Reply to Reviewer 1

The study demonstrates the potential of assimilating L-band InSAR-derived snow depth retrievals to enhance snowpack simulations. The results showcase improved spatial and temporal resolution in capturing snowpack properties, leading to more accurate predictions of SWE/Snow Depth compared to open loop or control simulations. The paper is well-written and grounded in established scientific principles. However, some aspects of the methodology and analysis could benefit from further elaboration to strengthen the overall validity of the findings.

We are appreciative of the Reviewer's comments. Below, we address the Reviewer's specific comments (in bold blue).

Major comment

The study utilizes L-band InSAR data to estimate changes in snow depth, which necessitates knowledge of snow density. As I understand it, the authors propose using an average density derived from bias-corrected ASO data. After evaluating the error associated with different time periods, they opt to employ this information to calculate incremental snow depth. By combining absolute ASO information with incremental L-band data, the data assimilation process is transformed into a traditional data assimilation problem. However, Equation 3, based on Liens et al. (2015), derives incremental SWE without requiring snow density but necessitates calibration of the parameter alpha. In the present form, I am not totally sure about the rationale behind assimilating Δz instead of ΔSWE. A clearer explanation of this decision would help to avoid potential confusion.

The use of InSAR data to estimate changes in snow depth requires snow density to estimate the bulk snowpack permittivity (Eq. 1). We do not use the snow density from ASO data, which is only available for specific dates. For InSAR retrievals, as stated in Section 2.3, we use average bulk snow density between two repeat pass dates from MSHM reference runs.

Eq. 3 provides a linear relationship between SWE and InSAR phase change, but still requires the calibration of parameter alpha (Liens et al. 2015). Besides, InSAR retrieval using this method will only provide change in SWE and to convert it into absolute SWE, we will need prior SWE measurements or snow depth and snow density. In general, snow depth measurements are more readily available (e.g. lidar or ground measurements) and also since we evaluate the model results with snow depth measurements from the lidar and ground-based measurements, we employ Eq. 1 for the estimation of snow depth change in this study. We have added the following sentences in the revised manuscript for clarification:

Ln 234 – "Also, InSAR retrievals only provide change in snow depth or SWE, so even for SWE changes, one would require prior SWE measurements or snow depth and snow density to obtain absolute SWE for assimilation. In general, snow depth measurements are more readily available (e.g. lidar or ground measurements) and models with data assimilation can provide close estimate of snow density, so this study also provides a framework for using InSAR snow depth change for data assimilation."

Specific comments

L101: NISAR mission was already introduced at L80

Corrected.

L112: This sentence is unclear due to the lack of context about the dataset. Please provide more details or consider removing it from the introduction.

The sentence has been revised for clarification.

"The InSAR retrievals with the common first flight date (but different repeat pass) with airborne lidar measurements of snow depth were used for …"

L 228-230: for which snow density the value of 69 cm is valid?

This was for $\rho_s = 300\ kgm^{-3}$. It has been added to the sentence.

Section 2.4 the source of the density estimation used to derived the snow depth from the InSAR data has to be clearly stated.

For InSAR retrievals, as mentioned in Section 2.3, we use average bulk snow density between two repeat pass dates from MSHM reference runs.

Section 3 To enhance the clarity and conciseness of the results, consider summarizing the key findings in a table (now only in the text as numbers). This will make it easier for the reader to compare different scenarios and draw conclusions.

We have added the following tables in the revised manusript.

Table 2: Coherence for treeless and forested environment for different retrievals.

