**General Comments**

This manuscript presents the retrieval of snow depth using several UAS platform designs with variable optical sensor setups over three landcover types, to test the efficacy of snow depth retrieval in harsh arctic environments. The experiment is exhaustively described and well documented, with very articulate description of the drone setup, the weather during flights, and set up of the ground control points vs RTK positioning. The major result of the paper that is echoed by the community of UAV snow depth retrievals is that with increasing complexity in the canopy, the reduction of accuracy of the retrieved snow depth. However, the authors also present that whether GCPs are included in the scene or RTK positioning correction Is used that the different in error is negligible.

Response: We thank the anonymous referee for the review of the manuscript and the helpful comments and suggestions. Point by point responses (in blue font) are given below.

Overall the paper is very well written and I believe will be ready for publication with minor revision. There are some improvements that can be made with respect to the presentation of results, especially when having the capability to compare individual measurements to their associated snow depth retrieval. In the results section box plots are used to compare UAV-derived snow depths to the DoDs and snow line measurements. There is a place for including the box plot as it does present the distribution of the data, however it is difficult to ascertain where the most disagreement is occurring in the retrievals. I would expect to see a 1:1 scatterplot that compares the UAV-derived snow depths to in-situ observations. That would allow the reader to understand where the highest deviations are occurring. Are they occurring in areas of thin snowpack, or deep snow? I would also recommend showing the deviations on the map – are they locally clustered? Or homogeneously distributed about the map?

Response: Thank you for the suggestion. We can certainly provide the 1:1 scatterplots comparing UAS-derived and in-situ observation snow depths in the supplement. However, the suggested maps are somewhat problematic since it is difficult to display multiple values on one map. Due to the dataset containing three different UAS, two different baselines (snow-free models) for each, three different subplots, and four different dates, the amount of maps would be quite high. We would suggest that plotting the difference between UAS-derived snow depths and in-situ observations ($\Delta hs_t$) against the stake number (or alternatively distance along the transect) would essentially provide the same information. This, while still being able to provide more information on one graph and retaining readability. See examples below.

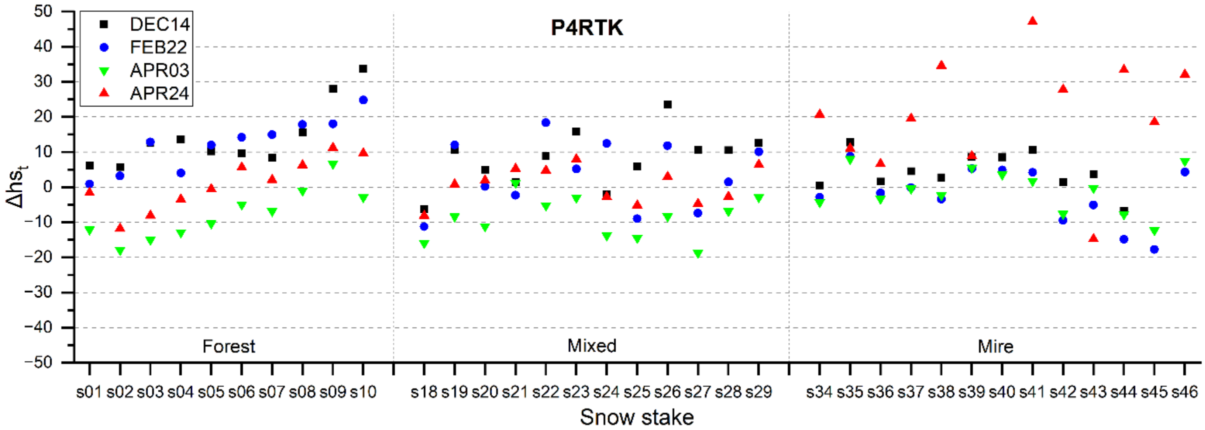

Fig. Difference between UAS-derived snow depths and in-situ observations ($\Delta hs_t$) against the stake number with P4RTK.

