Review of “Identifying Snow-Covered Areas from Unoccupied Aerial Systems (UAS)

Visible Imagery: A Comparison of Methods” (Moradi et al. 2025)

General Comments

In their study, the authors present a method of mapping snow cover using drones with different sensor systems and land surface classification algorithms. The study's primary objective is to conduct an inter-comparison experiment to evaluate the performance of various snow classification methods in relation to different acquiring sensors, study sites, and training data availability. This is an interesting setup because it addresses common challenges associated with drone surveys and aims to provide recommendations on the most suitable methods for specific acquisition conditions or user applications. The manuscript is well written and mostly well structured. The scientific objects are within the scope of The Cryosphere.

Thank you for the positive words.

However, I think the analysis of the differences between the methods is a bit shallow. On the one hand, the experimental setup is relatively simplistic (i.e. binary classification of “snow” vs. “no- snow”), but on the other, it may be flawed by some uncompensated differences in data acquisition (image dates), as outlined in my general comments below. Therefore, I believe that major revisions are necessary before this study can be published. Nevertheless, expanding the experimental setup could provide a much deeper understanding of the differences between the methods and their underlying causes.

Overall Response

We appreciate the constructive feedback from both reviewers. Reviewer 1 and Reviewer 2 were generally supportive of the study concept but noted that the experimental design and data acquisition were not originally intended to support robust cross-sensor or cross-site comparisons. Both reviewers also pointed out opportunities to extend the analysis, including examining performance in forested environments (Reviewer 2) and incorporating lidar-based estimates of snow-covered area (SCA).

We agree with these assessments and have revised the manuscript accordingly. First, we removed the cross-sensor and cross-site comparisons from the manuscript to ensure that the scope of the analysis aligns with the original experimental design. Second, we added a new analysis in the Results section that uses UAS lidar observations to estimate SCA, providing an additional and independent perspective on snow detection. Finally, we expanded the Discussion to address SCA estimation in forested environments. This section explores both the potential transferability of the RF approach developed in open areas and the limitations of optical imagery in forested terrain.

We believe these revisions better align the manuscript with the strengths of the dataset while also acknowledging the challenges and opportunities associated with extending these methods to more complex environments. Our intent is to maintain the original goal of presenting results under controlled and straightforward conditions while clearly articulating limitations, challenges, and next steps for applications in more complex forested landscapes.

• Experiment images: The dataset used is certainly interesting, covering a wide range of snow cover conditions. However, a significant limitation of the cross-sensor and cross-site comparison is that the images were not acquired on the same day. It is likely that there is some kind of date-related bias between the images because the properties of the snow, the illumination and the shadows are likely to be different. Ideally, images from different sensors at the same site would have been acquired on the same day, immediately after each other. As far as I can see, these potential sources of difference are not currently accounted for in the experiments. Were the images acquired at the same time of day? What about weather conditions (e.g. cloudiness)? To what extent are the observed cross-sensor and cross-site differences actually related to acquisition dates?

Response
We agree that because the data collected by different sensors were not acquired simultaneously and environmental conditions differed between acquisition dates, the analysis would not fully support the intended comparison. For this reason, we decided to remove Experiments 2 and 3 from the manuscript.

To provide additional context for the readers, a table summarizing the weather conditions during image acquisition was added to the supplementary material

