Brief communication: Improving botanical monitoring in proglacial areas with high-resolution UAV data

Review by Jan Komarek

We thank the reviewer for their challenging suggestions and comments which helped us to improve the manuscript.

Dear authors,

Your paper addresses a relevant and timely topic of optimising botanical monitoring in rapidly changing proglacial environments. However, in its current form, the manuscript lacks the technical depth and thematic clarity required to bridge the gap between remote sensing and field ecology.

We might have miscommunicated the goal of the article. "Bridging the gap between remote sensing and field ecology" is a valid but very ambitious goal that we did not aim to tackle. We believe that the interest of this work lies in the detailed and innovative comparison of UAV observation and expert in situ observation, leading to a better understanding of the complementarity and differences between traditional human-based observation and modern technology-based observation.

The abstract is very general and does not present results. The conclusion here is vague. After reading, there is no need to read the paper. Sell your paper and attract readers' attention.

We modified the abstract to take into account this comment (common to all three reviewers), keeping in mind the 100-word limit:

"Proglacial areas are undergoing rapid ecological and land cover changes as glaciers retreat globally. Unmanned Aerial Vehicles (UAV) are particularly suited for monitoring such complex terrains but vegetation retrievals, such as vegetation cover, have rarely been evaluated against traditional botanical surveys. In this study, we compared the vegetation cover derived from UAV images at various resolutions with vegetation cover measured in situ by experts below the Aneto glacier (Pyrenees). UAV retrievals accurately captured the altitudinal gradient in vegetation cover and the spatial variability between vegetated plots. While increasing image resolution reduced biases, they remained significant in areas with sparse vegetation."

The introduction should also be more specific. RS (drones included) has supported vegetation mapping for years. What are the gaps? There is a known misunderstanding between ecological mapping and RS surveys based on thematic resolution, etc. I would appreciate it if you were more specific about the gaps and potential methods. All three RQs are vague, could be answered immediately. Read state-of-the-art and name specific gaps. Which vegetation cover did you map? Why such? Which strategy did you use for bridging in-situ and RS?

The article is focused on UAV vegetation cover in proglacial context. We purposefully do not include a review of all remote sensing methods used to map the vegetation cover. In this specific *context, a proper evaluation of UAV vegetation cover retrieval has never been presented, to the best of our knowledge. Following this comment, we looked further for such article and could only find works using UAV retrievals used as reference for evaluation of satellite products or vegetation cover mapping with UAV in different environmental setup.*

Thus, we disagree that our research questions could be answered immediately. If so, we would be glad to be pointed to the articles that we missed in our research.

Taking this comment into account, we modified the introduction and, for instance, now state that UAV has been used for vegetation cover mapping for a long time:

"Unmanned Aerial Vehicles (UAV) have long been used for vegetation mapping in other context (Li et al., 2023) but have recently emerged as a convenient solution for very high resolutionhigh-resolution mapping of complex terrain such as proglacial areas (Vidaller et al., 2024; Corte et al., 2024)."

We also made the research questions less vague:

"Here, we aim at filling this gap and answering the following questions: how do UAV vegetation cover estimates compare to in situ observations in a proglacial environment? What is the impact of the UAV images resolution on the vegetation cover retrieval? Additionally, we explore how can UAV data be used to complete or improve in situ sampling strategies in this adverse environment."

Finally, we agree that we did not comment enough about potential thematic mismatch. We see the value of this idea and expanded the parts referring to it in the discussion.

Methods. The study area is impressive. But I miss any information about vegetation cover. Perhaps you should introduce typical species (and their common sizes, as these are probably crucial for understanding your results).

We added details in the text:

"A total of 65 species were identified during the botanical survey with the most abundant across all plots being Leucanthemopsis alpina, Linaria alpina, Poa alpina, Saxifraga moschata and Veronica alpina. These species are characteristic alpine conditions found in periglacial environment. They formed continuous cover at lower elevation (A1) and occurred as isolated individuals at higher elevation (A4-6)."

I also wonder about the 3rd polynomial georeferencing. Any reason for doing this? Do you have a quantitative control of georeferencing? The systematic shift caused by GNSS should be treated as a 1st-order at maximum at most once the imagery is aligned using SfM.

The best georeferencing method was determined empirically. We agree that a pure horizontal geolocation error should be corrected with 1st order transformations. The improvement brought by the 3rd polynomial algorithm likely highlights orthorectification errors in the UAV data. We observe that this does not have a large impact on the final land cover classification and thus, the results of *the article. We do not have additional quantitative evaluation besides the residuals of the control points already described in the text.*

Note that we found a small error in the text while answering this comment. The 3rd order polynomial transform was actually used for the high-resolution products while the linear transform was used for the low-resolution product, inversely to what was stated first:

"The low-resolution images (70 m flight height) were georeferenced using a ~~polynomial of degree 3~~ linear transform and further used as reference to georeference the high-resolution images (30 m and 3 m flight height) with a polynomial of degree three ~~linear~~ transform."

