Exploring the potential of forest snow modelling at the tree- and snowpack layer scale – Response to reviewer # 2

General Comments

The paper presents recent work by the authors to combine the strengths of two state-of-the-art snow models: the forest canopy representation from an intermediate complexity snow model (FSM2), and the detailed multi-layer snowpack model (Crocus). After outlining the two models and the process of combination into a new model (FSMCRO), two forested testing sites are introduced. Qualitative comparison of FSMCRO simulations is made to observations and baseline simulations with FSM2. The FSMCRO simulations are then interrogated at a series of points and transects to highlight differences in snow microstructure driven by location within the forest stand. The multi-physics ensemble capabilities of FSMCRO are used to investigate how robust the simulated spatial differences in snowpack microstructure are to model uncertainty. Finally, the variability in microstructure from a multi-physics ensemble driven with domain-average meteorology is compared to the variability in microstructure produced by a deterministic high-resolution simulation at over same domain.

The work is novel and showcases new modelling capabilities that are undoubted state-of-the-art. The processes simulated are relevant to readers of *The Cryosphere*. However, there are several areas that require further description, results or discussion:

The showcase-style manuscript, where different capabilities are presented and described, doesn't necessarily demonstrate significant advances in knowledge provided by the new system. I expected to see more examples of quantitatively analysis of the multi-dimensional data - e.g. evolution of CV of different parameters over time. As well, I expected the manuscript to begin to draw relationships between the (modelled) stratigraphy, the spatial structure, and the physical processes, to provide some hypotheses for future observational and/or modelling work.

The manuscript needs to be clearer about the link between the simulated results and reality. This could be achieved by presenting further quantitative statistics from the observations presented, as well as attempting to validate against microstructural observations. Similarly, the manuscript needs to provide more commentary on whether the patterns shown in the simulations are likely to be real or not, referring to available observational studies.

From a methodological point of view, the manuscript needs more discussion on differences between the FSM2 canopy model with what is implemented in FSMCRO, the reasons for the trade-offs, and discussion of the potential impact of these differences on the simulations. Also, while not the focus of the paper, the difficulties encountered when attempting to couple the models at the snow surface are mentioned, and it would be insightful to briefly expand on some of the issues encountered that led to the choice to develop a 0-layer model instead.

The paper should make a valuable contribution to *The Cryosphere* with revision.

We would like to thank the reviewer for the positive assessment of our work, the detailed review, and the constructive suggestions. Based on the general comments above, we plan two major additions in a revised version of a manuscript: 1) A comparison of model results with snow pit observations from Sodankylä, even though these were taken outside of our study domain and in other years than this considered here. For more details, please see our reply to reviewer 1, main comment #1; 2) Expand the current Sections 3.3. ('Horizontal variability of snowpack stratigraphy along a forest discontinuity at different points in time') and 3.4 ('Spatial patterns and fractional partitioning of snow stratigraphy from fully distributed simulations') to include more quantitative analysis and variability metrics, and comment on the underlying processes in Section 4.2 of the discussion ('new insights on the impact of canopy structure on snow stratigraphy'.

However, maintaining the proof-of-concept style of the paper remains important to us for two reasons: 1) its length is already at the higher limit for a TC article; and 2) a thorough validation against snow pit observations, suggestions for model improvements, and a full analysis of all dimensions of variability covered by the simulations shall provide material for a separate study. While we agree that all these aspects are interesting, they cannot all be covered by one article. Note, for instance, recent work by Bouchard et al., who dedicated one paper separately to each of these aspects using the model SNOWPACK. In general, attempts of validation at snow pits are very rare, extremely challenging, and so far consisted of targeted efforts for existing rather than new models (e.g. Calonne et al. 2020, https://doi.org/10.5194/tc-14-1829-2020; Leppäenen et al. 2017, https://doi.org/10.3189/2015JoG14J026). The goal of this manuscript here is (and should remain) to bring together detailed canopy and snowpack representation for the first time, and to demonstrate that snow stratigraphy is sensitive to this heterogeneous forest structure. To our knowledge, this has never been done before and we believe it provides sufficient content for one article.

