RC2

This article compares various methods to downscale satellite snow cover (SC) map (30 m Harmonized Landsat-Sentinel to 500 m MODIS) to 3 m resolution matching the PlanetScope imagery. The benefits of increasing the complexity of the downscaling approach and of including external satellite products are evaluated. As the last sentence of the abstract nails it : downscaling of MODIS-VIIRS products with random forest trained on PlanetScope derived SC map is often equivalent to downscaling of Landsat-Sentinel SC maps with simpler probabilistic approach. The article is quite methodological, rather dense with results which are meticulously presented. The discussion brings interesting points about the resolution and the scale of the snowpack variability. In general, the work provides precise guidance for future application of such downscaling.

We want to thank you for your feedback. We are excited about how this manuscript will be improved by addressing these points.

L24 Is the sentence grammatically correct? I do not understand how « but also when... » connects with the previous part of the sentence.

Thanks for pointing this out. We believe the phrase "but also" may have made this more difficult to read when our intent was to present two contrasting points: 1) that commercial observations can provide context about the distribution of fine scale snow cover from coarse-scale fSCA, ***but conversely*** *2) using finer-scale fSCA from HLS without context from commercial observations performs similarly using only information about the terrain. We will revise this sentence to make it clearer that these are contrasting findings.*

L24 Why « in many regions » ? Could it be deleted ? Weird grammar

Yes, this will be deleted. Our intent here was to not overstate the results since this study only focused on 7 domains. However, we think that this was specified well by the main text and agree that this is not necessary for the abstract.

L25 « may » => « can »

"May" will be changed to "can".

L27 The introduction is well written and presents well the context of SC maps downscaling. I only wonder if such downscaling approaches has been actually used in other works as a way to produce high-resolution SC maps before using them as input for the rest of the work.If such works exists, please highlight them. (B1)

This is a great idea. We looked into this but ultimately struggled to find examples of where these specific approaches were used as the first step for broader hydrological or biological modeling or observing studies. That said, these downscaling approaches are built on the concept of fine-scale snow pattern repeatability, which bears repeating. There are also examples where fine-scale snow patterns have been used to simulate/estimate snow distribution by combining with statistical and process based models. After the questions numbered in the introduction, we will revise the last paragraph to touch on this using text along these lines:

> *"Again, the tested downscaling approaches assume some degree of annual repeatability in snow cover, which has been used by multiple studies for modeling*

and reconstructing the fine-scale distribution of snow (e.g., Brauchli et al., 2017; Pflug and Lundquist, 2020). This suggests that these downscaling approaches can be extended to include multiple satellite sensors (e.g., Premier et al., 2023) as the basis for estimating hydrologically, biologically, and safety-relevant snow information at spatial resolutions much finer than common operational satellite retrievals. Despite this, no study has compared these downscaling approaches, nor have used them as the basis for inferring snow cover at 3 m spatial resolutions as a precursor for estimating hydrological and biological snow-driven impacts. This experimental design uniquely focuses on determining the degree to which downscaled snow cover estimates are altered by sensor accuracies, snow cover pattern references (either terrain-based or observationally based), and different downscaling approaches across seven different domains and four different years."

L30 « species » twice, redundant.

Thanks for catching this. Our first mention of "species" will be removed.

L34 « Information about the distribution of snow cover and snow structure **at footprints relative to** wind-driven » this is not completely clear (for a non-native speaker)

We agree that this text was not as clear as it could have been. We will revise this sentence to clarify that viewing snow cover at resolutions capable of resolving the snow cover heterogeneity driven by processes like wind redistribution is important for applications like avalanche forecasting.

L50 This review article could be of interest for this introduction : Gascoin S, Luojus K, Nagler T, Lievens H, Masiokas M, Jonas T, Zheng Z and De Rosnay P (2024) Remote sensing of mountain snow from space: status and recommendations. Front. Earth Sci. 12:1381323. doi: 10.3389/feart.2024.1381323

This is a great suggestion. Thanks for pointing us to this manuscript. We will add a passage to the third paragraph of the introduction specifying that Gascoin et al. (2024) identified the coarse spatial resolution of snow cover observations from current operational platforms as one of the most prominent limitations for resolving snow melt out patterns and snow-driven ecology in mountainous terrain.

L122 About VIIRS images : explain somewhere why the pixels appear as tilted diamonds (Fig. 3 and 6.)