Date	Site	Sensor	Start Time (local time)	End Time (local time)	Mean air temp (during flight) [°C]	Stdev air temp (during flight) [°C]	Mean SW_{in} [W/m^2]	Stdev SW_{in} [W/m^2]	Sky Condition	Field camera snow depth (cm)	Judd mean snow depth (cm)	Lidar mean snow depth (cm)
12/23/2020	TF	P4	9:26:00	11:10:00	-1.05	0.48	304.76	38.26	Partly Cloudy	7	9.21	12.8
1/25/2021	TF	P4	10:43:00	11:40:00	-1.86	0.56	428.25	12.55	Cloudy	0	0	No data
2/20/2021	TF	P4	12:04:00	12:44:00	-1.1	0.28	479.63	132.73	Cloudy	20	17.4	24.5
2/23/2021	TF	P4	12:16:00	12:55:00	5.01	0.23	541.13	144.44	Mostly Clear	15	13.8	17.6
2/24/2021	TF	P4	10:41:00	11:20:00	7.55	0.24	590.63	18.59	Clear	10	11.5	13.3
2/26/2021	TF	P4	11:57:00	12:38:00	0.16	0.18	632.75	3.41	Cloudy	7	9.06	8.7
2/28/2021	TF	P4	12:21:00	13:01:00	5.64	0.09	261.63	16.71	Clear	7	9.56	12.5
3/1/2021	TF	P4	16:37:00	17:16:00	5.36	0.16	29.5	17.02	Cloudy	4	7.72	No data
3/3/2021	TF	P4	12:00:00	12:39:00	3.24	0.34	459.63	177.73	Partly Cloudy	3	7.25	5.4
3/7/2021	TF	P4	11:25:00	12:06:00	-2.1	0.25	679.33	2.74	Clear	0	6.38	4.5
3/10/2021	TF	P4	13:50:00	14:34:00	7.84	0.17	541.67	27.6	Clear	0	3.23	No data
3/11/2021	TF	P4	15:12:00	15:51:00	18.89	0.35	265.13	62.83	Clear	0	1.65	0
4/2/2021	TF	P4	14:19:00	14:58:00	2.49	0.21	406.5	94.09	Clear	0	0	No data
1/12/2022	TF	P4	14:47:00	15:26:00	0.43	0.09	19.5	10.07	Clear	4	4.2	No data
1/26/2022	TF	P4	12:50:00	13:32:00	-6.03	0.18	490.11	15.76	Clear	0	0	No data
1/18/2024	KF	P4	13:39:00	14:43:00	-3.25	0.1	162.31	42.14	Partly Cloudy	8	6.4	No data
1/23/2024	KF	P4	14:26:00	15:17:00	3.21	0.03	86.1	23.04	Cloudy	6	6.7	No data
1/27/2024	KF	P4	13:46:00	14:26:00	3.4	0.16	94.25	17.3	Mostly Clear	4	4.6	No data
2/6/2024	KF	P4	14:36:00	15:34:00	3.69	0.14	259.45	40.05	Clear	4	5.4	No data
2/8/2024	KF	P4	13:19:00	13:59:00	6.25	0.14	436	17.19	Clear	2	3.8	No data
2/12/2024	KF	P4	15:40:00	16:27:00	6.12	0.8	80.9	59.04	Partly Cloudy	0	0	No data
2/26/2024	KF	P4	10:45:00	11:25:00	5.34	0.88	491.11	139	Mostly Clear	0	0	No data
4/5/2024	KF	P4	15:16:00	16:02:00	5.61	0.13	95.78	22.5	Cloudy	15	11.9	No data
4/8/2024	KF	P4	14:41:00	15:27:00	16.48	0.76	246.78	96.5	Clear	0	0	No data
12/27/2024	KF	MR	13:25:00	14:10:00	3.5	0.12	165.3	34.01	Cloudy	No data	No data	No data
1/13/2025	KF	P4	15:38:00	16:31:00	0.86	0.3	0	0	Cloudy	No data	No data	No data
2/3/2025	KF	MR	12:15:00	13:00:00	2.81	0.17	356.3	8.43	Partly Cloudy	No data	No data	16.6
2/19/2025	KF	MR	12:25:00	13:15:00	-4.84	0.29	424.27	8.67	Mostly Clear	No data	No data	40.7
3/9/2025	KF	MR	14:35:00	15:20:00	4.35	0.24	192.9	27.29	Partly Cloudy	No data	No data	6.9
3/10/2025	KF	MR	16:45:00	17:25:00	12.31	0.28	196.56	44.89	Mostly Clear	No data	No data	4.3

Experiment sites: Similarly to the selection of images, I am not entirely convinced by the study area to which the image mosaics are masked (Figure 1). It looks like you have removed all areas except for the almost flat fields with sparse vegetation. While this certainly improves the separation of snow and bare ground, it means that you cannot draw any conclusions about the performance of the applied algorithms in more complex terrain, such as slopes and forests. Conversely, I suspect that the overall good performance of your classifiers (almost all approaches yield reasonable results) is due to the exclusion of challenging areas. This is made even more apparent by the fact that, in your manuscript, you mention that satellite snow cover struggles in complex terrain or forest areas compared to drone surveys (Line 59), yet your own dataset is restricted to areas where a satellite image would also most likely produce good classification results (at a coarser resolution of course). I think this is a missed opportunity, as your analysis could be much more informative if you assessed the accuracy of different methods in different landscapes.

Response: The reviewer's point is well taken; we agree that evaluating SCA mapping methods in more complex environments such as forests and sloped terrain would provide valuable insight. The overarching goal of this work was to start with the most straightforward conditions (fields, flat terrain with sparse vegetation) and establish a baseline comparison of commonly used SCA mapping approaches under controlled conditions at shallow and intermittent snow regimes and then conduct future work to move more challenging conditions. This would allow us to isolate differences among the tested classification methods without the additional confounding effects introduced by complex terrain, forest canopy, or steep slopes. We advocate for looking at the simple case first because even under these relatively homogenous surface conditions, not all tested methods produced high accuracy or provided similar performances. Evaluating the methods in a controlled setting therefore provides a necessary baseline for understanding their behavior before extending the analysis to more complex environments.

At the same time, we recognize that the dataset provides an opportunity to examine UAS-based SCA detection in forested environments, and we agree that offering some insight into these conditions would add value for readers. To address the reviewer's comment in more detail, we have expanded the Discussion to include results related to SCA mapping in forested areas. These additions highlight that extending the analysis to forested and more complex terrain represents an important direction for future work and outline the challenges associated with relying solely on optical imagery in such environments.

"Although this study focuses on open-field environments to enable a controlled comparison of SCA mapping approaches, extending the analysis to forested environments remains an important direction for future work. To explore this possibility, we masked the forest canopy and applied the RF, SVM, and MLE models developed in the open-field analysis to canopy gaps within the forested portions of the Kingman Farm study site orthomosaics (11 dates; see Supplementary Material for methodological details). When applied to canopy gaps, the models' accuracy, balanced accuracy, and F1 scores were comparable to those obtained in the open areas, while Cohen's kappa values decreased modestly from ~0.90 in open terrain to ~0.80 within forest openings (Figure 10).