	Period (days)	Coherence (HH)	
		Treeless	Forested
Feb 12-19	7	0.71 ± 0.15	0.65 ± 0.18
Feb 19-26	7	0.6 ± 0.18	0.5 ± 0.2
Feb 1-12	11	0.48 ± 0.18	0.47 ± 0.19
Feb 26 – Mar 12	15	0.49 ± 0.17	0.39 ± 0.18
Feb 1-19	18	0.47 ± 0.18	0.43 ± 0.19
Feb 1-26	25	0.46 ± 0.18	0.43 ± 0.18
Feb 1- Mar 12	40	0.39 ± 0.16	0.36 ± 0.17

Table 3: InSAR retrievals of snow depth change with different polarizations for Feb 1-12

Polariation	Coherence	% missing	Avg. Change (cm)
VV	0.51	8	-1.13
HH	0.46	11	-0.1
HV	0.39	36	2.61
VH	0.39	54	-2.67

Table 4: Root mean square error (RMSE) for modeled snow depth (cm) over treeless environments with reference to pit (IOP and TSD) and snow pole measurements.

	Pit	W1A	W1B	W3A
CTRL	35.2	17.6	21.2	27.2
OL	36.0	11.57	14.80	30.8
DA	18.3	8.1	21	20.8
DAU	18.0	8.5	22.2	19.2

L448 please revise the sentence

The sentence was revised.

"The OL run shows similar tendency as InSAR retrievals,"

L492 To enhance clarity, please provide a more detailed explanation of how the density value was determined. This will help to avoid confusion and strengthen the overall understanding of the methodology.

In this study, the average bulk density (depth weighted density) between two repeat pass dates from the reference MSHM model runs were used for the InSAR retrievals.

Code availability: Some of the links provided in the code availability section are not functional. I strongly recommend making the data and code used to reproduce the results openly accessible. This includes model outputs and InSAR-derived snow depth data. Given that many of the results in this paper were made possible by the open access nature of the data used, it would be beneficial to maintain this level of transparency and openness.

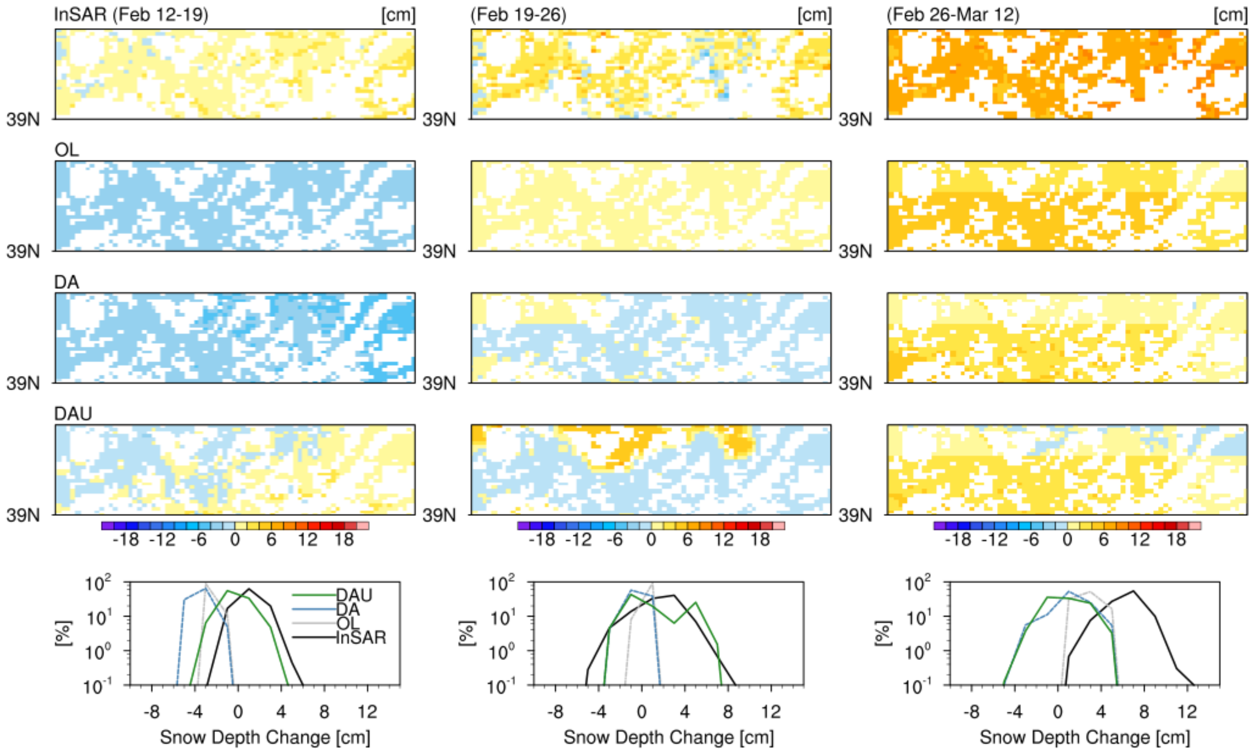
The data and code used to reproduce the results are openly accessible. The source codes for MPDAF will be made open with the publication of Shrestha and Barros (2024, WRR) which is still under review.

