns:

- First, I am curious as to why authors used massif features instead of watershed features. This might seem counterintuitive to readers, for a study on hydrological impacts. I would think that watersheds are more appropriate features for studying floods, but also landslides, as they may increase the debris input in the hydrographic network, leading to infrastructure damages downstream. I think this choice needs to be addressed/explained better, especially since the authors do refer to watersheds in a couple of occasions in the manuscript.
- Second, while the authors account for total rainfall amount (as well as duration) of ROS events, they don't seem to consider at all the underlying snowmelt contribution. As a result, I'd be more cautious in attributing a flood to ROS, instead of just rainfall, since the contribution of the snowmelt fraction of runoff (which is even accelerated during the ROS event, e.g., Yang et al. (2023)) is unknown. To me at least this appears to be a substantial limitation, which I am sure the readers would like to see addressed, or at the very least discussed.

Since the snow reanalysis used provides daily SWE data at 6am UTC (L111), I imagine that snowmelt could have been easily approximated. Even following a very simplistic approach, the authors could have derived snowmelt rate by averaging gridded SWE_{daily} values across the massif feature, calculating the daily ΔSWE , and dividing it by 24h (making the assumption that snowmelt rate is homogenous, and regrettably but understandably disregarding the diurnal melt cycle). This could then be integrated in the subsequent analysis.

Yang, Z., Chen, R., Liu, Y., Zhao, Y., Liu, Z., and Liu, J.: The impact of rain-on-snow events on the snowmelt process: A field study, *Hydrological Processes*, 37, e15019, <https://doi.org/10.1002/hyp.15019>, 2023

Introduction

Minor

- L18-20: Are the authors describing all ROS events, or maybe just the impactful ones? Would there be a different definition for ROS events where the existing snowpack is not yet close to saturation? I think that that definition given by the

authors in L69 is more appropriate and therefore I would encourage a slight rephrasing here.

- L30: This may be a little early to mention “In our study area” as it hasn’t yet been explicitly defined. The authors may want to consider rephrasing to “In the French Alps” or similar.
- L38-40: Although reanalysis solves the temporal resolution problem, the authors may wish to – at least briefly – discuss the limitations of coarse gridded products, including underestimations in precipitation amount or their inability to often correctly resolve the precipitation phase due to different representation of orography (and therefore air temperature). The authors describe this in L62-65, but I think it’d more fitting here instead of at the end of the introduction when the study’s aim and objectives are outlined.
- L60-62: I would consider removing this sentence, and rely on stating the study aim and objectives (as you do later in this paragraph) instead of having the readers assume them here.
- L66: The RTM acronym should be written out in full in its first mention.
- L72: Could the authors please provide some more clarity as to what is the difference between “Impactful” and “Heavy-rainfall” ROS events? Are the latter also impactful? In the following lines (L73-76) only “all” and “impactful” ROS events are mentioned – “Heavy-rainfall” ones aren’t – causing confusion as to what is the definition, and purpose, of that last category.

Data

Major

- L85: My main concern regarding the study area and subsequent analysis, is that uses massif features instead of catchment features. This seems like a counterintuitive choice for an analysis of a hydrologically relevant topic, such as flooding (as linked with the RTM database later). While I would strongly encourage this, I understand that changing this now would take a certain amount of effort & time. If the authors chose to keep their analysis at massif scales, then please provide an explanation regarding this choice, as well as its possible implications for the results.
- L104-106: It is unclear if the snowpack is calculated with or without taking into account the aspect, which would have a substantial effect on the snowpack properties due to different shading & exposure to incoming shortwave radiation, affecting its melting, which would be highly relevant to the prediction of a ROS event.
- Section 2.3: At this point it is still unclear which S2M properties are validated against the observations, as well as whether the S2M reanalysis was assimilated

with any of the same weather stations, rendering them unfit for use in the validation stage. Shouldn't the validation happen against an independent dataset? This point is also addressed by the authors at the beginning of the results section (L247-250) but it would be better if it were moved here or in the methods section.

Minor

- L106: Please add a comma in "[...] was used, as the RTM [...]". Also, the use of the word 'flat' is somewhat confusing, and I would encourage a slight rephrasing.
- L148-161: The authors do a very good job in outlining all sources of potential bias to the dataset. Although later in L208-209 it is mentioned that the dataset is only used after more systematic reporting began, I think this information should be moved to here, in order to make it clear from the start. The authors may want to consider also outlining here any other pre-processing steps taken toward limiting the RTM dataset bias.