Although the models perform well in canopy gaps, reliable SCA mapping in forested environments depends on the availability of a consistent and accurate canopy masking approach. This is particularly challenging in landscapes dominated by deciduous forests. Using a k-medoids clustering algorithm, we generated canopy masks when the forest floor was snow-covered; however, we encountered several limitations that are likely common to most unsupervised classification approaches. In particular, these methods cannot reliably distinguish branches lying on the ground from branches that remain elevated within the canopy. In deciduous forests, branch structures are sufficiently fine that modest shifts in canopy position due to wind or snow loading, as well as minor alignment differences between UAS orthomosaics, prevent masks generated for one date from being reliably applied to other dates. These *issues highlight an important limitation of relying solely on optical imagery— the focus of this study—for SCA mapping in forested environments.*

Additional challenges arise from branches and litter present on the snow surface, which can reduce classification reliability and complicate identification of true canopy gaps where snow accumulates on the ground. These effects are particularly pronounced in deciduous forests under leaf-off conditions. Furthermore, snow intercepted and retained on branches cannot be easily distinguished from snow on the ground using optical imagery alone. Together, these limitations highlight the challenges of relying on optical data for snow mapping in forested landscapes and motivate the exploration of existing satellite fSCA and NDSI methods which use canopy cover information and integration of complementary observations, such as LiDAR, which can provide additional information about sub-canopy snow distribution.''

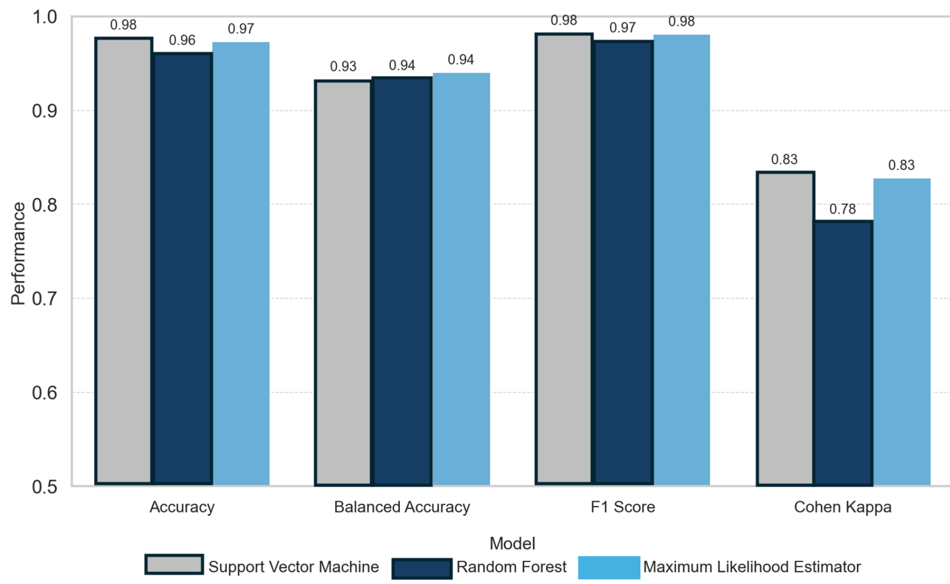
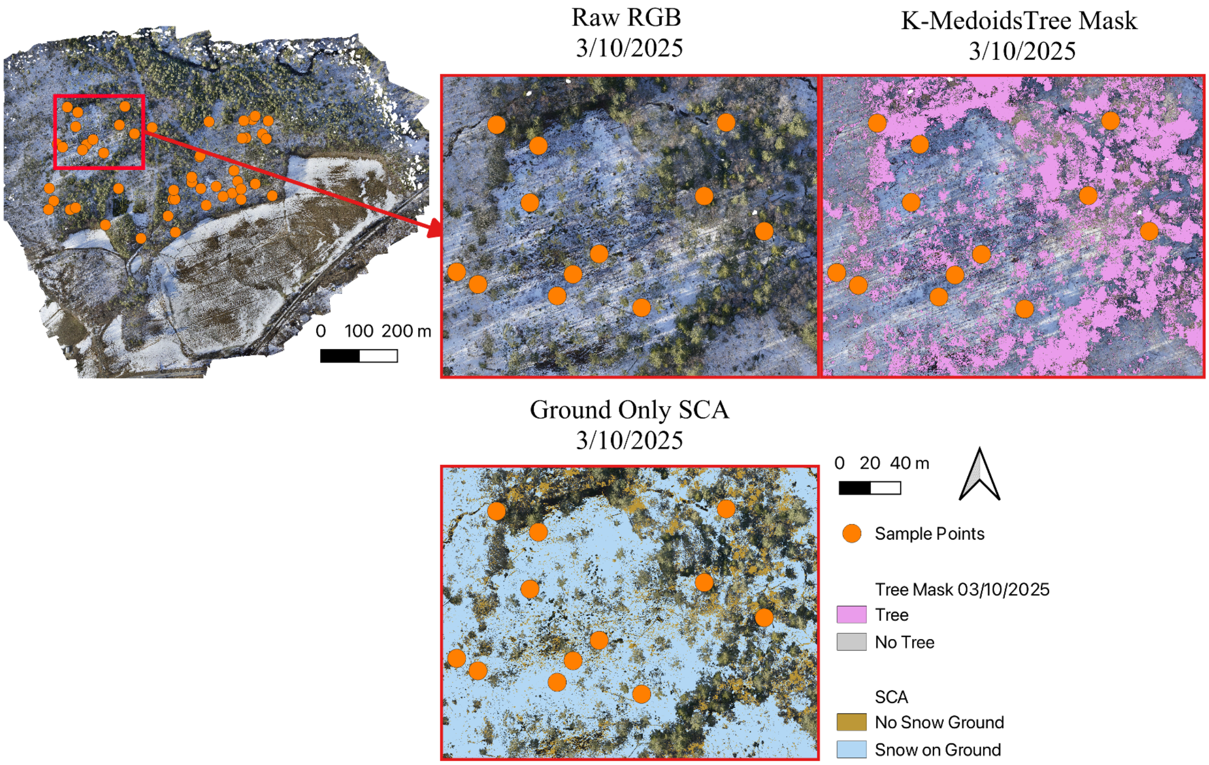