Figure 5 a The legend seems a bit unclear. Could you please provide more details or clarify the meaning of the different colors used?

The colorbar was choosen just to highlight the spatial heterogenity in coherence.

Fig 9 To better highlight the spatial variability, consider including a zoomed-in view of a specific region, similar to Figure 5.

The spatial variability shown in Fig. 5 is at 3-5 m resolution. In Fig. 9, the spatial variability is shown at 90 m resolution. The figure below shows a zoomed portion of Fig. 9 near the centre of the domain. At this scale, much of the small scale variability is filtered out, so a zoomed-in view does not add much information for the discussion.



InSAR (Feb 12-19)
[cm]
(Feb 19-26)
[cm]
(Feb 26-Mar 12)
[cm]
39N
OL
DA
DAU
-18 -12 -6 0 6 12 18
DAU
DA
OL
InSAR
[%]
Snow Depth Change [cm]

Reply to Reviewer 2

This paper presents, to my knowledge, the first example of data assimilation over a spatial domain with L-band InSAR snow retrievals, and is an important step forward toward the operational use of L-band InSAR snow products. This is of particular importance due to the upcoming launch of NISAR. This paper demonstrates the usefulness of assimilation of L-band InSAR snow depth retrievals, and shows how it improves the accuracy of the author's modeling results over flat topography on Grand Mesa, using data from the 2020 NASA SnowEx UAVSAR timeseries. The challenges with retrievals in forest are shown, pointing to more work needed in these environments. It would be worth mentioning more work is needed in complex topography too. This work is a logical next step in the use of L-band InSAR for improving snow estimation at large scale, and it is important that this work is published. There are however some major revisions I feel are needed before this work is ready for publication, as there are important details missing, many of which are outlined in detail with line-by-line comments below.

I have a few big picture concerns that I think are most important to address before publication, followed by a list of more minor things.

We are appreciative of the Reviewer's comments. We agree that more work is needed for retrievals in forested environments and complex topography, which we now mention explicitly in the Conclusion section. Below, we address the Reviewer's specific comments (in bold blue).

- The paper in general lacks details throughout on the snowpack model used, and some details on the data assimilation steps, how state variables are changed, etc. The paper "Shrestha and Barros, WRR, 2024" is often referenced, however this doesn't have a complete citation and I couldn't find it on the WRR site. Is this in review? If so, details in the WRR paper are needed for this TCD paper to be published. This is also really critical since the code used to do all of the modeling and DA in this paper is not currently available. There are many very widely used snowpack models that are open source and contain very detailed physics – how is this model different? This model appears to not have been used outside this research group; if it was made open to the community as it appears the authors intended, I think it would likely be useful for many other researchers. As an example of more detail: L272: "part of the data is used for assimilation, rest for evaluation". What is the split? Randomly selected? Or different regions? How exactly are the snow depth changes assimilated? What kind of error covariance model? Is mass changed? Is density changed? Just a few examples, more below.

The paper "Shrestha and Barros (2024, WRR)" is still under review . We have included additional detail about the snowpack model and data assimilation steps in the revised manuscript.

