We thank the anonymous referee #2 for his/her thorough review with constructive comments and suggestions that certainly will improve the manuscript. In the following, we will address the referees' comments point by point. We mark "black" the comments given by the referee, and our responses in "blue".

Comment on tc-2022-124

Anonymous Referee #2

Referee comment on "Topographic and vegetation controls of the spatial distribution of snow depth in agro-forested environments by UAV-lidar" by Vasana Dharmadasa et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-124-RC2, 2022

The study by Dharmadasa et al. explores snow depth distribution at three Canadian sites (agro-forested and boreal forest) based on snow depth maps derived from UAV-LiDAR. Scaling behavior is investigated, and random forest models are used to assess the importance of various topographic and vegetation controls on these snow depth distributions. The topic of the study is of interest to the community, and in general, the manuscript is neatly organized, and the figures are well made. However, I have some methodological concerns that need to be addressed before this manuscript can be considered for publication. A major issue is certainly that the data at hand may not be sufficient to reach the conclusions drawn in the study – in the current state, the key findings and novelties of the study as well as the potential impacts of these findings are not highlighted well enough. Please find my major and minor comments detailed below, as well as some suggestions that I believe would make the study more novel and convincing.

We thank referee #2 for the challenging remarks and appreciate the overall potential he/she sees in our study. We acknowledge that the key findings, potential impacts of the findings and novelties were not highlighted well enough, and we will address this thoroughly in the revised manuscript. We will also take into consideration the reviewer's suggestions in the revised version that will significantly improve the quality of the manuscript.

Major comments:

1. In the methods, the authors should state more explicitly that the datasets are published already. Some redundancy with their article published earlier this year is unavoidable but should be kept to a limit. There are instances where the text could be shortened and simply refer to Dharmadasa et al. 2022, I have pointed these out in the minor comments. Watch out for self-plagiarism - Figure 1 and Table 1 are almost identical to Dharmadasa et al. 2022, and need to include a proper reference (e.g. 'adapted from').

   Thanks. We will add "adapted from" for Figure 1 and Table 1 and change the text accordingly when we refer Dharmadasa et al. (2022).

2. I am not convinced by the choice of aggregating the variables to different resolutions (10-20m) for the RF modelling and related analysis, in my opinion this raises some problematic questions:

- Firstly, I don't understand why the aggregation of the vegetation parameters is needed at all (L200ff). 'Canopy height' and 'Tree height' is not necessarily the same thing, and I see no problem with computing canopy height if the pixel size is smaller than the tree crown. Doing so would actually allow extracting the small canopy gaps that have been shown to be the main source of forest snow variability in some other studies (cited in this manuscript), while these gaps are averaged out if the pixels are aggregated to 10-20m resolution. This averaging is likely masking some dependency of snow depth distribution on forest structure.

- Likewise, averaging topographic variables will average out most of the micro-topographic variability that I understood to be the focus of the study. Again, some dependency between snow depth and these variables may be masked by this aggregation. It should be clarified whether the authors are trying to quantify micro-topography or topography.
- I find it particularly problematic to use an aggregation that is larger than the scale break identified in Section 3.2. Doesn't this mean that the variability that you are trying to explain is averaged out?
- Finally, aggregating leaves you with a rather small sample size, as the surveyed areas are quite small. Especially in the case of Montmorency, the field landcover type covers a very limited area only.
- I would find the analysis more convincing if it was conducted at the 1.4 m resolution, I suggest doing that in view of a resubmission.

  The rational for the initial analysis was to study the snow cover heterogeneity at the scale of single-tree. But we agree with your comment that aggregating does average out the variability smaller than the single-tree scale we considered. Considering your comments and suggestion, we repeated the analysis at the 1.4m resolution. Results show an overall improvement of the RF model performances at all sites (field+forest R^2 in Sainte-Marthe is 0.66, Saint-Maurice is 0.46 and Montmorency is 0.3 now). The importance of the variables at each site were only slightly changed. At agro-forested sites, windward forest edge effect and microtopography still have the highest impact on snow depth variability and in coniferous site, it is still the forest structure variability (LAI). But windward forest edge effect in forested areas at the two agro-forested sites are now evident compared to the coarse scaled data. We will replace the whole RF analysis by the updated analysis done at 1.4m resolution results in revised manuscript.

