

Review Report

Manuscript: *Comparing 3 m resolution snow cover downscaled from MODIS, VIIRS, and HLS using commercial satellite imagery and terrain information*

Authors: Justin M. Pflug, Kehan Yang, Nicoleta Cristea, Emma T. Boudreau, Carrie M. Vuyovich, Sujay V. Kumar

Journal: The Cryosphere

Recommendation: Minor revisions

General assessment

The authors present a comparison of different downscaling methods to extract hyper-resolution snow cover information from widely available but (much) coarser spaceborne observations of fractional snow-covered area. The comparison is based on three published methods of increasing complexity and input-data demands, investigating the added benefits (and limitations) of using commercial data. Their results suggest that the most complex downscaling approach (based on a random forest model) generally outperforms simpler, either solely terrain- or climatologically driven methods when estimating snow cover evolution at a 3 m resolution. However, for a target resolution of 30 meters or more, these simpler methods that do not rely on additional commercial data may be sufficient. While the study focuses on relatively small alpine meadow domains, it can provide guidance on selecting appropriate downscaling methods depending on the intended application and data availability. The manuscript is generally well written: the methodology and results are presented comprehensively and in a logical order, and the Discussion addresses the relevant open questions and potential limitations of the presented findings.

My only general remark concerns the figures and their introduction in the text. First, the manuscript would benefit from guiding the reader more directly through the results, e.g., by explicitly introducing each figure and explaining what it shows. Currently, several figures are referenced only in passing, as a bracketed reference at the end of a statement. Moreover, not all subpanels (e.g., the right panel of Fig. 9) or supplementary figures (e.g., Fig. S13) are discussed in the text, and Fig. S20 (L567) is referenced before Fig. S19 (L604). Second, several figures (e.g., Figs. 1, 2, and 7) are small and densely packed, so that text, tick labels, and colorbars are difficult to read without zooming in considerably. I would therefore suggest enlarging these figures and increasing font sizes, splitting or removing less central panels where necessary (also consider whether all 22 supplemental figures are necessary), and ensuring that every panel shown in the main text is also discussed there.

Overall, this is a relevant contribution to the field suitable for publication in *The Cryosphere* after minor revisions. I'd like to compliment the authors on their work. Please find my specific comments below, some of which may reflect personal stylistic preference.

Line comments

- While certainly a stylistic choice, in my opinion, the use of "1) ..., 2) ..., and 3) ..." makes the text more difficult to read in places.
 - L151ff: The references to the corresponding subsections already separate the three items sufficiently. Moreover, the numbering conflicts with the Abstract, where 1) is the probabilistic method and 2) the RF approach.

- L190, L218f, L254ff, L478, L489, L563: With only two elements, numbering is unnecessary. At L190, if kept, it should read "resolved 1) the spatial pattern ..., and 2) the overall bias ...".
- L332ff, L640: The numbering interrupts the flow of the text rather than adding clarity.
- L424ff: Listing the core findings seems appropriate here, but consider "first, second, and lastly" instead.
- Please avoid to have neighboring closing and opening brackets "(...) (...)" as in L121, L140, L141, L152, L214, L253
- L12: Characterizing the fine-scale snow distribution as "... important for avalanche forecasting and biological refugia" might be too limited. Consider adding snow hydrological processes (and others you find relevant) as you did in the introduction.
- L20: Consider rephrasing the sentence "The approach used to ..." as currently it is formulated slightly confusingly.
- L104: If "model architectures" here refers to the downscaling approaches, consider rephrasing (as it is the only location in the manuscript where this phrase appears). Otherwise, please explain it to avoid ambiguity.
- L106ff: In contrast to the following paragraph, I could not find any information about the (average) revisit times or (average) number of observations per season for the PlanetScope data.
- L116ff: Consider already hinting at the leave-one-out approach here. While you introduce the details below, the question about independent validation immediately came to mind while reading this sentence.
- L125ff: Again, consider hinting at the method used to convert HLS data to fSCA here. Even though it is explained a few lines below, this sentence/paragraph left me wondering.
- L138: Consider removing "Finally", as it implies a sequential order of the data, which was not apparent to me.
- Figure 1: The figure is rather small, especially the text/sources in DAN and the tick labels. Removing the border outside the model domains might also improve visibility.
- L155: "in detail" is written without a hyphen.
- L162: Consider adding the corresponding reference (probably Pflug et al. 2024) to support the statement about the "high accuracy of these snow cover maps". Do the same for the second sentence in the Conclusion (L634ff).
- L164ff + all equations: Multi-letter variables, e.g., $\mathbf{FN}$, could be misunderstood as multiplication, like $\mathbf{FN} = \mathbf{F} \times \mathbf{N}$. Consider reformatting these, e.g., using FN in LaTeX. This applies to all equations and most variables (except $k, z_0, \underline{z}, w$) in your manuscript. At any rate, at least ensure consistency in variables, e.g., by also using italicized $\mathbf{FN}$ in the text and not only in the equations, using either $\mathbf{F1}_{\text{score}}$ or "F1 score", etc.
- L165ff: This paragraph could benefit from some improved phrasing and additional explanation.
 - "For each date, the F1 score is calculated twice, ..." could be misinterpreted as "two times a day" rather than two different values, one for "snow presence" and one for "snow absence".
 - Consider rephrasing "the number of snow present and snow absent cells", e.g., to something like "the respective number of cells with and without snow".