Ln 277 – "The snow microstructure evolution is simulated using a detailed microphysical scheme based on the CROCUS snowpack model (Vionnet et al., 2012). The bottom boundary conditions are kept constant during the cold season assuming frozen soils for snow-on

conditions and fixed deep soil temperature. Fresh snow density in this study is based on the parameterization of Hedstrom and Pomeroy (1998) and wet bulb temperature is used as threshold to partition precipitation as rain or snow (Wang et al., 2019). Currently, the rain versus snow partitioning only allows for existence of rain or snow and does not allow for mixed form of rain/snow."

MSHM is able to simulate macrophysical and microphysical properties of snowpack similar to other multilayer snowpack models. MPDAF which contains the interface to couple snow hydrology model with radiative transfer code and DART will be made freely available to the community after the publication of Shrestha and Barros (2024, WRR). MPDAF can be used with any snow hydrology model that produces the same snowpack stratigraphy as MSHM. The MSHM source code is available upon request sent directly to the second author (barros@illinois.edu). We are completing a User's Manual for DCHM (beta version already available) at which time MSHM (the snow component of DCHM – Duke Coupled surface-subsurface Hydrology Model) will be made publicly available along with DCHM. We welcome collaborators to work with us to test it before open release.

Out of 7 retrievals over Grand Mesa, we used three (out of 4) retrievals with the same first flight (Feb 1) to obtain absolute snow depth for data assimilation using the ASO lidar snow depth measurements. The fourth retrieval (40-day span repeat pass) had lots of missing data, so it was not used. Then we used the remaining 3 retrievals for evaluation. For assimilating snow depth change, we first convert them to absolute snow depth by adding the actual snow depth on the first flight date. We assigned an observational error of 10 % for the snow depth retrievals at 50 m resolution, which is consistent with the errors from the InSAR retrievals using the UAVSAR data in this study. Then we assimilate this snow depth using the EAKF filter in DART. The assimilation method is now discussed in more detail.

Ln 295 - "For ensemble Kalman filters, ensemble size fewer than 20-30 can lead to statistical error and larger ensemble size take longer to run with very little benefit. So, 30-100 ensemble size is recommended to use with ensemble Kalman filters in DART. In this study, we have only two DART state vectors and use 48 ensemble members which is also constrained by the computational and data storage requirements for hyper-resolution runs. In MSHM, vertically discretized snow depth and SWE are prognostic variables. However, we only use the integrated quantities of these prognostic variables (i.e. total SWE and total snow depth) as state vectors in DART, which are updated by the ensemble filters (e.g. EAKF used here). The ensemble filter assumes a Gaussian relation between the variables in the joint state space prior distribution. First, update increments are computed for each ensemble sample of the observation variable, which is then used to solve for the increments for each state variable. This requires prior cross covariance of each state variable with observation variable along with variance of the observation variable (see Anderson 2003 for details). Then, we use a newly developed repartition algorithm to distribute the increments in the vertical profile with mass conservation (snow density is updated). For bulk or single layer snow hydrology models, such repartioning is not needed. DA directly impacts the top layers of the snowpack where snow is added or removed based on assimilation increments. The lower layers are only impacted by the modeled snow evolution after the assimilation, primarily due to addition of new layers on top or removal of existing snow layers. The localization setups were used such that the observations only impact model grid points where they are located. The goal is to reduce the impact to surrounding grids with different landcover characteristics, and

therefore different retrieval uncertainty such as in the case of forested and non-forested grid points."

This work leverages multiple recently released open source software packages. Many of these packages were produced with great effort by students in the community, and were part of their research – please make sure the student authors of these packages you are using are explicitly given credit, not just a link to a repo (uavsar_pytools, herbie, etc). Zenodo references would help here. Many of these resulted from collaborations built at SnowEx Hackweek as well, which would be great to acknowledge maybe in the Acknowledgements section, along with mentioning the NASA SnowEx campaign there, in particular the field and airborne crews who collected all the data used in this paper. Especially given all the open source software and community collected data that this paper uses, I strongly encourage the authors to make their model and data assimilation software open, as it appears they intend to (although link currently broken and repo appears to be private).