3. Some parts of the results chapters require more detail / explanation, and the discussion needs to be more convincing. Some examples:

- Section 3.1: The large overlap of forest and field histogram makes me wonder whether the difference between the two distributions is statistically significant – testing this would be appropriate (see e.g. Currier et al. 2018 for examples of such tests). The inter-site comparison is problematic because data acquisition did occur in different years.

  We will test the statistical significance of field and forest snow depths and include them in the revised version. Initial plan for the study was to acquire data in Montmorency in 2020 as well. But in compliances with the COVID-19 regulations that were in effect during that time, we were not allowed to access the site for surveys. However, the long term ECCC data shows that Montmorency always receives much higher snowfall than the two agro-forested sites. Therefore, we decided to utilize the data available at hand for the analysis.

- Section 3.2: It is unclear how the scale breaks were identified.

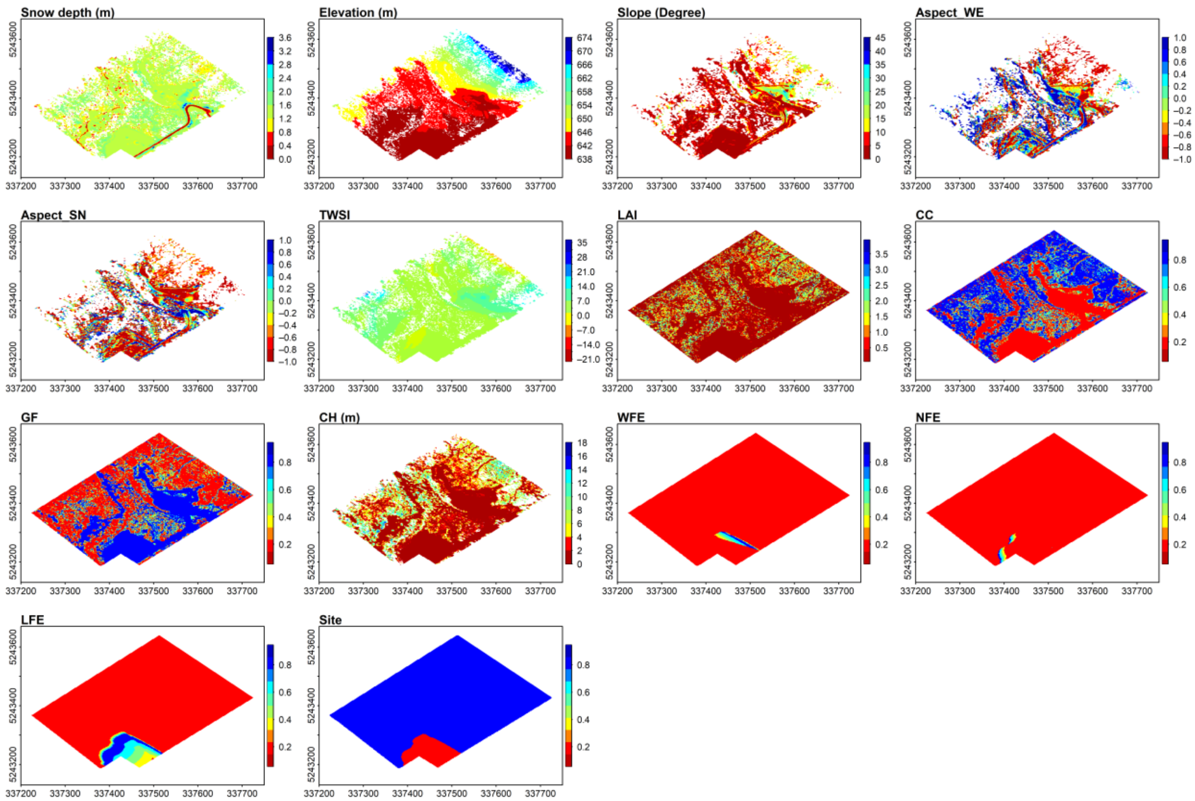
  We will update this section in the revised manuscript based on extensive comments made by reviewer #1 on this point. Please see our detailed response regarding the updated variogram analysis along with an explanation of identification of the scale breaks in AC1 report (response 1 to reviewer #1).

- Section 3.3.1: You need to better justify why you computed so many topographic and vegetation variables if most of these remain unused in the analysis / RF model. You also need to explain why you used the same predictors at all sites – for instance, NFE and WFE are homogeneous across the forested area in Montmorency forest so it is not surprising that they have no predictive power.

