Response to Editor

Dear Dr. Marsh

We are grateful to the four reviewers for their detailed and constructive comments on the manuscript and particularly to those reviewers who reviewed the original and revised submission.

In response to your specific requests, we have
- highlighted what are the main differences between the two study methods (in the introduction) and between their results (in the discussion). Details are provided in the first response to Reviewer #1.
- addressed referee #1s question concerning the use of the term “within cell variability” by simplifying the explanation, adding equations to explain our calculations and adding text that references Deems et al. (2013) to explain that the source of the variability is due to sampling limitations. We hope that this will allow readers understand this part of our paper,
- we have addressed referee #4s comments concern a common approach to report on accuracy that is used in much of the snow mapping papers to date that is an additional point that is similar to referee #1, and
- we have addressed the other numerous questions raised by all referees.

Our response to each comment is outlined below in **bold** and revised text is in red. The copy of the track changes manuscript is provided following our responses to the reviewers. We hope these responses are clear, and we look forward to submitting the revised manuscript.

Regards, Jennifer

Anonymous Referee #1

Thank you for the detailed comments and the opportunity to clarify that this article is the first to present snow depth maps measured with UAS-based lidar. We have provided detailed responses to the reviewer following each of the reviewer's comments.

Jacobs et al. provided substantial improvements in the manuscript structure and scientific content. Adding the comparison between forest type is especially interesting (Figure 6 and S3). However, some parts are still a bit confusing and would need some work to ease the readers' comprehension of the work. I think the article can be recommended for publication if the following minor points are addressed.
The two main points are :
1. The comparison with the work of Harder et al. 2020 seems a bit light. I think it is beneficial for the community to have several validation of similar methods published, but it would be interesting to highlight what are the main differences between the two study methods (in the introduction) and between their results (in the discussions).
We reviewed the earlier manuscript and concur that our manuscript is the first UAS-based lidar snow depth mapping manuscript when it was reviewed and during our revision. Their manuscript was not available when we originally submitted our contribution in February 2020. Only days before our resubmission, on June 15th, the Harder et al. 2020 was published. We appreciate the opportunity to now revise our manuscript to include the requested comparison.

The introduction was revised as follows:
Harder et al. (2020) compared snow depth estimates between lidar versus SfM techniques using in-situ snow depth observations in mountain and prairie environments, focusing on sub-canopy snow, which has been a challenge to measure in the snow remote sensing community. Using a considerably more expensive UAS-based lidar system (~$300K Canadian), they found that the lidar system tends to have lower errors than the SfM to capture sub-canopy snow distributions at moderate depth of snowpack (up to 2 m and 1 m of the maximum depth for mountain and prairie areas, respectively). In this study, we assess the ability of a more modest cost UAS lidar system (~$70K U.S. dollars) to map snow depth focusing on shallow and ephemeral snowpack (< 20 cm).

We also add a comparison of this study's results to their results in the discussion section as follows:
This study's lidar snow depth performance metrics are comparable to those from the more extensive lidar surveys made Harder et al. (2020), In the field, our snow depth errors, 1 cm bias and 1.2 cm RMSD, were modestly better than those from their open sites snow depth 3 cm bias and RMSE values on the order of 10 cm. In the forest, our snow depth errors, 7 cm bias and 10 cm RMSD, were also modestly lower than those from their forest sites 9 to 13 cm bias and 15 cm RMSE. While it is difficult to make direct comparison across different study sites, snow conditions, and ground validation approaches, these early findings indicate that UAS lidar has the capability of mapping snow depths in open and forested regions and has improved performance as compared to previous SfM results particularly for vegetated surface. It is also noteworthy that this

study's mapping was conducted using the Velodyne Puck series, a laser scanner adapted from the assisted and autonomous vehicle applications, rather than the specialized Riegl miniVUX-1UAV used by Harder et al. (2020) resulting in a complete mapping system that was approximately one-third the costs of their Riegl system.

2. I appreciate the clarification brought by the authors about the terms « precision » and « accuracy » as requested in my previous review. However the use of the «within cell variability » term is still a source of confusion for me. The « local scale variability »(L349)/« within cell variability » (L351) is measured with the « standard deviation of the lidar elevation values ». What is this standard deviation ? Please provide the formula or a clear description of it in the methods. Is there one standard deviation for each DTM or is it a pixel-based metric ? Is it related to sigma_on and sigma_off used to calculate the one-sided confidence interval (equation 1) ? If they are related, please make clear why one is the combination of the lidar accuracy and surface elevation variation (one-sided confidence interval, L234) and the other is the local scale variability of the ground surface elevation (L349).
A. The authors appreciate the reviewer's recommendations for clarification. A description of the standard deviation was added. The standard deviation is a pixel-based metric and is equivalent to the sigma_on and sigma_off used to calculate the one-sided confidence interval. The term terms referenced above were replaced by a single term "grid cell variability" which is now used consistently throughout the manuscript. The text now reads "The snow-on and snow-off standard deviation is calculated for each individual grid cell using all the individual snow-on and snow-off lidar ground return elevations, respectively. These standard deviations, measures of the variability of the snow-on and snow-off lidar ground returns within a grid cell, are referred as the grid cell variability. This variability depends on the lidar instrument's relative accuracy (Maune and Nayegandhi, 2018), which includes intra-swatch accuracy (i.e., precision or repeatability of measurements) and inter-swath accuracy (i.e., differences in elevations between overlapping swaths), as well as surface elevation variations. The contribution from the individual sources of variability was not assessed."

Below are minor points.
Title : Why was « unmanned » replaced with « unpiloted » ? I cannot find the term « unpiloted » in the snow depth UAS literature cited in the article (De Michele et al., 2016, Bühler et al., 2016, Adams et al., 2018, Harder et al., 2020, Eberhard et al., 2020).
A. The author team appreciates the comment. Perhaps it is a minor thing, but we would prefer to keep unpiloted. There has been a movement away from this term (see https://www.planetary.org/articles/10050900-finding-new-language for a discussion regarding NASA's policies). It seems harmless to go with unpiloted over unmanned, in our opinion. At worst it demonstrates our interest in engaging in inclusivity in the science world (even if the argument can be made that there's no implication of exclusivity in the word unmanned). There also is precedent in the scientific community (Davis et al. 2020; Wagner et al. 2019, among others).

Davis, J., Blesius, L., Slocombe, M., Maher, S., Vasey, M., Christian, P. and Lynch,

P., 2020. Unpiloted Aerial System (UAS)-Supported Biogeomorphic Analysis of Restored Sierra Nevada Montane Meadows. *Remote Sensing*, *12*(11), p.1828.

Wagner, M., Doe, R.K., Johnson, A., Chen, Z., Das, J. and Cerveny, R.S., 2019. Unpiloted Aerial Systems (UASs) Application for Tornado Damage Surveys: Benefits and Procedures. *Bulletin of the American Meteorological Society*, *100*(12), pp.2405-2409.

L61 : the citation of Deems 2013 and Lopez-Moreno 2017 suggests that they made a review of remote sensing methods for snow depth mapping while these article focus respectively on ALS and TLS. I would remove these citations as the list of the remote sensing methods is given in the next sentence.
A. Citations were removed.

L62 : ALS is airborne laser scanning, fine. L196 : « UAS laser scanning » becomes ALS. Confusing, especially in discussion : « the UAS lidar surveys presented in this study have key differences from previous ALS surveys » L458 and in the following paragraph. I would use two acronyms for airborne laser scanning (ALS since at least Deems et al., 2013) and UAS laser scanning (ULS ?...)

Deems JS, Painter TH and Finnegan DC (2013) Lidar measurement of snow depth : a review. *J. Glaciol.* **59**(215), 467–479 (doi:10.3189/2013JoG12J154)

A. Good point. After review, it appears that despite our statement "UAS lidar surveys (here after noted as ALS measurements)", we did not in fact use ALS to refer to the UAS lidar surveys. Harder et al. (2020) did not use the term ALS either for the UAS lidar surveys. Thus the statement "UAS lidar surveys (here after noted as ALS measurements)" was removed.

L76 : This sentence is too long. Split or reduce. Name the « snowpack features » ?
A. Modified as recommended.

L100 : now we want to know more about this paper...what do they conclude ? What does your work add ?
A. As noted above, the introduction was revised as follows:
Harder et al. (2020) compared snow depth estimates between lidar versus SfM techniques using in-situ snow depth observations in mountain and prairie environments, focusing on sub-canopy snow, which has been a challenge to measure in the snow remote sensing community. Using a considerably more expensive UAS-based lidar system (~$300K Canadian), they found that the lidar system tends to have lower errors than the SfM to capture sub-canopy snow distributions at moderate depth of snowpack (up to 2 m and 1 m of the maximum depth for mountain and prairie areas, respectively). In this study, we assess the ability of a more modest cost UAS lidar system (~$70K U.S. dollars) to map snow depth focusing on shallow and ephemeral snowpack (< 20 cm).

L141 : interesting information, but please explain the « to 3.6 V per cell » Cells are never mentionned before.

A. **The authors understand why confusion would arise here. Lithium polymer batteries are composed polymer cells in series. The batteries used on our system were a pair of batteries, each with 6 cells in series. A single cell at full charge has 4.2 V per cell, and 3.6 V at a safe discharge that the rotors can maintain altitude of the system. Rather than distract from the focus of the paper, we opted to remove the specifics about battery discharge, and note the total flight time from ascent to descent of the aircraft.**

L150 : please expand a bit on the geo-referencement. Is it adjusting one point-cloud relatively to the other ? Or are they completely individually geo-referenced ?

A. Lidar returns were individually georeferenced by synching timestamps of returns from the lidar sensor with timestamps of position and attitude data from the post-processed INS data. Georeferenced point clouds were produced and output to LAS files using Headwall Photonics, Inc.'s LidarTools software. This is now explained in the text as follows:

The bare-earth and snow-on point clouds were georeferenced solely using the INS data respective to each flight. The point clouds were not co-registered to each other as there were no reliable common ground control points between surveys. For UAS lidar snow depth surveying, co-registration between point clouds would likely be unattainable due to insufficient common ground control. We determined results would be more meaningful when bare-earth and snow-on point clouds were processed solely relying on the capability of the INS.

Note: A similar reason for not co-registering bare-earth and snow-on point clouds was presented by Goetz and Brenning (2019).

Goetz, J., & Brenning, A. (2019). Quantifying uncertainties in snow depth mapping from structure from motion photogrammetry in an alpine area. *Water Resources Research*, 55, 7772– 7783. https://doi.org/10.1029/2019WR025251

L174 : How are the boundaries of the field and forest « known » ? External dataset ?

A. Modified to be more specific.

L175 : Maybe remove the second « (Figure 1) ».

A. Removed.

L173-185 : I am a bit confused : « intact canopy » means with needles/leaves in winter ? Is it that the needless branches backscatter less the laser, making the deciduous tree canopy appear lower ? Add a sentence or two to explain that, although Figure S3

We thank the reviewer for suggesting to tighten this section up. We tried to make our methodology a little clearer and moved some sentences around as well as changing wording. The following section of reads as follows:

The snow-off DTM was used to develop a 1 m resolution map of slope (Horn, 1981). Vegetation cover type (field/forest) was determined from optical imagery. We developed a Canopy Height Model (CHM) by subtracting the DTM produced using ground-classified points from the DSM produced using all lidar points. This results in a digital model consisting solely of canopy heights with no topography. The CHM generation used raster images with a 1 m resolution. The forested area was further classified as coniferous or deciduous for the study region. Within the forested area, the CHM was used to distinguish the upper canopy that did not lose needles/foliage from other forested regions with trees with no leaves using our snow-off survey that was collected with leaf off in the spring. A 3 by 3 maximum convolve filter was used to enhance the edges of canopy crowns and expand smaller regions that might have just one pixel of an intact canopy or a hole in a larger canopy (Palace et al., 2008). A 15 m threshold was used to differentiate between the upper level intact coniferous canopy and canopies that had lost their leaves. CHM pixels that were below this threshold were deemed deciduous canopies (see Figure S3 in supporting information for intermediate figure). The 5.6 ha forested area has a forest type that is 65% deciduous and 35% coniferous.

L186 : Why extract 5000 points and not use all available points ? Even at the highest resolution (0.1 m) the 0.1 km2 raster should represent ~107 points which is not computationally unbearable. To be considered in future work.
A. A 0.1 km^2 area is 100,000 m^2.

L186 : I think you can simplify. « Once the vegetation forest type was classified, the raster binary image was vectorized. » is not really necessary. Something like : « Three sets of 5000 points were extracted respectively in the field, in the eastern forest and in the western forest. At each of these random points... »
A. Modified as recommended.

L191 : Has this test ever been used in geophysics studies ? It would be good to cite other papers using it. It is otherwise really hard to assess the relevance of this test without extensive statistical background. If it is relevant, this is a very welcome technic to estimate significant snow depth differences.
A. The Steel-Dwass test is a common test in environmental science and has been used in geophysics. It is a non-parametric test that is often used when data is not normally distributed. We have added the following text and citation to the manuscript as suggested by the reviewer.

The Steel-Dwass test has been previously used in geophysical work to examine non-parametric datasets (Slotznick et al., 2019).

Slotznick, S.P., Sperling, E.A., Tosca, N.J., Miller, A.J., Clayton, K.E., van Helmond, N.A.G.M., Slomp, C.P. and Swanson-Hysell, N.L., 2020. Unraveling the mineralogical complexity of sediment iron speciation using sequential extractions. *Geochemistry, Geophysics, Geosystems*, *21*(2).

L229 : Consider giving the one-sided confidence interval equation.

A. Reviewer #4 also gave specific suggestions regarding this section and terminology. Modifications were also made in response to Reviewer #1 and Reviewer #4. The paragraph now reads as follows:

The one-sided width of the 95% confidence limits ($CI_{95\%,+/-}$) for each grid cell's lidar derived estimate of the mean snow depth is a measure of uncertainty. The $CI_{95\%,+/-}$ values are used to compare the reliability of the snow depth estimates among cells. The $CI_{95\%,+/-}$ values were calculated using each grid cell's bare-earth and snow-on pooled sample standard deviation (s_d) and the number of bare-earth and snow-on lidar returns (n and m respectively) (Helsel and Hirsh, 2002).

$$CI_{95\%+/-} = t_{crit} s_d \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)} \tag{1}$$

A cell's pooled sample standard deviation (s_d) was calculated as

$$s_d = \sqrt{\frac{(n-1){s_{off}}^2 + (m-1){s_{on}}^2}{(n+m-2)}} \tag{2}$$

where s_{on} and s_{off} are the standard deviations of the snow-on and snow-off lidar ground return elevations, respectively. The s_{on} and s_{off} values are a measure of the grid cell variability. This variability depends on the lidar instrument's relative accuracy (Maune and Nayegandhi, 2018), which includes intra-swatch accuracy (i.e., precision or repeatability of measurements) and inter-swath accuracy (i.e., differences in elevations between overlapping swaths), as well as surface elevation variations and terrain induced errors (Deems et al., 2013). The contribution from the individual sources of variability was not assessed in the current study.

L229 : « of the snow on and snow off elevation » is confusing. Could you remove it ?
A. Removed as recommended.

L233 : « of the snow-on and snow-off **ground** returns » ?
A. Added the word "ground" to the two sentences noted.

L264 : «**absolute** low bias ». In absolute yes, in relative no.
A. Changed to "absolute low bias"

L266 : « mean height profiles » does not seem right. The heights are not averaged, they are normalized. It is rather a distribution of the normalized elevation, or something along that.
A. The text in section 3.2 was reworded to clarify:
To provide insight to differences between the forest and field observations, height profiles of classified returns were calculated for 25 m^2 square regions centered on all forest (n=12) and field (n=7) study plots from lidar data. Height profiles were averaged for each site type, from here on referred to as mean height profiles (Figure 4).

L272 : bring « respectively » earlier in the sentence.
A. Modified as recommended.

L321 : confidence interval **decreases**
A. Corrected.

L346 : « the ability to capture the mean snow depth » I disagree: one-sided confidence interval still includes natural infra-pixel variability. You might perfectly capture the mean of a pixel with high infra- pixel variability and still have a large confidence interval.
A. Point taken and we agree. The sentence now reads "In addition to the lidar point cloud density, the ability to narrow the confidence interval of the mean snow depth also depends on the ground surface variability within a cell as well as the lidar performance."

L359-L363 : this paragraph is a bit confusing. It would benefit from i) defining better the standard deviation (see main comment 2. above) and ii) explain the relationship between point density per cell and cell size. The last sentence of the paragraph implies that reducing the cell size reduces the ground return density. This should be explained.

L359 : **point cloud density** instead of « the point per cell »
A. Modified as recommended.

L360 : **snow depth map resolution** instead of « spatial scales »
A. Modified as recommended.

L387 : « reduced accuracy of the GNSS » : in forest ?
A. Modified as recommended.

Paragraph from L410 : This paragraph is a bit hard to read. It should be simplified and the conclusions better highlighted. See some suggestions below.
L412 : « While this result is not entirely surprising » : little added value of this sentence. I suggest removing it.
A. Modified as recommended.