Figure 10 - Performance metrics of the three best-performing models from the reference SCA analysis, evaluated in forested area canopy gaps using 11 orthomosaics collected at the Kingman Farm study site.

Also, additional details describing our workflow for estimating SCA in forest gaps have been added to the supplementary material.

''In this study, we mapped snow cover in forested environments by first identifying and masking tree canopies. Tree canopy masks were generated using an unsupervised clustering approach applied to the orthomosaic imagery. A representative subset of pixels was randomly sampled from each image to capture the range of surface conditions present in the scene. The sampled pixel values were transformed into the CIELAB color space and reduced using principal component analysis (PCA) to improve feature separability and reduce dimensionality. K-medoids clustering was then applied to classify pixels into multiple clusters representing different surface types. The cluster corresponding to tree canopy was identified through visual inspection and used to generate a binary mask that removed forested areas from each orthomosaic. After masking the canopy, 50 validation points were randomly sampled from the remaining gaps between tree canopies and labeled as snow or no snow. The models trained in the open-field experiments presented in this manuscript were then applied to estimate snow cover at these validation points, and their performance was evaluated. An example of the classification of tree canopy and subsequent SCA classification of the ground is shown in the following figure.''



The reviewer is also correct that an earlier version of the manuscript included discussion in the Introduction related to forest environments that was not directly supported by the analyses presented. Because the present study does not explicitly target forested terrain, we have revised the Introduction to remove discussion of forest- and terrain-specific challenges associated with satellite SCA products. Instead, the Introduction now focuses on the conditions directly addressed in this study, namely shallow and spatially discontinuous snow cover.

"Additionally, UAS-based snow mapping may be less affected than coarse-resolution satellite snow products in shallow or patchy snow conditions because centimeter-scale imagery reduces mixed-pixel effects and allows better detection of fine-scale snow heterogeneity, whereas satellite products can exhibit increased omission errors when snow cover is thin or spatially discontinuous"

Data availability: Your data availability statement only refers to meteorological data which is available elsewhere. What about your survey data? The raw images and your intermediate classification results should also be made available via an open data repository.

Response:
Thank you for the comment. This is an extensive dataset that we would like to make available to the community. We have requested approval from our funding agency (DoD) to make the dataset publicly available. Unfortunately, the release of the dataset must remain pending until this approval is granted.

Specific Comments

Line 50-52: Unoccupied aerial systems (UAS) are emerging as a reliable, low-cost method for collecting snow data at finer spatial resolution than satellites and are not limited by orbital revisit times.“ => While it is true that drones provide much better resolution than non-commercial satellite images, the logistical effort required depends on the use case. For example, the surface reflectance products of the Landsat/Sentinel missions have a revisit time of a few days. They are distributed free of charge by the ESA and NASA, and they include infrared channels, which are great for mapping snow. Conversely, drones require an operator in the field, surveys of large areas take time, and they are limited by weather conditions. So, all in all, there are specific applications where one system is superior to the other. If you want to compare satellites and drones in the introduction (which makes sense), you should briefly summaries the advantages and disadvantages of drones and satellites, and which data is useful for which application.

Response:
Thank you for this suggestion. We agree that the relative advantages of UAS and satellite imagery depend strongly on the application and that UAS should not be presented as universally "low-effort" or superior. We revised the Introduction to provide a more clarity on specific applications of UAS as follows:

"*Satellite products (e.g., Landsat and Sentinel) provide free, repeatable multispectral coverage suitable for regional snow monitoring, but their coarser spatial resolution and susceptibility to cloud cover and mixed-pixel effects can limit their application. In contrast, unoccupied aerial systems (UAS) offer much finer spatial resolution and flexible, on-demand data acquisition, enabling improved monitoring of snow in heterogeneous environments. These capabilities make UAS particularly well suited for capturing highly dynamic snow conditions, such as ephemeral or transitional snow cover. However, their application may be constrained by weather conditions and demands from operational logistics (Sturm, 2015; Verfaillie et al., 2023).*"

Line 57-60: "… does not appear impacted to the same extent as satellite snow cover products in forested areas, complex terrain and shallow snow conditions …" => add brief explanation why there is less impact

Response:
Thank you for the suggestion. We have added a brief explanation clarifying why UAS-based snow mapping may be less impacted than coarse-resolution satellite snow products in heterogenous and shallow snow conditions. The primary reason is that satellite snow mapping errors often arise from (i) mixed pixels at coarse spatial resolution, (ii) reduced detectability when snow is thin or sparse. For example, Hall and Riggs (2007) note that omission errors are greatest when only trace amounts of snow are present and that snow/cloud confusion (including cloud-shadowed land) contributes to errors under thin/sparse snow conditions. In contrast, UAS imagery is acquired at centimeter-scale resolution and can therefore better resolve fine-scale snow heterogeneity (reducing mixed-pixel ambiguity). We revised the text accordingly to reflect these mechanisms more explicitly.