We are thankful for the suggestion – we have now included the Zenodo references. We also mentioned the NASA SnowEx campaign, SnowEx Hackweek and the field and airborne crews in the acknowledgement section. The model and data assimilation software (MPDAF) will be made open source (see comments above).

- There are several uses of snow pit data, taken in different locations, that need to be interpreted with caution. For example, changes in snow depth between two pits in different locations, when it is on the order of less than 10cm, can easily be due just to spatial variability, even when only a few meters apart. Density changes between Feb 1 and Feb 12 from snowpit observations in different locations is likely due to spatial variability, not temporal change. I know this trend is also in the ASO modeled and bias-corrected densities, however those will show the same trend as the field data, due to the bias correction applied to ASO SWE.

We agree that the density change between Feb 1 and Feb 12 from snow pit observations could also be contributed by spatial variability alone, but the model also suggests that compaction increases snow density over the 11 day period, so we can not neglect the impact of ongoing compaction in the snow pit. The ASO SWE uses the bias corrected modeled snow density based on the pit data, so it has contribution from both compaction and spatial variabilty. We added the following sentence to highlight this caveat.

Ln 383 – "Here, we have to note that the snow density change in pit data could also be attributed due to spatial variability besides compaction, which could be contributing to higher differences."

- Density used in the L-band InSAR inversion. The authors show that their reference model has significantly higher densities than the field data, but still use it in the inversion – why isn't a field-measured density used here? This was also a really hard detail to find in the paper – consider stating that you are using reference model densities more clearly and earlier in the paper – i.e. in the section on the UAVSAR retrieval. Given the time series of depth and SWE from the nearby meteorological stations, it is unlikely there was much settlement of snow below the Feb 1 surface between Feb 1-12. It seems therefore more appropriate to use the density of the upper part of the snowpack for the inversion, since this is the part that changed and caused the phase change in the UAVSAR data. No modeled density profiles are shown, although it is mentioned that it was basically two layers. Consider showing a comparison with field-measured density profiles, to help the reader see that the modeled density profiles are reasonable.

Field density data are very limited and even in this study are not available for all the periods when InSAR retrieval is done. Also, for future NISAR mission, field measurements alone will not be able to provide the snow density for all grid locations. Thus, a strategy is need for global applications. We use modeled snow density obtained by running MSHM from the beginning of the 2020 water year continuously with atmospheric forcing from HRRR. HRRR is one of the Numerical Weather Prediction (NWP) models used by the National Weather Service in the US. Elsewhere, other NWP models can used to drive models like MSHM, and thus snow density estimates can be made available everywhere. We trust this is now more clearly explained in Section 2.3.

Ln 198 – "Field measurements of density data are generally sparse and may not be available for all the periods. Also, for future NISAR mission, field measurements alone will not be able to provide the snow density for all grid locations. So, we use modeled snow density (MSHM reference run) driven by atmospheric forcing from an operational numerical weather prediction (NWP), which can be applied everywhere to obtain the bulk snowpack permittivity."