L414 : « in snow depth due to pockets of duff and woody debris, and due to higher variability in subnivean terrain in the forested areas of the study site » are duff and woody debris not part of the subnivean terrain ?
A. Modified.

L414 : is it really related to « variability in snow depth » ? or variability in snow-off terrain since this is the only thing discussed after.
A. Modified as recommended.

L416 : Long sentence. Consider shortening with « drive higher confidence interval **in these areas.** »

L416 : High relief terrain is defined by high elevation variability over short distances. Also, I am not sure what this sentence is stating.

A. Modified. The sentence now reads "On steep slopes, there is more variability in ground return elevations over shorter distances, which would partially drive higher confidence intervals of ground surface elevation for pixels located in high relief areas." This is similar to the point made about L346.

L418 : remove « of the study site »

A. Modified as recommended.

L419 : Interesting result. Please repeat which is higher than which (deciduous, coniferous, field). Cite literature showing the canopy interception effect. Is it not surprising that deciduous and coniferous trees have the same impact interception ?

A. We have stated the mean snow depth earlier in the paper for each of the cover types. But in the section L 419, we believe the author is referring to on line 419, we are examining significant differences. We do find it interesting that coniferous regions had lower snow depth and could be likely to interception. What is additionally interesting is that the CI is higher under coniferous regions, highlighting the influence that lower ground returns due to intact canopy has on the estimate of snow depth. Figure 6 highlights the findings and statistical analysis. There is an evolving literature on forest canopy snow interception and impacts on snowpack distribution. However, most of the literature focuses on deciduous or coniferous forests rather than the mixed forests of this study region. The manuscript was modified to include references on snow interception as follows:

This indicates the possible influence of tree canopies on snow accumulation due to enhanced snow interception in forests (see reviews in Clark et al., 2011), and particularly in conifer stands, but also could be the result of an under-sampled ground surface in forested areas relative to field areas. Despite challenges with sampling in the forest area, some degree of coherence for snow depth in the forest is apparent. The forest interception effects may be captured on average through forest structure parameters such as canopy closure and leaf area index that have traditionally used in snow models with canopy-snow interactions (see reviews in Snow model inter-comparison project – SNOWMIP2 by Essery et al., 2009; Rutter et al., 2009). However, the finer scale heterogeneity may benefit from additional parameters such as the mean distance to canopy and total gap area (Moeser et al., 2016) or modifications that reflect variations in canopy structure (Mazzotti et al., 2019).

Essery, R., Rutter, N., Pomeroy, J., Baxter, R., Stähli, M., Gustafsson, D., Barr, A., Bartlett, P., and Elder, K.: SNOWMIP2: An evaluation of forest snow process simulations, Bulletin of the American Meteorological Society, 90, 1120-1136, 2009.
Mazzotti, G., Currier, W. R., 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 Resources Research, 55, 6198-6216, 2019.

Moeser, D., Mazzotti, G., Helbig, N., and Jonas, T.: Representing spatial variability of forest snow: Implementation of a new interception model, Water Resources Research, 52, 1208-1226, 2016.
Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., and Deng, H.: Evaluation of forest snow processes models (SnowMIP2), Journal of Geophysical Research: Atmospheres, 114, 2009.

L426, L440, you mention large UAV and heavy payloads. L372 the UAV seemed to be light (<25 kg) and small in size. Please homogenize.
A. This is a reasonable point. These sections were rewritten to homogenize and to differentiate the high lift UAVs capable of carrying a lidar sensor package from a UAV that supports SfM.

L463 : SD ?
A. Modified to read "snow depth".

L475 : please say what area is typically covered by airborne laser scanning campaign.

Figure 4 : In a and c a certain amount of points classified as non-ground is centered on 0 m height. This seems to be points which are wrongly classified. Is this assumption true ? Are they at specific locations ? Consider adding something about this in the results.
A. The data presented in these figures is from a large footprint around each plot. Because the peak of non-ground returns near 0 m is actually slightly above 0 m height, it is possible that the returns at and around that peak are misclassified ground returns, returns from low-lying vegetation (which is especially prevalent below deciduous canopy at our study site), or some combination of both. This warrants further exploration and could be examined by calculating relative height profiles at finer scales closer to our DTM (~1m), but was outside of the scope of this study.

Figure 6 a and b: there are some negative snow depth. This is possible due to the uncertainty of snow-on and snow-off DTM but should be presented in the results.
A. The text was modified based on the reviewer's comment to read:
Figure 6a also reveals that there are some negative snow depths in the two forest types that is due to the uncertainty of the snow-on and snow-off DTMs.

Figure 8. a.: See main comment about the standard deviation of lidar elevation. Here it seems to be a general metric : one standard deviation per snow depth map.
A. Agreed. The reviewer makes the point that the text needs to be clearer regarding what is presented. Figure captions were reviewed to ensure clarity about when the individual cell confidence intervals are presented versus when and how those values are being aggregated.

Anonymous Referee #2

Thank you for your input and detailed comments. We have considered this reviewer's comments in light of the comments from the other three reviewers.

This paper presents a snow depth map methodology for a UAV and LIDAR combination to measure thin snowpack of ephemeral snow. The study site contains a low vegetated field with a mixed forest which is useful to evaluate vegetation interaction. The main question focuses on the capability of LIDAR for snow depth mapping mounted on UAV because the sensor combination of LIDAR and UAV had not been extensively published yet. Most of the paper evaluates the accuracy of the map with respect to point cloud density (link to DTM resolution and LIDAR returns), vegetation cover and slope. A substantial amount of methodology and flight experience makes this paper focus on technical and methodological issues rather than informing us on snow processes of thin and ephemeral snow with vegetation interaction.

I do not think another methodological paper on snow depth mapping would be beneficial to the Cryosphere community. Differential snow depth mapping from LIDAR dates to the beginning of the century (Deems et al., 2006; Hopkinson et al., 2004) with airborne data from plane quickly became a more efficient tool to map large areas. A numerous numbers of article have used airborne LIDAR data for snow depth mapping (Currier et al., 2019; Grünewald et al., 2013; Hopkinson et al., 2012, 2004; Mazzotti et al., 2019; Nolan et al., 2015; Painter et al., 2016) with development on data processing (see review from Deems et al., 2013) and topographic and vegetation induced errors on LIDAR DEM (Spaete et al., 2011). Since the UAV platform only change the altitude and the coverage area, I do not see the point of having another methodological article. Regarding flight experience with UAV, a substantial amount of paper regarding snow mapping has already been published for smaller area with multi-rotor systems (Buhler et al., 2016; Cimoli et al., 2017; Fernandes et al., 2018) and larger area with fixed wing systems (De Michele et al., 2016; Harder et al., 2016; Redpath et al., 2018).

The high error in forested environment clearly need to be more investigated. It is stated that the magnaprobe measurements overestimate the snow depth in forested environment by penetrating low lying vegetation and soil. So, is the LIDAR mean depth of 6cm in more representative than magnaprobe (15 cm) and federal sampling tube (13 cm) which both showed larger depth for forested than field? Some forested areas can have deeper snowpack (Trujillo et al., 2009) due to wind reduction but snow interception by canopy will decrease snow accumulation on the ground as forest cover increases (Varhola et al., 2010). This needs to be sorted out especially when it is know that differential snow depth mapping underestimate snow for small vegetation as the snow off DEM is higher than the bare ground but when snow is measured for truth scene, the same vegetation is compressed by the weight of the snow (Buhler et al., 2016; Nolan et al., 2015).

The paper lacks a novel or in depth analysis of relations or processes regarding snow derived from a snow map that would be beneficial to the community like for e.g. a comparison with SfM photogrammetry in forested environments (Harder et al., 2020), statistical analysis of snow depth (Grünewald et al., 2013; Wainwright et al., 2017) or spatial analysis with variogram or fractal analysis between open field and forested environments (Deems et al., 2006; Redpath et al., 2018; Trujillo et al., 2009). To do this, I would suggest trying to improve with additional dataset over wider areas, only one site is not enough or perhaps different seasons to explore temporality and resubmit to this journal after next field season or go with this version towards a technical journal regarding UAV.

A. The author team posits that there are significant potential advances in snow hydrology as drones and lidar become more cost-effective, easier to use, and safer to fly. While there is excellent work using optical sensors on a UAV and the airborne lidar work (e.g., references by the reviewer), the use of lidar on a UAV platform is not trivial. Thus, before analyses of snow depth relations or processes can occur, the first challenge is obtaining research grade lidar data. The previous revision modified the submission to support the target audience seeking to collect UAS lidar data for snow hydrology. In addition, a more extensive discussion of the contribution appears in the previous response to reviewer #2.

Based on the reviewer's comments regarding observations in the forested environment, the manuscript has been modified to include a more extensive discussion of forest canopy and snow interactions as well as lidar performance.

We also agree with the reviewer's comments that "The high error in forested environment clearly need to be more investigated." As new remote sensing platforms lead to higher resolution snow depth measurements, the ability to conduct in situ ground sampling at resolutions needed to validate remote sensing observations will need to be sorted out. We are currently conducting a new study that focuses on sorting out the issues discussed by the reviewer.

Anonymous Referee #3

Thank you for the original review and the recommendations for minor revisions. We have provided detailed responses to the reviewer in ***bold*** *following each of the reviewer's comments.*

In my original review, I commented on the limited field data used to produce the paper. I appreciate the effort the authors went to place their study in context (their Table R1), and while I would like to have seen their analysis encompass far more results, I grant that the paper as now re-written has a distinct and useful purpose related to pushing new technology ahead. I particularly found the discussion of UAS/lidar use and limitations quite helpful. The discussion of forest and slope errors is also far more mature and comprehensive. I recommend publication with some minor revisions.

• GENERAL: There are some inconsistencies between how things are discussed in the text and how they are labeled in the figures. The authors should go through and try to make these consistent. For example, cell size vs. DTM resolution, and I think terrain slope vs. terrain variability.
A. All figure captions and axes were reviewed. Figures 2a, 7b, and 8a x-axes were changed to "Cell Size (m)". The y-axis in figure 8a was changed to "Cell Stdev (cm)". The terrain slope was clarified throughout the manuscript. The term "terrain variability" was modified to "ground surface variability" where appropriate in the text.

• In the abstract the authors say that in the forest there was " ....modestly reduced performance", but the degradation is 4x. Delete "Modest".
A. Modified as recommended.

• Perhaps it is that I have been doing research and spent 9 years as a sub- and chief editor of a journal, but I still found the introduction longer than need be. Anyone reading this paper is likely to be deeply involved in snow research and knows snow changes a lot, that our model are good but drift without real data to guide them, etc. I suggest just getting to it in the introduction: delete the first 3 paragraphs of the introduction....or at least make the point we need high quality depth data more concisely.
A. After reviewing the introduction, the authors agree. The first three paragraphs were removed.

• Line 78-80: This sentence is run-on and awkward.
A. Also noted by reviewer #1. This sentence was rewritten.

• Line 87: My understanding is that now DGPS level positioning on UAS's removes the need for ground control points. If that is true, perhaps moderate this statement.
A. There is an active, informal group of snow hydrologists who meet routinely to discuss SfM for SD mapping (led by Dan McGrath, Colorado State University). To date, that group's consensus is that despite the DGPS available on the newer and more expensive DJI systems, at least two to three ground control points are still

needed. The authors have chosen to keep the sentence.

• Figure 1: Much better and very useful and handsome.
A. Thank you.

• Figure 2: Last word (bottom). Not sure what this means.
A. The words in parentheses should have been “left” and “right”. Because the figures are labeled as “(a)” and “(b)”, these references were removed.

• Section 3.2: Well done.
A. Thank you.

• Line 309: Compare this to the statement in the Abstract.
A. The abstract was modified to match the referenced text.

• Figure 5: Great figure now even better.
A. Thank you.

• Line 351: Insert “the”
A. Modified as recommended.

Anonymous Referee #4

*Thank you for the original review and the recommendations for minor revisions. We have provided detailed responses to the reviewer in **bold** following each of the reviewer's comments.*

This is an interesting paper on using UAS based lidar map snow depth. This is an emerging tool that will undoubtedly advance observation of fine scale snow distributions and hopefully understanding of small –scale snow processes ultimately. Overall this paper provides a detailed description of a campaign in 2018/2019 which is important to understand the accuracy of this instrument and operational, considerations. I do have some concerns in terms of its placement in the context of the current literature, the communication of some of the accuracy results, and some of the more technical aspects of the writing. I would consider these minor revisions and once resolved look forward to its publication! I will summaries some main points that need addressing followed by some technical notes.

Literature context. The introduction could be made more punchy and to the point. Large scale context of snow is important but we're only looking at how to measure it at a fine scale in this paper so making it more concise on the progression in instrument techniques that have been employed would be an improvement.
A. After reviewing the introduction, the authors agree with this reviewer and reviewer #3. The first three paragraphs were removed and the paper now starts with the instrument techniques.

The Harder et al 2020 paper is cited as the first example of a UAS-lidar snow depth study. This is a clear comparable study but besides being mentioned nothing is else is mentioned about it in the intro. What are the common comparison points that can be used to relate the two studies? What differentiates the two (ie. Sensor quality/cost to get comparable results as this is labeled as a modest cost system (define?) versus Harder et al who use a higher cost Riegl system)?
A. Thank you for this comment. The Harder et al. (2020) paper was published only a few days before our revised article was submitted last summer. Based on this reviewer's and reviewer #1's comments, additional comparisons between this study and Harder et al. (2020) were added. We appreciate the opportunity to now revise our manuscript to include the requested comparison.

The introduction was revised as follows:
Harder et al. (2020) compared snow depth estimates between lidar versus SfM techniques using in-situ snow depth observations in mountain and prairie environments, focusing on sub-canopy snow, which has been a challenge to measure in the snow remote sensing community. Using a considerably more expensive UAS-based lidar system (~\$300K Canadian), they found that the lidar system tends to have lower errors than the SfM to capture sub-canopy snow distributions at moderate depth of snowpack (up to 2 m and 1 m of the maximum depth for mountain and prairie areas, respectively). In this study, we assess the ability of a more modest cost UAS lidar system (~\$70K U.S. dollars) to map

snow depth focusing on shallow and ephemeral snowpack (< 20 cm).

We also add a comparison of this study's results to their results in the discussion section as follows:
This study's lidar snow depth performance metrics are comparable to those from the more extensive lidar surveys made Harder et al. (2020), In the field, our snow depth errors, 1 cm bias and 1.2 cm RMSD, were modestly better than those from their open sites snow depth 3 cm bias and RMSE values on the order of 10 cm. In the forest, our snow depth errors, 7 cm bias and 10 cm RMSD, were also modestly lower than those from their forest sites 9 to 13 cm bias and 15 cm RMSE. While it is difficult to make direct comparison across different study sites, snow conditions, and ground validation approaches, these early findings indicate that UAS lidar has the capability of mapping snow depths in open and forested regions and has improved performance as compared to previous SfM results particularly for vegetated surface. It is also noteworthy that this study's mapping was conducted using the Velodyne Puck series, a laser scanner adapted from the assisted and autonomous vehicle applications, rather than the specialized Riegl miniVUX-1UAV used by Harder et al. (2020) resulting in a complete mapping system that was approximately one-third the costs of their Riegl system.

Accuracy results: There are many ways to present the results of this works but a common approach in much of the UAS-SFM, (and lidar), ALS and TLS is to consider the RMSE and Bias metrics prominently and these are reported in the literature review. This paper does present these values in Figure 3 and the first paragraph of section 3.2. The metrics reported RSMD of 1.22cm and 10.5 cm in open field and forest respectively are quite comparable to other work but these connections are not made. Rather only the 1.22 cm open field RMSE is reported in the abstract – no mention of the larger forest error or anything in the conclusion. In lieu of direct comparison with insitu data there is a focus on the one sided confidence intervals that are reported at < 4cm. Confidence intervals can be important to isolating areas uncertainty but are not fully suitable to assess the skill/accuracy of the product. If the results could be clarified to emphasize the RMSD and MAD more clearly and the meaning/role of the confidence intervals to be clarified more strongly would make this better.

Technical writing Writing is good throughout but there are some sections and word choices that need work – specific examples given below but not exhaustive. Specificity of word section and consistency throughout will make things much clearer.

Technical edits
Line 33:"differentially" – word choice. Awkward with "differences" earlier in sentence
A. The line was in an introductory paragraph that was removed when the introduction was modified.

Line 48: "High vertical resolution snow mapping" word choice. Perhaps "Mapping snow with high spatial resolutions and vertical precisions...."
A. The line was in an introductory paragraph that was removed when the

introduction was modified.

Line 51: “Despite importance” -> “Despite the importance”
A. The line was in an introductory paragraph that was removed when the introduction was modified.

Line 54-56: “ Using traditional, precise point measurements with a limited sample size, the experimental design requires a balance between the sampling extent and sample spacing (Clark et al. 2011). >: “ Using a traditional experimental design with precise point measurements of a limited sample size requires a balance between sampling extent and sample spacing”
A. The line was in an introductory paragraph that was removed when the introduction was modified.