Why was some land cover masked manually, and others using RF?

We decided to mask manually the snow and water covered area to keep a small number of classes. One might find it best to use the same classification methodology for all land covers but we decided to focus on the most relevant one for botanical studies: vegetation and bare ground.

Results. The results show a mixture of RF potential misclassification, with altitude as a confounding factor. The results indicate that a 10+ cm plant can be identified accurately using sub-centimetre imagery.

"Only two plots were available in the 1 cm map, which did not allow for robust statistical calculations." So why did you acquire them?

The UAV acquisition plan and the in situ sampling plan had to be adapted on the spot due to unplanned field conditions, leading to this small number of plots available. We now state it more clearly:

"Due to on-site modifications of the sampling strategy to adapt to the field, all the in situ plots were not within the the UAV acquisitions, resulting in 14, 12 and 38 points available for comparison for the UAV resolution of 0.1 cm, 1 cm and 5 cm respectively."

Figure 2 suggests that only 0.1 cm makes sense, while others underestimate. Well, perhaps it is connected with MMU, which probably should be defined at the very beginning, together with the botanists.

The in situ land cover was determined by dividing 10 units between the different cover classes per subplot of 20 cm by 20 cm (e.g. soil:1, rock:2, canopy:7). Assuming aggregated cover distribution, it corresponds to a minimal unit of 40 cm² or a square (pixel) of ~6 cm which is larger than any of the UAV products used here. However, due to the discontinuous distribution of vegetation, 1 unit of vegetation does not mean that a 6 cm square could be identified as vegetated in the UAV map. Thus, no direct relationship can be established between the in situ minimal unit size and the UAV resolution.

However, we agree that this is a relevant information and now include it in the method:

"Each plot was divided into 25 cells of 0.2 x 0.2 m, in which the proportion occupied by vegetation (including mosses), soil and rock were estimated by experienced observers with a 10% resolution.

This means that the minimal detectable cover area would be a ~6 cm square if the vegetation was aggregated."

Figure 3 presents the expected decrease in vegetation cover. It seems there is no overlap between in-situ and UAV data at A4-A6. Consider changing the whole chart. It would also be nice to present a table with a quantitative evaluation.

There is overlap between the minimum-maximum range of in situ and UAV data. To enhance this, we increased the thickness of the boxplots whiskers and made transparent the boxplots with no overlap between in situ and UAV data.

We are unsure in what way we should "change the whole chart" and would need guidance about this request.

Table S1 provides the statistics (min, max, mean, median, standard deviation, RMSE, Pearson correlation coefficient) of the vegetation cove from in situ observation, from UAV images and the difference of both. We added a reference to Table S1 in the legend of Figure 2 which is focused on the evaluation of both datasets.