Since the present study does not target forested environments, forest-specific challenges associated with satellite SCA mapping were removed from the text.

"Additionally, UAS-based snow mapping may be less affected than coarse-resolution satellite snow products in shallow or patchy snow conditions because centimeter-scale imagery reduces mixed-pixel effects and allows better detection of fine-scale snow heterogeneity, whereas satellite products can exhibit increased omission errors when snow cover is thin or spatially discontinuous (Ault et al., 2006; Hall and Riggs, 2007; Gao et al., 2010; Frei et al., 2012)."

Line 164-192: You provide a lot of technical details for each drone/sensor system. Instead, why not add a table that summarises the technical differences between the systems, such as sensor specifications, camera settings and flight line parameters? You could then focus the text on the key differences between the systems. However, if you decide to stick with the current version, that is fine too.

Response:
After reviewing the technical details, the only major difference is the resolution and auto exposure determination between the two cameras. We plan to put a table of parameters in supplemental materials and reduce the text in the body of the manuscript.

• Line 176 & 189: Please specify which DEM was used, and whether the same DEM was used to process each dataset

Response:
The text has been updated for clarity as follows:
"A subset of the global Shuttle Radar Topography Mission (SRTM) 1 arc-second digital elevation model (DEM) was uploaded and referenced to ensure a consistent flight altitude was maintained throughout each flight."

• Line 208-209: "… most images collected during complete snow cover could not be effectively stitched together and were excluded from this study." => I would imagine that it is not a significant issue for the purposes of your study, as separating snow from bare ground in images with continuous, thick snow cover is likely easier than in cases with thin, fractional snow cover (as you demonstrate later). Nevertheless, for an actual drone mapping campaign to monitor snow cover evolution, this would present a significant challenge, as the creation of orthomosaics would fail for certain acquisition dates. At the very least, I would mention in the discussion that your test dataset is not representative of complete snow cover applications.

Response:
We appreciate the reviewer highlighting this point. The difficulty in generating orthomosaics during periods of complete snow cover arises from the lack of stable visual features required by structure-from-motion algorithms to reliably match tie points between images. As the reviewer notes, this issue is unlikely to affect the core objective of the present study, which focuses on classification performance under shallow and spatially heterogeneous snow conditions where distinguishing snow from bare ground is more challenging.

We agree that this limitation is important to acknowledge in the broader context of drone-based snow monitoring. To address this comment, we are adding text in the Discussion noting that the dataset used in this study is not representative of conditions with continuous, thick snow cover and that orthomosaic generation can become unreliable in such cases due to the absence of persistent surface features with citations to the structure-from-motion snow literature. We also note that this limitation primarily affects image reconstruction rather than snow classification itself and that future drone mapping campaigns targeting continuous snow cover may require alternative acquisition strategies (e.g., inclusion of ground control features, lower flight altitudes, or complementary sensors).

• Line 209: What approximate snow depths do you mean by "uniform snowpack"?

Response:
Good point, uniform implies spatially consistent. For clarity, the manuscript is updated as follows:

"Of the 36 dates, 10 represented very shallow/patchy snow conditions with observed snow depths less than 2 cm, 4 were completely snow free, and the remaining 22 dates had snowpacks with snow depths ranging from ~5 to ~45 cm within the open fields. The deepest snow depths were observed on surveys conducted during the 2024-2025 winter season. For the other winters, the average snow depths were ~10 cm.

• Line 220-222: "Field cameras were installed following the method used in NASA's 2020 SnowEx field camera campaign … " => You refer to personal communication here, but do not explain what was actually done. Could you please provide a description of the approach that was taken?

Response:

The appropriate citation is added to the manuscript along with additional details:

"*On dates when lidar was not flown, snow depth observations were acquired from either automated ultrasonic snow depth sensors (Judd Communications LLC) or from Moultrie Wingscapes Birdcam Pro Field Camera time lapse images of snow stakes (i.e., PVC poles spray-painted red and marked with 1 cm and 10 cm increments), both deployed in the open field at the respective site and sampled every 15-minutes. Field cameras were installed following the approach used in NASA's 2020 SnowEx field camera campaign in Grand Mesa, Colorado (see Proulx et al., 2022 for details).*"

The methodological details provided in the cited manuscript (Proulx et al., 2022) are as follows: ""Moultrie Wingscapes Birdcam Pro Field Cameras were used to capture images of the snowpack relative to a 1.5 meter marked PVC pole following the method used in NASA's 2020 SnowEx field camera campaign in Grand Mesa, CO (personal communication, 16 November 2020). Three cameras were used; one was in the open field, one was in the coniferous forest, and one was in the deciduous forest (Fig. 1). Each camera was mounted approximately 0.85 m above the ground and placed approximately 5.5 m from its respective PVC pole. Each camera's field of view included the entirety of the PVC pole, some of the ground surface below the pole, and some open area above the pole. Each PVC pole was spray-painted red and was marked with 1 cm and 10 cm increments. The cameras captured images of the poles every 15-minutes for the duration of the study period. Snow depth was derived by manual inspection of the photos and recorded to the nearest cm."

