Assessing spatially distributed snow simulations with MEB-Crocus in subalpine forests through modelling experiments – *Mazzotti et al., submitted to HESS*

Reply to REVIEW 2

Mazzotti et al. present novel results from a comparison of two operational snow-hydrological models. The model of interest, Crocus, was experimented with using different land cover datasets and parameterizations. Differences in model outputs, relative to the benchmark model FSM2oshd, are used to attribute the sensitivity of Crocus to these modelling decision. The study provides an important contribution to hydrological science by quantifying the relative influence of land cover data and model parameterizations on model sensitivity. I have some comments for the authors to consider, regarding justifying the benchmark model, better describing the parameterizations, quantifying model differences, and some reworking of the text.

Thank you very much Dr. Alex Cebulski for the positive assessment of our work and for the detailed comments. Please find our answers below each comment in blue.

General Comments

1. Evaluation of the benchmark model

While those familiar with snow-hydrology modelling will generally know FSM2 to have impressive model scores, the average reader might need some evidence to back up the use of FSM as a benchmark model. I think some minor reworking to the introduction or methods section could help support the use of FSM as a benchmark.

We have extended the justification of our use of FSM2oshd in Section 2.2, including some results from the existing evaluations of the model (Haagmans et al. 2026, Mazzotti et al. 2021, 2023).

2. More details on the model parameterizations

It would be helpful to know how the snow interception and ablation processes were represented in the models. Since the study attributes some of the difference in model performance to these parameterizations, interpretation of the results would be improved by providing some context on these equations, especially since the study referenced with these equations is not yet published. I see some of these equations are in the supplement, but are incomplete (i.e., the initial snow interception parameterization is missing). I would recommend listing the equations for the MEBCroXT_all setup in the manuscript, if length permits, and then describing how these differ from MEBCro_def and FSM. Based on this, in the discussion it would be interesting to link the interpretation of your results (i.e., Crocus to overestimate SWE in denser forest) to the process representation differences between the models. Related to this it would be revealing to have a figure diagnosing canopy snow ablation terms (i.e., cumulative sublimation, melt/drip) for the various models to better get at process representation/diagnosis. Some additional length needed for this description could be gained from removing some of the descriptions of statistics and other model details.

We have added a table contrasting the forest snow process parametrization for the two models and pointing to the relevant sources in the Methods section. The reason why we did not dig deeper into the individual canopy snow processes is that these were not as well validated at the process level for FSM as were the sub-canopy energy transfer processes. Mazzotti et al. (2020a,b) found that below-canopy accumulation patterns were generally well reproduced over the course of a season, but did not validate individual canopy snow mass balance processes such as sublimation and unloading individually. Benchmarking at the canopy snow process level thus seemed beyond scope, particularly given that such model variables are not available by default from either of the models. Furthermore, we should note that Pauze et al. (submitted, preprint to become available soon) have a strong focus on validating these processes in MEB-Crocus. We will point to this limitation of our study and to findings from Pauze et al. (2026) in the Discussion of the revised manuscript.

3. Transferability of results to other sites and climates.

The model was tested in Switzerland, however, the plan is to operationalize the model in France. How might the change in location of future applications of the Crocus model influence model performance? Are there differences in land cover type and climate that might influence model performance or transferability of parameters. Moreover, how might future climatic conditions influence model performance. For example, as forest snowfall regimes transition from being dominated by intercepted snow sublimation losses to more unloading and drip in a warmer climate. I know this isn't the main topic of this study, but it would be great to hear the authors thoughts on these ideas through the addition of a few sentences in the discussion.

Thank you for raising this point – we considered it when designing the study, but we agree it was not sufficiently addressed in the manuscript. Land cover type and snow climates are comparable in the French and Swiss Alps (see also Fig. 8 in Sturm et al. 2021; https://doi.org/10.1175/JHM-D-21-0070.1). You may notice there that our study site is representative of a montane climate, while other parts of the French and Swiss Alps are classified as maritime. The years of our analysis that feature low snow accumulation are indicative of this maritime climate. We have added these considerations on results transferability (in space and time) to Section 4.4, which addresses limitations of our model comparison study.

4. Quantifying model differences

Much of the paper centres around comparing different model results. Currently this is done quantitatively, quite nicely in the figures, but in the text I think it would be helpful to have some additional statistics (i.e., percentage difference) to support claims and/or be more specific regarding phrasing such as "approximately equal" see L492 below and others throughout.

This is a fair point. Quantitative statistics have been added in multiple instances of the Results section.

5. Conclusions

The concluding remarks in Section 5 regarding the attribution of land cover and parameterizations to model sensitivity differs from those given in the abstract. Some reworking of Section 5 could better emphasize these important results. This information is also hard to interpret in Figure 12, so some work on the specific comments listed below related to this could help better emphasize these results.

Thank you for this remark, which made us realize that our initial presentation of this result was misleading. We have reworked Section 4.1 and Figure 12 (see comments below) and ensured consistency between Abstract and Conclusions, with process parametrizations and land cover data presented as sources of model discrepancies of similar importance in both sections.

