

Review of “Spatio-temporal snow data assimilation with the ICESat-2 laser altimeter”  
by Mazzolini and others  
submitted to The Cryosphere

### Summary

The article presents the results of three data assimilation studies that incorporate (1) fractional snow-covered area from Sentinel-2, (2) snow depth from ICESat-2, and (3) both fractional snow-covered area and snow depth to determine which approach has the greatest improvement on modeled snow depths. The study is performed for a site in the Spanish Pyrenes where multiple drone-derived DEMs are available to assess the performance of the data-assimilated modeled outputs. The authors find that the inclusion of both fractional snow-covered area within the catchment and ICESat-2 snow depths from outside the catchment improve the model’s ability to capture the distribution of snow depths in the catchment. Model performance is particularly improved during the snow accumulation season when ICESat-2 data are available, and degrades as the dominant processes that dictate snow distribution shift from accumulation to ablation processes.

The results of the study are interesting and the data assimilation approach appears to be a promising method to make the most use out of the sparse ICESat-2 tracks. I appreciate the detailed descriptions of agreement and disagreement between model outputs and observations. However, the writing can be a bit difficult to follow at times and I recommend that the authors make a number of revisions to the text and the figures in order to improve the manuscript.

### Major Comments

1. There are several places where references are located early in sentences and it is unclear if they apply to the entire sentence, or where there is no reference provided but it should be. I’ve listed a few lines here but please make sure references are clear throughout the text:
  - a. lines 17-18: Is Mott et al. (2018) for the entire sentence? If not, you need another reference to support everything that comes after its current location.
  - b. Lines 105-107: You say that most DA research on snow has focused on temporal data assimilation with a few exceptions. You cite the exceptions but not the “most”.
  - c. Lines 118-119: The ATL03 data product is still be validated? You need to provide a reference here or remove the comment.
  - d. Lines 129-133: You need some references here for information about the watershed, such as the fraction of precipitation that falls as snow and total precipitation.
  - e. Lines 147-154: Neuenschwander et al. 2020 (<https://doi.org/10.1016/j.rse.2020.112110>) showed strong returns over snow for the weak beams. You should cite them here and I recommend you re-examine your weak beam data.

2. The introduction is very long. I understand that the authors feel like they need to provide background on a number of topics in order to justify and explain their work, but the reader is left wondering where they are going with the work because the introduction is so long. I recommend that the introduction is shortened considerably. You could base each paragraph around the following topics: (1) Why it is important to know snow depths across watersheds, (2) ICESat-2 looks like it can be used to estimate snow depths along its flight tracks, albeit with fairly large uncertainties, but we need a way to spatially and temporally extrapolate, (3) data assimilation techniques have shown promise for extrapolation, (4) this study explores data assimilation of ICESat-2 snow depths and Sentinel-2 snow-covered area. Then you can move a lot of the extra detail on techniques to measure snow depth (currently lines 33-50 and then 51-63 on ICESat-2 details) and various data assimilation techniques (currently lines 76-114) to the supplement so readers who are not familiar with those topics have a resource to lean on without bogging down the reader who knows plenty about those topics.
3. This seems like something pretty minor but all figures should have letter labels. Right now you need to refer to some of them by location and it would be a lot easier if they were all consistently lettered.
4. When describing the use of ensemble members in the data assimilation section, you state that you perturb some forcing variables. Why were those specific variables perturbed? You describe the shape of the perturbation but not the magnitude. What were the ranges of perturbation magnitudes and how were they selected?
5. The correlation length scale is stated as 1.5 in line 283. That is a unitless number. What does that equate to in terms of meters? Is it 30 m (1.5 grid cells)? Does that mean there is no correlation more than two times that distance away based on your explanation in line 280? There has been a lot of research on spatial correlation of snow depth and you need to tie your choice for this parameter to the literature. Right now you state that it was chosen to make the “size of the resulting neighborhoods acceptable”. Acceptable to who or based on what? I recommend looking at <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020WR027343> and references therein regarding spatial correlation length scales. You could calculate variograms for your drone-based snow depths to determine the most appropriate scale for your study region.
6. The model has been described in more detail by the authors in their other publications that are cited in this manuscript, but it would be helpful to have a bit more detail in places. For example, in lines 286-287 it is stated that ICESat-2 snow depths are “spatially propagated” but fSCA information is not. What does this mean exactly? In the spatial propagation section you describe certain parameters that can be extracted from digital elevation models and how they are calculated over various distances. But there isn’t a clear explanation of how the ICESat-2 data from outside the drone domain are spatially propagated. There is also no description of how the data are actually assimilated. Are the downscaled ERA5 data used to estimate snow patterns and then FSM2 adjusts tunable parameters to better match the fSCA maps? Is this what you are trying to explain in line 302? Everything is fairly disconnected as is and the reader needs to have a

general idea of how the modeling works without having to go back and read multiple other journal articles.