Proulx, H., Jacobs, J.M., Burakowski, E.A., Cho, E., Hunsaker, A.G., Sullivan, F.B., Palace, M. and Wagner, C., 2022. Comparison of in-situ snow depth measurements and impacts on validation of unpiloted aerial system Lidar over a mixed-use temperate forest landscape. The Cryosphere Discussions, 2022, pp.1-20.

• Figure 2 & 3 & 4: The images are relatively small, and it is not always easy to tell the difference between those with similar snow conditions. I suggest moving these figures to the supplementary material. In the main text, you could show a figure with fewer, larger panels instead. For example, select one or two images with distinctly different snow conditions for each site and each camera. This would allow readers to compare the differences between sites and cameras with and without a large snow cover fraction.

Response:

Figures 3 and 4 have been revised to display a smaller set of representative images at a larger size. The original versions of these figures have been moved to the supplementary material. Figure 4 will be revised to include an outline of the forested area analyzed in the discussion.

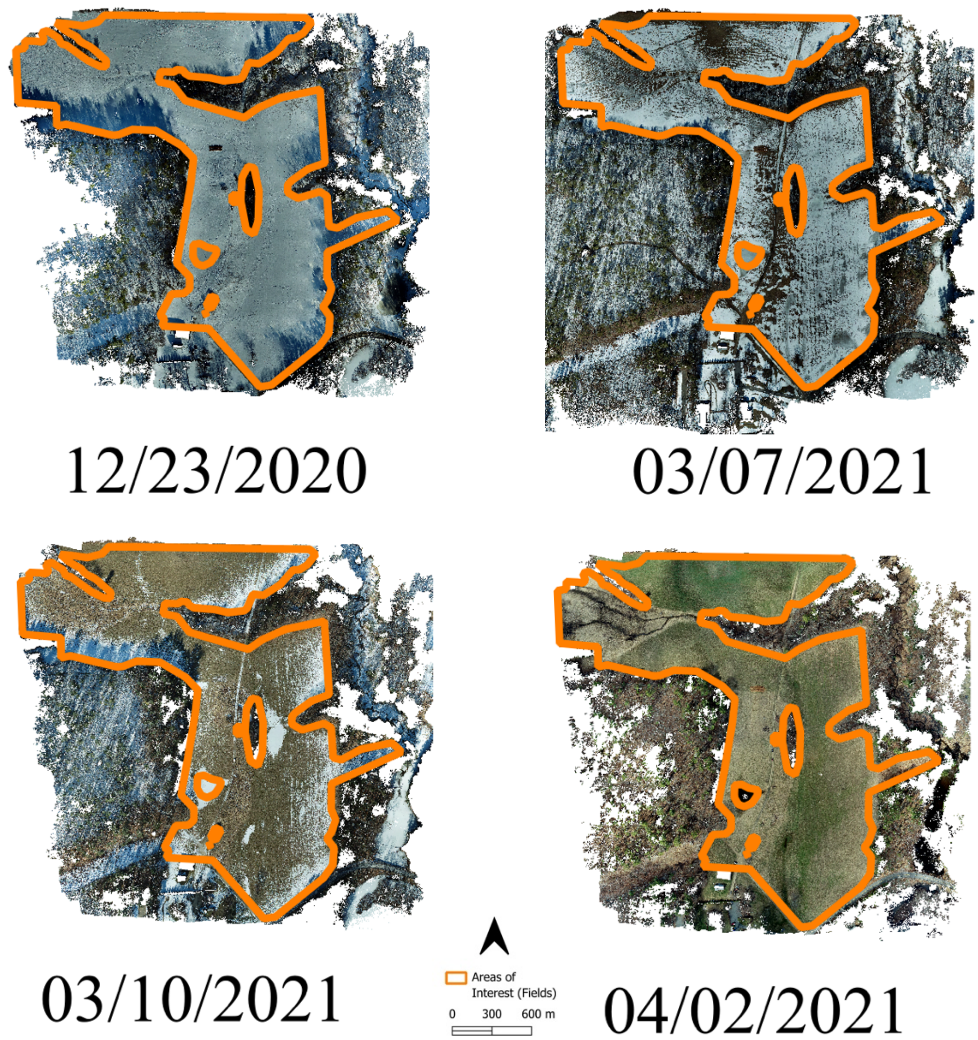


Figure 3. Example RGB orthomosaics of Thompson Farm acquired with the Phantom 4. Survey dates are shown below each orthomosaic. Areas of interest are outlined in orange.