Specific comments

L47-49 - Much of this paragraph focuses on spatial distribution, but I think the key for this paper is getting both the spatial and temporal distributions. So some reworking of this paragraph could be helpful to mention both dimensions.

This is a valid point, we have slightly modified the paragraph to mention the temporal dimension as well (for instance, we now refer to latency of meteorological products and specify that these applications are in near-real time, at daily scales and throughout the entire season).

L58 - Some recent large-scale acquisitions now also available in western Canada (Bisset et al., 2026).

Bisset, R., Floyd, W. C., Menounos, B., Bishop, A., Marchenko, S., Fisk, G., Cebulski, A., Marshall, P., Heathfield, D., Beffort, S., White, R., Arriola, S. G., & Viner, N. (2026). Snow Water Storage Within Eight Pacific Coastal Watersheds in British Columbia (Canada) Inferred From Four Years of Airborne Lidar Data. Water Resources Research, 62(5), e2025WR040837. https://doi.org/10.1029/2025WR040837

Thank you for pointing out this relevant study to us, which was published after our initial submission. We have extended the statement to North America and included the reference.

L76-77 - I think its fine to use FSM2oshd as a benchmark as it permits this interesting model evaluation. However, I think other readers that are not familiar with snow modelling might need more convincing. I think some text here could be warranted to quantify the existing performance of FSM2oshd to justify the use as a benchmark.

See answer to major comment #1

L80 - I see Q1 was discussed in 4.1 which I think is an informative discussion, but it is not mentioned again in the conclusions, is this needed as a specific research question?

We would like to keep this as a research question as we believe it to be an important aspect of the paper. We now more specifically refer back to it in the conclusion: *'Only additional targeted model experiments allowed unraveling the origin of these model discrepancies. They present a possible approach to evaluating model performance based on comparison with another model.'*

Note that further down the conclusion reads: '*While these model experiments highlighted the difficulties associated with making such model experiments meaningful'*

L127 - Since a main theme of this paper focuses on how process representations influence the results, I think a description of the mass balance parameterizations for intercepted snow is warranted here.

This will be taken care of – see answer to major comment #2

L172-181 - See general comment #1.

See answer to major comment #1

L189 - The Quéno et al. (2024) study is not referenced again, consider whether this ref is needed here.

We have removed the reference.

L284-290 - These lines introduce the figures but do not provide interpretation of the results, consider whether this paragraph is really needed.

We agree that this paragraph was redundant with the figure legend, we have shortened and rearranged it.

L193 - Some more detail on the actual tree species would be useful here. Snow interception model development studies are limited to specific species. So additional information on the tree species here could help interpret the results.

We have added the names of dominant species.

L212 - Change word <Thank> to <Thanks>. Also consider numbering strategy here, starting with 1 might be more appropriate.

Has been corrected, and the numbering was replaced by bullets.

L243 - Is <defined as in Mazzotti et al. (2023)> needed here?

We wanted to highlight the consistency with our previous work given that the definition of SSD and PSD involves a choice of a threshold. We have however removed this reference as the threshold is now addressed in discussion (following a comment from the first review)

L259 - How is drainage of rainfall or canopy snow drip handled in the subcanopy snowpack? Is it included in the melt term? Please make sure this mass balance is valid.

Rain and canopy snow drip are routed through the snowpack and added to melt at its bottom (refreezing may occur). The mass balance has been checked.

L269-271 - <However, given the many dimensions...model results is indispensable.> Is this sentence necessary? I think the next sentence still reads okay without it.

We have removed the sentence as suggested.

L330 - First sentence missing figure reference?

Figure references are provided in the subsequent sentence.

L336 - Check wording on <reveal the two models to simulate>.

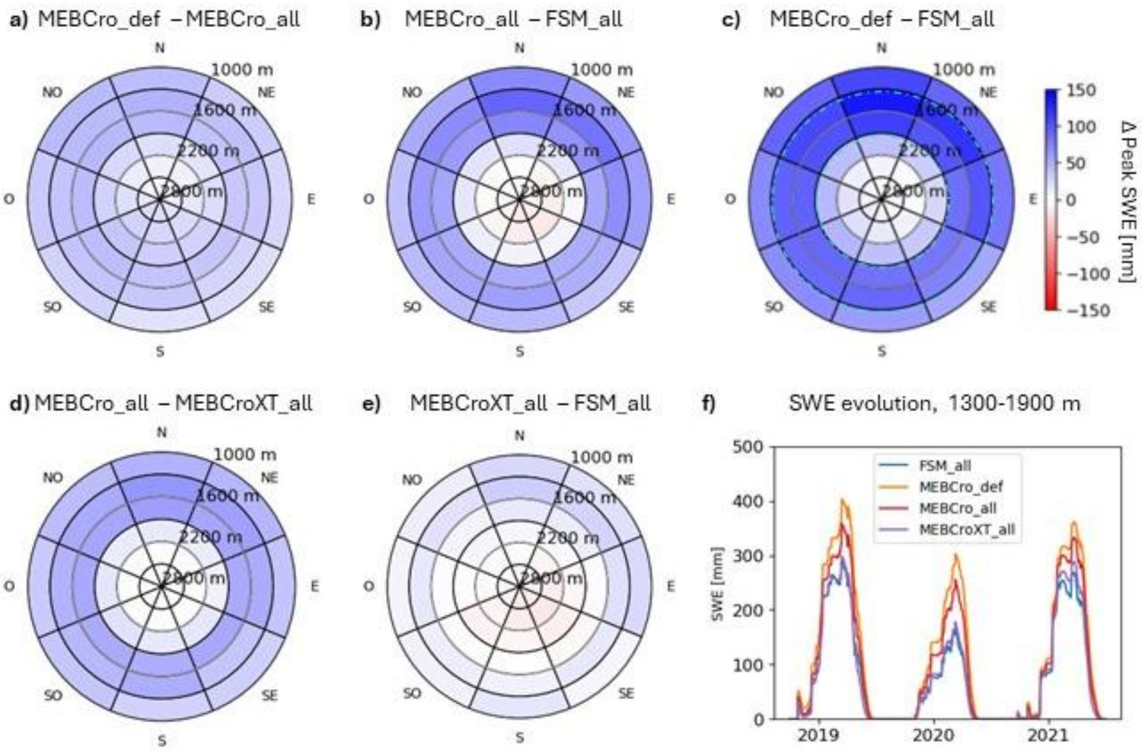
We have modified this to *'reveal that the two models simulate'*