We discarded Elevation from the analysis since the elevation difference at any of the three sites was too small (Table 1) to produce any meaningful local orographic effect on precipitation, or adiabatic effects on air temperature, e.g., Mazzotti et al. (2019), and could mask other local topographic effects on accumulation related to slope, aspect and terrain roughness (wind sheltering), due to collinearity. And, irrespective of the variable type, we excluded collinear variables prior to building the RF models using the variance inflation factor (VIF) function in R.

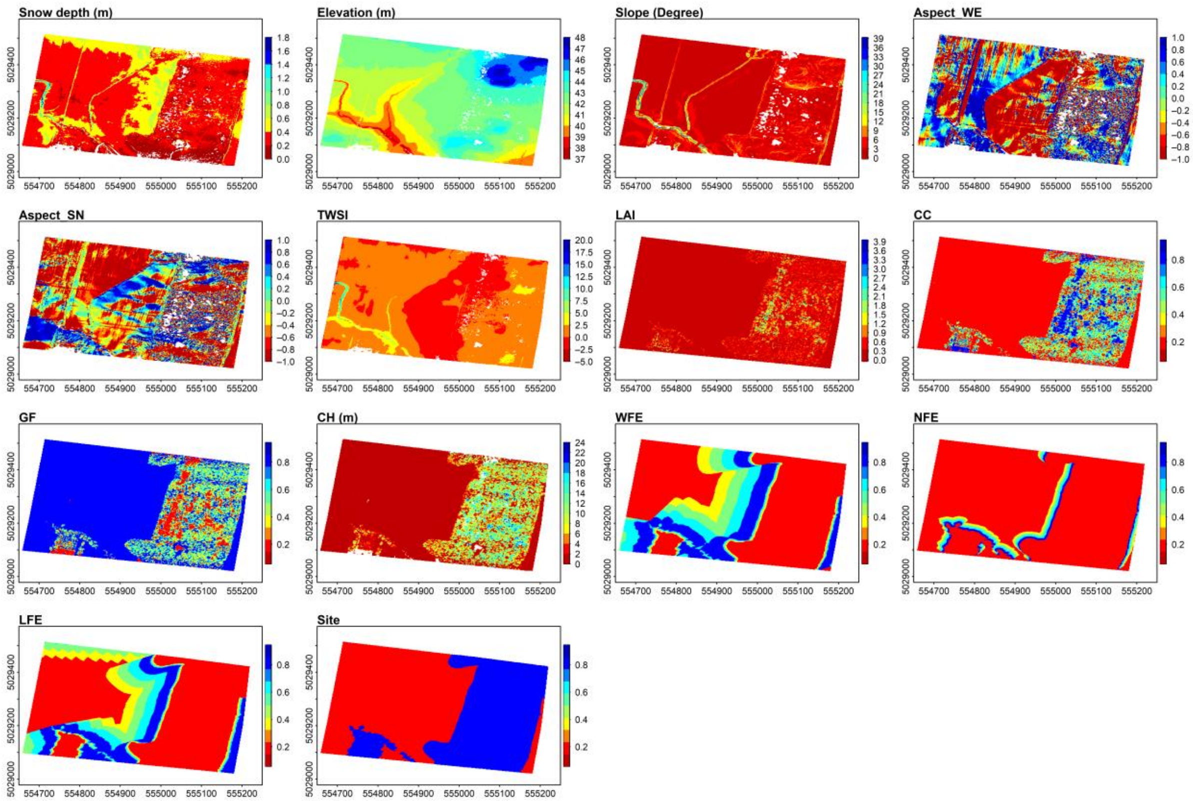
Vegetation descriptors (LAI, CC, GF, and CH) were strongly intercorrelated (with r of 0.84–0.97) and could not be used together in a predictive model (at least not without compromising the interpretation of variable importance in the RF model).Therefore, LAI was selected to use in RF analysis as it has been shown to be a strong predictor of snow accumulation (Hedstrom and Pomeroy, 1998; Broxton et al., 2015). In the revised manuscript we added a new sensitivity analysis which shows that the choice of forest structure descriptor has a negligible impact on the performance (R^2) of the model.