Throughout: Are we using UAV or UAS? Title and intro start with UAV and then it switches to UAS at line 82.
A. Jennifer et al. We use UAV to indicate the vehicle and UAS to indicate the combination of the UAV and the sensors. The manuscript was reviewed for its use of UAV and UAS to be consistent with those definitions.

Line 140-146: I see there were batteries changes in this mapping mission. How long does it take to survey this area? Important information to present and perhaps bring into discussion at end on how scalable this technology is with actual times/areas to benchmark a discussion.
A. Good point. The 2 hr mapping time as needed to address UAS lidar scalability was added to section 4.3.

Line 151: “antiparallel” = perpendicular?
A. Antiparallel means the same line, but in the opposite direction. The sentence was modified to read “Boresighting calibration was performed using returns from the first two parallel flight lines that were collected in opposite directions (i.e., antiparallel).”

Line 152-154: The boreshighting and roll-pitch offsets need to be done for each flight? Or are these done once and they are stable thereafter?
A. The boreshighting and roll-pitch offsets were done once for the entire survey.

Line 155-156: What and how are points filtered?
A. From unclassified, georeferenced point clouds, non-ground returns are filtered out of the data set using the Progressive Morphological Filter. The resulting data set contains only ground surface returns. The ground returns were coded according to LAS specifications, then merged with the non-ground returns. We have attempted to clarify in the text:
Briefly, the PMF operates iteratively on sets of two parameters, window size and elevation thresholds to erode and dilate point cloud data sets to estimate surface topography. The result of the PMF is that non-ground returns (i.e., trees, shrubs, and

noise) are filtered out of point cloud data sets, so that only returns from ground surfaces remain. The two data sets, non-ground returns and ground returns from the original point cloud, are coded according to LAS specifications and merged. For a full explanation of PMF, see Zhang et al., (2003).

Line 174-175: (Figure 1) at end of sentence implies that it contains a methodology which it does not and is a little confusing.
A. The text in the referenced sentence "(Figure 1)" was removed.

Line 173-184: This section could be tightened up. Landscape classes where derived from CHM and simply > 15m was coniferous, deciduous was < 15m and field was what? 0m?
A. We thank the reviewer for suggesting to tighten this section up. We tried to make our methodology a little clearer by editing the section. The section now reads as follows:
The snow-off DTM was used to develop a 1 m resolution map of slope (Horn, 1981). Vegetation cover type (field/forest) was determined from optical imagery. A Canopy Height Model (CHM) was developed by subtracting the DTM produced using ground-classified points from the DSM produced using all lidar points. This results in a digital model consisting solely of canopy heights with no topography. The CHM generation used raster images with a 1 m resolution. The forested area was further classified as coniferous or deciduous for the study region. Within the forested area, the CHM was used to distinguish the upper canopy that did not lose needles/foliage from other forested regions with trees with no leaves using our snow-off survey that was collected with leaf off in the spring. A 3 by 3 maximum convolve filter was used to enhance the edges of canopy crowns and expand smaller regions that might have just one pixel of an intact canopy or a hole in a larger canopy (Palace et al., 2008). A 15 m threshold was used to differentiate between the upper level intact coniferous canopy and canopies that had lost their leaves. CHM pixels that were below this threshold were deemed deciduous canopies (see Figure S3 in supporting information for intermediate figure). The 5.6 ha forested area has a forest type that is 65% deciduous and 35% coniferous. "

Line 211-217: Is this section necessary? At this stage of snow depth accuracy evaluation do we even need to know that the soil was frozen?
A. This section was added in response to an earlier reviewer who requested ancillary data on soil frost conditions to support the discussion of the magnaprobe. This section was retained.

Line 227-233: I found this section/ equation/term definitions to be confusing and incomplete.
A. Thank you for pointing out. Reviewer #1 has also gave specific suggestions regarding this section and terminology. Modifications were also made in response to Reviewer #1. The paragraph now reads as follows:

The one-sided width of the 95% confidence limits ($CI_{95\%,+/-}$) for each grid cell's lidar derived estimate of the mean snow depth is a measure of uncertainty. The $CI_{95\%,+/-}$

values are used to compare the reliability of the snow depth estimates among cells. The $CI_{95\%,+/-}$ values were calculated using each grid cell's bare-earth and snow-on pooled sample standard deviation (s_d) and the number of bare-earth and snow-on lidar returns (n and m respectively) (Helsel and Hirsh, 2002).

$$CI_{95\%+/-} = t_{crit} s_d \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)} \quad (1)$$

A cell's pooled sample standard deviation (s_d) was calculated as

$$s_d = \sqrt{\frac{(n-1){s_{off}}^2 + (m-1){s_{on}}^2}{(n+m-2)}} \quad (2)$$

where s_{on} and s_{off} are the standard deviations of the snow-on and snow-off lidar ground return elevations, respectively. The $s_{\square n}$ and s_{off} values are a measure of the grid cell variability. This variability depends on the lidar instrument's relative accuracy (Maune and Nayegandhi, 2018), which includes intra-swatch accuracy (i.e., precision or repeatability of measurements) and inter-swath accuracy (i.e., differences in elevations between overlapping swaths), as well as surface elevation variations and terrain induced errors (Deems et al., 2013). The contribution from the individual sources of variability was not assessed in the current study.

Figure 7: seems low res in my pdf.
A. Figure 7 was replaced.

Line 380: surveying in locations prior to snow-on (and presumable mark with a stake?) risks modifying snow processes at that locations and poetically biasing the snowpack which may have relevance certain processes being studied, and risks destructive sampling impacts if same location is being repeatedly visited over a season.
A. The author team agrees entirely with these comments. The team has had numerous internal and external discussions with other UAV researchers about sampling strategies that minimize the impacts, yet allow validation to occur. This is a challenge for validation high resolution snowpack measurements. To elucidate the downside of a priori surveying, the following section was rewritten to read "Surveying and marking sample locations prior to the winter season might reduce this effort. However, the use of sampling stakes risks modifying snow processes at the sample locations and potentially biasing the snowpack and incurring destructive sampling impacts if same location is being repeatedly visited over a season."

Line 411-412: Confidence intervals for snow depth are high on slopes? What slope angles are we seeing this relationship? In the literature snow depth errors tend to be higher on slopes – see terrain induced errors section in Deems 2013 Deems, Jeffrey S., Thomas H. Painter, and David C. Finnegan. "Lidar measurement of snow depth: a review." Journal of Glaciology 59.215 (2013): 467-479.
A. Yes, the confidence intervals were larger on slopes within a single cell. Slopes that exceeded 20^{o} had a notable increase in their confidence intervals. While there was limited very steep terrain, there were a few large boulders in the forest. A reference

to Deems et al. (2013) discussion of terrain induced errors.

Line 416: “higher” word choice- “larger” may be better
A. Modified as recommended.

Line 453: “antiparallel” = “perpendicular”
A. Antiparallel is now defined in section 2.2, so this later language should now be clear.

Snow depth mapping with unpiloted aerial system lidar observations: A case study in Durham, New Hampshire, United States

Jennifer M. Jacobs[1,2], Adam G. Hunsaker[1,2], Franklin B. Sullivan[2], Michael Palace[2,3], Elizabeth A. Burakowski[2], Christina Herrick[2], Eunsang Cho[1,2]

[1]Department of Civil and Environmental Engineering, University of New Hampshire, Durham, NH, 03824, USA
[2]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, 03824, USA
[3]Department of Earth Sciences, University of New Hampshire, Durham, NH, 03824, USA

Correspondence to: Jennifer M. Jacobs (Jennifer.jacobs@unh.edu)

Abstract. Terrestrial and airborne laser scanning and structure from motion techniques have emerged as viable methods to map snow depths. While these systems have advanced snow hydrology, these techniques have noted limitations in either horizontal or vertical resolution. Lidar on an unpiloted aerial vehicle (UAV) is another potential method to observe field and slope scale variations at the vertical resolutions needed to resolve local variations in snowpack depth and to quantify snow depth when snowpacks are shallow. This paper provides some of the earliest snow depth mapping results on the landscape scale that were measured using lidar on a UAV. The system, which uses modest cost, commercially available components, was assessed in a mixed deciduous and coniferous forest and open field for a thin snowpack (< 20 cm). The lidar classified point clouds had an average of 90 and 364 points/m^2 ground returns in the forest and field, respectively. In the field, in-situ and lidar mean snow depths, at 0.4 m horizontal resolution, had a mean absolute difference of 0.96 cm and a root mean squared error of 1.22 cm. At 1 m horizontal resolution, the field snow depth confidence intervals were consistently less than 1 cm. The forest areas had reduced performance with a mean absolute difference of 9.6 cm, a root mean squared error of 10.5 cm, and an average one-sided confidence interval of 3.5 cm. Although the mean lidar snow depths were only 10.3 cm in the field and 6.0 cm in the forest, a pairwise Steel-Dwass test showed that snow depths were significantly different between the coniferous forest, the deciduous forest, and the field land covers ($p < 0.0001$). Snow depths were shallower and snow depth confidence intervals were higher in areas with steep slopes. Results of this study suggest that performance depends on both the point cloud density, which can be increased or decreased by modifying the flight plan over different vegetation types, and the grid cell variability that depends on site surface conditions.

1 Introduction

Over the past two decades, remote sensing methods, providing spatially continuous, high-resolution snow depth maps at local and regional scales, have greatly advanced the ability to characterize the spatiotemporal variability of snow depth over earlier work using snow probes. Spaceborne photogrammetry (e.g. Marti et al. 2016, McGrath et al. 2019, Shaw et al. 2020), airborne laser scanning (ALS) (Deems et al., 2013; Harpold et al., 2014; Kirchner et al., 2014), terrestrial laser scanning (TLS) (Grünewald et al. 2010; Currier et al. 2019), and structure-from-motion photogrammetry (SfM) (Nolan et al., 2015; Bühler et al., 2016; Harder et al., 2016) have emerged as viable methods to map surface elevations with snow-off and snow-on conditions in order to differentially map snow depths.

ALS and TLS both rely on well-established lidar (light detection and ranging) technology. TLS, applied from a fixed ground position, is able to measure snow depth with high vertical accuracy (Fey et al., 2019), and has the advantage of being

Jennifer Jacobs 1/19/2021 10:39 AM
Deleted: and heavily vegetated

Jennifer Jacobs 1/19/2021 10:09 AM
Deleted: modestly

Jennifer Jacobs 1/19/2021 10:39 AM
Deleted:

Jennifer Jacobs 1/19/2021 10:37 AM
Deleted: with typical confidence intervals within 4 cm.

Jennifer Jacobs 1/19/2021 10:41 AM
Deleted: was

Jennifer Jacobs 1/18/2021 7:21 PM
Deleted: within

Jennifer Jacobs 1/19/2021 10:49 AM
Deleted: ... [1]

Jennifer Jacobs 1/18/2021 4:54 PM
Deleted: (see reviews in Deems et al., 2013; López-Moreno et al., 2017)

relatively low-cost and portable, making repeat observations possible. However, TLS uncertainties are caused by large incident angles, occlusion from hills and trees that can cause data gaps in forested domains (Currier et al., 2019; Palace et al., 2016), and challenges to provide a stable scanner position for the tripod in snow-on conditions (Schweizer et al., 2003). ALS technology such as that deployed on the Airborne Snow Observatory (ASO) (Painter et al., 2016) has the advantage of being able to cover large areas, but it is extremely expensive and has limited availability and flexibility of deployment, which impacts its use for most studies. ALS also has issues with observation gaps in forested regions (Broxton et al., 2015; Currier and Lundquist, 2018; Mazzotti et al., 2019) but possibly to a lesser extent than TLS (Currier et al., 2019). The typical vertical accuracies from these platforms are on the order of 10 cm (Kraus et al., 2011; Deems et al., 2013) with a relatively low return density (~10 returns/m^2) (Cook et al., 2013). These accuracies and densities may not be adequate to observe spatial variations at point scales (0 to 5 m) to hillslope and field scales (1-100 m) and to detect snow depth changes over short time scales due to single events, densification, wind redistribution, sloughing of snow-off slopes, trapping of snow by vegetation, and forest canopy interception (Clark et al., 2011; Mott et al., 2011; Mott et al., 2018).

SfM can create a digital surface model (DSM) from photographs taken using a standard consumer-grade digital camera. When using an unpiloted aerial system (UAS), which deploys a camera on an unpiloted aerial vehicle (UAV). SfM is a low cost method that has the capacity for routine snow depth monitoring (Adams et al., 2018; Bühler et al., 2016; De Michele et al., 2016; Harder et al., 2016; Vander Jagt et al., 2015). Reported accuracies range from 8 to 30 cm using UAS SfM (Adams et al., 2018; Bühler et al., 2016; Goetz and Brenning, 2019; Harder et al., 2016; Meyer and Skiles, 2019; Harder et al., 2020). The primary drawbacks of UAS SfM as compared to lidar for mapping snow depth are that the DSM needs to be georeferenced using ground control points (GCPs) with known coordinates and may require significant manual steps (Tonkin et al., 2016; Meyer and Skiles, 2019), although new techniques are emerging that may reduce field data collection time (Gabrlik et al., 2019; Meyer and Skiles, 2019). Dense canopy or vegetation can reduce performance when snow compresses the vegetation relative to the snow-off imagery or when above-canopy vegetation is falsely interpreted to be the snow surface (Bühler et al., 2017; Cimoli et al., 2017; De Michele et al., 2016; Fernandes et al., 2018; Harder et al., 2016; Nolan et al., 2015). Canopy effects impact SfM snow mapping capability in regions where snowpacks are masked by dense forest canopies. The inability to sense portions of the ground/snow surface beneath dense canopies results in fine scale variations in snow depth, such as tree wells, not being accurately represented in UAS SfM snow depth products (Harder et al., 2020).

UAS lidar, a UAV-mounted laser scanning system, has been widely used in forest-related research (e.g. canopy height and forest change detection) (Wallace et al., 2012; 2014) and appears to offer the advantages of both the UAS SfM and lidar for snow depth mapping. UAS lidar also eliminates many of the drawbacks that arise from ALS and TLS systems discussed earlier. However, to date there is only one previous study that estimates snow depth using UAS-based lidar (Harder et al., 2020). Harder et al. (2020) compared snow depth estimates between lidar versus SfM techniques using in-situ snow depth observations in mountain and prairie environments, focusing on sub-canopy snow, which has been a challenge to measure in the snow remote sensing community. Using a considerably more expensive UAS-based lidar system (~$300K Canadian), they found that the lidar system tends to have lower errors than the SfM to capture sub-canopy snow distributions at moderate depth of snowpack (up to 2 m and 1 m of the maximum depth for mountain and prairie areas, respectively). In this study, we assess the ability of a more modest cost UAS lidar system (~$70K U.S. dollars) to map snow depth focusing on shallow and ephemeral snowpack (< 20 cm). The pilot study described here serves as a proof-of-concept for providing a high vertical resolution snowpack dataset in open terrain and forests in the north-eastern United States. Snow depth magnitude and variability are mapped and analyzed for differences by land use and slope. The study highlights results from the 2019 winter season that provide insights as to the potential for UAS lidar mapping of snow depth as well as details about the

Jennifer Jacobs 1/18/2021 5:18 PM
Deleted: For some snowpack features,

Jennifer Jacobs 1/18/2021 5:18 PM
Deleted: t

Liz Burakowski 1/22/2021 2:56 PM
Comment [1]: accuracies … are

Liz Burakowski 1/22/2021 2:56 PM
Deleted: is

Jennifer Jacobs 1/18/2021 5:18 PM
Deleted: ,

Jennifer Jacobs 1/18/2021 5:18 PM
Deleted: , as well as

Jennifer Jacobs 1/19/2021 10:46 AM
Deleted: or

Jennifer Jacobs 1/19/2021 10:47 AM
Deleted: .

Jennifer Jacobs 1/24/2021 9:59 AM
Deleted: deployed

Jennifer Jacobs 1/24/2021 9:57 AM
Deleted: system

Jennifer Jacobs 1/24/2021 9:57 AM
Deleted: S

Jennifer Jacobs 1/24/2021 9:57 AM
Deleted: platform

Jennifer Jacobs 1/24/2021 9:48 AM
Deleted: -based

Jennifer Jacobs 1/24/2021 10:00 AM
Deleted:

Jennifer Jacobs 1/24/2021 10:06 AM
Deleted: A

Jennifer Jacobs 1/24/2021 10:06 AM
Deleted: platform

Eunsang Cho 1/22/2021 12:31 AM
Deleted:

Jennifer Jacobs 1/24/2021 10:22 AM
Deleted: snow which

Jennifer Jacobs 1/24/2021 6:19 PM
Deleted: seek to

Jennifer Jacobs 1/24/2021 10:06 AM
Deleted: -based

Jennifer Jacobs 1/24/2021 10:22 AM
Deleted: XX

Jennifer Jacobs 1/24/2021 6:19 PM
Deleted: provide

Jennifer Jacobs 1/24/2021 6:20 PM
Deleted: mapping

Eunsang Cho 1/22/2021 12:35 AM
Deleted: The purpose of this paper is to assess the ability of a UAS platform to provide snow depth using a modest cost UAS-based lidar.