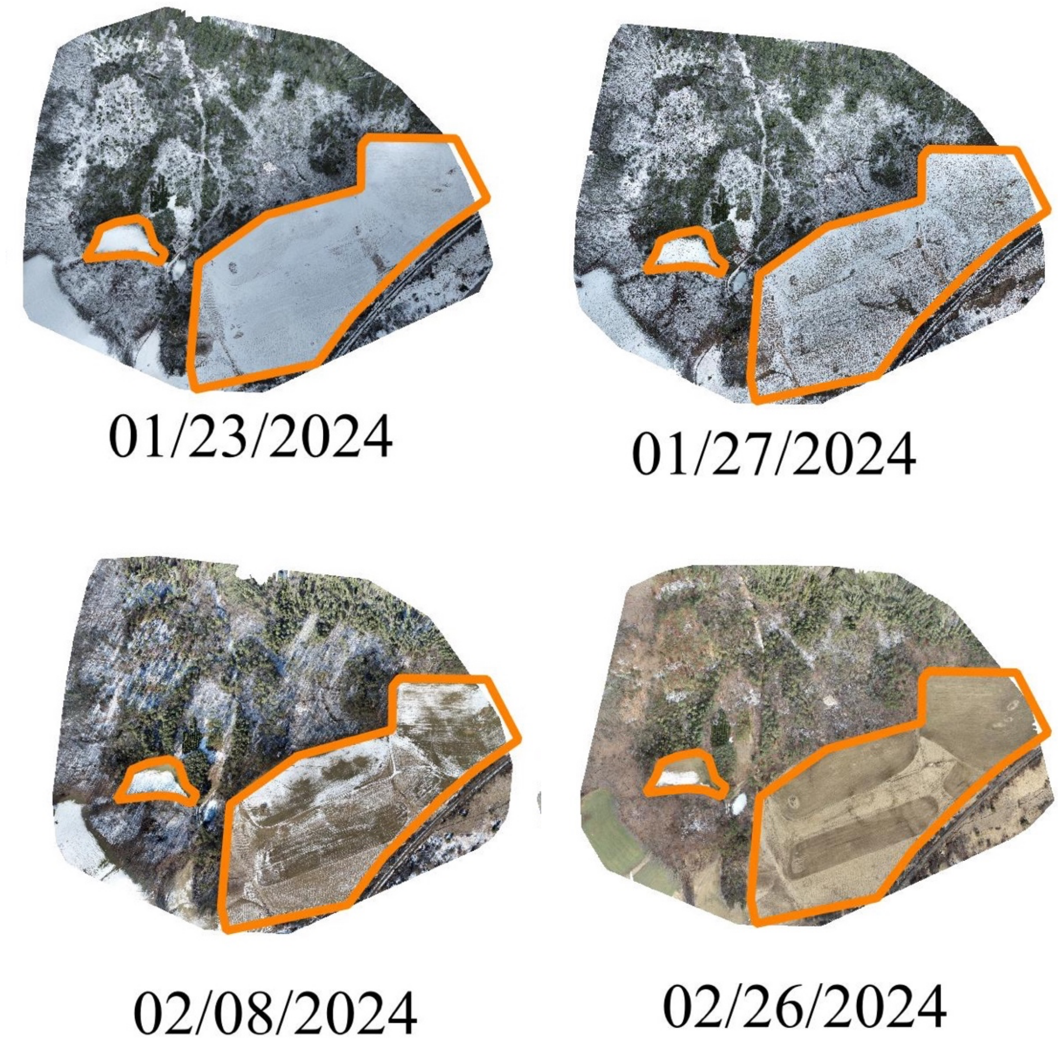


Figure 4. Example RGB orthomosaics of Kingman Farm acquired with the Phantom 4. Survey dates are shown below each orthomosaic. Areas of interest are outlined in orange.

Line 250: "… the field area of interest process was isolated … " => What do you mean?

Response:

The text has been revised for clarity:

"*For each orthomosaic, site-specific masks were applied to exclude surrounding forested areas, leaving only the field area for analysis*."

Line 251: "… forested areas was removed …" => see my general comment about selection of areas

Line 287-289: It is tedious for readers to have to look up the threshold value in the supplement. If the value is fixed, just mention it here and refer to the supplement for details.

Response:
A list of the tested thresholds has been added to the manuscript to improve readability:
"*The optimal threshold selection was guided by 5-fold cross-validation on a range of candidate thresholds (95, 105, 115, 127, 135, 145), using balanced accuracy (i.e., the average accuracy of both the positive and negative classes) as the evaluation criterion*"

Line 302: The results section is divided into subsections for each experiment. I suggest doing the same here, creating subsections for the 'base experiment', 'second experiment', and so on.

Response:
Subsections were added with a new naming convention (see the next comment).

Line 302: To improve readability, you could replace "base experiment", "second experiment", "third experiment", "fourth experiment" with a more (mathematically) name convention, e.g. SCA reference

Response:
Thank you for this suggestion. We changed the two experiments names to "Reference SCA" and "Held-out imagery SCA" both in the methods and results sections.

Line 404: "… is rooted in differences in snow conditions." => see my general remark about image dates and conditions

Response:
To address this concern, the cross-sensor experiment has been removed from the manuscript.

Line 440: In the conclusion you state that 12 images are the minimum requirement for a good training set (based on Figure 8?). Please elaborate why you conclude that 12 images is the threshold.

Response:
Twelve images were adopted as the minimum sample size in this study (Figures 8 and S1). Although the SVM classifier was generally more robust, maintaining good performance with as few as 10 images, the performance of the MLE and RF classifiers deteriorated across a larger number of image combinations when fewer than 12 images were used. In these cases, model behavior became less stable, as evidenced by an increased number of outliers with classification accuracy below 80%.

Line 521: "… identify one or more classifiers that can rapidly …" => Remove "rapidly" because you do not analyze processing speed.

