In "Intercomparison of photogrammetric platforms for spatially continuous snow depth mapping" several different photogrammetric approaches are tested in a sparsely vegetated study area in Switzerland. This manuscript is a thorough and useful comparison of state of the art photogrammetric tools for snow depth mapping. It does somewhat read like a commercial for the eBee+ UAS, however the authors have stated no conflicting interests and have demonstrated the eBee's advantages over the other platforms. I would recommend this manuscript for publication subject to a few minor changes.

Dear Dr. Bair

We thank you very much for your review. We absolutely have no conflict of interest concerning the eBee+ UAS. The institute from which most authors are affiliated uses from a range of manufactures. After testing multiple platforms for operational use in alpine environments and for mapping snow depth we determined the eBee+ had the best performance. However, we will change the text in such a way that the name eBee+ is mentioned less.

1) The vegetation issues are discussed, but a solution is not presented. I suggest identification of these regions and spatial interpolation may be the best approach, but there are other solutions. Note that negative snow depths can occur in glacierized areas or areas with persistent snow cover as well with snow on/off differencing approaches.

   Vegetation errors are not easy to eliminate, especially when generated by grassland with various length and uncertain compaction between snow-free and snow-covered conditions during mapping. We agree with the related comment added directly to our manuscript that this limitation of the photogrammetric method maybe insurmountable. In principle, we also agree that a DTM may produce a small improvement for some vegetation, but not for all vegetation. LIDAR in particular is more capable to mitigate this effect thanks to penetration through vegetation as we stressed in Section 6.5. Nonetheless, we disagree that interpolation from open areas onto regions classified as vegetated are the best approach in general. While it may be true on relatively uniform terrain, we believe that in mountainous regions the variability in topography maybe too great to be accurately captured by interpolation from remote open areas. In our opinion, it is of rather minor importance for this specific publication as the investigation is focused primarily above tree line. But we will take this point into consideration and add a comment in section 6.5 of the paper.

2) There are numerous grammatical and stylistic errors. I suggest an English language service be used prior to publication.

   The paper will be proof-read by a native English speaker.

3) The manual validation effort is impressive in scope but seems unnecessary. It seems to me that the resources used could have been better used on snow-covered and snowfree lidar flights, perhaps along with the Ultracam. And why weren't any forested areas sampled manually?

   We thank the reviewer for his appreciation of our efforts towards validation, although we respectfully disagree with the comment that it was unnecessary. We agree that in principle, repeated lidar flight would have been desirable for validation. However and despite what the reviewer may assume, the associated cost remains very high and would have exceeded the project funding, especially given that an additional manned flight in summer would have been

required. We did attempt a winter TLS scan, but due to various recording problems including suboptimal viewing angles, we could not georeference the TLS scan correctly. Therefore, the manual measurements are an important and necessary independent control for the snow depth maps. The large area comparison is there to further assess the Pleiades data, recognizing that the Ultracam is significantly better on the small scale test area, and hence can be used as reference. If "ground truth" is really what bothers the reviewer, then you could propose to rephrase with "reference" and adding the above element as justification.

We did not sample any forest areas manually due to lack of time and manpower. We wanted to concentrate on the Schürlialp area with manual sampling.

4) I don't agree that the Ultracam showed robust enough performance, especially inforested areas, to be called ground truth.

   We agree with you when specifically considering forested terrain. However, when considering the Ultracam DSM in non-forested terrain the RMSE was 0.17 m and the NMAD was of 0.12 m (raw data), which, we believe is sufficiently accurate to be compared with the Pleiades snow depth map over a larger area. The large area comparison is there to further assess the Pleiades data, recognizing that the Ultracam is significantly better on the small scale test area, and hence can be used as reference. We agree that ground-truth is not the right word here and will replace it with reference.

Responses to questions in the paper:

Page 6:
Figure 2:
What type of automated sensors were used, how often were the mesurements recorded, and most importantly, how were the depths filtered? Snow depths from acoustic sensors never look that flat. I'm guessing a single time is shown from each day?