Jennifer Jacobs 1/19/2021 11:03 AM
Deleted: terrain

Eunsang Cho 1/22/2021 12:31 AM
Deleted:

system, its deployment, and operational and validation challenges. We explore the capability of UAS through the comparison of contemporary field-based snow depth measurements collected in a landscape containing fields and forests.

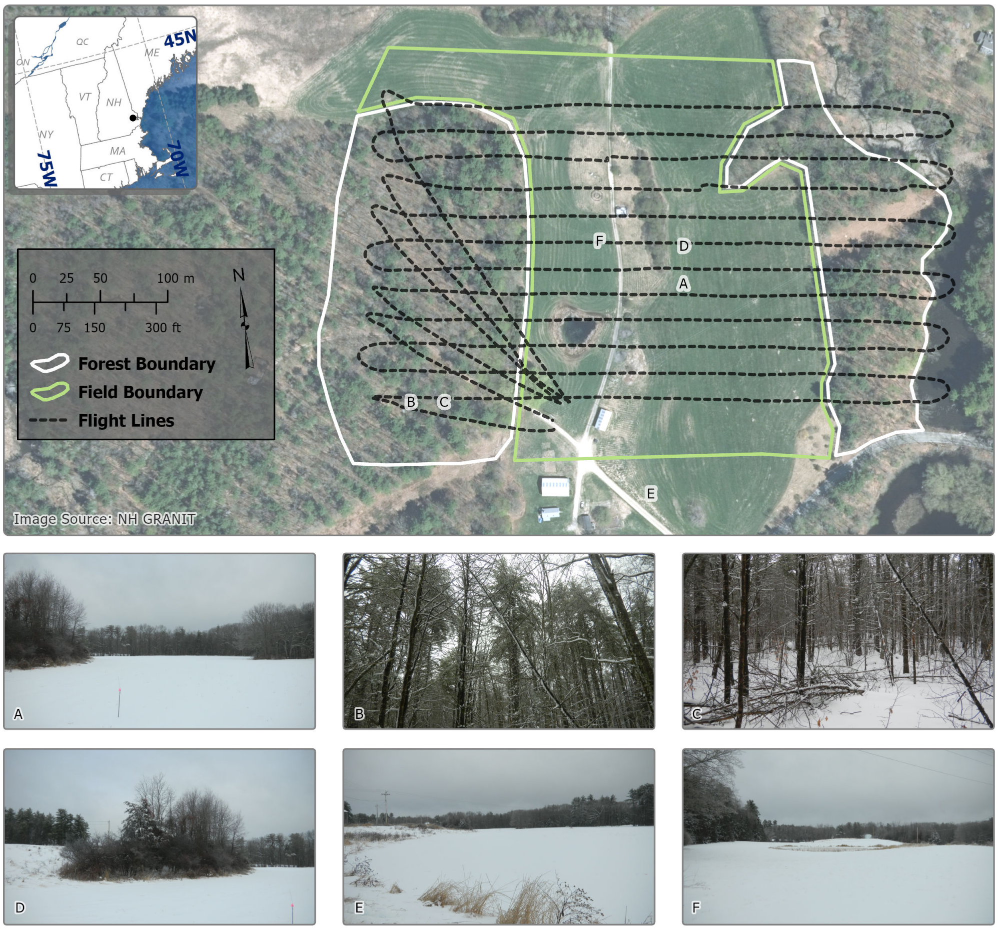


Figure 1. 2015 aerial imagery of Thompson Farm, Durham NH showing both forest and field region with lidar flight lines (top). Ground imagery (a to f), collected in December 2019, locations are noted on the top map and show the surface and leaf off forest conditions (bottom panels).

2 Site, Data, and Methods

2.1 Site

The test flights were conducted at the University of New Hampshire's Thompson Farm Research Observatory in southeast New Hampshire, United States (N 43.10892°, W 70.94853°, 35 m above sea level, ASL), which was chosen for its mixed hardwood forest and open field land covers (Burakowski et al., 2015; Burakowski et al., 2018) that are characteristic of the region (Figure 1). Thompson Farm has an area of 0.83 km^2 and little topographic relief (Perron et al., 2004). The agricultural fields are actively managed for pasture grass. The mixed deciduous and coniferous forest is composed primarily of white pine (*Pinus strobus*), northern red oak (*Quercus rubra*), red maple (*Acer rubrum*), shagbark hickory (*Carya ovata*), and white oak (*Quercus alba*) (Perron et al., 2004). There are two logging access roads that run north-south through the pasture and into the western forest section.

Liz Burakowski 1/22/2021 3:02 PM
Deleted: “wood

Liz Burakowski 1/22/2021 3:02 PM
Deleted: ”

2.2 UAS Laser Scanning

A series of UAS lidar surveys were conducted over approximately a 0.1 km^2 (9.8 hectares) area (430 by 225 m) within the farm during the winter 2018/2019 (Figure 1). Here, we focus on the snow-on flight conducted on January 23, 2019 and the snow-off flight conducted on April 11, 2019. We selected the January 23 flight because it had snowed approximately 11.5 cm with 1.8 cm of snow water equivalent from January 19th to January 20th and the air temperature was persistently below freezing prior to the flight. For the April 11, 2019 snow-off flight, the deciduous component of the canopy and understory were both dormant.

We used an Eagle XF UAS manufactured by UAV America, which carried a small, light-weight lidar sensor (Velodyne VLP-16) suitable for UAS deployment (see Table S1 in supporting information). The VLP-16 is a 16-channel lidar sensor with a 30-degree vertical field of view with rotating lasers that are spaced evenly between -15 to +15 degrees. Each channel rotates to provide a horizontal field of view of 360-degrees. The VLP-16 collects up to 300,000 points per second with an accuracy of +/- 3 cm at a range of 100 m. The sensor was mounted with the vertical field of view parallel with the ground. The payload is equipped with an Applanix APX-15 UAV inertial navigation system (INS), which has 2-5 cm positional, 0.025-degree roll and pitch, and 0.08-degree true heading uncertainties following post-processing. The INS has a measurement rate of 200 Hz, allowing for a timestamp to associate each lidar pulse with the closest data for latitude, longitude, altitude, and perspective information (roll/pitch/yaw), which is required for georeferencing returns.

Flights were conducted to maximize spatial coverage while conserving batteries due to the limited flight time of the Eagle XF (approx. 9 minutes flight time from ascent to descent). Because of the limited flight time, flights were conducted at an altitude of 81 m for greater spatial coverage and multiple return flight lines were necessary for battery exchanges (Figure 1). Automated flights were conducted using UgCS flight planning software. Flight speed was 7 m/s, with a total of 12 parallel flight lines with targeted overlap of 40 percent. A complete survey of the study area took approximately 2 hrs. This includes the time required to calibrate the INS and set-up and break-down the UAS. Because of degrading accuracy at distances >100 m with the VLP-16, returns acquired outside of +/- 30 degrees of nadir view angles in the horizontal field of view were filtered to limit target distance and improve overall accuracy.

Applanix APX-15 INS data were post-processed to a Smoothed Best-Estimate Trajectory (SBET) file using POSPac MMS UAV (v. 8.2.1), resulting in approximately 3 cm positional accuracy for both the snow-on and snow-off flights. Lidar returns were individually georeferenced by synching timestamps of returns from the lidar sensor with timestamps of position and attitude data from the post-processed INS data. Georeferenced point clouds were produced and output to LAS files using Headwall Photonics, Inc.'s LidarTools software. The bare-earth and snow-on point clouds were georeferenced solely using the INS data respective to each flight. The point clouds were not co-registered to each other as there were no reliable common ground control points between surveys. For UAS lidar snow depth surveying, co-registration between point clouds would likely be unattainable due to insufficient common ground control. We determined results would be more meaningful when bare-earth and snow-on point clouds were processed solely relying on the capability of the INS. Boresighting calibration was performed using returns from the first two parallel flight lines that were collected in opposite directions (i.e., antiparallel). A roll offset was determined using 10 m cross sections along the flight lines over flat terrain, and a pitch offset was determined using 1 m cross sections across the flight lines over terrain with moderate relief (see Figure S2 in supporting information). Resulting LAS (LASer) point clouds were generated for the entire study area and projected in WGS84 UTM Zone 19N (EPSG 32619). Flight and filtering parameters of the raw point cloud resulted in return densities of approximately 150 returns/m^2 for each of the two flights.

Sullivan, Franklin 1/22/2021 11:21 AM
Deleted: to discharge to 3.6 V per cell

Sullivan, Franklin 1/22/2021 11:29 AM
Deleted: Lidar returns were matched to post-processed INS data and georeferenced using Headwall Photonics, Inc.'s Lidar Tools software.

Jennifer Jacobs 1/24/2021 10:17 AM
Deleted: UAV Lidar

Jennifer Jacobs 1/19/2021 11:50 AM
Deleted: directions.

2.3 Lidar Classification and Gridding

Three-dimensional point clouds were processed using the progressive morphological filter algorithm (PMF), in the lidR package (https://github.com/Jean-Romain/lidR) of R (v. 3.4, Team, 2018) to identify ground returns. Briefly, the PMF operates iteratively on sets of two parameters, window size and elevation thresholds to erode and dilate point cloud data sets to estimate surface topography. The result of the PMF is that non-ground returns (i.e., trees, shrubs, and noise) are filtered out of point cloud data sets, so that only returns from ground surfaces remain. The two data sets, non-ground returns and ground returns from the original point cloud, are coded according to LAS specifications and merged. For a full explanation of PMF, see Zhang et al., (2003). For ground classification, point clouds were chunked into 100 m square tiles with a 15 m buffer on all sides using catalog options in lidR to ensure returns near tile edges were classified. Processing was distributed across 8 computing cores to improve efficiency. PMF was parameterized using a set of window sizes of 1, 3, 5, and 9 m, and elevation thresholds of 0.2, 1.5, 3, and 7 m, which were determined by varying value sets and assessing digital terrain models (DTMs) to determine the parameter sets that produced a visually smooth surface over a dense grid (*sensu* Muir et al., 2017). Following ground classification for each tile, returns within the 15 m tile buffers were removed, and all resulting 100 m square ground classified tiles were merged. The resulting point clouds for each data set included both the classified ground returns and the non-ground returns. Snow-on and snow-off ground point clouds were gridded at 0.1, 0.2, 0.4, 0.5, and 1.0 m spatial resolutions using the average of all grid points within each grid cell (Currier et al., 2019). Gridded products for each data set were forced to the same coordinate grid to generate DTMs as raster files.

2.4 Slope and Vegetation Cover Classification and Analysis

The snow-off DTM was used to develop a 1 m resolution map of slope (Horn, 1981). Vegetation cover type (field/forest) was determined from optical imagery. A Canopy Height Model (CHM) was developed by subtracting the DTM produced using ground-classified points from the DSM produced using all lidar points. This results in a digital model consisting solely of canopy heights with no topography. The CHM generation used raster images with a 1 m resolution. The forested area was further classified as coniferous or deciduous for the study region. Within the forested area, the CHM was used to distinguish the upper canopy that did not lose needles/foliage from other forested regions with trees with no leaves using our snow-off survey that was collected with leaf off in the spring. A 3 by 3 maximum convolve filter was used to enhance the edges of canopy crowns and expand smaller regions that might have just one pixel of an intact canopy or a hole in a larger canopy (Palace et al., 2008). A 15 m threshold was used to differentiate between the upper level intact coniferous canopy and canopies that had lost their leaves. CHM pixels that were below this threshold were deemed deciduous canopies (see Figure S3 in supporting information for intermediate figure). The 5.6 ha forested area has a forest type that is 65% deciduous and 35% coniferous.

Once the vegetation forest type was classified, three sets of 5000 points were extracted respectively in the field, in the eastern forest and in the western forest (Palace et al., 2017). At each of these random points, slope, vegetation type (field, deciduous, coniferous), snow depth, and snow depth confidence interval values were extracted. Because of missing values in the raster images, not all random points extracted values. Slope was assigned to one of three categories: 0-10 degrees, 10-20 degrees, and greater than 20 degrees. Because the extracted datasets (i.e., snow depth, confidence interval, and slope) were not normally distributed, the non-parametric Steel-Dwass Method test was used to test for differences. The Steel-Dwass test has been previously used in geophysical work to examine non-parametric datasets (Slotznick et al., 2019). This non-parametric method is useful when sample numbers are large and groups are small, because it allows type I errors to be controlled (Dolgun and Demirhan, 2017).

Sullivan, Franklin 1/22/2021 12:43 PM
Deleted: T

Sullivan, Franklin 1/22/2021 12:37 PM
Deleted: as an approach to filter out

Sullivan, Franklin 1/22/2021 12:39 PM
Deleted: from

Sullivan, Franklin 1/22/2021 12:44 PM
Deleted: (Zhang et al., 2003)

Jennifer Jacobs 1/24/2021 10:24 AM
Deleted: The snow-off DTM was used to develop a 1 m horizontal resolution map of slope (Horn, 1981). Vegetation cover type (field/forest) was determined from the known boundaries of field and forest. The forested area was further classified as coniferous or deciduous for the study region using the following methodology (Figure 1). Within the forested area (Figure 1), a Canopy Height Model (CHM) was used to distinguish the intact upper canopy from other forest cover using our snow-off survey, collected with leaf off in the spring (Sullivan et al., 2017). The CHM was generated by subtracting the DTM produced using ground-classified points from the DSM produced using all lidar points. This results in a digital model consisting solely of canopy heights with no terrain or topography. The CHM generation used raster images with a 1 m resolution. A 3 by 3 maximum convolve filter was used to enhance the edges of canopy crowns and expand smaller regions that might have just one pixel of an intact canopy or a hole in a larger canopy (Palace et al., 2008). A 15 m threshold was used to differentiate between the upper level intact coniferous canopy. CHM pixels that were below this threshold were deemed deciduous canopies (see Figure S3 in supporting information for intermediate figure). The 5.6 ha forested area has a forest type that is 65% deciduous and 35% coniferous.

Jennifer Jacobs 1/18/2021 6:09 PM
Deleted: the raster binary image was vectorized. Within the forest and field regions of our study, a subsample was created from the entire image of 5000 random points in the field and 5000 random points in each of the eastern and western forested areas

2.5 *In Situ* Observations

A magnaprobe (Sturm and Holmgren, 2018) was used to compare to the UAS lidar survey over two transects. The first transect consisted of 12 sample locations in the field and 5 locations in the eastern forest of our study site. The second transect consisted of 11 sample locations in the western forest. Sample locations were separated by approximately 10 m. The field transect follows the prevailing westerly wind direction with its west side at the foot of a modest depression (approximately 3-4 m below the land further to the west) and the east side transitioning into a wooded area. Following (Harder et al. 2016) and (Bühler et al. 2016), each sample location includes 5 samples in a cross pattern with the four ordinal directions sampled approximately 20 cm from the center sampling location in the cross. The five samples are used to provide a measure of snow depth central tendency and variation over a 0.4 x 0.4 m pixel. Because the magnaprobe GPS has an absolute accuracy of 8 m, a Trimble© Geo7X GNSS Positioning Unit with Zephr™ antenna was used to collect each sampling location's center point with an estimated horizontal uncertainty of 2.51cm (standard deviation σ 0.95 cm) and 4.17cm (σ 4.60 cm) for the field and forest, respectively after differential correction. Along the same forest and field transects, a federal snow tube sampler was used to collect a single sample of snow depth and snow water equivalent (SWE) at each magnaprobe sample location for a total of 12 field samples and 16 forest samples. SWE was measured by inserting the aluminium tube vertically into the snowpack and a core was extracted and weighed using a spring scale.

Jennifer Jacobs 1/24/2021 6:01 PM
Deleted: s

Jennifer Jacobs 1/18/2021 5:12 PM
Deleted: (hereafter noted as ALS measurements)

An independent study collected soil frost depth from three locations at the Thompson Farm Research Observatory using Gandahl-Cold Regions Research and Engineering Laboratory (CRREL) style frost tubes. The frost tubes have flexible, polyethylene inner tubing filled with methylene blue dye whose color change is easy to differentiate when extruded from ice (Gandahl 1957). A nylon string housed inside the polyethylene tubing affixes ice during periods of thaw. The outer tubing consists of PVC pipe installed between 0.4 to 0.5 m below soil surface (Ricard et al., 1976; Sharratt and McCool, 2005). Prior to the January 19th and 20th, 2019 snowfall event, soil frost was 23.5 to 25.5 cm in the field and 5.5 to 8.5 cm in the west forest.

2.6 Snow Depth Uncertainty Assessment

The snow depth accuracy was assessed by comparing the lidar snow depth measurements to the magnaprobe measurements. Here, accuracy is the measure of the agreement of the lidar snow depth measurements relative to the in situ measurements (Eberhard et al., 2020; Maune and Nayegandhi, 2018). Error statistics were calculated and the results were summarized by forest and field locations. At each magnaprobe location, the average and standard deviation of the five magnaprobe samples were calculated. The average lidar snow depth was determined for a 0.4 x 0.4 m cell centered on the center magnaprobe location. The mean absolute difference (MAD) and root mean square deviation (RMSD) were used to characterize the differences between the magnaprobe snow depths and the lidar snow depths.

