Our point-by-point replies (in black) to the question and comments by Reviewer 2 (in blue).

The manuscript "Permafrost degradation at two monitored palsa mires in north-west Finland" by Verdonen et al. presents an in-depth analysis of palsa degradation at two sites, relying on multi-year field data from a variety of sources. The manuscript is well written and I recommend publication after addressing the following comments:

We thank Reviewer 2 for insightful comments and many good questions, and are grateful for their time and effort in providing valuable feedback. We believe that addressing the issues raised by Reviewer 2 will substantially improve our manuscript, particularly regarding the descriptions of the methods applied.

1. Sect. 3.2: Please provide further details on the DEM generation, i.e. the number of GCP's employed, the accuracy in lateral and vertical direction as provided by the photogrammetry software. Please also provide details on how consistency in time was ensured, i.e. are there any stable points in the DEM's that could be used to check whether there are global offsets between individual years? The authors write "Variations in the UA systems, settings used in the data collection and the devices used to collect the coordinates of the GCPs resulted in discrepancies in the DEMs of different years. Therefore, we used the palsa polygons as delineated from the 2016 orthomosaic to extract only the areas of the main palsas from the DEMs. We then used the minimum value within that area as the base altitude for the respective year." It would be nice to motivate this procedure (which I don't question) from the uncertainties inherent in the DEM generation procedure, at least to some extent.

Thank you for pointing out missing details regarding generation of the DEMs. We agree that these procedures and uncertainties should be described better. Reviewer 1 also asked for more details about the aerial data used in this work. Therefore, we will add more information about the UAS DEM processing and uncertainties due to vegetation cover on the palsas in the first paragraph of Section 3.2, and add a table with details including acquisition dates, spatial and spectral resolutions, number of GCPs, and XYZ errors in the Supplementary materials. We will also change DEMs to DSMs when referring to the UAS-based elevation models as advised by Reviewer 1.

Unfortunately, we do not yet have fixed GCPs, which would allow a more accurate comparison of the different UAS DSMs (as suggested by Fraser et al., 2022). Thus, the only objects that could potentially be used as "stable" reference points at our sites were boulders around the mire areas. The issue with them is that they are located at the edges of the areas covered by the UAS surveys, where the "edge effect" causes uncertainties in the elevation.

2. 162 ff: Please explain in more detail why SnowModel is a suitable tool to reproduce snow dynamics in the extremely challenging environment on top of a palsa, i.e. present some key elements of the model physics, in particular on wind redistribution. The validation provided in favor of the model relies on an unpublished master thesis which I wasn't able to access with a quick Google search. Please provide more details on this work in the manuscript, i.e. include the main findings of the thesis in this study. From the information provided, it is not possible to assess whether the modeled snow data allow for a sounds assessment of long-term trends, thus also

affecting the Results section. Additional validation on snow onset and melt-out could possibly be obtained from remote sensing data, e.g. Sentinel-2, at least for years with infrequent cloud cover in the respective periods.

In addition to the validation presented in the Master's thesis of A. Störmer (2020), we are confident that the SnowModel (Liston and Elder, 2006) is suitable (although not perfect) tool to estimate snow dynamics on palsas because it takes into account different meteorological variables and climatic processes, local topography and vegetation height. In short, the SnowModel is a compilation of four sub models: (1) SnowPack to simulate changes in snow depth based on temperature, precipitation, melting and sublimation processes, (2) SnowTran-3D to simulate the effects of wind, suspension, saltation, snow erosion at slopes, accumulation at the end of hill slopes, and vegetation, (3) MicroMet to simulate the basic distribution of meteorological input for the investigation area, and (4) EnBal to simulate the energy balance of the surface based on the meteorological inputs of the MicroMet model.

We agree that the description of the SnowModel is too short in the current version of the manuscript. We did not describe it in more detail, as it was a complementary addition to the analysis to check whether the local snow depth values would show better correlation with the ALT compared to the values measured at the meteorological station in Kilpisjärvi. We also agree that the unavailability of the thesis, which is referred to, makes it impossible for the reader to check further details. Thank you for pointing this out! To address this issue, we will include a more detailed description of the SnowModel, the key elements of the model physics, the parameters used, potential issues, and a summary of the results from A. Störmer's thesis, which are relevant to the presented work, in the Supplementary materials of the revised manuscript.