In the original manuscript we used the same variables at all sites for the RF analysis to allow for an easier intercomparison of the effect of driving variables between the sites. In the updated analysis (RF analysis with 1.4m grids), we decided to select the forest edge matrix guided by the landscape setting at each site. E.g., in Montmorency, we now include the leeward forest edge (LFE) matrix instead of windward forest edge (WFE) matrix because the leeward edge seemingly has more influence on snow depth variability with its larger extent than the windward edge (as shown in following figure). This is also more logical as the open areas in Montmorency constitutes a large gap within an overall forested environment, so deposition is expected leeward of the forest edge with little remobilization (erosion) within the gap. It is also confirmed by the improved R^2 of 0.56 in RF field in Montmorency. Also, at 1.4m scale, windward, leeward and northward forest edge pixels in the forest are more evident and variable than in the previously used coarse scale at all sites. The variables used in the updated RF analysis in Montmorency now include *Slope, Aspect_WE, Aspect_SN, TWSI, LAI, LFE, and NFE.*



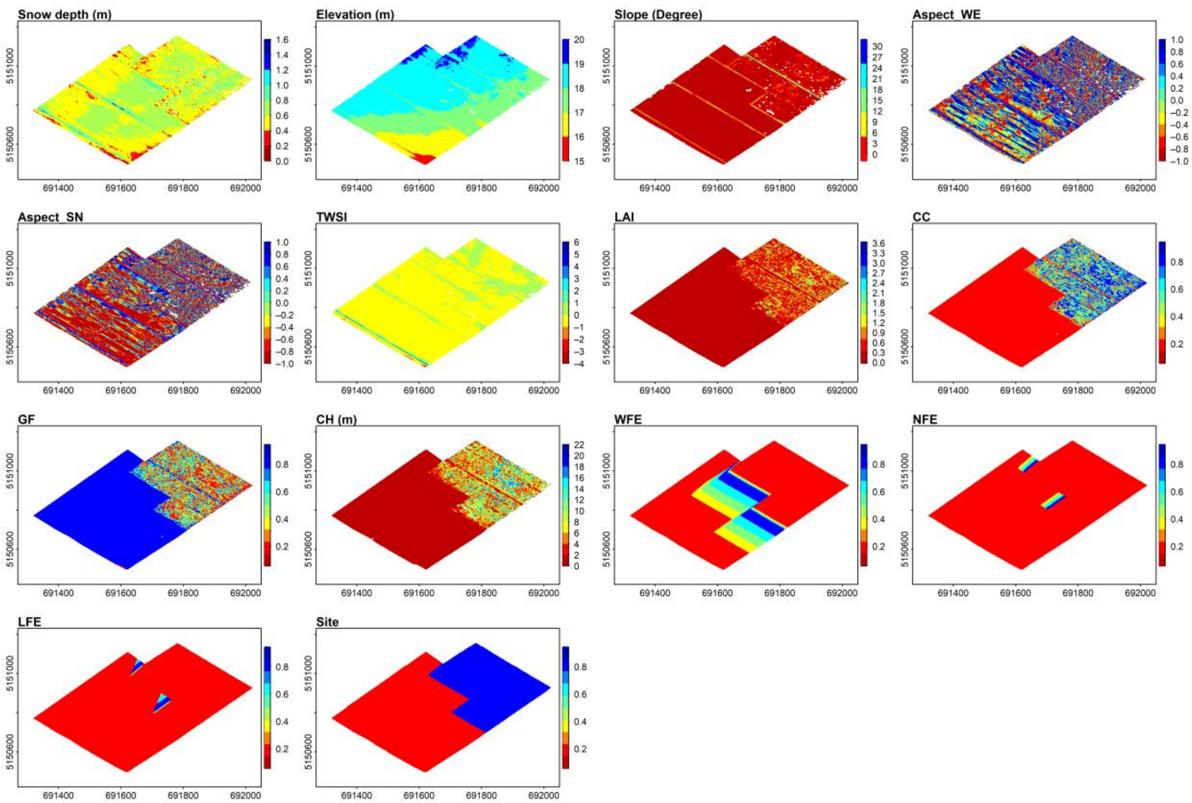
Montmorency variables

In Sainte-Marthe, both WFE and LFE have large extents (following figure) but are collinear due to the two dominant and opposed wind directions. Including both variables in the RF model would thus compromise the interpretation of the variable importance; collinearity generally reportedly makes it difficult to separately evaluate the predictive power (variable importance) of the predictors (Bair et al., 2018). Hence, we opted to use the WFE only in the final RF analysis and will now discuss this choice on the interpretation of forest edge effect at Sainte-Marthe. The variables used in the RF analysis at Sainte-Marthe are *Slope, Aspect_WE, Aspect_SN, TWSI, LAI, WFE, and NFE.*