The one-sided width of the 95% confidence limits ($CI_{95\%,+/-}$) for each grid cell's lidar derived estimate of the mean snow depth is a measure of uncertainty. The $CI_{95\%,+/-}$ values are used to compare the reliability of the snow depth estimates among cells. The $CI_{95\%,+/-}$ values were calculated using each grid cell's bare-earth and snow-on pooled sample standard deviation (s_d) and the number of bare-earth and snow-on lidar returns (n and m respectively) (Helsel and Hirsh, 2002).

$$CI_{95\%+/-} = t_{crit} s_d \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)} \quad (1)$$

where t_{crit} is the critical value of the Student's t-distribution with a significance level of 0.05 and s_d is the cell's pooled sample standard deviation which was calculated as

$$s_d = \sqrt{\frac{(n-1)s_{off}^2 + (m-1)s_{on}^2}{(n+m-2)}} \quad (2)$$

where s_{on} and s_{off} are the standard deviations of the snow-on and snow-off lidar ground return elevations, respectively. The s_{on} and s_{off} values are a measure of the grid cell variability. This variability depends on the lidar instrument's relative accuracy (Maune and Nayegandhi, 2018), which includes intra-swatch accuracy (i.e., precision or repeatability of measurements) and inter-swath accuracy (i.e., differences in elevations between overlapping swaths), as well as surface elevation variations and terrain induced errors (Deems et al., 2013). The contribution from the individual sources of variability was not assessed in the current study.

3 Results and discussion

3.1 Snow Depth Survey

The snow-on and snow-off lidar ground returns yielded an average point cloud density of 90 and 364 points/m^2 in the forest and field, respectively, with 6.7% of the 1 m^2 forest cells and 0.03% of the 1 m^2 field cells having less than 5 point/m^2 (Figure 2). There was a wide range of the point cloud densities (Figure 2b). The highest point cloud density occurred for those cells sampled by both the regular flight lines and the multiple return flight lines conducted for the three battery exchanges. The vast majority of field cells (82%) have more than 100 points/m^2. Only 1% of the field cells had less than 25 points/m^2 and most of those cells were in shrubbery or dense vegetation surrounding the small pond in the center of the study site (Figure 1). In contrast, 41% of the forest cells had more than 100 points/m^2 and nearly 20% of the forest cells had less than 25 points/m^2 with 8% having fewer than 10 points/m^2 (Figure 2b). Only 0.086% and 0.95% of the 1 m resolution field and forest cells, respectively, had no ground returns. The number of points per cell decreases with decreasing cell size (Figure 2a). In the field, reducing the gridded resolution from 1 m to 0.5 m lowers the mean cell return count to 91 points per cell on average. Thus a 0.5 m field cell has approximately the same number of returns as a 1 m forest cell. At a 0.2 m spatial resolution, the mean number of ground returns is 14.6 and 3.6 in the field and forest, respectively.

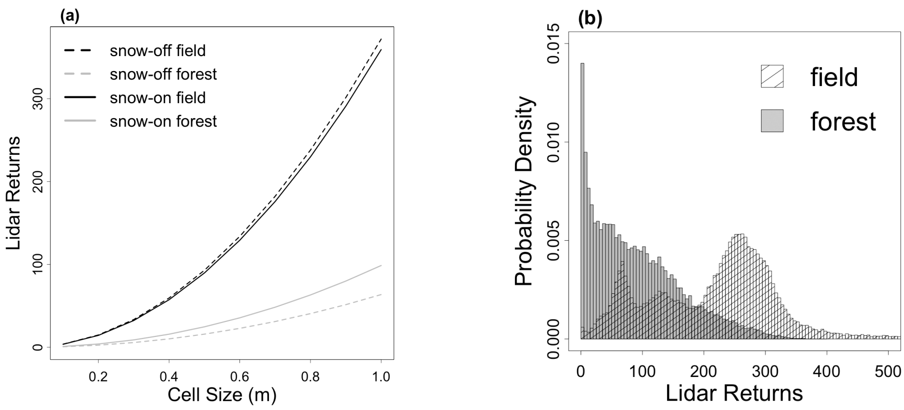


Figure 2. (a) Average lidar point cloud density of the ground returns versus cell size by land cover, and snow-on and snow-off state. (b) Probability density function for the number of lidar ground returns by square meter for the forest (gray) and the field (white).

Jennifer Jacobs 1/24/2021 8:00 PM
Formatted: Subscript

Liz Burakowski 1/22/2021 3:11 PM
Deleted: is

Jennifer Jacobs 1/19/2021 10:23 AM
Deleted: (top)

Jennifer Jacobs 1/19/2021 10:23 AM
Deleted: (bottom).

3.2 Lidar and *In Situ* Snow Depth Comparison

Based on the magnaprobe snow depth and UAS-mapped snow depth measurements, the accuracy of lidar snow depth measurements differed between field and forest cells (Figure 3). In the field, the mean snow depth from the magnaprobe (12.2 cm ±0.56 cm) was only slightly greater than that from the lidar (11.2 cm ± 0.72 cm) and the MAD and RMSD values were 0.96 cm and 1.22 cm, respectively. In the forest, the mean snow depth from the magnaprobe (15.2 cm ± 2.3 cm) was twice as large as the lidar snow depths (7.8 cm± 6.3 cm) and the MAD and RMSD were 9.6 cm and 10.5 cm, respectively. The mean snow depth from the Federal snow tube was (12.9 cm ±0.71 cm) and (13.1 cm ±1.9 cm) in the field and forest, respectively. There is a notable absolute low bias in the lidar forest snow depth relative to the magnaprobe and snow tube for west forest in particular with the exception of one site.

To provide insight to differences between the forest and field observations, height profiles of classified returns were calculated for 25 m^2 square regions centered on all forest (n=12) and field (n=7) study plots from lidar data. Height profiles were averaged for each site type, from here on referred to as mean height profiles (Figure 4). To do this, all lidar returns were extracted from the bounding box of each plot, then the mean elevation of ground returns was calculated within each plot. Return height profiles for each plot were determined by subtracting the mean ground elevation of the plot from each return, then the normalized return elevations were binned in 0.1 m height increments. Within the forests, an average of 2142 and 2889 returns were classified as ground and non-ground, respectively, in snow-free conditions for each 25 m^2 plot. Snow-on conditions had a comparable number of ground returns (2218), but fewer non-ground returns (1721). In field plots, an average of 5666 ground returns and 154 non-ground returns in snow-free conditions were obtained for each 25 m^2 plot, with 7567 ground returns and 25 non-ground returns in snow-on conditions. Figure 4 also shows that there is a greater range of ground return elevations in the forest as compared to the field. In forest plots, ground return elevations had an average standard deviation of 0.157 m and 0.154 m in snow-free and snow-on conditions, respectively, while in field plots, ground return elevations had standard deviations of 0.058 m and 0.050 m in snow-free and snow-on conditions, respectively.

Sullivan, Franklin 1/22/2021 11:47 AM
Deleted: mean

Sullivan, Franklin 1/22/2021 11:47 AM
Deleted: a

Jennifer Jacobs 1/18/2021 6:16 PM
Deleted: , respectively

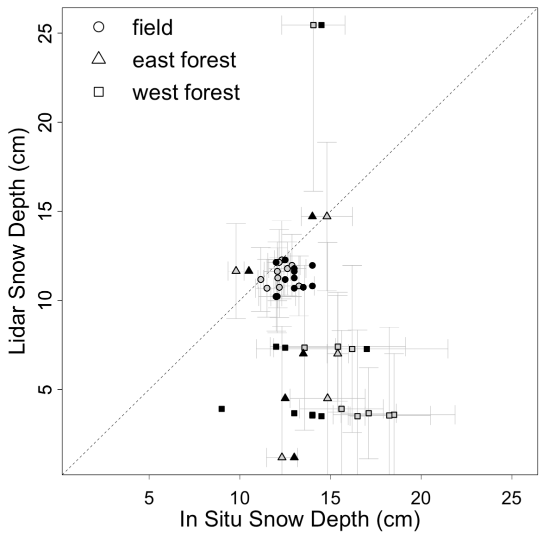


Figure 3. Comparison between the magnaprobe (gray fill) and snow tube (black fill) versus the lidar snow depth measurements by location. The mean and 95% confidence intervals were calculated using the five magnaprobe snow depths and the lidar snow depths, equation 1, averaged over a 0.4 x 0.4 m grid cell. Single snow tube snow depth measurements are shown without confidence intervals.

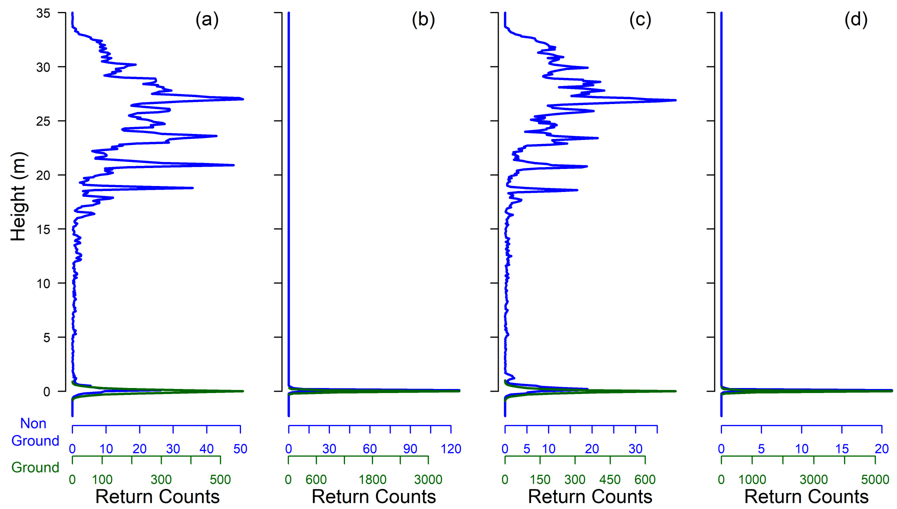


Figure 4. Mean height profiles for all ground (green) and non-ground (blue) lidar returns within a 5 m x 5 m region centered on each transect plot in snow-free conditions (a, b) and snow-on conditions (c, d) in forest (a, c) and field (b, d) study plots.

3.3 Snow Depth Maps from UAS Lidar

The UAS-mapped snow depth, mapped by subtracting snow-off DTMs from snow-on DTMs, revealed a shallow snowpack whose depth ranges from less than 2 cm to over 18 cm (Figure 5). The mean lidar snow depth was 10.3 cm in the field and 6.0 cm in the forest. Despite the shallow conditions, spatially coherent patterns are readily discernible. The field snowpack depth had higher spatial variability than the west forest snowpack and more spatial organization. In the field, the deepest snow was in the low-lying northeast areas that are sheltered from westerly winds. A relatively moderate and consistent snowpack occurred in southern part of the east field and west of the small pond. The shallowest snowpack was found in the center portion of the field, which is slightly elevated and, unlike most of the field, was not mowed. Lower snow depth at the forest edge distinguishes the field to forest transition. A non-parametric Steel-Dwass test found significant variation for the mean snow depth among the two forest types and field ($p < 0.0001$) (Figure 6a). Figures 6a and 6b also reveal that there are some negative snow depths in the two forest types that is due to the uncertainty of the snow-on and snow-off DTMs. A pairwise Steel-Dwass test showed that snow depths were significantly different between the three pairs of field and forest types ($p < 0.0001$). When comparing just field and forest as categories, the test also found significant differences for snow depth ($p < 0.0001$). Snow depth was also determined to be significantly different among the three slope group categories using the Steel-Dwass test where regions with a limited slope (Group 1) had more decidedly different snow than steeper regions ($p < 0.0001$) (Figure 6b).

The one-sided confidence interval values of the mean snow depth estimate are remarkably consistent in the field and typically are between 0.5 to 1 cm regardless of snow depth (Figure 5b). Modestly larger confidence intervals occur adjacent to the north-south road where the fields were not mowed prior to winter as well as the northern and southern extents of the flight lines likely due to the reduced sampling density. The forest had an average one-sided confidence interval of 3.5 cm, which was considerably higher than the field. Where the forest is predominantly comprised of deciduous trees, the typical

Liz Burakowski 1/22/2021 3:13 PM
Deleted: s

Liz Burakowski 1/22/2021 3:14 PM
Deleted: s

Liz Burakowski 1/22/2021 3:14 PM
Deleted: is

Liz Burakowski 1/22/2021 3:14 PM
Deleted: s

Liz Burakowski 1/22/2021 3:14 PM
Deleted: is

Liz Burakowski 1/22/2021 3:16 PM
Deleted: is

one-sided confidence intervals of the mean snow depth were as low as 1 to 2 cm. The largest one-sided confidence interval values occurred in the middle of the field where there was dense shrubbery, at the edge of the fields, and in clusters within the forest where the forest sections were dominated by coniferous trees (*Pinus strobus*). The nexus of flight lines in the take-off and landing area resulted in a local area with very high confidence. A non-parametric Steel-Dwass test found significant variation for confidence intervals of the mean snow depth among the two forest types and field ($p < 0.0001$) (Figure 6c). A pairwise Steel-Dwass test showed that confidence intervals were significantly different between the three pairs of field and forest types and ($p < 0.0001$). Confidence intervals were also significantly different among the three slope categories as determined using a Steel-Dwass test ($p < 0.0001$) (Figure 6d).

3.4 Point Cloud Density, Spatial Resolution, and Canopy Profiles

Confidence intervals for the mean snow depths by grid cell were examined in light of the point cloud density and the spatial resolution at which lidar returns were aggregated. The confidence interval width for a mean snow depth of a 1 m^2 area decreased dramatically as the lidar point cloud density increased (Figure 7a). Except for the cells with fewer than 10 point/m^2, forest cells had larger confidence intervals for the mean depths than field cells for a given sample size. When the density exceeds 25 point/m^2 in the field and 50 point/m^2 in the forest, confidence intervals were typically 2 cm. The cells with the highest point cloud densities had one-sided confidence intervals of about 1 and 1.5 cm for the field and forest cells, respectively. The field cells with more than 50 point/m^2 did not have noticeably smaller confidence intervals, but the increased density did reduce the number of cells with anomalously small confidence intervals. Given the high lidar point cloud density for the field cells, it is possible that reasonable estimates of snow depth can be made at scales finer than 1 m (Figure 7b).