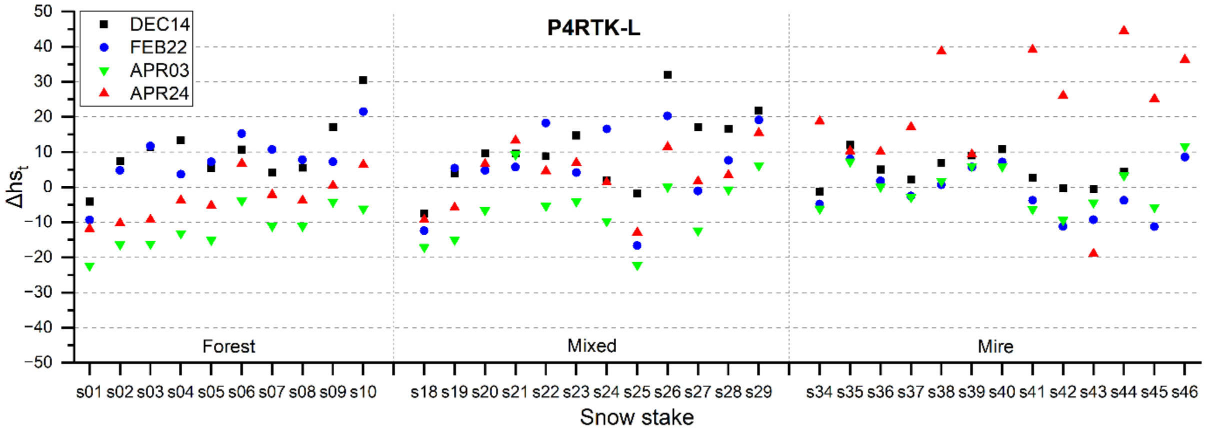

Fig. Difference between UAS-derived snow depths and in-situ observations ($\Delta hs_t$) against the stake number with P4RTK-L.

The Discussion section was well written, however quite long. Section 4.2, specifically lines 434 to 470 read more like a review of issues to consider when doing topographic or snow mapping with UAVs, discussion lens focal length, light conditions, horizontal/vertical accuracy, battery life, etc. I would recommend either shortening or removing this section for clarify of the results. However, lines 471 to 499 are relevant to the research being conducted in this manuscript, so I would suggest keeping it.

Response: We will shorten the suggested part and describe the issues more briefly while relying on providing suitable references for further info. Referee #1 asked for more emphasis and focus on recommendations and best practices, such as “the best platform for accuracy/ease of use (RTK vs. GCPs), general recommendations on GCP use, environmental operating suggestions (cold temps and wind), appropriate baselines, and operations in low light conditions”. Some of the aspects discussed in lines 434 to 470 are still relevant in that regard.

**Specific Comments:**

Page 2 Line 30: “snowlines” I see this more commonly referred to as snow transects in snow literature. I’m not sure if this is something that requires changing.

Response: In our view, both terms are used, but perhaps this is a more common practice in the Nordic countries. Nevertheless, all instances of “snowlines” changed to “snow transects” as suggested.

Page 2 Line 36: “Homgren” Should be Holmgren.

Response: Misspelling, fixed.

Page 2 Lines 45 – 49: There have also been some UAS work on freshwater lake ice to retrieve snow depth from structure from motion (Gunn et al., 2022), which could also be included in your description of UAS work in arctic conditions on line 54.

Gunn, G. E., Jones, B. M., & Rangel, R. C. (2021). Unpiloted aerial vehicle retrieval of snow depth over freshwater lake ice using structure from motion. Frontiers in Remote Sensing, 2, 675846.

Response: Thank you for the suggestion, citation added.

Page 3 Line 77: “mire” This is the first time you refer to “mire” areas. What are mire areas (aka bogs, wetland).

Response: Mire is commonly used terminology for areas with wet waterlogged soils forming peat. It includes wetlands, peatlands (bogs, fens, aapamires, etc.) and many other types. We kept the terminology ‘mire’ as it is commonly used and standardised terminology. Sentence now changed to give short description:

“One is an open treeless area with waterlogged peat soils, here referred to as mire area (approx. 14.4 ha).”

Page 6 Line 146: “placed an average of 52 meters apart” – One thing to keep in mind is that the stakes are placed quite far apart – how do you validate the spatial heterogeneity of snow depth when the spatial autocorrelation is typically around 50 meters? Or is that the reason that you’re choosing 50m as an average distance apart.

Response: We used a standardised snow line/transect measurement system following the protocol of the Finnish Environmental Institute (Kuusisto, 1984; Lundberg and Koivusalo, 2003). By measuring snow depth in every 50 m, the measurement contains different landcover types and randomly selected measurement points. This reduces autocorrelation in the measurements.

Kuusisto, E.: Snow accumulation and snowmelt in Finland, Publications of the Water Research Institute 55, National Board of Waters, Helsinki, 149, 1984.

Lundberg, A. and Koivusalo, H.: Estimating winter evaporation in boreal forests with operational snow course data, Hydrol Process, 17, 1479–1493, https://doi.org/10.1002/hyp.1179, 2003.