L492 - <Vegetation data and process parametrizations contribute in approximately equal>. They do not look equal in Figure 12, can you clarify? A percent difference calculation could be helpful here and throughout this paragraph to support claims on the degrees of model sensitivity.

We meant that the differences were in the same order of magnitude, but we have rephrased this as the paragraph has been reworked to better clarify our findings, including quantitative measures of model discrepancies. See also our reply to the next comment.

L493-495 / Figure 12 - It would be helpful to have lettering on the panels of Fig. 12 as in the other figures. The bottom left plot seems out of place, based on the model descriptions this plot would represent vegetation dataset and parameterization differences, and yet the column name suggests it is vegetation effects. The change between the <MEBCro_def-MEBCro_all> and <MEBCro_def-MEBCroXT_all> comparisons seems to suggest there is a large sensitivity in changing the parameterizations, which is not in agreement with the final sentence referenced here. Please clarify. It might be useful to add in a <MEBCro_all-MEBCroXT_all> comparison as well to better illustrate the parameterization change sensitivity, although this might just be expected to be similar to the <MEBCro_all-FSM_all> comparison?

Thank you for pointing out this source of confusion, you are right that the comparison MEBCro_def-MEBCroXT_all does not serve the purpose as it includes changes in both process parametrizations and vegetation data. We have replaced it with a MEBCro_all-MEBCroXT_all comparison as you suggest as we agree this better illustrates the parametrization change sensitivity. Lettering has been added on the panels, see modified figure below. The entire paragraph will be revisited for the sake of better clarity.



L520-524 - See major comment #2. Can you support this claim with a process diagnosis for the different models showing different levels of throughfall, unloading, sublimation, and melt/drip? Was there any attribution of a chance in the initial loading of snow in the canopy as well? I agree without measurements, direct assessment of the parameterizations is difficult, but it is still possible to do a process diagnosis which I think would be interesting.

See answer to major comment #2. While we cannot offer a process diagnosis, the discussion of this point will be expanded.

L524-525 - There is a new study from Cebulski et al., (2026, hydrological processes) where this new approach shows improvements in continental, subarctic, and especially maritime needleleaf forests where melt/drip dominates.

Cebulski, A. C., Pomeroy, J. W., & Floyd, W. C. (2026). An Evaluation of New Snow Interception and Ablation Parameterisations in Continental, Subarctic and Maritime Needleleaf Forests. Hydrological Processes, 40(6), e70573. https://doi.org/10.1002/hyp.70573

Thanks for pointing us to this study, which was unavailable at the time of submission. We now refer to it with the following sentence: *Recent attempts to explicitly compute the energy balance* *of intercepted snow provide a way forward (Cebulski and Pomeroy, 2026), having already shown improvements across snow climates but especially where canopy snow drip dominates (Cebulski et al., 2026).*

L543 - I think its worth mentioning here that sub-canopy snow accumulation measurements are rarely undertaken (and more are needed) and also not possible from regional-scale remote sensing methods. This can also help justify the use of the benchmark model.

We have modified the sentence to: *'Unarguably, continued efforts to obtain experimental data on forest-snow processes and sub-canopy snow accumulation remain valuable and are still needed, as remote sensing-based solutions are still out of reach.'*

L545 - Remove double brackets from citation.

Thanks for the catch, done.

L568 - <yielding larger peak SWE than FSM2oshd> What processes contributed to this difference?. Fix word <fore>.

Thank you for catching the typo, which has been fixed. We have added *'mostly due to lower canopy snow losses'* to the sentence.

Several Figures - Several figures are quite blurry. Please increase the quality / resolution.

We apologize for this; appropriate resolution will be ensured with the final submission.

Thank you for the opportunity to review this manuscript. The comments above are yours to consider.

Thank you for taking the time to review our work and for the careful and constructive comments that will surely help improve the paper.