Liz Burakowski 1/22/2021 3:17 PM
Deleted: is

Liz Burakowski 1/22/2021 3:17 PM
Deleted: are

Jennifer Jacobs 1/18/2021 6:20 PM
Deleted: increases

Liz Burakowski 1/22/2021 3:17 PM
Deleted: s

Liz Burakowski 1/22/2021 3:17 PM
Deleted: s

Liz Burakowski 1/22/2021 3:18 PM
Deleted: ve

Liz Burakowski 1/22/2021 3:18 PM
Deleted: are

Liz Burakowski 1/22/2021 3:18 PM
Deleted: ve

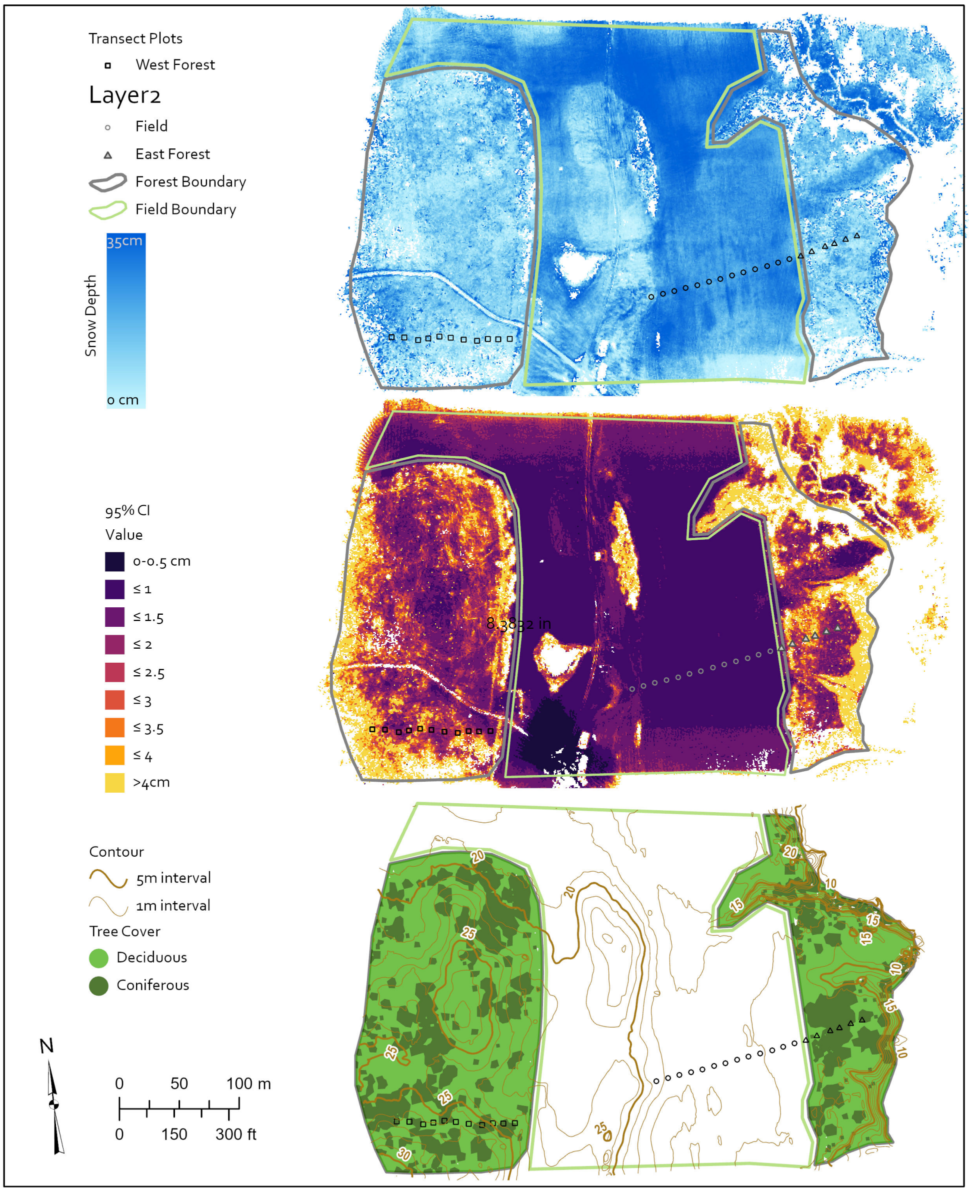
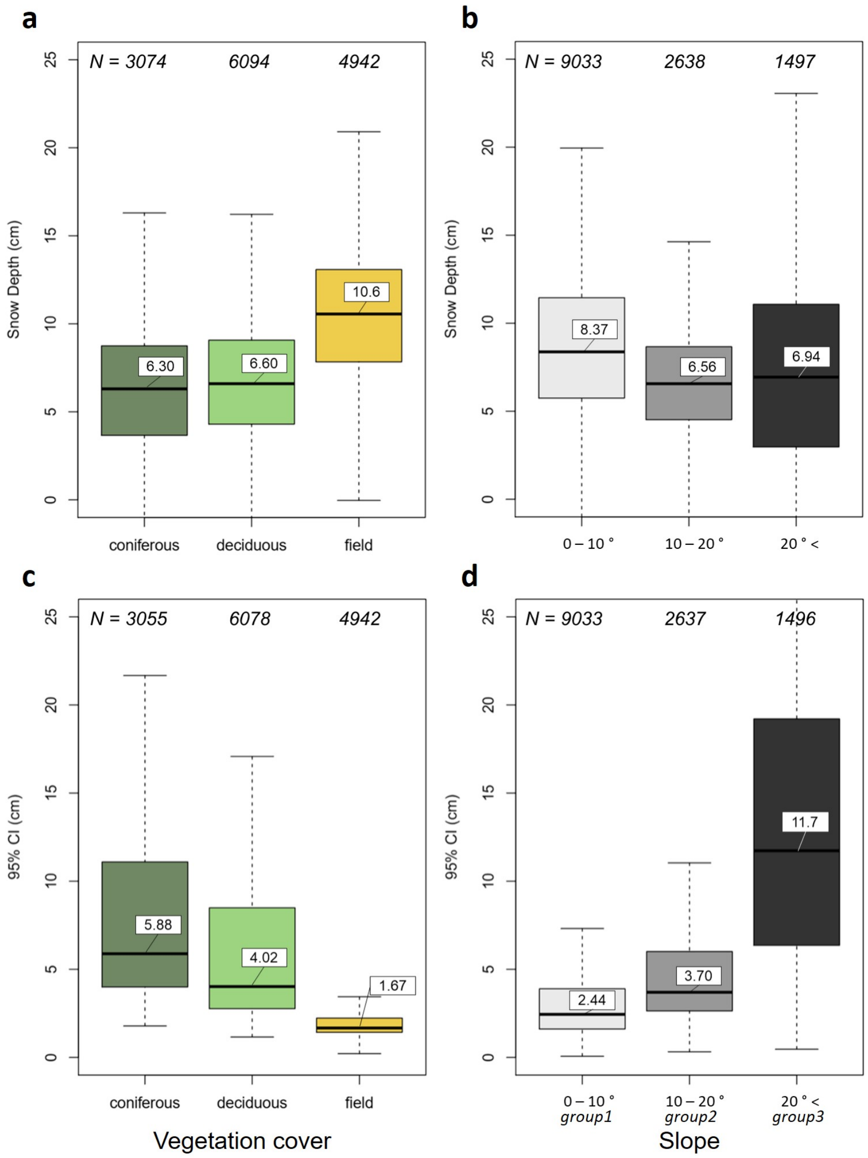


Figure 5. Average (top) snow depth values, (middle) one sided confidence intervals, and (bottom) topography and forest cover type. Snow depth and confidence intervals calculated from the snow-on and snow-off lidar point clouds for 1 m^2 cells at Thompson Farm, Durham, NH. Topography and forest cover type determined from snow-off lidar point clouds on snow-off flight for 1 m^2 cells conducted on April 11, 2019.



440 **Figure 6.** Snow depths (a,b) and their one sided confidence intervals (c,d) from the random sample points of the field and forest at Thompson Farm, Durham, NH on January 23, 2019 from the individual cells for 1 m^2 cells by vegetation cover (a,c) and slope group (b,d). Boxplots show the lower quartile, median, upper quartile, and whiskers with the median value noted. Because of missing values in the raster images, not all random points extracted values and resulted in different numbers of samples points for vegetation cover classes.

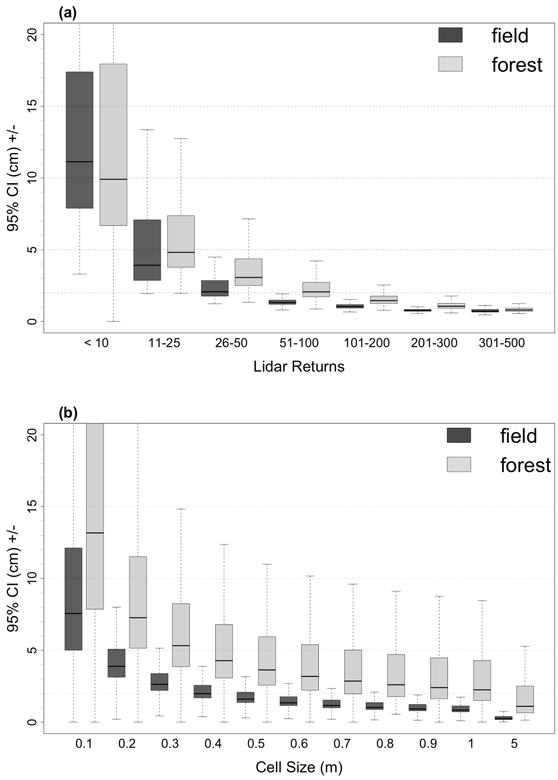


Figure 7. One sided confidence intervals of the mean snow depth values in the field and forest at Thompson Farm, Durham, NH on January 23, 2019 from the individual cells for 1 m^2 cells by land cover and point cloud density (a) and for grid resolutions ranging 0.1 to 5 m (b). Boxplots show the lower quartile, median, upper quartile, and whiskers.

In addition to the lidar point cloud density, the ability to reduce the confidence interval of the mean snow depth also depends on the ground surface variability within a cell as well as the lidar performance. For this site and its shallow snowpack, the grid cell variability of the ground surface elevation, estimated by calculating the standard deviation of the lidar elevation values, and found to depend primarily on the cell size and, to a more limited extent, on land cover and snow cover (Figure 8a). Snow cover reduced the grid cell variability in the field by about 1 cm, but has a limited effect in the forest. It is possible that the modest snowpack was able to flatten the higher grass in the field, while the forest's vegetation and ground surface features that dominate the grid cell variability were only minimally compacted by the snow. Within the 1 m grid cells, snow depth variability was much lower in the field than the forest (Figure 8b). Both distributions had a positive skew. Typical standard deviations of the lidar surface elevation values within a 10 cm cell were on the order of 1.5 and 2 cm for the field and forest, respectively. That variability doubled for a 20 cm cell. The grid cell variability increased gradually to about 3 to 4 cm in the field, and to about 6 cm in the forest.

Thus, confidence intervals largely depended on the point cloud density in the lidar cloud because the standard deviation of a cell's surface elevation is relatively constant for snow depth resolutions from 0.5 to 1 m (Figure 8a). In the field, reducing the cell size from 1 m to 0.5 m still yields about 100 points/m^2 and provides snow depth estimates within +/- 1.5 cm. Because the forest cells required a higher ground return density to capture these snow depths within a 1 cm, any reduction in cell size below 1 m greatly increased the cell mean snow depth's confidence intervals.

Jennifer Jacobs 1/18/2021 6:21 PM
Deleted: capture
Jennifer Jacobs 1/18/2021 7:20 PM
Deleted: local scale
Jennifer Jacobs 1/18/2021 7:25 PM
Deleted: was
Liz Burakowski 1/22/2021 3:20 PM
Deleted: s
Jennifer Jacobs 1/18/2021 7:20 PM
Deleted: within
Jennifer Jacobs 1/19/2021 11:06 AM
Deleted: terrain
Jennifer Jacobs 1/18/2021 7:20 PM
Deleted: within
Liz Burakowski 1/22/2021 3:20 PM
Deleted: are
Liz Burakowski 1/22/2021 3:20 PM
Deleted: are
Liz Burakowski 1/22/2021 3:21 PM
Deleted: s
Jennifer Jacobs 1/18/2021 7:20 PM
Deleted: within
Liz Burakowski 1/22/2021 3:21 PM
Deleted: s
Jennifer Jacobs 1/18/2021 6:33 PM
Deleted: per cell
Jennifer Jacobs 1/18/2021 6:33 PM
Deleted: spatial scales
Liz Burakowski 1/22/2021 3:21 PM
Deleted:

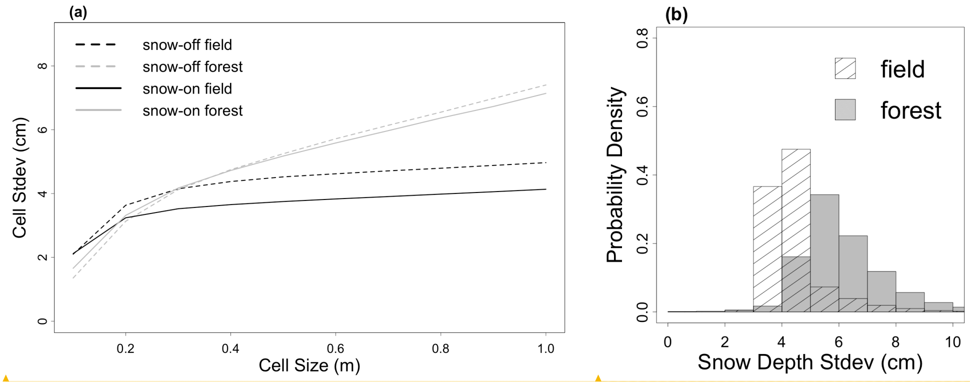


Figure 8. Lidar surface elevation standard deviations by (a) cell size and land cover. Cell standard deviations are the average of the individual cells standard deviation values calculated for cells with side dimensions ranging from 0.1 to 1.0 m. (b) Probability density function of the pooled snow depth standard deviation for each 1 m^2 cell in the forest (gray) and field (hashed).

4. Challenges and Recommended Improvements to UAS Lidar Snow Depth Mapping

Despite UAS-based lidar's increasing use in the natural sciences and capacity to make high-resolution snow maps, there are many operational and technical challenges that require consideration prior to successfully conducting UAS-based lidar surveys that produce research grade, high-resolution snow depth data. For the lidar surveys, the hardware and supporting software analysis tools can be expensive and require trained pilots and lidar data analysis specialists. In this section, we present some general considerations regarding validation of the lidar snow depth maps, selection and deployment of a lidar sensor on a UAV for snow depth mapping as well as specific insights that we experienced when using our system.

4.1 *In Situ* and UAS Sampling

This study's lidar snow depth performance metrics are comparable to those from the more extensive lidar surveys made Harder et al. (2020), In the field, our snow depth errors, 1 cm bias and 1.2 cm RMSD, were modestly better than those from their open sites snow depth 3 cm bias and RMSE values on the order of 10 cm. In the forest, our snow depth errors, 7 cm bias and 10 cm RMSD, were also modestly lower than those from their forest sites 9 to 13 cm bias and 15 cm RMSE. While it is difficult to make direct comparison across different study sites, snow conditions, and ground validation approaches, these early findings indicate that UAS lidar has the capability of mapping snow depths in open and forested regions and has improved performance as compared to previous SfM results particularly for vegetated surface. It is also noteworthy that this study's mapping was conducted using the Velodyne Puck series, a laser scanner adapted from the assisted and autonomous vehicle applications, rather than the specialized Riegl miniVUX-1UAV used by Harder et al. (2020) resulting in a complete mapping system that was approximately one-third the costs of their Riegl system.

While UAS-based lidar surveys can measure snow depth to within a centimeter at high spatial resolutions, validation of those observations is challenging. A time consuming collection of high accuracy GNSS survey points was required to co-locate magnaprobe and lidar observations. Surveying and marking sample locations prior to the winter season might reduce this effort. However, the use of sampling stakes risks modifying snow processes at the sample locations and potentially biasing the snowpack and incurring destructive sampling impacts if same location is being repeatedly visited over a season.

Unknown
Formatted: Font color: Black

Unknown
Formatted: Font color: Black

Jennifer Jacobs 1/24/2021 1:37 PM
Deleted: Even though the UAVs are modest in size (i.e., weighing less than 25 kg),

Jennifer Jacobs 1/24/2021 7:25 PM
Deleted:

Jennifer Jacobs 1/19/2021 2:16 PM
Deleted: in

It is also challenging to make *in situ* snow depth measurements that provide centimeter accuracy. In this study, the magnaprobe *in situ* snow depth observations made in the forest were considerably higher than the lidar observations as compared to the open field where the magnaprobe and lidar measurements were within 1 cm. Previous studies also found that snow depth observations from ALS measurements are biased lower than those from snow-probe observations in the forest (Hopkinson et al., 2004, Currier et al., 2019; Harder et al., 2020). In past studies, the causes of these differences have been partially attributed to the snow probe's ability to penetrate the soil and vegetation, human observers tending to make snow depth measurements in locations with relatively high snow (Sturm and Holmgren, 2018) and the reduced accuracy of the GNSS in the forest. Our study suggests additional issues in forest sampling including enhanced ground surface variability in forested areas relative to adjacent field areas and reduced lidar returns in forested areas as compared to field areas combine with sampling issues to contribute to the higher uncertainty in the forest snow depths observed in our study.

In this study, the cold temperatures and snow-free conditions prior to the January 19^{th} and 20^{th} snowfall event resulted in deeper frozen soils (23.5 to 25.5 cm) in the field and shallower soil frost depth (5.5 to 8.5 cm) in the west forest, which would have limited the probe penetration into soils at both sites. However, the forest has a 1-4 cm thick organic leaf litter layer that may have been penetrated by the magnaprobe. The average Federal snow sampler tube depths (13.1 cm) were not as deep as the magna probe (15.2 cm) and thus more closely match the lidar snow depth (7.8 cm; see Figure 3), though a considerable low bias (~5.3 cm) similar to that found by Harder et al. (2020) persists in the lidar snow depth relative to the federal snow sampler snow depths. Additional factors such as downed logs, thick understory, and fine-scale topographic features (ie: small boulders and hummocky terrain) as well as reduced ground return density may contribute to the lidar snow depth errors in a forest, whereas these factors are absent in the field.

An improved understanding of forest canopies impacts on lidar returns is also warranted. Recent work has demonstrated that lidar pulses are "lost" at a much higher rate in forest canopies than open ground due to interception, absorption, and scattering through canopy transmission, with the loss ratio largely influenced by the range of the target from the sensor (Liu et al., 2020). The data that we presented in this paper were acquired using constant flight speed and at consistent altitude above target areas. Because of this, it is feasible that forest canopy conditions and variable understory vegetation density may have resulted in lost pulses and increased uncertainty in our data set. Indeed, we did observe lower return densities for both ground and all returns in forested areas in our data set (Figure 4).

