

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
<https://doi.org/10.5194/hess-2021-281-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-281

Anonymous Referee #2

Referee comment on "Mapping snow depth and volume at the alpine watershed scale from aerial imagery using Structure from Motion" by Joachim Meyer et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-281-RC2>, 2021

Comments on " Mapping snow depth and volume at the alpine watershed scale from aerial imagery using Structure from Motion" HESS-2021-281

Meyer et al. present an evaluation of a snow depth map calculated from airborne visible images with the Structure from Motion (SfM) method. The snow depth map is compared with a synchronous snow depth map calculated from airborne LiDAR, considered as a reference. The authors suggest that the SfM snow depth map is suited for future large scale campaign as they find a little bias (<0.1 m) and a satisfying accuracy (NMAD <0.2 m) over 300 km² of mountain terrain partially covered with snow.

It is an interesting work which completes previous studies presenting new methods to map snow depth at high spatial resolution in mountain terrain using unmanned vehicle, airplanes or satellites combined with lidar or photogrammetry. The authors have a unique and rich pair of datasets at hand and they have made efforts to extract relevant information for the snow science community.

However, many points in the article need improvement. I listed below six major points which should be addressed. I know the amount of work which is implied by these remarks, but I believe it is necessary to make this article suitable for publication in HESS.

Major comments

Data and methods

1. The classification used is not clearly described. What date? What method is used? What are the different classes? It becomes very hard to understand l.144 : "Finally, the snow volume was compared, grouped by different surface classifications (snow, rock, vegetation; Figure 2a)"

It sounds like there is snow on pixels classified as rock and vegetation. But what is the category "snow" then? In the same way, Figure 7b is hard to understand : we are looking at snow depth on vegetation, rock and...snow? Related to the classification : how are calculated the snow cover area (SCA) map? Is there an independent SCA from SfM data and one from lidar data? Otherwise, how is there a difference between both SCA (l.182)?

2. What is the interest for this study to calculate the SWE? Here, this work focuses on mapping snow depth. Showing that SfM snow depths match lidar measurements is enough to show that valuable SWE maps can be further derived. Plus, as long as it is not clear how the SCA maps are calculated, the difference between SfM SWE and lidar SWE is hard to interpret. Using a single density factor and comparing it to a complex spatialized model could be the topic of another entire study.

Results

3. One main result is that there is no bias between SfM and lidar snow depth (Table 2, l.180). However, this is not what is suggested by Figure 5.a which is commented l.199 : "As a whole, SfM showed an underestimation of snow depth compared to ASO, using the 1 m resolution pixel-by-pixel values (Figure 5a)."

In this figure, one can see that the heat map is not centered at all on the one-to-one line.

It is divided into two populations : one large (presumably from what can be inferred of the color scale) in which SfM snow depths are inferior to lidar and one in which lidar snow depth are between 0 m and 1 m and SfM snow depth are greater. Thus, it seems inaccurate to conclude that there is no bias in SfM snow depth. Besides, the latter population (where ASO snow depth is between 0 m and 1 m) needs more explanations. Is it related to errors in the classification? To errors in the lidar/SfM snow-on/snow-free DEM? To the downsampling of the ASO map?

4. Another main result is that SfM snow depths are less accurate for shallow snowpack. Is its accuracy relative to the snow depth? I cannot really imagine a reason for absolute accuracy to be worse for shallow snowpack. The surface of a shallow snowpack should appear just the same as the surface of a deep snowpack on an airborne image. Please give us your opinion on this point.

Comparison to existing studies and novelty of the work

5. The article is not strongly embedded in the existing literature. The need for more evaluation of SfM snow depth is justified by a single sentence in the introduction (l.56-l.59). This seems a bit short as three previous studies calculated snow depth map from airborne SfM. These studies are marginally used in the discussion. Bühler et al. (2015) is not even used any further in the article, although they calculated a 145 km² snow depth map with airborne SfM. The authors should clarify what is the added value of their study with respect to the current knowledge.