Sainte-Marthe variables

In Saint-Maurice, LFE has only a few pixels (following figure) and was hence omitted from the RF analysis. The variables used in the RF analysis in Saint-Maurice are *Slope, Aspect_WE, Aspect_SN, TWSI, LAI, WFE, and NFE.* We will improve section 3.3.1. by better explaining the selection of variables at each site in the revised version.



Saint-Maurice variables

- Section 3.3.3: This section is a bit lengthy, and the key messages don't quite come through. Maybe it would be better to show fewer subpanels in Figure 6 and focus on the variables that actually exhibit interesting relationships to snow depth.

  Thanks. We will revise this section.

- In the discussion, you need to comment on the actual benefit of such an RF approach. The model performance shown in 3.3.4 is pretty poor, and it's not straightforward to extract the interesting information from the plots in Figure 6. In fact, I felt like the most insightful Figure to grasp the snow depth variability was Fig 3, where increased accumulation in sheltered locations is visible by eye.

  RF notably allows capturing non-linear relationships between snow accumulation and landscape variables. We will discuss the benefits of using RF approach in the updated discussion section, notably compared to traditional multiple linear regression models.

4. Given that the datasets have already been presented and used in an earlier study by the same work, the added value of the analysis presented in this manuscript seems a bit limited and the novelty of the study is not emphasized enough. What are the key findings, and where will they be useful? I realize that it's not easy to get more out of such a limited amount of data, but I tried to make sume suggestions that the authors could consider in view of a re-submission.

The earlier study solely focused on the accuracy of the UAV-LIDAR system, a necessary methodological step, and did not analyze the spatial heterogeneity of the derived snow depth maps. We thank the referee #2 for his/her insightful suggestions to improve the quality of our manuscript. We will adopt the suggestions in the revised manuscript as much as possible.

- If the authors have the opportunity to acquire more data in the upcoming winter, the authors should consider postponing the final analysis to after the upcoming season. Repeated flights over the same site would allow for a much more insightful analysis. It is very difficult to draw conclusions on individual processes based on snow distribution data from one acquisition only, which I think is part of the reason why the discussion does not seem very conclusive. For instance, it could be interesting to see if the snow depth maxima at the forest edges are a recurring feature, and if their effect persists throughout ablation (that's just an idea, but one could do much more). It would also be good to survey all sites in the same year to allow a more convincing comparison between sites.

  We agree that having repeated flights in the same areas in several winters would be better. But unfortunately, due to logistical reasons, we will not have the opportunity to acquire more data in the upcoming winter in these areas. This caveat will be mentioned in the discussion. However, a new, ongoing study will explore the temporal variability of the snow depth in a coniferous site. Still the analysis presented here is thought to largely reflect the typical conditions at the sites and to portray key differences between agro-forested environments and boreal forests. To place the studied years in climatological context, we will compare the climatological conditions in the study year (snowfall, air temperature, wind speed, and wind direction) with historical conditions at the sites and include in the revised version.

- Since the authors analyzed many more terrain and vegetation variables than they ended up using in the RF model, it could be interesting to dedicate a section on the physiographic variables themselves, attempting to identify a set of variables useful to characterize this sort of landscape. Maybe in comparison to variables that have been related to snow distribution in other studies. For example: Elevation has been found to exert a main control on snow depth in complex terrain in other studies, but it is not a really 'useful' predictor at the sites used in this study – is this a consequence of the site choice, or is this site representative of the terrain found in the whole ecoregion?

  One particularity of our sites (also related to the scale of the analysis) is the negligible elevation range. Many studies conducted in mountainous environments have shown the preponderant influence of elevation on the distribution of snow cover. While the elevation range becomes important over larger extent on the Canadian shield (Montmorency-type physiography), the low elevation St-Lawrence lowlands (Sainte-Marthe and Saint-Maurice) remain mostly flat and local topography (terrain roughness) and land cover and land use are expected to control the spatial distribution of the snow cover. We will add a section in the discussion in the revised version to discuss the potential predictors at similar landscapes.