Specific Comments

Ln 108 - either in the methods of discussion section, it would be useful to reflect on how much the results depend on the specific model choices, noting that there are some subjective choices here.

This is a valid point, and we will comment on the choice of the two models and potential impacts on the result in Section 4.2 of the discussion ('new insights on the impact of canopy structure on snow stratigraphy'). In terms of canopy representation, important model choices are parameters related to interception and unloading as well as the choice of radiative transfer schemes. The suitability of these model choices was demonstrated in the process-level validation by Mazzotti et al. (2021). For the snow representation, current microstructure-resolving model applications basically rely on either Crocus or SNOWPACK. Uncertainty in process representation choices in Crocus are captured by using the ensemble framework, which makes the model an adequate choice for our study.

Ln 148 – perhaps add "hereafter referred to as "FSM2" after (FSM2.0.3) to distinguish the enhanced canopy model from the standard FSM2 model – see next comment.

Will be added for clarity, thank you for catching this.

Ln 159 – "The model has so far been used for research purposes" - the original FSM2 has been used in many research and operational applications - make it clear you mean the canopy version here.

This will be clarified. Please note that the 'original FSM2' does have a canopy implementation as well, which is however unsuitable for meter-resolution simulations.

Ln 205 – please provide a short commentary in the methods section on which methods remain the same as FSM2 and the extent to which others have been modified. A table would be a very handy reference for the reader.

Thank you for this suggestion, we will add a table summarizing similarities and differences to the original implementation in Appendix A1 (following the current equations).

Ln 242 – 254 – it would help the reader if the numbering and ordering aligned with the order that results are presented.

We will rearrange the order of the presented simulations to match the presentation of results, note however that this will require splitting the current point 3 into two separate points.

Ln 267-272 – this largely repeats the preceding section (2.2.3) and could be removed or combined with the above.

We agree; this paragraph will be shortened to avoid redundancy.

Ln 283 – it would be useful to report some basic quantitative stats from the validation here (e.g., overall bias, RMSE, R, CV) to give confidence in this application.

We will add overall bias values in the main article, please note that the detailed quantitative stats are already available and referenced in the Supplementary Material. We would like to keep most of the validation in the Supplementary Material to avoid distracting from what we consider to be the main aspects of this study (Sections 3.3 and 3.4)

Ln 290 – while it is understandable that the irregular and uncertain location of snowpit observations may limit a full quantitative evaluation of the FSMCRO simulations, it would be instructive to present some of the observations here if only to highlight the shortcomings of the available observations, motivate hypotheses that could be interrogated with FSMCRO and comment on how these may be validated with new observations. Not including observation of microstructure substantially reduces the readers confidence that new model system is simulating real patterns.

Some comparison to snow pit data will be included and discussed in the Supplementary Material of a revised version of the manuscript. Note however, that snow pit data is only available starting WY 2019, and related comparisons are problematic for many reasons (see your major comment above).

Ln 305 – “formation of surface melt forms/crusts (red) happens ca. 10 days earlier under-canopy than in the canopy gap” – this is not immediately clear from the figure – please add the dates to show the specific period intended.

Dates will be added.

Ln 343 – here and elsewhere (including Ln 377 and figures) it would be easier for readers in both northern and southern hemispheres if 'sun-exposed edge' and 'shaded edge' were used in place of 'south-exposed edge' and 'north-exposed edge’. Either way, please be consistent throughout the text and figures with the terms used (e.g., next sentence has ‘sun-exposed edge’, figures have ‘n-facing’).

This is a good point. We will replace instances of ‘south’ and ‘north’ with ‘sun-exposed’ and ‘shaded’ throughout the manuscript.

Ln 351 – “does not overestimate the variability of snow stratigraphy.” please be specific - do you mean that in the accumulation season, vertical variability is large, whereas in the ablation season, horizontal variability is large? If so, please state this.