Further, natural snowpacks are characterized by strong spatial variability and multi-layer vertical stratigraphy with varying snow density. The phase delay is an integral of the phase delay over these multiple layers which could change due to multiple physical processes depending on the temporal baseline of the repeat pass. This could vary from case to case as also suggested by the Reviewer for Feb 1-12 retrieval. So, for a generalized application, we use the average bulk snow density between the two repeat pass dates from MSHM reference runs. Below, we show the comparison between reference run (CTRL) and pit snow density during SnowEX'20 IOP for few pits. Note that HRRR underestimates snowfall and consequently MSHM generally underestimates snow depth. Also, despite spatial interpolation, the spatial resolution of HRRR over GM (3 km) is dominant and therefore there is a scale gap between MSHM effective resolution and the point-scale of individual pits. Finally, MSHM does not describe depth hoar formation, which is apparent at the bottom of all snow density profiles. A detailed analysis of MSHM simulation of snow density profiles is available in Cao and Barros (2020). One important difference is that in the former fresh snow density was assumed to have density of 30 kg/m^3 consistent with falling snow, whereas the density of fresh snow in the current application of the MSHM is following the empirical parametrization of Hedstrom and Pomeroy (1998) for fresh deposited snow, which tends to be higher than 60 kg/m^3, which should contribute to higher snow density at the top of the snowpack. Because the MSHM reference simulations for each pixel start in fall of 2019 and the model runs independently until February of 2020, any discrepancies are integrated over time in the lower and older layers of the snowpack independently of depth hoar processes. Nevertheless, with regular assimilation of snow data from the onset of snowpack, model can produce improved profile of snow density (Shrestha and Barros 2024, WRR).

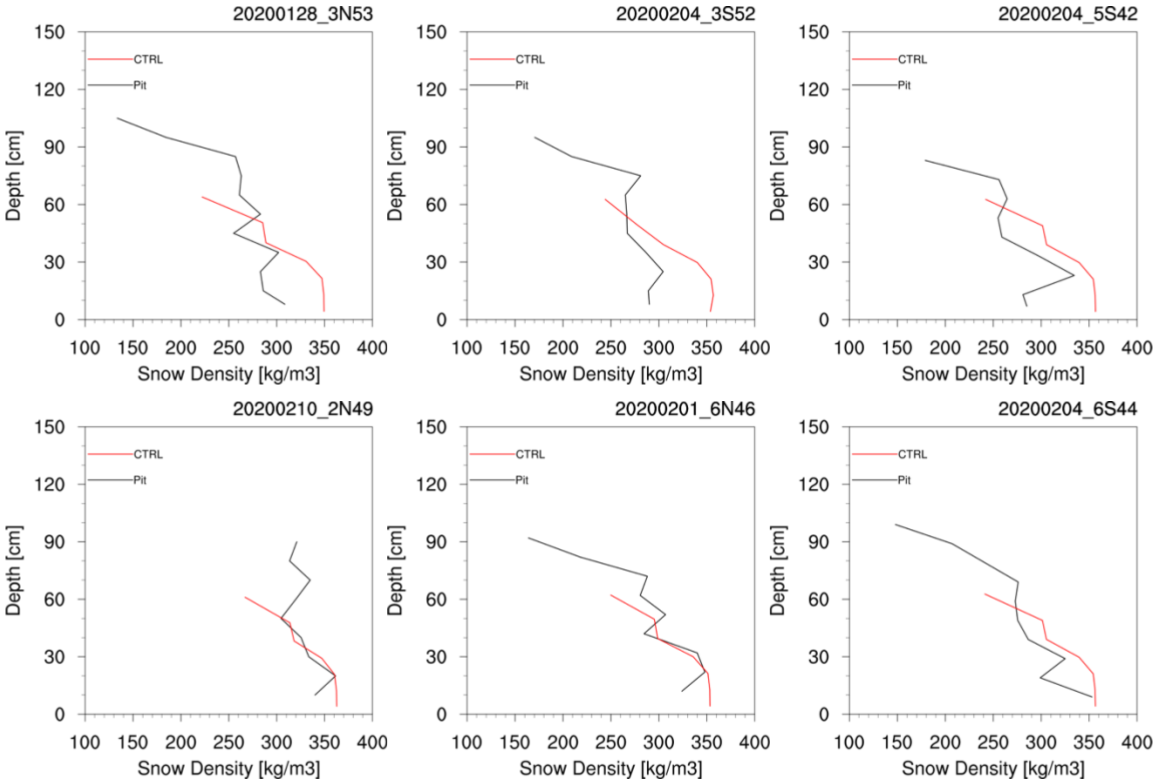


Figure R1: Vertical profile of modeled (CTRL -HRRR forcing) and measured snow density. For pit, average snow density between the two measurements is used.