Methods

Major

- L181-182: Once again, I am puzzled about the authors' choice to not analyse the data per hydrological catchment, and I think this requires further explanation in the Section 2.1. Even in L212-213, the authors point out that impact may have been reported on watershed scales, but without explanation, they analyse data in massif scales.
- L235: Which statistical test was used to calculate the trends, and which test was used to compare the two 30-year periods? This should also be made clear when reporting results later in the manuscript (especially in figure captions) to improve clarity. In the results section later, I have commented on the authors' somewhat misleading use of the word 'trend' when comparing absolute numbers between two periods, instead of running a formal trend analysis across the time series (e.g., Mann-Kendall, Theil-Sen).

Minor

- L168-171: The process described here by the authors became clear to me after re-reading this a couple of times. I would kindly encourage them to rephrase this section in a simpler way, in order to improve the flow for the readers. Moreover, in the part about the "moving-window", please explicitly state the length of the window. It would also be helpful to add a clarification in the part about the 12h separation between events regarding the absence of rainfall: e.g., "[...] a minimum separation of 12 hours with no rainfall occurring was used to separate different events", or similar. I would also add this clarification in L177.
- Fig. 2: On top of the changes requested above, the addition of a horizontal line between the two events indicating that the window is >12h and therefore the

events are separate, would also help readers understand the ROS event definitions better.

- L183-193: I think that this entire section could benefit from some rewriting/rephrasing to improve flow.
- L195-196: I imagine that the letters “T”, “I”, and “G” attributed to “torrential floods”, “floods”, and “landslides”, respectively, are taken from the equivalent terms in French? However, as they are less intuitive for non-French speaking readers, could you please consider changing to the first letter of the respective English term across the entire manuscript?
- L220-221: Could you please expand further on how this approach accounts for the potential extent of the contributing watershed, as it is currently unclear to me?
- L229: Could you please provide some further detail and explanation on how you came about selecting the 50% threshold and what this exactly means for the subsequent analysis?
- L232: They have been referred as such since L72, so the word “hereafter” is not necessary.
- L236: Please replace “yyyy” and “yyyy+1” with “ yr_i ” and “ yr_{i+1} ” respectively.
- L240: It is more standard to say “[...] if p-value > 0.05” or “[...] if confidence level is lower than 95%”.

Results

Major

- L330 & Figure 8: The results do appear very clear on the boxplots, but it would be a good idea to also perform a statistical test here, to test for significant differences in the means/variability between all and impactful ROS, for rainfall amount and event duration respectively.
- L368: The authors need to rephrase the “downward trend” part. As far as I understand it, Fig. 12 doesn’t report results from a trend analysis (e.g., Mann-Kendall or Theil-Sen), which should be reported as a function of time (x decline/increase per year/decade) but a comparison between two periods (reported in absolute numbers). Please correct this everywhere where it’s needed, across both text and figure captions (12,13,14,15).
- I think there is a small (& hopefully easy to fix) discrepancy in the hatching (statistical significance) of Figures 13 and 15. Take for example snow cover at the lowest elevation band, in August or July. Is there (or was there) snow cover at that time of year at those altitudes? If not, it might be worth making sure that

your code correctly filters data for your prescribed SD & SWE thresholds and excludes them from statistical analysis. These can then appear as 'greyed out' areas on the graphs to improve clarity. Same goes for the ROS panels. This might once again be an issue caused by the choice of statistical test, a mention of which I am yet to come across in the manuscript...

Minor

- L242-250: This section describes the comparison between S2M and weather stations, and should therefore be moved to the Methods section as no results are yet reported here. However, given that the dataset is not independent, it is unclear to me, what the purpose of this comparison is, and if the sensitivity analysis is indeed necessary. Afterall, wouldn't the temperature and precipitation fields have been modelled across the elevation gradient originally for S2M, using these very observations? How is it expected that the point-scale observation (which would be at a specific elevation and aspect) would be representative for all elevations and aspects of that massif?
- L267: "Temporal biases [...] over time." – I think this should also be moved to (or at least mentioned in) the methods, as I don't think it currently is.
- L290: Here it would be very interesting to know the % coverage of those elevation bands.
- L309-314 & Figure 6: Here the authors present very interesting results on the overall links between ROS events and impact on the ground. I am wondering if the authors had the opportunity to also investigate results separately per category (torrential flood, flood, landslide), to see whether they all follow the same temporal distribution. I really think it would be a very valuable addition to include in the manuscript (including updating figure 6 accordingly).
- L320: I am wondering why Maurienne massif has the highest rate of impact from ROS, considering it doesn't have the highest rate of ROS events. I would be interested to see the authors' thoughts on this disconnect.
- L340: "as announced" → "as described"

Appendices

Minor

- Figure C1: It would be useful to have maps of means for other variables too as a reference point (e.g., snow cover duration etc.)
- Figure C2: Please add the units (hours I imagine?) on the y-axis or figure caption.