Yes, this is the case, and it is already stated a few lines further up. With this sentence, we meant that the qualitative differences depicted by the discrete/categorical grain type parameter are not giving an unrealistic picture of variability, as the more quantitative/continuous parameter SSA backs up the variability patterns seen in the grain type plots. We will rephrase the sentence for it to be clearer.

Ln 363 – “In contrast, snow depth variability between ensemble members at each location is in the same order of magnitude as differences between the two locations.” does this mean that the structural differences are more likely real? and that the snow depth differences are not? or just that the model behaves in the same way for the same forcing? Please comment.

This result implies that prediction of structural differences is more robust than the prediction of snow depth differences, because all ensemble members agree on the structural differences (surface crust yes vs. no), while the difference in snow depth between the two locations is within the uncertainty represented by the ensemble simulation at each location. We will rephrase this to make our point clearer.

Ln 383 – “This finding provides strong evidence of the substantial impact of canopy structural heterogeneity on modelled snow stratigraphy, suggesting that the resulting variability by far exceeds model uncertainty.” - was the ensemble system was validated against forested as well as open-site locations? this would be needed to conclude that model uncertainty is fully captured by the ensemble, and thus that model uncertainty is less than the explicitly resolved spatial variability.

ESCROC represents uncertainty in surface and internal snow processes and has been evaluated in a large range of environments and climatic conditions (see Lafaysse 2023; https://theses.hal.science/tel-04130109/). We therefore have confidence that model uncertainty is well represented even when near-surface atmospheric conditions are modified to account for the effect of canopy, as done in this study. Obviously, ESCROC does not represent forest-snow interactions uncertainties, but the above statement only links the impact of forest processes (and their variability) to the uncertainty in snow process representation, which is therefore appropriate.

Ln 395 - is there indirect ways to validate these sorts of results - e.g. surface temperature from thermal imaging?

Thermal imaging, especially from drone-based platforms, would certainly provide interesting datasets for validation, yet such datasets were not available within the context of this study. Note that the use of such datasets for validation is not straightforward due to the strong temporal dynamics of surface temperature (one image provides a temporal snapshot, while model forcing data was available at hourly resolution). We will add a comment on potential validation approaches as suggestion for future work in a revised version of the discussion (Section 4.4).

Ln 414 – "snowfall events" – please give dates or use annotations on figure to highlight period being referred to.

Snowfall periods will be marked in the figures.

Ln 415 "co-exist at the surface" – again please be specific about what periods are being referred to.

The period will be specified.

Ln 415 "The ensemble thus does not capture variable metamorphism rates that are tightly linked to specific canopy structure" – would we expect it to? Please comment.

No, we do not expect it to, because the variability in canopy structure is (currently) not included in the ensemble. This is why hyper-resolution simulations as shown here provide added value, and it is also the reason why we suggest that canopy processes should be added in an ensemble if uncertainty coming from canopy structure metrics is to be accounted for. We will rephrase the relevant paragraph to make this point clearer.

Ln 435-438 – a description of the issues encountered when trying to fully couple the models is not provided, so it is hard to assess how the result of Cristea et al.2022 are relevant here, and what we are to learn from these comments. Do you mean that these pitfalls should be documented in future studies? Or that this justifies your choice of a zero-layer approach? Or that Crocus is free from these issues. Please revise.

The 'numerical pitfalls' are mainly related to instabilities in surface temperatures that are common in models that allow thin layering. The results of Cristea et al. (2022) are relevant in this context because they showed that the mere choice of the number of layers in a snow model had a strong impact on model results, which does point to such numerical issues. In the case of canopy coupling, instabilities happen because the surface temperature is part of both the canopy and the snow energy balance, and the two are solved sequentially rather than in a fully-coupled way in FSM2 (as is the case in most other Land Surface Models as well). Choosing a zero-layer approach allowed us to avoid running into this issue, because thermal diffusion and any updates to snow temperatures are only treated in the snow routine. We will detail this reasoning in a revised version of the manuscript.

