**Review**

**Improved workflow for customized ICESat-2 ATL06 elevations captures seasonal mountain snow depths at sub-kilometer scale**

The paper explores how ICESat-2 satellite data can be used to estimate mountain snow depth more accurately by comparing satellite elevations with high-resolution snow-free terrain models. The authors show that with careful processing—such as reducing positioning errors and adjusting for terrain effects—ICESat-2 measurements can closely match ground and airborne observations. The study finds that the satellite performs best in areas with moderate slopes and deeper seasonal snow, and suggests that many mountain regions have conditions suitable for reliable ICESat-2 snow-depth observations. This approach could broaden the use of ICESat-2 for monitoring snowpack and supporting water-resource modeling.

**General comments:**

I find that the paper is well written and that thorough analysis has been performed, identifying limitations in using ICESat-2 for generating snow depth in mountainous areas, while also showing where it can be used. However, I would like to see more descriptions of the general way the processing is done, especially the generation of the hybrid ATL06 product. That description is currently lacking in my view but can easily be fixed. It would also be of interest to include surface classification directly from the number of return photons inside each segment. That would avoid, in my opinion, fully relying on imagery as I understand it, and instead use the inherent physics of the measurements to suppelemt the analysis.

**Line-by-line comments:**

L53: Should this not be 17 m instead of 11 m?

L56: “Comparing ICESat-2 data to an independently collected snow-free DTM introduces additional geolocation errors.” Can you state more specifically what you mean and why?

L70: 17 m or 11 m?

L73: “ICESat-2 returns have a geolocation uncertainty of ~4.4 m.” Add the error, which is ±6 m, and the fundamental product you are referring to.

L98: I would like some more details of the hybrid data product, as this is important for the study. I think at least a paragraph or two should be dedicated to that purpose to explain how the data is generated.
L176: “The snow-free ICESat-2 height residuals, h_residual, are the difference between ICESat-2 and DTM ground elevations when and where snow was not observed in near-coincident satellite imagery.” How were the snow-free conditions determined from the satellite imagery?

L79: ~11 or 17 m?

L181: Why is the “n_fit_photon” not used to calculate when you have snow or snow-free conditions, or used in combination with the imagery? The classification will be quite clear, as the number of return photons can be used to easily separate the two types of returns.

L192: Can you provide some more justification for why “h_mean” is used and not “h_li,” for the reader to get a better grasp of why it’s important to use it?

L206: Same question as before—can you use the photon count for each segment to determine snow-free conditions?

Figure 2: The text in the figure is very small, so I suggest increasing the font size to make it more visible.

L224: Can you mention the methods that were tested, so the reader does not need to go into the supplement?

L226–L234: Are these co-registrations different from the ones in the appendix?

L276: “Which is more than double the expected precision (4.4 m) of ICESat-2 geolocation.” The estimated standard deviation of the error is, however, 6 m, which would still fit within the 1-sigma error. I would not expect you to find an expected precision of 4.4 m, especially in regions of steep terrain.

L284: "We find that ICESat-2 snow depth has a negative bias of ~0.6 m and uncertainty of ~1 m regardless of co-registration approach." So, is there a need to apply the co-registration if these biases still exist?

L305: I would highly suggest that you perform a simple correlation-length analysis of the differences to get an idea of what the optimal comparison radius would be. That would better inform the maximum distance at which you can calculate statistics. Or at least provide a figure of the statistics as a function of your smoothing length (100 m, 500 m, 1000 m, and 5000 m). The optimal smoothing length would most likely be correlated with the average slope magnitude at each site.

L360: Could it also be related to the fact that applying time-variant co-registration reduces the number of samples available and biases the dataset toward specific slope/topographical regions, increasing the noise in the registration? Maybe looking at the number of return photons can help reduce this issue by reducing the impact of mixed surface types where snow and snow-free terrain overlap.

L416: How large are these negative values? To reduce the risk of biasing the snow depth when removing SD < 0, could you allow for smaller negative values to be kept, perhaps within some limit or error?

L430: I think grouping them into elevation zones rather than horizontal distance bins would be a more effective approach, as you will increase data density. That's why I suggested calculating the spatial autocorrelation: you can use that to first get all data within that distance and then group them in elevation bands.