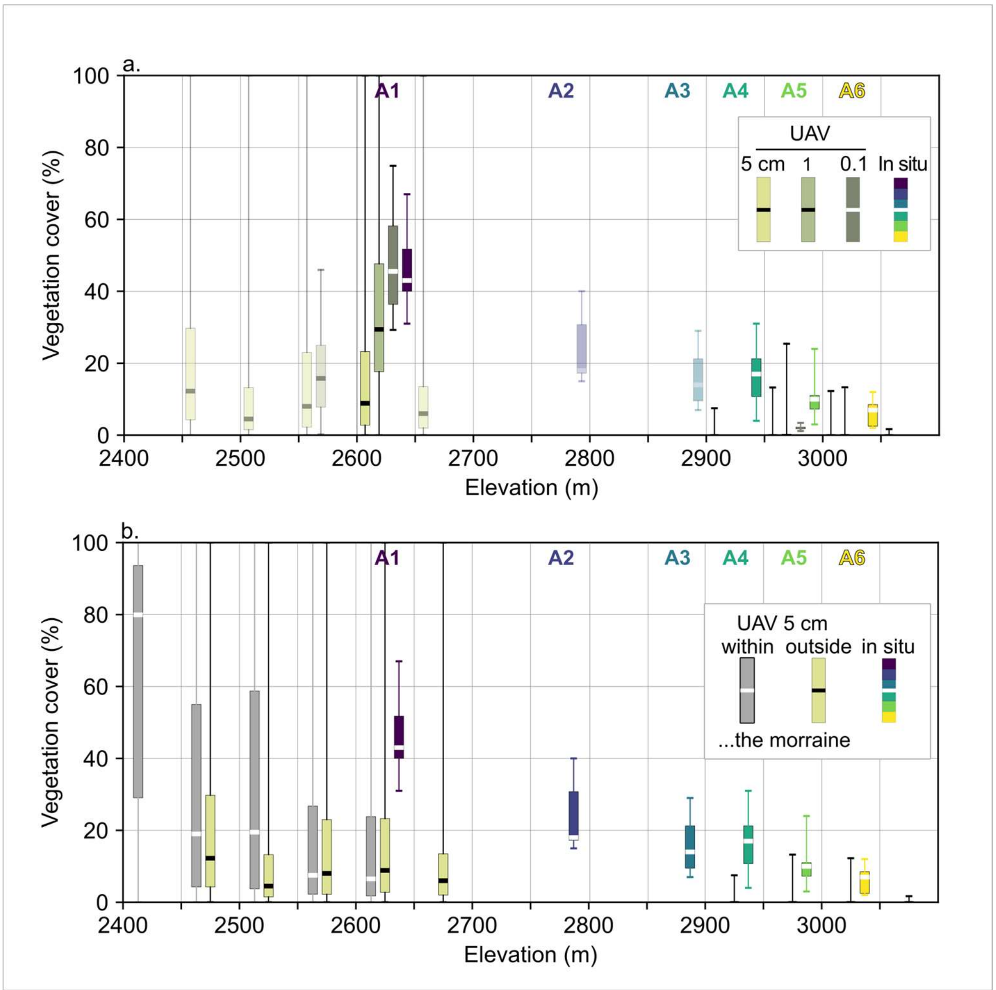

***Figure 3**. Distribution of the vegetation cover derived from UAV 5 cm, 1 cm and 0.1 cm images within the moraine extent and from in situ observation at the sampling points A1 to A6 (a). Boxplots were in situ and UAV data are not available are shown slightly transparent. Distribution of the vegetation cover derived from UAV 5 cm images within the moraine extent (i.e. once ice-covered*

during the LIA) in green, out of the moraine extent (i.e. not ice-covered even during the LIA) in grey and from in situ observation at the sampling points A1 to A6 (b).

Discussion. You conclude that the higher elevation would require higher spatial resolution UAV imagery. Without information about specific species and their size, it's kind of an empty proclamation. A figure from the ground may help readers to better understand. It is a bit difficult to understand the environment in which 0.1 cm is insufficient for vegetation cover detection. Sub-centimetre resolution is no longer a mixel.

We agree that 0.1 cm is already a very high-resolution and thus, we do not advocate blindly for higher resolution. We rather suggest that sensors with more pixels would enable to cover larger areas at a given resolution. We modified the text to make this idea clearer:

"The best agreement between in situ and UAV was obtained for 0.1 cm high-resolution images. However, they were acquired in static mode and could only be acquired over small areas. Since it is a very high resolution able to resolve most details of proglacial environment, there is no obvious improvement foreseeable by increasing further the images resolution. A good compromise between the area covered and the accuracy of the vegetation cover estimate could be achieved with UAV images with a spatial resolution of 1 cm to 5 cm, corrected with a few in situ vegetation cover measurements or a few higher resolution images. Furthermore, more accurate vegetation cover over larger areas could be expected with better sensors. The UAV of this study was chosen due to its low weight, but at the cost of a medium resolution of 20 megapixels. Sensors with 50 to 100 megapixels are becoming increasingly available and would increase the area covered in this study at a given resolution."

The shadow effect is not discussed at all. What was the proportion of shaded vegetation? How did this influence the results? That could mention bias.

Thank you for this suggestion. Combined with the confusion matrix suggested by another reviewer. This led us to present the proportion of terrain in the shade.

"Larger classification errors are expected in shaded areas but the proportions of rock in the shade is generally low (e.g. under 8% at 1 cm around A4-6) and could, thus, only partially explain the vegetation cover bias. Besides, the confusion matrix showed an overestimation of the vegetation in the shade due to the confusion with rock in the shade, which should actually compensate the UAV-in situ difference."

The in-situ usually includes mosses and lichens, which tend to be spectrally indistinguishable from the rocks or shadows. This creates a thematic mismatch. The reported underestimation may be due to the detection of vascular plants alone. This may introduce another bias.

The in situ includes mosses but not lichens. Thus, it is coherent to classify rocks and lichens together in the UAV images. The confusion between these two classes cannot explain the vegetation cover biases. Errors related to shaded areas is discussed in the previous comment.

Conclusion. I would not claim that “UAV vegetation cover showed clear bias compared to in-situ”. You need to deal with or discuss the bias first. There is obvious disagreement over what is considered vegetation by botanists and RS. And it was not explained, actually.

We believe that we improved the way we discussed the biases by answering the previous comments. We also modified the conclusion in that sense:

“However, UAV vegetation cover showed ~~clear~~ biases compared to in situ observations leading to a general underestimation of the cover. This bias is reduced with higher resolution images but at the cost of a reduced area covered. Misclassification in shaded areas could not completely explain this bias which could then result from thematic mismatch and errors in mixed pixels.”