Also, the conclusions and findings of this article (Meyer et al., 2021) should be compared with the ones from Meyer and Skiles (2019). The main innovations of this article (2021) are that a snow-free DSM is used and have to be correctly geo-located before differencing snow-on and snow-off DSM. What was learned from that? Is the accuracy measured in this article in line with the accuracy measured in Meyer and Skiles (2019)? Since a snow-free DSM is used : are there larger errors in the snow-on or snow-free DSM?

Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P., and Ginzler, C.: Snow depth mapping in high-alpine catchments using digital photogrammetry, 9, 229–243, <https://doi.org/10.5194/tc-9-229-2015>, 2015.

Meyer, J. and Skiles, S. M.: Assessing the Ability of Structure From Motion to Map High-Resolution Snow Surface Elevations 420 in Complex Terrain: A Case Study From Senator Beck Basin, CO, *Water Resour. Res.*, 55, 6596–6605, <https://doi.org/10.1029/2018WR024518>, 2019.

Language

6. Many sentences do not read easily. Some are too long. It often brings confusion. I am not a native english speak myself but I feel that the manuscript should be carefully revised before next submission. For examples, see specific comments on l.22, l.60, l.104, l.118, l.142-143, caption of Figure 4...

Specific comments

l.20 : "certain environment" Please precise.

l.22 : "Snow depth and SWE" are "applied"?

l.27 : remove "differentially"

I.31 : "which resulted in varying degrees for monitoring snow depth accuracy" what does it mean?

I.32 : This article, as stated I.28, focuses on raster-based products. Treichler and Käab did not use rasters. I suggest removing its citation.

I.35 : To my knowledge, Marti et al. (2016) was the first study to calculate snow depth from photogrammetry satellite. To my understanding, no snow depth is calculated in Shean et al. (2016). They rather discuss how to "limit the influence of points acquired when seasonal snow was present on bedrock surfaces". Consider swapping both references.

I.44 : "up to alpine catchment size" do alpine catchment have a typical upper area boundary? A quick google scholar search of "alpine catchment" provides article studying alpine catchment of more than 1 km².

I.49 : Avoid "/", maybe instead : "high resolution and high accuracy"

I.49 : "over any desired target area" This is arguable. Some facilities are needed nearby.

I.59 : "the coincidental collection with this study"?

I.60 : "the coincidental collection [...] as well as precluding measurement errors due to manual recording such as snow probes" shorten this sentence (It enables precluding => it precludes), and check the meaning : it is not the coincidental collection which precludes errors in snow probes measurements.

I.61 : "the data **used**"

I.62 : "Meyer & Skiles (2019) showed that accurate DEMs can be generated from imagery collected from piloted aircraft over bright snow surfaces using SfM" It sounds a bit like this is first time it was done, in contradiction with I.56. Rewrite this paragraph to enhance the novelty of this work.

I.71 : "E E"

I.99 : Is there a reason for this difference in acquisition interval?

I.102 : It sounds like only an ASO snow depth map was used. At least the ASO snow-on DEM and the classification seemed to have been used. Include them clearly in the Data section.

I.103 : "the identical difference principle" what is this? If it is not a usual term, please delete or rephrase.

I.104 : "subtracted with" reformulate : subtracted from

I.117 : consider rephrasing with a more direct structure. Maybe something like : "Controle surfaces of the ASO snow-on acquisition flight was used as a reference."

I.117 : "ASO snow-on acquisition flight" : it is not the flight which is used. Is it a point-cloud? The 3 m DSM? Please explain in the data section.

I.118 : "have" instead of "are"?

I.118 : "identified" how? See comments on classifications.

I.122 : "An added advantage of co-registering of the SfM point clouds" delete "of"

I.133 : "3 m and 50 m."

I.135 : "with SfM supporting the higher resolution through the previously mentioned high point density," not necessary, long sentence, grammatically peculiar.