Validation of the snow-on and snow-off dates with freely available optical satellite data, such as Sentinel-2 or even Landsat, would be a good addition. Unfortunately, frequent cloud cover over the study areas resulted in very few images that could be compared with the dates derived from the snow depth data at the Kilpisjärvi Weather Station. The few available images for the snow-off dates confirm what could be also expected based on the lower accumulation of snow on the palsa surface and game camera recordings at Peera from autumn 2016 until end of summer 2017, the actual snow-off dates are generally earlier for palsas than observed at the weather station. We will include this point in the description of the snow parameters in Section 3.3. Fewer images are available for the snow-on dates at our sites, but based on the game camera images, the difference is much less pronounced.

3. Sect. 4.1 The negative trend for the second site is very interesting – please add 1-2 sentences to highlight the procedure again, in particular that only values from the TOParea, i.e. the still stable part in later years, are compared. It is easy to miss this as a casual reader.

We will add clarifications that the values in this case are top-of-palsa ALTs. We will also include the results using all ALT values ≤ 1m to evaluate the effects of using only TOP values on temporal trends and correlations with climatic parameters. Because of this addition, we will state clearly throughout the manuscript whenever we refer to the top-of-palsa ALT.

4. Fig. 4: It is not really clear to which site the regression parameters and the R2 values belong, the one on the top also has a different color in some of the plots?

We agree that the colours of the texts within Figure 4 are confusing. We will edit the colours of the equations and R^2-values so that the ones for Peera will be in black, and the ones for Laassaniemi will be in light blue similar to the point symbols.

5. Sect. 4.2 I don't think it makes sense to present correlations that are not statistically significant, even if there is a trend. This is exactly the point of a statistical significance test. So for me the main conclusion of this section is that ALT is not strongly controlled by any of the tested parameters, except for the ones pointed out by the authors as significant. But also for these, it would be good to discuss the level of significance some more. This in itself is a very important result, in particular that the clear decrease in ALT for the second site does not seem to be controlled by larger-scale climatic drivers, but more by local factors which the authors cannot quantify at this point. I do not question the analysis (except perhaps the snow data, see above), but I think this section needs to be rewritten to some extent.

We kindly disagree about the relevance of including non-significant correlations, such as relationships between the ALT and air temperatures (Fig. 4 a, b, d, and e) in this case. The air temperatures have been found to be important factors controlling the ALT in palsas in other studies (Åkerman and Johansson, 2008; Sannel et al., 2016; Mamet et al., 2017), which makes it interesting that the relationships with these parameters were not that strong at our sites. And that is why we think it is important to present them in the results. We will also include this point in the discussion, as it has not been addressed in the submitted manuscript.

In the revised version, we will also include the R^2-values in the main text, as advised by Reviewer 1, and rewrite Section 4.2, so that it states more clearly, which correlations were statistically significant.

6. Table 2: can you add the corresponding data, i.e. 2016 and 2021, from the dGPS surveys to this table! The difference in absolute values seems to be significant between the two methods, so having a direct comparison of the same time slices is important.

The UAS DSMs, which were used to derive the values presented in Table 2, cover larger areas than the RTK-GNSS surveys from the same years. For this reason, we do not include the RTK-GNSS DTM values for 2016 and 2021 in the same table. Instead, we decided to add a new Section 4.5, where we will compare palsa height changes 2016–2021 based on the two methods. In this new section, we will also include the RMSEs for the elevation and height values between the RTK-GNSS and UAS DSMs.

7. Fig. 6: why are there two color legends (one in meter, the other in cm)? I think it could also be good to adjust the color scales and not use confining max-min-values. Right now one mainly sees the areas of full collapse, but it is equally important to be able to assess to what extent the main areas of the palsa have subsided. Furthermore, the authors should clarify to what extent the increases near the palsa edge are due to vegetation (i.e. is there vegetation of such height at all? Is the first survey taken after leaf-fall and the second before?), or the result of consistency issues between the DEM's, like global shifts, tilts or rotations (see also comment on DEM accuracy above).

We will enlarge the legend text and rearrange Figure 6 so that it is more clear that one legend refers to the differences in metres and the other refers to the differences in per cent. We wanted the same colours to indicate the same values for both palsas. To achieve this, we used 25 classes with 10 cm intervals (except for the min and max classes). The continuous legend for the changes in metres is a compromise, as showing all the 25 classes in the figure legend does not seem reasonable to us, and using fewer classes hinders changes at Laassaniemi. We acknowledge that it is difficult to assess height changes in the core areas of the palsas, however. For the revised version, we will look more into the options and will either include contours or adjust the colours to address this issue.