- Exploring the relationships between snow and physiographic variables at different spatial aggregation levels could be interesting.

  Thanks. The updated manuscript will replace the initial aggregated scale by the high-resolution 1.4m analysis. We plan however to include a section in the revised paper to compare the RF results at 1.4m scale to single-tree scale results (coarse scale we used earlier).

- Applying additional modelling approaches (statistical, or even physically based) to compare with the RF model could be insightful, especially to draw conclusion on the utility of these findings for later work or practical applications.

  Thanks. We will compare the RF results with multiple linear regression at 1.4m scale and add them in the revised version.

- Adding an application of the model results would be a nice addition – e.g. suggest tiling approaches, extend to larger area or entire watershed, etc.

  Thanks. We agree that this will be a great addition to our study. However, we anticipate that the revised manuscript will have two or more pages in addition to the current 29 pages after including new sections and plots. As much as we appreciate this suggestion, we think addition of this suggestion would make the revised paper lengthier. Moreover, we think we will be able to highlight the novelty of the study by implementing the suggestions 2,3, and 4.

Minor comments (including wording/language suggestions)

L37 'topography and vegetation type, and density' -> you mean vegetation density? Sentence doesn't read very well, consider rephrasing

Yes. It is vegetation density. We will rephrase this sentence.

L46: a very nice and comprehensive paper on the topic: https://doi.org/10.1029/2011WR010745 - I suggest including this reference

Thanks. We will include this reference.

L51 'a short scale break is reflected by interception' -> I would say it's the other way round?

We meant that interception is causing the short scale break. We will change the sentence as "For instance, these studies emphasized that canopy interception causes a short scale break distance in forested areas (9–12 m) where the effect of wind redistribution is minimal".

L62: You should refer to much more recent developments of process-based models, as some of these models now actually do resolve small-scale variability due to heterogeneous canopy structure. See Broxton et al. 2015 (already cited elsewhere) and Mazzotti et al. (https://doi.org/10.1029/2019WR026129 and https://doi.org/10.1029/2020WR027572).

Thanks. We will refer to recent developments of process-based models and amend the text accordingly in the revised version.

L93: 'one of the earliest results […]'. I think this is not a very fair 'selling argument' for this study, since the data used for the analysis has already been presented in another paper.

Thanks. We will remove this sentence.

L108: incomplete sentence (WMO's station network?)

Thanks. It is WMO's station network. We will correct this.

Table 1: Winter season -> snow cover period?

Yes.

L136 is this vertical or horizontal accuracy, or both?

It is the vertical accuracy.

L164: what do you mean by 'multipath effect'?

Internal reflection of GNSS signals against obstacles (trees) before reaching the receiver.

L165: you just said the accuracy is comparable to previous studies, so what is the improvement? I would omit the entire end of the paragraph from 162 onward and just refer to Dharmadasa 2022.

We will remove the section from L162 onwards and refer to Dharmadasa et al. (2022).

Section 2.2.2-2.2.4: Please specify that maps of the variables are found in the supplementary material, I was missing those maps here and found them only much later. Note that the figures in the supplement should include the units for all variables.

Thanks. We will correct this in the relevant sections and add the units in supplement figures.

L197-199 Is GC = 1-CC? and at what resolution are these metrics calculated, also 1.4m?

Yes. They were also calculated as the same size as the LAI. We will add this in the revised version.

L201: How did you estimate crown diameter?

We calculated the crown diameter in LiDAR360 software. LiDAR360 uses Li et al. (2012) point cloud segmentation algorithm to segment individual trees. Crown diameter is one of the output we get from this segmented trees in the software (Greenvalley-International, 2020). We will describe this in the revised manuscript.

L224: This approach seems a combination of Currier & Lundquist and the DCE presented by Mazzotti et al 2019 (which however has no notion of search distance contrary to your method). Maybe worth noting?

Thanks. We will include this in the revised version.

L256: It is not very clear how you define the variable 'Site', or at least it wasn't to me when looking at the descriptor maps in the supplementary material and comparing with the other vegetation metrics. I think this is quite crucial for understanding the edge metrics, hence I more detail is needed here.