Response:
Thank you for pointing this out. The word "rapidly" has been removed from the text for clarity.

Line 685-688: How do your results validate “satellite snow cover products, converting satellite NDSI observations of fSCA, and downscaling satellite SCA observations”? Explain this or delete.

Response:
Previous studies have used higher-resolution satellite snow maps (e.g., Landsat-derived snow cover or fSCA) as reference data to validate coarser-resolution snow products and to diagnose mixed-pixel behavior under patchy snow conditions. For example, Salomonson and Appel (2004) used Landsat ETM+ as higher-resolution reference information to derive the relationship between MODIS NDSI and fractional snow cover within 500-m pixels. Similarly, Rittger et al. (2013) evaluated MODIS snow-cover mapping methods using Landsat-based snow maps as reference. Additional work has validated MODIS fractional snow cover retrievals using temporally and spatially coincident Landsat scenes across multiple regions.

UAS-based optical imagery provides centimeter-to-decimeter spatial resolution, substantially finer than publicly available satellite imagery (e.g., Sentinel-2 at ~10 m) and even typical commercial optical imagery (~1–3 m). As a result, UAS snow-covered area (SCA) maps can serve as a more detailed reference dataset for validation and mixed-pixel analyses, particularly in landscapes where snow cover is highly heterogeneous and prone to patchiness. To date, UAS SCA maps have been less commonly used for satellite snow-cover downscaling than higher-resolution satellite products, but examples do exist (Johnston et al., 2025). Instead, several studies have primarily used UAS-derived snow maps to support MODIS-scale (500 m) fractional snow cover (fSCA) estimation by training machine-learning models with UAS observations as reference data (Liang et al., 2017; Liu et al., 2020). Our study offers a guidance on effective approaches to map SCA using UAS observations that potentially can be used for the above-mentioned applications.

Line 705: Data availability => Drone images and classification masks should be made available to reproduce results. Technical Corrections

Response:
Thank you for the comment. This is an extensive dataset that we would like to make available to the community. We have requested approval from our funding agency (DoD) to make the dataset publicly available. Unfortunately, the release of the dataset must remain pending until this approval is granted

Line 62: Here you could add a reference to large-scale snow cover products based on NDSI (e.g. https://essd.copernicus.org/articles/11/493/2019/)

Response: We added the Gascoin et al. (2019) reference.

Gascoin, S., Grizonnet, M., Bouchet, M., Salgues, G. and Hagolle, O., 2019. Theia Snow collection: High-resolution operational snow cover maps from Sentinel-2 and Landsat-8 data. *Earth System Science Data*, *11*(2), pp.493-514.

Line 67: replace „blue-channel thresholding“ with „blue-channel“ reflectance (or similar)

Response:
The text has been revised as suggested:

“*Promising methods for differentiating between snow and non-snow in UAS RGB imagery include machine learning (ML) classifiers such as maximum likelihood, random forest, support vector machine, k-means clustering (Liang et al., 2017, Belmonte et al., 2021, Johnston et al., 2025; Niedzielski et al., 2018), and thresholding approaches such as static or dynamic thresholding of blue-channel reflectance*.”

Line 159: There is no citation of MODIS data.

Response:
The appropriate reference has been added.

Line 226: missing reference NOAA's publicly available archive (ADD-REFERENCE)"

Response:
The appropriate reference has been added:

"*Subhourly air temperature and shortwave radiation data was obtained from the USCRN stations located at the study sites, through NOAA's publicly available archive (NOAA NCEI, 2025) Air temperature (°C) and incoming shortwave radiation (W/m²) from 2020 to 2025 at five-minute temporal resolution were used to calculate the mean and standard deviation during each survey*."

NOAA National Centers for Environmental Information (NCEI). U.S. Climate Reference Network (USCRN) Data. National Oceanic and Atmospheric Administration. https://www.ncei.noaa.gov/products/land-based-station/us-climate-reference-network; Accessed: April 01, 2025.

Figure 1: The panel on the bottom right appears to be a Google Earth background. I think there should be a reference to this somewhere.

Response:
The appropriate citation has been added:

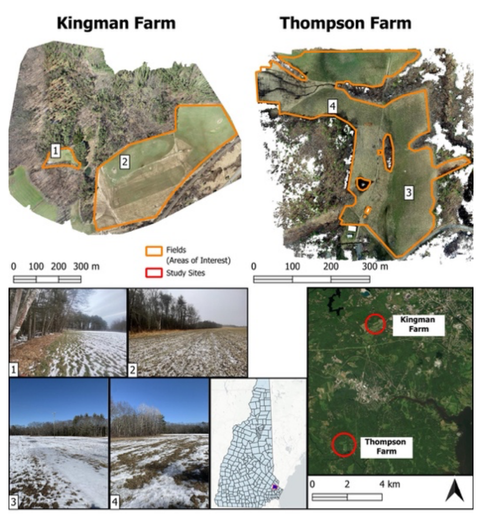


Figure 1. The two study sites, Kingman Farm and Thompson Farm, located at southern New Hampshire (basemap source: Esri, 2024).

Esri, 2024, . World Imagery [Basemap]. ArcGIS Online, https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer

Figure 5 and other bar charts: Do not include both red and green bars in the same graph.

Response:

The figure is updated with new set of colors