7. Figure 4-6: I really like that all the ensemble member's basin-averaged snow time series are shown in these figures, I like the color palette for the maps, and I like that the colors from the maps carry over into the histograms. That said, I think it is a bit of wasted space to keep showing the done map and histogram in every figure, especially since the map is also in Figure 1. I recommend showing a different map in Figure 1 to provide some added context and then merging these three figures into one multi-panel figure. In the merged figure, you could have the first column contain the legend for all the basin-averaged ensemble time series (which should be the same for all experiments but it is not) and then the drone peak snow map in the middle and drone peak snow histogram at the bottom. Then columns 2-4 would be the basin-averaged ensemble time series on top, snow depth map in the middle, and snow depth histogram on the bottom for experiments (C), (D), and (J). This would minimize redundancy and allow the reader to visually compare results a lot more easily.

### Minor Comments

- Line 17: "by strong" and a comma after "processes"
- Line 18: "and metamorphism"
- Lines 23-24: Either remove this sentence with the Dozier reference or rephrase. Currently it doesn't fit with the rest of the paragraph.
- Lines 34-37: Something odd seems to have happened with the formatting here. The Foster reference seems to be thrown into the middle of this very long sentence and the sentence does not make sense.
- Lines 36-37: "where mountain regions are masked out" hangs on at the end of this sentence like an afterthought but it is an important point. Rephrase to emphasize.
- Line 49: The coarse footprint of ICESat?
- Line 52: Significantly better sensor characteristics than what?
- Line 56: Rephrase to "measurement error on flat terrain" instead of having part of the description in parentheses.
- Lines 64-83: There is a lot of extra information packed into parentheses in these paragraphs. Revise the sentences so that most information is written into the sentence. The use of parentheses makes it more difficult to read. For example, just say "obtaining statistically optimal estimates" on line 77.
- Line 89: Either keep "despite" or "thanks to" in the sentence but do not include them both with one in parentheses.
- Line 110: Rephrase to "In contrast, Alonso-González et al. (2023) have shown..."
- Line 113: What is "(see comment 6 ?)"?
- Line 114: Add a space after ICESat-2.
- Figure 1: The use of dashed lines to show the zoomed in areas is confusing because the ICESat-2 tracks are also dashed lines. I recommend using solid lines for the zooms or the tracks.

- Lines 141-160: There is a lot of information repeated here that was already in the introduction. You don't need to provide all the details of ICESat-2, just the ones that are important for your work. Then you only need to include them in one place.
- Line 170: Replace "harvests" with "hosts"
- Line 184: Why select 60%? How sensitive are your results to a different threshold?
- Line 195: Why bring up the orbit of the satellite here? What do you mean by "footprint"?
- Line 197-198: This is an incomplete thought. You filter out the cells with steep slopes?
- Figure 2: I like the idea of this figure but I cannot see the gray "ground photons" in the top panel. Consider revising the figure so the very top panel shows all the data, a middle panel shows all the photon differences with respect to the DEM, and the bottom panel stays as is. You would remove the right panel.
- Line 211: "20 m spatial resolution of the simulations 3.6"? I think the 3.6 should be totally removed but this also makes me realize that you describe all the data you will assimilate before you really describe the basic model. You might want to flip that order, moving 3.3 to the top of the methods, because you refer to the spatial resolution of the simulations before you describe them.
- Lines 213-217: These sentences on the uncertainty are very confusing. You list a sigma of 0.34 on line 213 and then again on line 217. Are these the same uncertainty metric or are they different metrics that miraculously have the same value? If they are the same, only list it once.
- Line 222: Replace "7" with "seven"
- Line 222: Do you mean that you select the most spatially detailed versions of parameterizations that you can use? Or the most mathematically complex? Or something else?
- Lines 223-226: For all of these parameterizations, I would simply say "as a function of" rather than "depending on", "influenced by", "diagnosed by", etc.
- Line 232: Why is 400 appropriate?
- Line 236: I am a bit confused by this sentence. If you are looking at figure 1, do you mean that you downscale for each cell in that large spatial domain in the left-most map? Based on your number of cells it seems unlikely. Your description does not sound like you only cover the small drone-based area. Do you also downscale to all the cells underlying all the ICESat-2 tracks or just those two highlighted tracks?
- Line 239: This is the first mention of "the prior". Presumably this means the model simulation with zero data assimilation. That needs to be defined either in this data assimilation section or in the more generic modelling section.
- Line 242: "log-normal" instead of "logit-normal"?
- Line 247: How is it "clearly non-linear"? Is that an interpretation based on manual inspection? Is that explained in previous literature?
- Figure 3: Rearrange the panels, and add letter labels, so that they go from left to right according to the order that each variable is described in the text: TPI, Sx, then CSMD.
- Line 311: Explicitly state their measurement uncertainty.

- Line 321: “selected by the median operator”? Does this mean the map with the median snow depth out of all ensemble members?
- Lines 409-242: Here is where “the prior” comes up repeatedly but it was never well defined. Please address my earlier comment so the reader can more easily follow this discussion.
- Lines 475-490: I am not a data assimilation expert so found the repeated use of “hyper” to be confusing here. You mention hyperparameters as earlier in the text, which is fine, but you also refer to a hyperprior and say the experiments are almost hyper. Does this all just mean high spatial resolution?