*We agree that this was missing. In the revised manuscript, we will add text to the second paragraph of the methods section to specify that for each sensor used to provide coarser resolution fSCA (MODIS, VIIRS, and HLS), we create a mapping where we identify the 3 m cells defined by the PS grid that fall within each of the fSCA pixels. In doing so, we connect the area mapped by an fSCA pixel with the collection of 3 m pixels that we used to downscale binary snow presence and absence. We will specify that this means that we use the native resolution **and** projection of each fSCA dataset (in other words, we don't regrid and/or reproject), which is clearly evident by both the size and shape of each pixel for each sensor and fSCA product. We will also use this text to clarify that it is important to do this in order to use the fSCA values from each sensor explicitly instead of*

potentially adding impacts from merging nearby cells with differing amounts of biases from forest cover or viewing angle issues.

L154 « changes of **in** performance » ?

Here, "to" will be revised to "in".

L160 « While using PS snow cover... » Long complicated sentence, maybe reword of cut it in two. Besides, the last part is a loop thinking : ...the high accuracy of these snow cover maps...make these retrievals a high benchmark for accuracy.

Thanks for pointing this out. We agree that this sentence should be split in two. We will revise the first portion of this to much more clearly state that if annually persistent snow cover biases appear in PS, then the performance of the downscaling might be overestimated slightly since these appear in the data used to both inform and evaluate the downscaling. However, we still want to keep the second portion of this passage that specifies that despite this potential impact, PS snow cover is still among the most accurate remote-sensing based estimates of snow cover that is globally scalable.

L235 « the frequency of snow cover observed at that location, divided by the number of PS observations (Fig. 2c).» frequency should already be the ratio between the number of a given event divided by the total number of events. This sentence suggests that frequency is divided by total number of event.

We agree that our use of the term "frequency" was confusing at times. This was also indicated by Reviewer 1. We will therefore update our terminology throughout to be much more explicit, both with respect to the probabilistic downscaling and some of the figures that show this information (e.g., Figure 3, 7, 9, and 10). For example, we will revise this text and replace "frequency of snow cover" with "number of days with snow cover", or "percentage of days with snow cover".

L291 Maybe state the code name of the basins around L106. These information are only in the legend of Fig. 1 at the moment.

Good idea. We will add an explicit mention of the name and code for each domain to the second sentence of Section 2.

L293 « This parameter set implicitly reflected the terrain roughness » not clear, maybe add commas to structure the sentence ? Or rephrase.

We agree that this was not said as clearly as it could have been. We will revise this sentence to more explicitly state that snow cover and the duration of snow for this SVI increased in small-scale terrain depressions but decreased in cells with uplifts in the terrain. We believe that this will also flow much better into the following sentence which describes the broader-scale impact of solar exposure and shading on the SVI map with the opposing parameters (using a coarser 30 m DEM and a 0.75 weighting factor favoring snowmelt impacts on snow duration).

L446 « by any other study to date » cite the resolution of the other studies, of some at least.

We apologize that this was not clear. This passage was intended to connect to the citations in the previous sentence, pointing to the studies that the downscaling

approaches that we tested came from. We will revise this sentence to make it clear that we are relating to these studies.

L450 « While the accuracy of downscaled snow cover varied by the domain and year, accuracies generally increased with increases to**in** the informative power of the inputs used for the three downscaling approaches tested here. »

"to" will be revised to "in".

L476 « HLS was often capable of viewing **the pattern of snow cover from PS** much more explicitly than MODIS » Weird phrasing. It sounds like the patterns are a property of PS but they are actually an objective natural object. Maybe something along : « HLS and PS were often capable of viewing pattern with a similar level of details/similar scale, closer to the true scale of pattern than MODIS »

Thanks for the comment. We originally phrased it this way since the semivariogram was calculated using the percentage of snow-covered days from PS. However, we agree that the point being made in this passage is that the natural distribution of snow across the terrain is observed to a similar degree by PS and HLS, and much more so than MODIS or VIIRS. We will revise this text to make that distinction clearer.

L477 « footprint » Could it be deleted here ? Looking up this term earlier in the article, I am getting confused. Here, it could mean « pixel » but earlier it sounds like a synonym for image acquisition ? Besides, I guess that all the acquisitions were clipped to a common footprint ? It may be stated more clearly in the methods.

Thanks for pointing this out. We understand how the term "footprint" was confusing at times and was sometimes used in different ways. Here, we intended it to reference each individual fSCA "pixel". We will go through the manuscript and edit/replace this term for clarity.

L477 « viewed total or nearly full (fSCA near 1.0) or snow absent (fSCA near 0.0) conditions more frequently » is not easy to understand, especially « total or nearly full» pixels ? Conditions ? Maybe the concept of mixed-pixel can be helpful here.