Ln 444 – please briefly outline the reasons provided by Nousa et al, (2023) in the text.

Nousu et al. highlighted the example of the turbulent exchange parametrizations available in ESCROC, where one of them (Martin and Lejeune 1998) applies a stability correction term that is not implemented in

MEB and therefore is not readily applicable to forest simulations done with MEB-Crocus (coupled canopy). We will mention this example in the text.

Ln 452 – further discussion on the limitations of the system is needed. Especially the potential impacts of trade-offs made in the zero-layer implementation (e.g. non-interactive canopy temperature – snow surface temperature – snowpack feedback).

We will expand this discussion, referring to studies that have shown that air temperature is a reasonable proxy of canopy temperature in most situations and discussing the implications of this assumption. We will further refer to other forest-snow models that do not account for snow surface temperature – canopy feedbacks (e.g. SNOWPACK).

Ln 464 – expected more analysis of how to quantitatively assess multi-dimensional data - e.g. evolution of CV of different parameters over time.

We will provide an example of quantitative analysis of variability, see our reply to your general comments above. Note, however, that a thorough analysis of the multi-dimensional data is beyond the scope of this study given the paper's current length.

Ln 464 – expected the manuscript to begin to draw relationships between the (modelled) stratigraphy, the spatial structure, and the physical processes, to provide some hypotheses for future observational and/or modelling work.

Some discussion of the relationship between modelled stratigraphy and spatial canopy structure is included in the subsequent paragraphs (465ff). We will expand the discussion of relationships between modelled stratigraphy and the physical processes in this paragraph and include hypotheses for future work in Section 4.4.

Ln 473 - "The second mechanism [insolation-microstructure relationships] has, to our knowledge, not been captured by any simulation prior to this study."- please comment on whether these patterns have been observed, and if so, provide references.

To our knowledge, no observational study targeting how snow microstructure varies with insolation patterns in discontinuous forest exists. We will mention this explicitly in the revised discussion and highlight how hyper-resolution simulations with FSMCRO can be useful for identifying processes (and process interactions) that have never been studied.

Ln 577 – as per comment on Ln 205 – would be useful either here or in methods section to have a list or table of which schemes are identical, which are slightly, modified and which required larger modification.

See our answer to your comment above, a table will be added to the Appendix.

Technical corrections

Ln 129 – please provide a reference for the standalone version of Crocus.

The standalone version of Crocus is still unpublished but will be included in a publication covering the latest developments of Crocus (Lafaysse et al., in preparation). The standalone version of the code is currently available at https://opensource.umr-cnrm.fr/projects/surfex_git2/wiki/Install_standalone_version_of_Crocus. Yet, because a migration to GitHub is planned in the next few months, including this link to the publication is obsolete. We will ensure that the correct link is listed and regularly updated in the GitHub repository of FSMCRO.

Ln 130 – please explain what SVS2 is, as the reference (Vionnet et al, 2022) does not mention it.

SVS2 is the Land Surface Model used at Environment and Climate Change Canada, we will specify this in the revised manuscript. In Vionnet et al. (2022), the model is presented in Section 2.2. The version SVS-2 including Crocus is now also used in Wooley et al. (preprint: https://doi.org/10.5194/egusphere-2024-1237), this citation will be added.

Ln 173 – add 'incoming' before 'short-'

Will be done as suggested.

Ln 288 – 'FMI' - please expand the acronym.

The acronym was already spelled out in line 217 but not defined, this will be corrected, thank you for the catch.

Ln 299 – 'SSA' please expand first use of acronym.

Will be done, thanks for catching this.

Ln 315 – "discontinuous forest transect" -> "transect through discontinuous forest". The former implies the transect is discontinuous.

Will be modified as suggested.

Ln 323 – "peak of winter" -> "peak winter accumulation"

We will adapt this and check other occurrences of 'peak of winter' throughout the manuscript.