- To me, the sentence “The resulting F1 score varied between 0 and 1” implies this is particular to your case, while the $F1_{\text{score}}$ always ranges from 0 to 1.
- Consider starting a new paragraph when introducing “Cohen’s kappa” on Line 173
- Similar to Line 183ff (i.e., “Put simply, ...”), consider adding a short, practical explanation of both scores.
- L182: Consider removing the qualifier “that snow cover was downscaled from”, as it seems unnecessary at this point.
- L165ff and L182ff: These two paragraphs describing the evaluation metrics seem a bit misplaced here, in between a general outline of the Methods section and the definition of the three downscaling methods. Consider putting them in their own subsection at the end of Section 3 (e.g., “3.5 Evaluation metrics”).
- Figure 2: Conceptually very nice figure outlining the three different downscaling methods. Consider reordering/grouping the panels so that it is instantly apparent that a) and b) refer to method 1, c) and d) to method 2, and e) to method 3. As for my general comment on your figures, Figure 2 is also rather condensed and small, and the text and labels are especially difficult to read without zooming in quite a bit.
- L196: This sentence makes it seem like the topographic downscaling method only looks for locations with above-average accumulation and thereby snow cover duration, while, from my understanding, it also targets locations with enhanced ablation and earlier melt-out. Consider clarifying this introductory sentence.
- L200: Why are you not using the already defined abbreviation “fSCA” here?
- L204: Consider hyphenating south-southwest for better readability.
- L203: It is unclear how the TPI values are normalized, as they have no predefined range.
- L225ff: Consider clarifying that the “coarser-resolution footprint” refers to a single pixel of the respective sensor, and that the ranking of the SVI values happens independently for each sensor footprint/pixel. For example: “filling the snow cover in that order until the fSCA estimate of the respective sensor pixel was matched.” I think this is pivotal for understanding the downscaling methods, and it was not immediately clear to me.
- L230ff: To me, it was not immediately and entirely clear what the term “frequency of snow cover” (used throughout the manuscript) refers to. This is especially true given the seemingly interchangeable use of “frequency”, “relative frequency”, and “probability of snow cover”, as well as the use of different units (“percentage of days”, “days”). Consider clearly defining the term here (or when introducing it on Line 84), or rephrasing it, e.g., to “number of days with observed snow cover”. In general, the term is understandable from the context, but more precise wording would improve readability.
- L291ff: As for my general comments on your images, consider introducing Figure 3 here more directly and guiding the reader by adding more detail (e.g., that the maps are based on VIIRS retrievals).
- Figure 3: Consider spelling out the “Three-meter snow cover frequency ...” at the beginning of the caption.
- Figure 4: I suspect the light lines in panel b) represent the individual domains, but it is not stated in the caption.
- L326ff: Consider rephrasing the sentence “This was driven by ...” for improved readability.

- L332ff: “decisions made by each downscaling approach about where to assign 3m snow presence and absence” seems like a very convoluted way of saying the downscaling approach itself.
- Figure 5: Consider grouping the boxplots in panels b) and c) by the sensor instead of the downscaling approach. This would align better with the text, which focuses more on comparing improvements between different downscaling approaches (from the same sensor) than between sensors. However, both possible visualizations/groupings are interesting and simply highlight different parts of your results.
- L343: I think this should be Fig. 5a, if I am not mistaken.
- L384: Consider guiding the readers' attention more directly by referring to “rightmost top and bottom panels in Fig. 6”.
- L393ff: Consider giving both R^2 values and reordering the bias values (MODIS first, VIIRS second), matching their introduction in the sentence before (L 390).
- L399ff: Consider adding a sentence about the decrease in R^2 in the RF-based downscaling of VIIRS data compared to the probabilistic method.
- Figure 7: The “spatial plot” is very small, and its colorbar is illegible.
- L430: As I understand it, domain, year, downscaling method, and sensor are all available degrees of freedom. Wouldn't that make the sensor the “least important” descriptor of variance rather than the “next-strongest”?
- L434ff: Consider adding a reference to where to find these values (i.e., Table 1 and Figure 5).
- L461ff: This is a very interesting and relevant paragraph. Maybe you can indicate the mentioned scale breaks (or ranges) in the left panel of Figure 9.
- L469: Missing “of” in “large portion of the snow cover pattern variability”
- L479: Missing “the” in “This reduced the impact ...”. You could consider rephrasing this (and previous/subsequent) sentence(s), e.g., to avoid the repetitive beginnings “This reduced ...” and “This was opposed ...”
- Figure 9: It is unclear where/if you refer to the “spatial plots” in the right panel in the text.
- L513: It is unclear how the distributions in Figure 5 were used to identify outlying poor performances.
- L518ff: Where is the “terrain-based downscaling benchmark” defined?
- L536ff: The “speckled patterns” in the terrain-based downscaling are interesting. If I understood you correctly, this happens when HLS grid cells have a fSCA close to, but not at, 100%, and hence a few (of the lowest SVI-ranked 3 m pixels) remain snow-free. Consider refining your formulation and how the RF method avoids such issues.
- Figure10: Consider adding the labels of the domains directly in the figure (not only in its caption) and coloring the titles of the second and third columns of the right panel “HLS + Terrain-based” and “HLS + Random forest” blue and red, respectively (to establish a direct visual link to the left panel). Also, the white dashed lines representing the median in the left panel are difficult to see.
- L564: It is not fully clear how you bias-corrected the MODIS fSCA data. Consider explaining this in more detail.
- L583: Consider specifying the somewhat generic statement about “implicitly accounted for spatial autocorrelation features and alternative snow depletion and accumulation processes”.
- L593: It is unclear to me how Figure 3 supports this statement.

- L630: It is unclear how this statement relates to Figure 2.
- L636: I think this statement should be in the past tense: "... we tested ..."