One possible outcome of these lidar sampling issues in forests was a significant difference in snow depth confidence intervals between field and forest types and among slope groups. Confidence intervals were highest in conifer stands and on steep slopes and lowest in the field. This is likely partially the result of lower ground return density in forests due to the combined effects of lost pulses and canopy occlusion in forested areas. Additionally, this observation may be driven by increased variability in snow-off ground surface due to higher variability in the subnivean terrain in the forested areas of the study site (e.g., pockets of duff and woody debris). On cells where slopes exceed 20^{o}, there is more variability in ground return elevations over shorter distances due to errors in horizontal directions and spreading of the laser spot (Deems et al., 2013), which would partially drive larger confidence intervals of ground surface elevation for pixels located in high relief areas. These relatively high slope areas were more common in forested areas of the study site, and the DTM uncertainty resulting when there are high slopes also carries through snow depth estimation. Snow depth was significantly different between field and forested areas, as well as between conifer and deciduous forest types, despite the relatively high uncertainty. This indicates the possible influence of tree canopies on snow accumulation due to enhanced snow interception in forests (see reviews in Clark et al., 2011), and particularly in conifer stands, but also could be the result of an under-sampled ground surface in forested areas relative to field areas. Despite challenges with sampling in the forest area, some

Jennifer Jacobs 1/19/2021 11:06 AM
Deleted: terrain

Jennifer Jacobs 1/19/2021 11:07 AM
Deleted: terrain

Jennifer Jacobs 1/18/2021 6:35 PM
Deleted: While this result is not entirely surprising, i

Jennifer Jacobs 1/18/2021 6:36 PM
Deleted: t

Jennifer Jacobs 1/18/2021 6:38 PM
Deleted: depth due to pockets of duff and woody debris, and

Jennifer Jacobs 1/18/2021 6:37 PM
Deleted: .

Jennifer Jacobs 1/18/2021 6:39 PM
Deleted: Areas of high terrain relief are expected to have

Jennifer Jacobs 1/19/2021 2:21 PM
Formatted: Superscript

Jennifer Jacobs 1/19/2021 11:54 AM
Deleted: higher

Jennifer Jacobs 1/19/2021 2:23 PM
Deleted: High

Jennifer Jacobs 1/19/2021 2:21 PM
Deleted: relief

Jennifer Jacobs 1/18/2021 6:41 PM
Deleted: of the study site

Jennifer Jacobs 1/19/2021 2:23 PM
Deleted: around

degree of coherence for snow depth in the forest is apparent. The forest interception effects may be captured on average through forest structure parameters such as canopy closure and leaf area index that have traditionally used in snow models with canopy-snow interactions (see reviews in Snow model inter-comparison project – SNOWMIP2 by Essery et al., 2009; Rutter et al., 2009). However, the finer scale heterogeneity may benefit from additional parameters such as the mean distance to canopy and total gap area (Moeser et al., 2016) or modifications that reflect variations in canopy structure (Mazzotti et al., 2019). Snow depth also was significantly different among the three slope groups, possibly due to wind-driven snow displacement and sloughing on slopes during accumulation.

4.2 Flight Planning

Because high lift UAVs capable of carrying a lidar sensor package have challenges that may differ from small consumer grade UAVs used mainly for optical photogrammetry surveying, a well-formulated flight plan that addresses weather conditions, logistics of flying at a proposed site, flight lines, UAS equipment, and personnel is clearly needed. Weather impacts operations. UAS surveys cannot be conducted when there is any type of precipitation or in dense fog/clouds because moisture can cause electronic components to malfunction and moisture build-up on the propellers can also adversely affect lift production. Depending on the UAS, wind speeds exceeding 7 to 10 m/s may make flights more difficult. This project's Eagle XF high lift capacity UAV cannot be flown comfortably in winds greater than 8 m/s. At the study site, wind speeds often exceeded this threshold in the days immediately following snowfall except early in the morning. High wind speeds can also significantly reduce battery life as well as impact the accuracy of sensor observations. Low air temperatures can cause batteries to rapidly discharge. For winter UAS surveys, all flight and operational batteries were kept warm in a building, vehicle, or insulated cooler prior to the UAS survey. This also applies to the computer used to upload flight lines and relay telemetry information. A MIL-STD-810 certified Panasonic Toughbook was used in this study to handle the anticipated cold temperatures. Additionally, cold temperatures can severely limit the dexterity of the person manipulating the flight controls.

These high lift UAVs also have the potential to cause significant damage to person and property. The selection of a survey site not only needs to meet the scientific objectives of the UAS lidar survey, but also must have the proper attributes for safe and legal UAV operation including permission to operate the UAV at the site. Visual line of sight (VLOS) of the UAS needs to be maintained throughout the flight. When it is difficult to maintain VLOS (e.g., flying over forested or mountainous sites), spotters can be used if there is constant two-way communication between the spotters and the person operating the flight controls. For this study, an on-site, walk up tower with a spotter was necessary while the UAS was flown over the forest.

The deployment of a UAS lidar requires additional flight patterns designed for boresighting to ensure that point clouds are aligned (Painter et al., 2016). Provided that GNSS data are accurate, the most common reason for misalignment of point clouds is boresight angle errors (Li et al., 2019). Boresighting is the process of calculating the differences between lidar sensor and IMU roll, pitch, and yaw angle measurements to correct those errors in point clouds. Due to battery flight time limitations, we were unable to complete the flight pattern that is commonly used for boresighting alignment. Because of this, we leveraged our first two antiparallel flight lines for boresighting calibration. Additional details on boresighting calibration, our technique due to the flight time limitations, and examples of roll and pitch alignment errors observed during this field campaign appear in the supplemental materials.

4.3 UAS Sampling Strategies

While lidar calibration and data post-processing requirements are quite similar for UAS and airborne surveys, the UAS lidar surveys presented in this study have key differences from previous ALS surveys. As noted above, UAS flight durations are

Jennifer Jacobs 1/24/2021 2:29 PM
Deleted:

Jennifer Jacobs 1/24/2021 1:32 PM
Deleted: larger

Jennifer Jacobs 1/24/2021 1:34 PM
Deleted: that can carry

Jennifer Jacobs 1/24/2021 1:33 PM
Deleted: heavier

Jennifer Jacobs 1/24/2021 1:34 PM
Deleted: payloads

Jennifer Jacobs 1/24/2021 1:35 PM
Deleted: UAV

Jennifer Jacobs 1/24/2021 1:35 PM
Deleted: UAS

Jennifer Jacobs 1/24/2021 1:34 PM
Deleted: H

Jennifer Jacobs 1/24/2021 1:34 PM
Deleted: capable of carrying a lidar sensor package

Jennifer Jacobs 1/24/2021 9:51 AM
Deleted: UAV

Jennifer Jacobs 1/24/2021 1:35 PM
Deleted: UAV

Jennifer Jacobs 1/24/2021 1:35 PM
Deleted: UAV

Jennifer Jacobs 1/24/2021 9:51 AM
Deleted: UAV

Jennifer Jacobs 1/24/2021 9:51 AM
Deleted: system

considerably shorter, resulting in limited spatial coverage as compared to previous ALS snow depth surveys. This study took approximately 2 hours to map a 0.1 km^2 area. An advantage of UAS over ALS surveys is that the average point cloud density is much higher and has fewer missing pixels in the forest. This study's sampling densities and the proportion of areas with no ground returns are quite different from previous airborne lidar snow depth studies. This study had ground returns of 90 and 364 points/m^2 in the forest and field, respectively, and had no ground returns in only 0.086% and 0.95% of the 1 m resolution field and forest cells, respectively. In contrast, ALS surveys typically report surface model densities between 8 to 16 points/m^2 (Broxton et al., 2015; 2019; Currier et al., 2019; Kirchner et al., 2014) and ground returns between 3 and 6 points/m^2 (Broxton et al., 2019; Kirchner et al., 2014). ALS derived snow depth maps have a much greater proportion of areas that are masked due to no ground returns, particularly under trees, with masking areas ranging from less to 10% to more than 23% (Harpold et al., 2014; Mazzotti et al., 2019). While gap filling is possible, interpolation using measured snow depth values to fill under tree can overestimate snow depth (Zheng et al., 2016). Based on our work comparing field and forest lidar collections from a UAS, we suggest testing alternative flight plans, including reduced flight speed over forest canopies to account for lost pulses and canopy returns to produce ground return density that is comparable to field ground return density and to further reduce the number of missing pixels in an acquisition area. It is worth noting that as the capabilities of UAVs, power supplies and lightweight sensors continue to advance at an accelerated rate, UAS platforms will shortly rival the spatial coverage attainable by manned aircraft while maintaining improved efficiency and cost effectiveness.

A well understood challenge exists when developing a spatial sampling strategy in which, for given resources, there is a trade-off between spatial extent and sampling density (Clark et al. 2011). Increasing flight altitude can expand the spatial extent of an aerial survey. However, flying at higher altitudes results in a decreased point density. In theory, a higher point density could be achieved by slower speeds and increased swath overlap. The targeted spatial extent of an aerial survey dictates whether a manned aircraft or a UAV should be used. If the targeted area has a limited domain then using a manned airborne platform is probably overkill and inefficient for many studies and the use of a UAV would be more cost effective. However, as the domain increases in size, additional batteries would be required, much of the battery power would be used to reach the outer limits of the domain, and the ability to maintain the required line of sight could be difficult. Thus, there are end-members for survey site or regions where it is self-evident as to whether a UAV or an airborne platform should be used, but that leaves considerable gray areas where an appropriate choice of UAV platform with a well-designed mission could stretch the domain. Future research and technological advances are needed to offer insights for snow science observation platforms and trade-offs.

5. Conclusions

This paper describes and demonstrates a UAS-based lidar survey for snow depth mapping using commercially available components. The snow depth map was assessed in a mixed deciduous and coniferous forest and open field with little relief over a thin snowpack. The UAS includes an Eagle XF UAV manufactured by UAV America, a small, lightweight VLP-16 lidar (Velodyne, Inc.), and an Applanix APX-15 UAV INS. The INS has a measurement rate of 200 Hz, allowing returns to be georeferenced without ground control points. Data, post-processed to a SBET) file, resulted in approximately 3 cm positional accuracy. Flights were conducted at an altitude of 81 m and flight speed of 7 m/s, with a total of 12 parallel flight lines with targeted overlap of 40 percent. Once the point clouds were classified as ground and non-ground points, the flights yielded an average of 90 and 364 ground points/m^2 in the forest canopy and field, respectively, with 6.7% of the forest and 0.03% of the field cells having less than 5 point/m^2.

Jennifer Jacobs 1/24/2021 10:32 AM
Formatted: Not Highlight

Jennifer Jacobs 1/24/2021 10:32 AM
Formatted: Superscript

Jennifer Jacobs 1/18/2021 7:03 PM
Deleted: SD

Jennifer Jacobs 1/24/2021 9:52 AM
Deleted: platform

Liz Burakowski 1/22/2021 3:31 PM
Deleted: well designed

Jennifer Jacobs 1/24/2021 9:52 AM
Deleted: UAV

Jennifer Jacobs 1/24/2021 9:53 AM
Deleted: lidar system

Jennifer Jacobs 1/24/2021 9:53 AM
Deleted: UAS

Liz Burakowski 1/22/2021 3:31 PM
Deleted: light-weight

The snow depth map, generated by subtracting snow-off from snow-on DTMs derived from the resultant point clouds, reveals a snowpack whose depth ranges from less than 2 cm to over 18 cm. For both snow depth and confidence intervals, differences were found between vegetation cover types and slope, indicating complex snow-vegetation interaction that can be observed by UAS lidar return numbers. For 0.4 x 0.4 m cells, the *in situ* and lidar mean snow depths in the field were nearly identical with the MAD and RMSD values of approximately 1 cm. In the forest, the *in situ* mean snow depths from a magnaprobe were twice as large as the lidar snow depths with a correspondingly high RMSD. These forest differences have numerous possible explanations; 1) the snow probe's ability to penetrate the soil and vegetation resulting in random errors, 2) higher uncertainty in areas with canopy cover, variable ground surface, and high slope that occur more commonly in forested areas, 3) reduced total and ground return density in forests due to occlusion and lost pulses. Nevertheless, the results support previous findings indicating that there are limits to lidar snow depth validation at high horizontal and vertical spatial resolutions in some land covers and conditions. Mapped at 1 m^2 cells, a 0.5 to 1 cm snow depth confidence interval was achieved consistently in the field with confidence intervals increasing to median values of 4.0 cm in the deciduous forest and 5.9 cm in the coniferous forest. In the field, snow depth can be mapped at finer spatial resolutions with limited reduction in performance when reducing the cell size to 0.5 x 0.5 m and still achieving snow depth confidence intervals of less than 5 cm for a 0.2 x 0.2 m. Performance depends on both the point cloud density, which can be increased or decreased by changing the flight plan, and the grid cell variability that depends on site surface conditions.

Acknowledgements

This material is based upon work supported by the Broad Agency Announcement Program and the Cold Regions Research and Engineering Laboratory (ERDC-CRREL) under contract number W913E5-18-C-005. Any opinions, findings and conclusions or recommendations in this material are those of the author(s) and do not necessarily reflect the views of the Broad Agency Announcement Program and the Cold Regions Research and Engineering Laboratory.

The authors are grateful to Lee Friess for providing a technical review of the draft manuscript, Mahsa Moradi Khaneghahi for supporting manuscript preparation, and Ronny Schroeder contributing to the field data collection.

Data Availability

The UAS-based lidar point clouds and in-situ snow observations are available from the corresponding author upon reasonable request.

Author Contributions

JJ, AH, FS, and MP designed research and performed analysis. JJ, AH, FS, MP, EB, and EC conducted field work to obtain lidar and/or in-situ snow observations. AH, FS, CH, and EC produced figures. JJ wrote the initial draft. All authors contributed to manuscript review and editing.

Competing Interests

The authors declare that they have no conflict of interest.

Jennifer Jacobs 1/24/2021 9:54 AM
Deleted: UAV

Jennifer Jacobs 1/19/2021 11:08 AM
Deleted: areas of

Jennifer Jacobs 1/19/2021 11:08 AM
Deleted: terrain relief

Jennifer Jacobs 1/18/2021 7:21 PM
Deleted: within

References

Aanderud, Z. T., Jones, S. E., Schoolmaster Jr, D. R., Fierer, N., and Lennon, J. T.: Sensitivity of soil respiration and microbial communities to altered snowfall, Soil Biology and Biochemistry, 57, 217-227, 2013.

Aase, J. K. and Siddoway, F. H.: Crown-Depth Soil Temperatures and Winter Protection for Winter Wheat Survival1, Soil Science Society of America Journal, 43, 1229-1233, 1979.

Adams, M. S., Bühler, Y., and Fromm, R.: Multitemporal accuracy and precision assessment of unmanned aerial system photogrammetry for slope-scale snow depth maps in Alpine terrain, Pure and Applied Geophysics, 175, 3303-3324, 2018.

Adolph, A. C., Albert, M. R., Lazarcik, J., Dibb, J. E., Amante, J. M., and Price, A.: Dominance of grain size impacts on seasonal snow albedo at open sites in New Hampshire, Journal of Geophysical Research: Atmospheres, 122, 121-139, 2017.

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

Broxton, P. D., van Leeuwen, W. J., and Biederman, J. A.: Improving Snow Water Equivalent Maps With Machine Learning of Snow Survey and Lidar Measurements, Water Resources Research, 2019. 2019.

Bühler, Y., Adams, M. S., Bösch, R., and Stoffel, A.: Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations, The Cryosphere, 10, 1075-1088, 2016.

Bühler, Y., Adams, M. S., Stoffel, A., and Boesch, R.: Photogrammetric reconstruction of homogenous snow surfaces in alpine terrain applying near-infrared UAS imagery, International journal of remote sensing, 38, 3135-3158, 2017.

Burakowski, E., Tawfik, A., Ouimette, A., Lepine, L., Novick, K., Ollinger, S., Zarzycki, C., and Bonan, G.: The role of surface roughness, albedo, and Bowen ratio on ecosystem energy balance in the Eastern United States, Agricultural and Forest Meteorology, 249, 367-376, 2018.

Burakowski, E. A., Ollinger, S. V., Lepine, L., Schaaf, C. B., Wang, Z., Dibb, J. E., Hollinger, D. Y., Kim, J., Erb, A., and Martin, M.: Spatial scaling of reflectance and surface albedo over a mixed-use, temperate forest landscape during snow-covered periods, Remote Sensing of Environment, 158, 465-477, 2015.

Cimoli, E., Marcer, M., Vandecrux, B., Bøggild, C. E., Williams, G., and Simonsen, S. B.: Application of low-cost UASs and digital photogrammetry for high-resolution snow depth mapping in the Arctic, Remote Sensing, 9, 1144, 2017.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land‐surface models: A review, Water Resources Research, 47, 2011.

Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., Masek, J. G., Ranson, K. J., Ly, V., and Montesano, P. M.: NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager, Remote Sensing, 5, 4045-4066, 2013.

Currier, W. R. and Lundquist, J. D.: Snow depth variability at the forest edge in multiple climates in the western United States, Water Resources Research, 54, 8756-8773, 2018.

Currier, W. R., Pflug, J., Mazzotti, G., Jonas, T., Deems, J. S., Bormann, K. J., Painter, T. H., Hiemstra, C. A., Gelvin, A., and Uhlmann, Z.: Comparing aerial lidar observations with terrestrial lidar and snow‐probe transects from NASA's 2017 SnowEx campaign, Water Resources Research, 2019. 2019.

De Michele, C., Avanzi, F., Passoni, D., Barzaghi, R., Pinto, L., Dosso, P., Ghezzi, A., Gianatti, R., and Della Vedova, G.: Using a fixed-wing UAS to map snow depth distribution: an evaluation at peak accumulation, Cryosphere, 10, 511-522, 2016.