The binary variable, site was derived according to the field and forest area boundaries manually mapped at each study area. If there is a forest patch in the field, we considered that patch as forest and vice versa. For instance, in Sainte-Marthe we considered forest patch located to the southwest in field as forest in addition to the large forested area. But in Saint-Maurice and Montmorency, derivation of site variable was more straightforward as field and forest patches were well separated and so easily distinguishable than in Sainte-Marthe (i.e., there were no forest patches in field as in Sainte-Marthe). Once we delineate the forest and field boundaries, we assign 0 to field and 1 to forest. We will add more details to describe the derivation of site variable in the revised version.

L264 Tenses are inconsistent

Thanks. We will correct this.

L266: Variable importance of a variable? Consider rephrasing

Thanks. We will change this as "importance of a variable".

Figure 3: specify whether you used the binary variable or the land cover classification to create the histograms.

We used the boundaries from site variable to create the histograms. Figure 3 will be updated in the revised manuscript indicating the forest and field boundaries.

L308: how did you come to this conclusion?

Depending on the correlation coefficient value. Please note that in the updated analysis at 1.4m grid resolution, we now included the leeward forest edge (LFE) variable in Montmorency RF model, since it has more influence on the snow depth predictions than windward forest edge variable (which was also confirmed from the larger extent of LFE raster than WFE).

L310 'collinearity analysis suggested discarding GF and CH in favor of LAI at the two agro-forested sites, while LAI was instead flagged as colinear instead of GF and CH in the coniferous site'. Please rephrase – 'LAI was flagged as colinear' is unclear (colinear to what?)

Thanks. We will correct this in the revised version.

Figure 6: Y-axis label missing (snow depth)

Thanks. We will correct this in the revised version.

L445-446: Unloading through branches should reduce spatial variability, no?

Yes. It could. Lower semi-variance value in temperate forests compared to coniferous forest in figure 4 suggests that the overall spatial variability of the snowpack is less in these forests compared to that in coniferous forest. Yet, this could still result in a smaller correlation length in snow depth because of the interwind nature of branches on deciduous trees.

L511-513: The counterintuitive […] stations at the site. -> I don't understand what you are trying to say here.

We wanted to say, in contrast to preferential deposition on northerly slopes in the northern hemisphere, Montmorency field accumulates more snow on southerly slopes. This could be due to the influence of the various meteorological stations at the site. We will change the sentence in the revised version.

L565ff: this section needs to acknlowledge hyper resolution process based (physically based) models (see earlier comment). There are ways to account for fine scale canopy structure, while I would say that the terrain roughness still represents a major difficulty.

Thanks. We will address this in the revised version.

L483: I think this should be ‘Hydrologic response units’

Yes. Thanks. We will correct this in the revised version.

Section 4.3: The work from Safa et al. (https://doi.org/10.1029/2020WR027522) needs to be included in this discussion – they applied RF models as well.

Thanks. We will add this reference in the revised version.

Bair, E. H., Abreu Calfa, A., Rittger, K., and Dozier, J.: Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan, The Cryosphere, 12, 1579–1594, doi: 10.5194/tc-12-1579-2018, 2018.

Broxton, P. D., Harpold, A. A., Biederman, J. A., Troch, P. A., Molotch, N. P., and Brooks, P. D.: Quantifying the effects of vegetation structure on snow accumulation and ablation in mixed-conifer forests, Ecohydrology, 8, 1073–1094, 2015.

Dharmadasa, V., Kinnard, C., and Baraër, M.: An accuracy assessment of snow depth measurements in agro-forested environments by UAV lidar, Remote Sensing, 14, 1649, doi: 10.3390/rs14071649, 2022.

GreenValley-International: LiDAR360 User Guide, GreenValley International, Ltd, Berkeley, CA, USA, 2020.

Hedstrom, N. R. and Pomeroy, J. W.: Measurements and modelling of snow interception in the boreal forest, Hydrological Processes, 12, 1611–1625, 1998.

Li, W., Guo, Q., Jakubowski, M. K., and Kelly, M.: A new method for segmenting individual trees from the lidar point cloud, Photogrammetric Engineering & Remote Sensing, 78(1), 75–84, 2012.

Mazzotti, G., Currier, W., Deems, J. S., Pflug, J. M., Lundquist, J. D., and Jonas, T.: Revisiting snow cover variability and canopy structure within forest stands: Insights from airborne lidar data, Water Resour. Res, 55, 6198–6216, doi: 10.1029/2019WR024898, 2019.