This was not worded as well as it could have been. We will revise this passage to more explicitly state the fSCA from HLS was near 1.0 and 0.0 more often than MODIS and VIIRS as a function of less-frequently having mixed pixles averaging retrievals covering multiple features, including snow-present and snow-absent ground, but also canopy overstories and features like steep rock surfaces that never or rarely accumulate snow.

L481 « While the goals of this study were to downscale snow cover to 3 m resolutions, the spatial autocorrelations also suggests that a lot of the snow cover information could be captured by downscaling **to resolutions coarser than PS observations** using HLS snow cover if 30 m snow cover is sufficient (e.g., Figs. S14 and S15). » This sentence suggest to downscale at 30 m resolution or at a resolution between 3 m and 30 m ? State clearly.

This is something that we have thought a lot about. Overall, the ideal resolution to downscale snow cover to depends largely on the application and the domain, so we hesitate to make any specific recommendation. For example, in less complex and windy terrain, snow cover patterns may be more homogeneous meaning that PS observations

may be unnecessary. Additionally, snow cover at finer resolutions may not be as important for estimating total snow water resources but would be much more crucial for resolving the prevalence and placement of drift habitat for species like Wolverine and Polar Bears. We will add a sentence to the end of this paragraph to establish this.

L488 « a higher frequency of full snow cover or total snow absence at fine **scale** (30 – 40 m) » why mention 40 m here ? Never used before, is it in Selkowitz et al. (2014) ?

Thanks for pointing this out. We originally derived this range using the semivariograms from Figure 9. However, a lot of this text was ultimately distracting and was removed. We will remove this parenthetical to make this text clearer. We will also add "scale".

Figure 3. Maybe add as text in each subset, a metric showing the quality of the fit of the downscaled map with the reference (F1 and/or *k*). Discontinuity at the border of low-resolution pixel (north of all maps) could be commented somewhere. Any guidance to improve the method and avoid this ?

Thanks for the suggestion. While we see the validity to this, the intent of this plot was to show the effect of the parameter combinations on the total summed pattern of snow cover, shown here as the percentage of days with snow cover plotted at 3 m resolutions. As opposed to Figure 6 which shows the 3 m distribution of snow cover for a single date with a corresponding F1 score and k statistic, these maps integrate downscaled snow cover maps for all dates. We believe that the results in Figure 4 show the distribution and median F1 score for multiple sensors in this domain (4a) and all domains (4b) much clearer. That said, we agree that we should include the coefficient of determination of snow cover frequency relative to PS. This will be added.

We also have a discussion of the discontinuities shown in this plot in the second paragraph of Section 5.2, specifically pointing out how these were biases driven by low-biased VIIRS fSCA near the forest edge. We believe that this paragraph details how the RF approach, which learns and corrects for these types of annually persistent biases may be one of the best approaches for improving this type of sensor-side bias.

Figure 4. Right column : the colored line is the average of all domain ? And the transparent thine line are individual domains ? For the latter, the weight is not represented in the line type, is it ? Please complete the caption.

Thanks for pointing this out. This was also noted by Reviewer 1. We will edit the caption to clarify that the light lines represent the different weights for the different domains, and the darker lines are the domain averages (for each weight).

Figure 5. No need to modify all the boxplots figures but a consideration for future work : I find it disturbing to show outliers as distinct points from the boxplots. In this case, they are defined implicitly by the plotting library. Actually, they are part of the distribution. I suggest that the box and whiskers show exact percentile of the complete sample, for example 5th, 25th, 75th, and 95th percentiles.

We agree. We would normally not plot these as outliers, as you suggest. However, in this case, we found it useful to plot these outliers since we used these to explicitly identify the downscaled snow cover maps with performance that was lower than the overall

distribution. This was the analysis that we presented in the third paragraph in Section 5.2. We will revise the topic sentence of this paragraph to establish that the outliers were identified by the scenes with the F1 score that was lower than the lower quartile F1 score by at least 1.5 times the F1 score interquartile range.

Figure 7. To be improved a bit. Hatch may not be necessary, it is impossible to see what is below the hatch on the map.

We agree. While the information in the spatial plot was interesting, we found that the point we made using this plot was not entirely necessary and potentially distracting. We will revise this plot to only present the binned scatterplots, like we have presented for the other domains in the supplementary.

Figure S13. I was about to say that the figure should be improved (GIN-HLS completely black plot with a fit through it is almost funny!) but this figure do not seem to be referenced in the text. To be deleted ?

Yes, thank you for catching this! The text discussing this figure was removed from a previous draft of this manuscript since the information was unnecessary and distracting. This figure was unintentionally retained in the supplemental. This will be removed and the supplemental figures will be relabeled accordingly.