For each day and each station one value is displayed. 5DF, 5MA and 5WJ are SLF observers reading the snow depth from a snow pole manually on a measurement field every day. The time for manual measurements is between 07:00 and 07:30. For the automatic stations, the newest value whose measurement falls within the time window 06:30 to 08:15 is displayed. The snow depth at the automatic stations is measured with ultrasound every 30 minutes. We will add this to the caption.

Page 8:
3: That's unusually accurate for a drone. Are there other UAS with similar geolocational accuracy?

Yes, current RTK UAS can attain such accuracies. Other systems with comparable accuracies are the WingtraOne or the Trimble UX HP.

Page 11:
7: Why wasn't the ALS company hired for winter acquisitions. That would have avoided the bent/missing poles and likely would have produced more accurate snow depths.

We agree with you that an ALS flight would have been useful for the comparison but unfortunately it was not feasible for the project as explained in our response to an earlier comment.

Page 14:
26: Any speculation as to why the RMSE in the vertical direction of the summer flight was half that of the winter flight?

At the time of the summer flight we had a better DGNSS (Stonex S800) available with an accuracy in position of 0.014 to 0.022 m and in altitude of 0.02 m. This of course affected the error calculation. We will specify this change of DGNSS in the revised paper.

Page 20:
These images are pixelated and I cannot see the violet stars in (d)

We will improve the image quality. The purple stars were unfortunately forgotten during the reproduction of the image.

Page 21:
10: Negative snow depths are errors. In addition to vegetation, glacierized areas or those with permanent snow can show the same effect. The vegetation problems are a significant limitation of the photogrammetric methods vs lidar. It deserves to be mentioned that the photogrammetric methods have sometimes insurmountable problems in vegetation. No well-versed reader expects one method to work well everywhere.

In many cases we agree with you, but dense vegetation (e.g. *Alnus alnobetula*) present in the study area can also pose a problem to LIDAR. Reutebuch et al., 2003 for example examined the accuracy of a lidar terrain model under a conifer forest canopy and found poorer accuracy for dense canopy. We are currently investigating this topic further by deploying LiDAR on a UAS platform, but this analysis will not be included in this study. But it is a fact that with photogrammetric techniques in dense vegetation you are at the limit at some point and you can never make a DTM like with a LIDAR. We will mention this in the discussion.

Reutebuch, S. E., Mc Gaughey, R. J., Andersen, H. E., & Carson, W. W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing, 29*(5), 527-535. doi:10.5589/m03-022

Page 24:
Panels a-d need to be captioned explicitly: Ultracam (a), Ebee +(b),....

We will add this to the caption.

Page 29:
5: I don't agree that the Ultracam was tested thoroughly enough to be used as ground truth, specifically its performance in forests

Please see the comment below.

9: All of this underscores the fact that photogrammetric methods are not suitable for estimating snow depth in forested areas, especially not as a ground truth. Without an accurate and independent validation datasets to compare to, e.g. snow on/off TLS or ALS, I suggest simply comparing the snow depths from the different sensors rather than claiming the Ultracam can be used as ground truth.

We agree with you on this to a certain extent and will reformulate this statement. However, the comparison of Pléiades and Ultracam is nevertheless a reasonable comparison when considering the accuracy of the Pléiades data over a large geographic area and in particular within non-forested regions, which makes up > 80% of the comparison.

Page 30:

Table: The average snow depth, based on the Schürlialp measurements is ~ 1.3 m. Thus, the uncertainty exceeds or is close to the average snow depth

Yes, good point. We will add this to the discussion of the results.

Page 34:

18: Vegetation is a problem for all remote sensing techniques in snow. Sometimes the best method is interpolation from open areas where you have more confidence in your measurements. Perhaps that is the best solution for filling these negative values?

We do not agree in this situation. We believe interpolation would give valid results in some areas and not others since snow depth is highly variable, even over short distances, and dependent on underlying topography.

Page 36:

10: And as you say, the snowpack needs to be deep. For example, the eBEE might be great for tundra snowpacks, except for the high windows and depths that are < 0.5 m.

Yes, good point. We will add a comment in the Conclusion.