I.136 : "downsampled" : How? Which method? Can we estimate what snow depth accuracy can be reached after this operation? This is important. It is not advised to downsample any raster data especially since there is a high-density point cloud that could be rasterized at 1 m.

l.142-143 : "The analysis[...] binned the depth by elevation to assess similarities in the vertical relief" I don't understand.

l.155 : "stable terrain"? "control surface"?

l.163 : "the raster products exported from the respective SfM point clouds" long sentence. Maybe replace with "the DSM rasterized from ..."

l.164 : "the remaining difference in the NMAD" : not clear. Difference between NMAD and SD?

l.165 : A bad co-registration could occur and still provide a zero median/mean. Compare rather NMAD before and after co-registration.

l.167 : This is often done but could be discussed in this article. If the snow-on and snow-off lidar DSM are available to the authors, the difference of both should be calculated, producing a difference of lidar DEM (DoD). This would be immensely interesting to provide the NMAD of the control surface of this DoD and compare it to the NMAD of the SfM products. Please consider providing these numbers.

l.169 : Long, convoluted sentence. I would suggest something around : "Overall, there was good agreement in both snow depth and snow volume where snow depth was measured in both the SfM and ASO depth maps (Figure 3)." Introduce "SfM" and "ASO" somewhere in the methods.

l.175 : missing "and"?

l.177 : do we need "in the distributed product through NSIDC" Clarify this in the Data&Methods so that it does not show up in the results.

l.179 : maybe remove capital letters of Mean Median Standard Deviation?

l.183 : "The SfM SCA coincided with ASO at the 1 m resolution by 72%, 73% at 3 m, and 64% at 50 m, showing a small difference between the lower and higher resolution" Please rephrase : "to coincide by N%" is not a clear formulation to compare two simple figures.

I.192 : "a similar range of **map** resolution"?

I.194 : median (here) or mean (I.141) density?

I.194 : "less variable" this should not be in results but in discussion.

I.196 : "m" of SWE sounds odd for total basin SWE. Is it a common term?

I.199 : "As a whole, SfM showed an underestimation of snow depth compared to ASO, using the 1 m resolution pixel-by-pixel values (Figure 5a)." In Table 2 and in the paragraphs commenting it, ASO and SfM mean/median snow depth agree within a few cm. This seems contradictory. See main comment

I.201 : "highly localized"? small areas?

I.205 : "coinciding area"?

I.205 : are we talking about the land occupation under the snow? See main comment.

I.205 : grammar.

I.207 : "capturing 92 %" unclear.

I.208 : This is hard to grasp "For ASO, snow volume was distributed differently across land surface types in the entire watershed; 69% in open snow, 15% in rock, and 16% in vegetation." What is rock/vegetation? Rock/vegetation covered with snow? But what is open snow then?

I.209 : "in part" what else?

I.212 : "starting"

l.220 : treatment of negative snow depth and gaps should be explained in method.

5.3. Confusing section. I understand that "gaps" are defined as SfM without measurements and "missed area" as SfM with negative snow depth. Then l.220, "missed snow depth" have a positive (!) mean/median. And l.224-225 you mention negative snow depth value related to where SfM failed (!) to map snow.

l.248 : grammar?

l.252 "and the image RGB information (Shaw et al., 2020) or near-infrared spectrum (Deschamps-Berger et al., 2020)" These studies use satellite images. Here it sounds like they used airborne SfM.

l.264 : what is "perspective information"?

l. 266 : the first application of "photogrammetric snow depth products from satellite image" was by Marti et al., (2016). It is missing in the list of cited work.

l.269 : "We hypothesize that the snow-free scene, with more exposed vegetation and ground cover, degraded the accuracy for SfM." I understand that for vegetation. But why would ground cover degrade accuracy for SfM?

l.272 : "it is further feasible to align the two models to each other and compute snow depth and volume in relative geo-location space." Cite the work who did that, otherwise it is an assumption.

l.272 : "Alpine areas benefit from having exposed control surfaces for multi-view image processing and co-registration, having identifiable features in both scenes." I am not so sure about that. A counter-example would be an image acquisition after a fresh snowfall reaching the tree line.

l. 276 : "image" the influence of the resolution of the source image is not discussed here. It is rather the resolution of the snow depth map. Please rephrase.

l.278 : "the overlapping pixel wise difference" : unclear.