Deems, J. S., Fassnacht, S. R., and Elder, K. J.: Fractal distribution of snow depth from LiDAR data, Journal of Hydrometeorology, 7, 285-297, 2006.

Deems, J. S., Painter, T. H., and Finnegan, D. C.: Lidar measurement of snow depth: a review, J. Glaciol., 59, 467-479, 2013.

Dickinson, W. and Whiteley, H.: A sampling scheme for shallow snowpacks, Hydrological Sciences Journal, 17, 247-258, 1972.

Dolgun, A. and Demirhan, H.: Performance of nonparametric multiple comparison tests under heteroscedasticity, dependency, and skewed error distribution, Communications in Statistics-Simulation and Computation, 46, 5166-5183, 2017.

Essery, R., Rutter, N., Pomeroy, J., Baxter, R., Stähli, M., Gustafsson, D., Barr, A., Bartlett, P., and Elder, K.: SNOWMIP2: An evaluation of forest snow process simulations, Bulletin of the American Meteorological Society, 90, 1120-1136, 2009.

Fernandes, R., Prevost, C., Canisius, F., Leblanc, S. G., Maloley, M., Oakes, S., Holman, K., and Knudby, A.: Monitoring snow depth change across a range of landscapes with ephemeral snowpacks using structure from motion applied to lightweight unmanned aerial vehicle videos, The Cryosphere, 12, 3535-3550, 2018.

Fey, C., Schattan, P., Helfricht, K., and Schöber, J.: A compilation of multitemporal TLS snow depth distribution maps at the Weisssee snow research site (Kaunertal, Austria), Water Resources Research, 2019. 2019.

Friggens, M. M., Williams, M. I., Bagne, K. E., Wixom, T. T., and Cushman, S. A.: Effects of climate change on terrestrial animals [Chapter 9], In: Halofsky, Jessica E.; Peterson, David L.; Ho, Joanne J.; Little, Natalie, J.; Joyce, Linda A., eds. Climate change vulnerability and adaptation in the Intermountain Region [Part 2]. Gen. Tech. Rep. RMRS-GTR-375. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 264-315., 375, 264-315, 2018.

Jennifer Jacobs 1/24/2021 2:55 PM
Deleted:

Jennifer Jacobs 1/24/2021 2:55 PM
Formatted: Indent: Left: 0", Hanging: 0.5", Tabs: 0.5", Left

Gabrlik, P., Janata, P., Zalud, L., and Harcarik, J.: Towards Automatic UAS-Based Snow-Field Monitoring for Microclimate Research, Sensors, 19, 1945, 2019.

Gandahl, R.: Determination of the depth of soil freezing with a new frost meter (text in Swedish). Report 30. Stattensaginsititut Stockholm (SIP 16347). 1957.

Gichamo, T. Z. and Tarboton, D. G.: Ensemble Streamflow Forecasting using an Energy Balance Snowmelt Model Coupled to a Distributed Hydrologic Model with Assimilation of Snow and Streamflow Observations, Water Resources Research, 2019. 2019.

Goetz, J. and Brenning, A.: Quantifying uncertainties in snow depth mapping from structure from motion photogrammetry in an alpine area, Water Resources Research, 55, 7772-7783, 2019.

Groffman, P. M., Driscoll, C. T., Fahey, T. J., Hardy, J. P., Fitzhugh, R. D., and Tierney, G. L.: Colder soils in a warmer world: A snow manipulation study in a northern hardwood forest ecosystem, Biogeochemistry, 56, 135-150, 2001.

Grünewald, T., Schirmer, M., Mott, R., and Lehning, M.: Spatial and temporal variability of snow depth and SWE in a small mountain catchment, The Cryosphere, 4, 215-225, 2010.

Harder, P., Pomeroy, J. W., and Helgason, W. D.: Improving sub-canopy snow depth mapping with unmanned aerial vehicles: lidar versus structure-from-motion techniques, The Cryosphere, 14, 1919-1935, 2020.

Harder, P., Schirmer, M., Pomeroy, J., and Helgason, W.: Accuracy of snow depth estimation in mountain and prairie environments by an unmanned aerial vehicle, The Cryosphere, 10, 2559, 2016.

Harpold, A., Guo, Q., Molotch, N., Brooks, P., Bales, R., Fernandez‐Diaz, J., Musselman, K., Swetnam, T., Kirchner, P., and Meadows, M.: LiDAR‐derived snowpack data sets from mixed conifer forests across the Western United States, Water Resources Research, 50, 2749-2755, 2014.

Hatchett, B. J. and Eisen, H. G.: Brief Communication: Early season snowpack loss and implications for oversnow vehicle recreation travel planning, The Cryosphere, 13, 21-28, 2019.

Helsel, D. R. and Hirsch, R. M.: Statistical methods in water resources, Elsevier, 2002.

Henry, H. A. L.: Climate change and soil freezing dynamics: historical trends and projected changes, Climatic Change, 87, 421-434, 2008.

Hopkinson, C., Chasmer, L., Young-Pow, C., and Treitz, P.: Assessing forest metrics with a ground-based scanning lidar, Canadian Journal of Forest Research, 34, 573-583, 2004.

Horn, B. K.: Hill shading and the reflectance map, Proceedings of the IEEE, 69, 14-47, 1981.

Isard, S. A. and Schaetzl, R. J.: Effects of winter weather conditions on soil freezing in southern Michigan, Physical Geography, 19, 71-94, 1998.

Jagt, B. V., Lucieer, A., Wallace, L., Turner, D., and Durand, M.: Snow Depth Retrieval with UAS Using Photogrammetric Techniques, Geosciences, 5, 264-285, 2015.

Keyetieu, R. and Seube, N.: Automatic Data Selection and Boresight Adjustment of LiDAR Systems, Remote Sensing, 11, 1087, 2019.

Kirchner, P., Bales, R., Molotch, N., Flanagan, J., and Guo, Q.: LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California, Hydrology and Earth System Sciences, 18, 4261-4275, 2014.

Kraatz, S., Jacobs, J. M., Schröder, R., Cho, E., Cosh, M., Seyfried, M., Prueger, J., and Livingston, S.: Evaluation of SMAP freeze/thaw retrieval accuracy at core validation sites in the contiguous united states, Remote Sensing, 10, 1483, 2018.

Kraus, K. and Pfeifer, N.: Advanced DTM generation from LiDAR data, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 34, 23–30, 2011.

Kwok, R., Kacimi, S., Webster, M., Kurtz, N., and Petty, A.: Arctic Snow Depth and Sea Ice Thickness From ICESat‐2 and CryoSat‐2 Freeboards: A First Examination, Journal of Geophysical Research: Oceans, 125, e2019JC016008, 2020.

Li, Z., Tan, J., and Liu, H.: Rigorous boresight self-calibration of mobile and UAV LiDAR scanning systems by strip adjustment, Remote Sensing, 11, 442, 2019.

López-Moreno, J. I., Revuelto, J., Alonso-González, E., Sanmiguel-Vallelado, A., Fassnacht, S. R., Deems, J., and Morán-Tejeda, E.: Using very long-range terrestrial laser scanner to analyze the temporal consistency of the snowpack distribution in a high mountain environment, Journal of Mountain Science, 14, 823-842, 2017.

Marti, R., Gascoin, S., Berthier, E., De Pinel, M., Houet, T., and Laffly, D.: Mapping snow depth in open alpine terrain from stereo satellite imagery, Cryosphere, 10, 2016.

Maune, D. F. and Nayegandhi, A.: Digital elevation model technologies and applications: the DEM users manual, Asprs Publications, 2018.

Mazzotti, G., Currier, W. R., 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 Resources Research, 55, 6198-6216, 2019.

McGrath, D., Webb, R., Shean, D., Bonnell, R., Marshall, H. P., Painter, T. H., Molotch, N. P., Elder, K., Hiemstra, C., and Brucker, L.: Spatially extensive ground‐penetrating radar snow depth observations during NASA's 2017 SnowEx campaign: Comparison with In situ, airborne, and satellite observations, Water Resources Research, 55, 10026-10036, 2019.

Meyer, J. and Skiles, S. M.: Assessing the Ability of Structure From Motion to Map High‐Resolution Snow Surface Elevations in Complex Terrain: A Case Study From Senator Beck Basin, CO, Water Resources Research, 55, 6596-6605, 2019.

Eunsang Cho 1/22/2021 1:25 AM
Deleted: 1992

Jennifer Jacobs 1/24/2021 2:55 PM
Formatted: Indent: Left: 0", Hanging: 0.5", Tabs: 0.5", Left

Moeser, D., Mazzotti, G., Helbig, N., and Jonas, T.: Representing spatial variability of forest snow: Implementation of a new interception model, Water Resources Research, 52, 1208-1226, 2016.

Monson, R. K., Lipson, D. L., Burns, S. P., Turnipseed, A. A., Delany, A. C., Williams, M. W., and Schmidt, S. K.: Winter forest soil respiration controlled by climate and microbial community composition, Nature, 439, 711, 2006.

Mott, R., Schirmer, M., and Lehning, M.: Scaling properties of wind and snow depth distribution in an Alpine catchment, Journal of Geophysical Research: Atmospheres, 116, 2011.

Mott, R., Vionnet, V., and Grünewald, T.: The seasonal snow cover dynamics: review on wind-driven coupling processes, Frontiers in Earth Science, 6, 197, 2018.

Muir, J., Goodwin, N., Armston, J., Phinn, S., and Scarth, P.: An accuracy assessment of derived digital elevation models from terrestrial laser scanning in a sub-tropical forested environment, Remote Sensing, 9, 843, 2017.

Nolan, M., Larsen, C., and Sturm, M.: Mapping snow-depth from manned-aircraft on landscape scales at centimeter resolution using Structure-from-Motion photogrammetry, Cryosphere Discussions, 9, 2015.

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., and Marks, D.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo, Remote Sensing of Environment, 184, 139-152, 2016.

Palace, M., Keller, M., Asner, G. P., Hagen, S., and Braswell, B.: Amazon forest structure from IKONOS satellite data and the automated characterization of forest canopy properties, Biotropica, 40, 141-150, 2008.

Palace, M. W., McMichael, C., Braswell, B. H., Hagen, S. C., Bush, M. B., Neves, E., Tamanaha, E., Herrick, C., and Frolking, S.: Ancient Amazonian populations left lasting impacts on forest structure, Ecosphere, 8, e02035, 2017.

Pauli, J. N., Zuckerberg, B., Whiteman, J. P., and Porter, W.: The subnivium: a deteriorating seasonal refugium, Frontiers in Ecology and the Environment, 11, 260-267, 2013.

Perron, C.J., Bennett, K., & Lee, T.D.: Forest stewardship plan: Thompson farm. NH: University of New Hampshire. Produced by Ossipee Mountain Land Company, West Ossipee, https://colsa.unh.edu/sites/default/files/thompson-farm-plan.pdf, 2004.

Prince, M., Roy, A., Royer, A., and Langlois, A.: Timing and spatial variability of fall soil freezing in boreal forest and its effect on SMAP L-band radiometer measurements, Remote Sensing of Environment, 231, 111230, 2019.

Reinmann, A. B. and Templer, P. H.: Increased soil respiration in response to experimentally reduced snow cover and increased soil freezing in a temperate deciduous forest, Biogeochemistry, 140, 359-371, 2018.

Ricard, J. A., Tobiasson, W., and Greatorex, A.: The field assembled frost gage. Technical Note. Cold Regions Research and Engineering Laboratory, U.S. Army Corps of Engineers, Hanover, New Hampshire, 1976. 1976.

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., and Deng, H.: Evaluation of forest snow processes models (SnowMIP2), Journal of Geophysical Research: Atmospheres, 114, 2009.

Schweizer, J., Jamieson, J. B., and Schneebeli, M.: Snow avalanche formation, Rev. Geophys., 41, 25, 2003.

Seyfried, M. S., Wilcox, B. P., and Cooley, K. R.: Environmental conditions and processes associated with runoff from frozen soil at Reynolds Creek Watershed, Frozen soil impacts on agricultural, range, and forest lands, 1990. 125-134, 1990.

Sharratt, B. S. and McCool, D. K.: Frost Depth. In: Micrometeorology in Agricultural Systems, Hatfield, J. L. and Baker, J. M. (Eds.), 2005.

Shaw, T. E., Gascoin, S., Mendoza, P. A., Pellicciotti, F., and McPhee, J.: Snow Depth Patterns in a High Mountain Andean Catchment from Satellite Optical Tristereoscopic Remote Sensing, Water Resources Research, 56, e2019WR024880, 2020.

Singh, P., Wu, J. Q., McCool, D. K., Dun, S., Lin, C.-H., and Morse, J. R.: Winter hydrologic and erosion processes in the US Palouse region: Field experimentation and WEPP simulation, Vadose Zone Journal, 8, 426-436, 2009.

Sorensen, P. O., Finzi, A. C., Giasson, M.-A., Reinmann, A. B., Sanders-DeMott, R., and Templer, P. H.: Winter soil freeze-thaw cycles lead to reductions in soil microbial biomass and activity not compensated for by soil warming, Soil Biology and Biochemistry, 116, 39-47, 2018.

Starkloff, T., Stolte, J., Hessel, R., and Ritsema, C.: Investigating the development of shallow snowpacks on arable land, using comprehensive field observations and spatially distributed snow modelling, Hydrology Research, 49, 41-59, 2017.

Slotznick, S.P., Sperling, E.A., Tosca, N.J., Miller, A.J., Clayton, K.E., van Helmond, N.A.G.M., Slomp, C.P. and Swanson‐Hysell, N.L.: Unraveling the mineralogical complexity of sediment iron speciation using sequential extractions. Geochemistry, Geophysics, Geosystems, 21(2), 2020.

Sturm, M. and Holmgren, J.: An automatic snow depth probe for field validation campaigns, Water Resources Research, 54, 9695-9701, 2018.

Sullivan, F. B., Ducey, M. J., Orwig, D. A., Cook, B., and Palace, M. W.: Comparison of lidar-and allometry-derived canopy height models in an eastern deciduous forest, Forest Ecology and Management, 406, 83-94, 2017.

Tonkin, T. N. and Midgley, N. G.: Ground-Control Networks for Image Based Surface Reconstruction: An Investigation of Optimum Survey Designs Using UAV Derived Imagery and Structure-from-Motion Photogrammetry, Remote Sensing, 8, 8, 2016.

Tucker, C. L., Tamang, S., Pendall, E., and Ogle, K.: Shallow snowpack inhibits soil respiration in sagebrush steppe through multiple biotic and abiotic mechanisms, Ecosphere, 7, e01297, 2016.

Jennifer Jacobs 1/24/2021 2:55 PM
Formatted: Indent: Left: 0", Hanging: 0.5", Tabs: 0.5", Left

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:Times New Roman, 10

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: EndNote Bibliography, Indent: Left: 0", Hanging: 0.5", Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tabs: 0.5", Left + Not at 0.39" + 0.78" + 1.17" + 1.56" + 1.94" + 2.33" + 2.72" + 3.11" + 3.5" + 3.89" + 4.28" + 4.67"

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold, Not Italic

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold, Not Italic

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:10 pt, Not Bold

Jennifer Jacobs 1/24/2021 10:30 AM
Formatted: Font:8 pt

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X., Vuyovich, C. M., Cosh, M. H., and Jacobs, J. M.: Remote sensing of drivers of spring snowmelt flooding in the North Central US. In: Remote sensing of hydrological extremes, Springer, 2017.

Vander Jagt, B., Lucieer, A., Wallace, L., Turner, D., and Durand, M.: Snow depth retrieval with UAS using photogrammetric techniques, Geosciences, 5, 264-285, 2015.

Wallace, L., Lucieer, A., Watson, C., and Turner, D.: Development of a UAV-LiDAR system with application to forest inventory, Remote Sensing, 4, 1519-1543, 2012.

Wallace, L., Lucieer, A., and Watson, C. S.: Evaluating tree detection and segmentation routines on very high resolution UAV LiDAR data, IEEE Transactions on Geoscience and Remote Sensing, 52, 7619-7628, 2014.

Watanabe, K. and Osada, Y.: Comparison of hydraulic conductivity in frozen saturated and unfrozen unsaturated soils, Vadose Zone Journal, 15, 2016.

Wever, N., Jonas, T., Fierz, C., and Lehning, M.: Model simulations of the modulating effect of the snow cover in a rain-on-snow event, Hydrology and Earth System Sciences, 18, 4657-4669, 2014.

Yi, Y., Kimball, J. S., Chen, R. H., Moghaddam, M., and Miller, C. E.: Sensitivity of active-layer freezing process to snow cover in Arctic Alaska, The Cryosphere, 13, 197, 2019.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., and Zhang, C.: A progressive morphological filter for removing nonground measurements from airborne LIDAR data, IEEE transactions on geoscience and remote sensing, 41, 872-882, 2003.

Zhu, L., Ives, A. R., Zhang, C., Guo, Y., and Radeloff, V. C.: Climate change causes functionally colder winters for snow cover-dependent organisms, Nature Climate Change, 9, 886-893, 2019.