Review for manuscript "Natural climate variability is an important aspect of future projections of snow water resources and rain-on-snow events"

Authors: Michael Schirmer, Adam Winstral, Tobias Jonas, Paolo Burlando, and Nadav Peleg

Journal: The Cryosphere

Summary

The authors use a model chain consisting of climate models, a weather generator, and an energy balance snow model to identify dominant uncertainty sources in future changes in snow-water-equivalent and rain-on-snow runoff. They show that changes in ROS events emerge till the end of the century despite large uncertainties while ROS events with substantial snowmelt contributions don't show a clear change signal.

General remarks

The study by Schirmer et al. builds on a complex model chain consisting of climate models, a weather generator and an energy balance snow model to assess the importance of internal variability on the detection of future changes in snow and rain-on-snow runoff events. I think that the combination of different model types to better describe internal variability is a generally a valid approach to determine the importance of internal variability in change assessments of snow-related quantities compared to other uncertainty sources. However, I see a substantial need for clarification regarding the research questions and methodology and think that the approach chosen to decompose uncertainty into different contributors needs refinement. Given the current 'incomplete' methods descriptions, it is difficult to assess the validity of the results. Furthermore, I think that the manuscript would profit from reorganization, i.e. restructuring the methods section following a more logical sequence and from separating the results from the discussion. Finally, the manuscript would in my opinion profit from a visualization of the most important modeling steps and their relationships and from refining figures by adapting color schemes and adding legends. Please find my more detailed review below.

Major points

1. Research questions: the research questions are not entirely clear and should be explicitly stated in the introduction. From how I understand the study it is something along the lines of: 'How does the importance of internal variability differ between temperature-driven snow resources and rain-driven rain-on-snow events' and 'When is the time of emergence of changes in snow availability and rain-on-snow runoff.'
2. Introduction: In addition to model-based studies looking at changes in rain-on-snow floods, there are also observation-based studies, which I think should be mentioned in the introduction. E.g. Sikorska and Seibert (2020; 10.1080/02626667.2020.1749761) or Cheggwidden et al. (2020; 10.1088/1748-9326/ab986f).
3. Methods section organization: The methods section does not seem to follow a logical order and could in my opinion be more logically organized by following a 'chronological' modeling order. E.g. Area, Climate models, Weather generator, Snow model, Rain-on-

snow definition, Change assessment, Uncertainty decomposition. Providing a flowchart linking the most important modeling and analysis steps might enable further improvements in communication. Furthermore, the methods section lacks important methodological detail, which makes it difficult to assess the validity of the results.

4. Glacier retreat: The study region is influenced by a glacier, which affects runoff formation. However, the glacier-related changes in flow are not represented in the modeling chain (l. 86-87). This does not seem to be justified and might explain why melt-influenced changes in ROS events are don't show up clearly.
5. Weather generator: The weather generator description (Section 2.3) lacks important detail and it is therefore difficult to assess the validity of the approach. E.g. how does the weather generator use the climate simulations, how does the weather generator work, how is the temporal downscaling performed (l. 122), how are the different variables generated (l. 122), how is the inter-variable consistence (l.124) evaluated?
6. Snow model and variables: It remains unclear to me how the weather generator output is used to derive different snow-related variables (Section 2.2). Was the analysis performed per grid cell? Which variables were exactly derived? How was the model calibrated (l.95)? And what does the 'unpublished' model adjustment (l. 95-96) do?
7. Bias correction: Section 2.5. suggests that some bias correction might have been necessary to adjust simulated to observed values. Was such bias correction performed and if so why?
8. Uncertainty partitioning: The uncertainty partitioning procedure described in Section 2.6. does not seem to properly separate internal variability (residuals) from the signal. Or at least I can not see how the different uncertainty components have been decomposed e.g. using a procedure such as the one proposed by Hawkins and Sutton (2009; 10.1175/2009BAMS2607.1). The procedure used to derive climate model uncertainty also seems to encompass internal variability (l.153-154) and the procedure used to derive internal variability also seems to include climate model uncertainty (l.156-157). Furthermore, it would be nice to compute fractional uncertainty contributions that add up to 1, which currently does not seem to be the case.
9. Validation: I think that the methods section needs a 'Validation' subsection describing how the different models were evaluated. E.g. how were the validation stations chosen? Which variables were validated, …
10. Rain-on-snow events: how have these events been defined? There is a section called 'rain-on-snow' definition, which does, however, not really explain what you understand by a 'rain-on-snow' event. How is the 'surface water input' computed?
11. Results: I would clearly separate the results part from the methods section and discussion. Some parts can be moved from the Results to the Methods section (e.g. l. 206-216) and other parts to a newly created Discussion section (essentially everything that compares the study's findings to findings of existing studies). Furthermore, it would be nice if the results section followed a similar structure as the methods section.
12. Figure 11: Would be nice to depict the fractional contributions of the individual uncertainty sources. Currently, all of the sources seem to not be clearly separated (probably also related to the issue risen in comment 8).
13. Discussion: I miss a discussion of alternative strategies that could be used to quantify the relative importance of internal variability to other uncertainty sources besides weather generators. Recently, quite a few studies have used single model initialized large

ensembles (SMILES) for uncertainty decomposition and I think it would be important to mention this (e.g. Maher et al. 2021; 10.5194/esd-12-401-2021 or Lehner et al. 2020; 10.5194/esd-11-491-2020). Other topics I would address in the discussion section include: model deficiencies, comparisons of results with other studies (as distributed throughout the results section), and the generalizability of the findings to other geographical and climatic contexts given that the study relies on one catchment only.

14. Figures: Figure design should be improved by using continuous color schemes for continuous variables (e.g. Figure 1c) and by avoiding the use of rainbow color schemes, which are not color-blindness friendly (Figure 8a and 8b). Furthermore, complete legends should be provided for all figures.

Minor points

- L. l. 40-41: I don't think that it is correct to say that 'no studies have previously included internal climate variability in their analysis' as internal climate variability is per definition part of every change impact assessment. However, it might be ok to say that 'the relative importance of internal variability compared to other uncertainty sources has not been previously assessed.'
- L.58-59: needs rephrasing.
- L. 61: 'they' will
- L. 63: that 'drive runoff'
- L. 72-73: repeats content already provided in l. 68-69 and can be removed.
- L. 140: How do you define 'frequency', number of events per year or non-exceedance probability or anything else?
- L. 259-260: Not sure whether the statement that 'the uncertainty range for current climate is, by definition, only determined by natural variability' is true. If the simulations for the current period are run with different climate models (which is as I understand it the case), climate model uncertainty might also be present.
- L. 279-280: Don't understand this sentence and think it needs rephrasing.
- L. 293: Sentence seems incomplete. Would be nice to repeat the pixel based criteria, which the reader might no longer have present at this stage.