I.293 : satellites do not really have a larger coverage area (300 km²), but : less images, less radiometric depth, slightly lower image resolution, less known attitude (jitter)...

I.280 : "0.05 m (ASO, personal communication)." there must be a publication justifying this number. Otherwise, use the two ASO DEMs to calculate the difference over control area.

I.287 : "50 m there" Where?

I.298-299 : "atmospheric features like clouds" what other atmospheric features do you think of?

I.300 : "On the smaller scale using RPAS platforms, the accuracy is higher" sounds odd.

I.313 : could 6.5 be shortened and merged with the conclusions?

I.314 : too convoluted.

I.317 : Consider adding Brauchli et al. (2017)

Brauchli, T., Trujillo, E., Huwald, H., & Lehning, M. (2017). Influence of slope-scale snowmelt on catchment response simulated with the Alpine3D model. *Water Resources Research*, 53, 10,723–10,739. <https://doi.org/10.1002/2017WR021278>

I.319 : .)).

I.320 : I am not sure of the benefit of the first lidar flight. Lidar and SfM products have similar coverage and accuracy. Pflug and Lundquist (2020) suggest combining much more different datasets like a continuous accurate lidar snow depth with repeated measurements on a small portion of the terrain (<4 % of total surface).

I.322 : "we are encouraged that SfM can be an option" grammar.

I.324 : "it is further feasible **to source the images**" is it a common formulation?

I.324 : "from space-borne platforms" which satellites?

I.327 : "at hand"?

I.330 : "It also **emphasized to keep** the manual intervention for data processing to a minimum to be scalable with area size **and a readily available for operational use.**" grammar?

I.333 : "at 1 m, 3 m, with 50 m showing the largest difference" grammar?

I.334 : "at the all resolutions"

I.339 : "high resolution spatially complete data" too long.

I.341 : "contribute to explaining the consequences of our changing environment" too general.

Figures and tables

Figure 4. I do not understand what you mean with "compensate". I guess that the amount of missing areas at 50 m depends on how the setup of the IDW algorithm, doesn't it? Plus, the end of the caption is not clear "Higher resolutions at 1 m and 3 m [...] and successfully measured more areas with snow depths at edges of vegetation were more distinct." Scale is missing.

Fig 5.a : There is a constant set of point ~ 0.2 m in ASO snow depth. Please comment on

that.

Fig 5.b : I agree that the ASO snow depth should be considered as the reference. But here, the SfM smooth shape of histogram looks more convincing than the discontinuous shape of ASO's. Can you please comment on the discrepancy between ASO and SfM for snow depth $< \sim 0.6$ m.

I guess the two "mean" lines perfectly overlap. Find a visual way to show it like using two colored dashed lines and shifting the vertical origin of one of them.

Figure 6 : this figure could be more informative. Most of the colored area is pitch black. We cannot really tell how the two distributions compare. Consider changing the colorbar.

Figure 8 : "median SfM snow depth" : further in the caption, it sounds like only the negative SfM snow depth selected.

There are three orange dotted line, they cannot all show -1 m limit.

It is hard to tell at this scale but it does not seem like a "linear trend" for slopes $< 55^\circ$.

Table 1: Add the year to the dates.

Table 2 : Confusing what the column "Difference" is.

From the text (l.144) "calculated by subtracting the SDSfM from SDASO" => difference = SDASO-SDSFm

1st row of the table: difference = SDASO-SDSFm

All the mean and median row : difference (0.01) = SDSfM (1.06) - SDASO (1.05)

Table 2 : Include the note in the caption. What is "overlapping area by SfM"?