The leaves were still on during the UAS surveys in 2016 and 2021. The vegetation height varies from a few centimetres to ca. 50–60 cm. The shrub cover is particularly high at the northern edge of the Peera palsa. Although some displacement of the vegetation may have occurred, the apparent increases at Peera are more likely due to differences in the quality of the UAS data and the resulting DSMs. This will be clarified in the main text of Section 4.4 and in the caption of Figure 6. At the northern edge of the Laassaniemi palsa, the change was indeed due to vegetation growth, as the common cotton grass has expanded rapidly within the thermokarst pond.

8. 259ff: I am not sure about these correlations, is there any statistical significance? Also, the authors write that a higher value of snow onset (=later snowfall) correlates with a higher degradation rate for the second site (true?), but there is no correlation to e.g. fall air temperature? A later snow onset should rather lead to more ground cooling, except when the air temperatures are above freezing. If the data are like that, it is important to state this result, but the authors should check to what extent such correlations are statistically supported.

We will rephrase the paragraph about the correlations between the annual area loss rates and the climatic parameters, so that it states more clearly the lack of significant correlations and that the regression results are rather indicative because of the low number of samples. The R^2-values of the correlations mentioned in the text will also be included. In addition, we will move Table A2 into the main text, as advised by Reviewer 1.

The revised paragraph will read as follows:

> "Because of the low number of samples (only four periods), the results of the linear regression analyses between the annual area loss rates and climatic parameters are only indicative. None of the correlations were statistically significant at the 95 % confidence level (Table A2). At Peera, three parameters related to the winter air temperatures had the highest correlation coefficients, and p-values < 0.1. These parameters were MWAT ($R^2 = 0.88$), √FDD ($R^2 = -0.88$), and the number of days with air temperature < -10 °C ($R^2 = -0.87$). At Laassaniemi, the area loss rates had very little correlation with the climatic parameters; the lowest p-values were only around 0.3 for snow cover onset ($R^2 = -0.44$) and snow cover duration ($R^2 = 0.46$)."

Indeed, the regression results in Table A2 imply that at Peera, the area loss rates are higher in periods with later snow onset, although the correlation is not significant. At Laassaniemi, this relationship is opposite and is in line with the negative correlation between the ALT and snow onset at both sites (thinner ALT after later snow onset in the preceding winter). The opposing relationships are in agreement with our interpretation of the results concluding that the palsa area loss at Peera is more related to the changes in air temperatures, and at Laassaniemi, the area loss is more related to the changes in snow cover.

We did not discuss the correlation with autumn air temperatures. However, quick analysis showed that the area loss rates did not correlate with autumn (Sept., Oct., Nov.) air temperatures at either site. The analysis showed also that the mean autumn air temperatures have increased significantly 1960–2021 (+ 0.3 C per decade, $R^2 = 0.1$, $p = 0.012$), and the snow cover onset correlates with autumn temperatures ($R^2 = 0.35$, $p < 0.001$).

References

Fraser, R., Leblanc, S., Prevost, C., and van der Sluijs, J.: Towards Precise Drone-based Measurement of Elevation Change in Permafrost Terrain Experiencing Thaw and Thermokarst, Drone Syst. Appl., https://doi.org/10.1139/dsa-2022-0036, 2022.

Liston, G.E. and Elder, K. A.: Distributed Snow-Evolution Modeling System (SnowModel), J. Hydrometeorol., 7, 1259–1276, https://doi.org/10.1175/JHM548.1, 2006.

Mamet, S., Chun, K.P., Kershaw, G.G.L., Loranty, M.M., and Kershaw, P.: Recent Increases in Permafrost Thaw Rates and Areal Loss of Palsas in the Western Northwest Territories, Canada, Permafrost Periglac., 28, 619–633, https://doi.org/10.1002/ppp.1951, 2017.

Sannel, A.B.K., Hugelius, G., Jansson, P., and Kuhry, P.: Permafrost Warming in Subarctic Peatland – Which Meteorological Controls are Most Important? Permafrost Periglac., 27, 177–188, https://doi.org/10.1002/ppp.1862, 2016.

Störmer, A.: Modelling snow distribution over discontinuous permafrost related to climate change in Kilpisjärvi, Finnish-Lapland, M.S. thesis, Faculty of Natural Sciences, Gottfried Wilhelm Leibniz University of Hannover, Germany, 2020.

Åkerman, H.J. and Johansson, M.: Thawing Permafrost and Thicker Active Layers in Sub-arctic Sweden, Permafrost Periglac., 19, 279–292, https://doi.org/10.1002/ppp.626, 2008.