- This paper shows the same comparison with the same datasets between L-band InSAR depth retrievals and repeat lidar for the western part of Grand Mesa, as shown in Marshall et al., 2021, but some additional details would help compare these two studies. Why was VV used for this comparison, but then HH used for the rest of the DA results (HH used in the previous work)? It would be better to evaluate the technique and then apply it with the same polarization. A scatterplot would be helpful in Fig 5, and more quantitative comparison (R-value, RMSE). The previous work found R=0.76, RMSE=4.7cm using the near surface field measured density observations in the retrieval – looks like similar results here, but hard to tell just from figure 5. How big of an impact did the atmospheric correction have? How exactly was this retrieval calibrated with the time-lapse snow pole depths? Which ones were used and where were they?

Higher coherence in VV polarization was one of the reasons for using it for the comparison with the study in Marshall et al., 2021. Since all polarizations were not available for other retrievals and HH also showed similar results to VV polarization (see Fig. 5d), HH was used for the remainder of the study. The ASO Lidar (3 m) and InSAR retrieval (approx. 5 m) are at similar resolution but differ in geolocations, so a quantitative spatial comparison which would require spatial interpolation was not used, but we rather explore the patterns and frequency distributions within the same extent. For the frequency distribution (Fig. 5d) between ASO and InSAR (VV), we find R=0.97 and RMSE=2.03 cm. This information was added to the manuscript.

The atmospheric correction was applied for all retrievals. The impact of atmospheric correction was minimal for this retrieval. With and without atmospheric correction, the average snow depth change for the InSAR retrievals (VV) were -1.28 cm and -1.13 cm respectively for the scene (Fig. 5).

The InSAR snow depth change for a 3x3 box (native resolution) was compared with snow pole measurements from treeless environment (W1A, W1B and W3A; linearly interpolated to the time of repeat pass flight) for calibration. The average calibration was obtained from the above three snow pole comparisons. For Feb 1-12 retrievals in HV and VH polarization (which had lots of missing data) snow pole measurements from W5A, W6B and W6C were used for calibration.

The snow pole measurements (W1A, W1A and W3A) used for calibration are also mentioned in Fig.1. And, we have added the following text in the revised manuscript:

Ln 257 - "The InSAR snow depth change for a 3x3 box (native resolution) was compared with snow Pole measurements from treeless environment (W1A, W1B and W3A; linearly interpolated to the time of repeat pass flight) for calibration, and the average value was used for the calibration. For Feb 1-12 retrievals in HV and VH polarization (which had lot of missing data) snow pole measurements from W5A, W6B and W6C were used for calibration. Further, we use VV polarization with higher coherence for spatial evaluation with lidar data, but since it was not available for all retrievals, and HH polarization also exhibit higher coherence and similar results to VV polariztion, HH is used for the InSAR retreivals and data assimilation."

Ln 431 – "The impact of atmospheric correction was minimal for this retrieval. With and without atmospheric correction, the average snow depth change for the InSAR retrieval (VV) were -1.28 cm and -1.13 cm respectively for the scene. The ASO Lidar and InSAR retrieval are in similar resolution but differ in geolocations, so a quantitative spatial comparison which would require spatial interpolation was not used, but we only explore the patterns and frequency distributions within the same extent. For the frequency distribution (Fig. 5d) between ASO and InSAR (VV), we find R=0.97 and RMSE=2.03 cm, and the previous study (Marshall et al. 2021) found R=0.76, RMSE=4.7cm using the near surface field measured density observations in the retrieval."

- I understand why you focus on InSAR snow depth change retrievals, rather than the more direct SWE change retrievals – so you can compare with the unique coincident repeat lidar discussed in 5) above. But once that is done, why not assimilate change in SWE? Then you remove the sensitivity to the density estimates.

Yes, we foucs on InSAR snow depth change for evaluation with Lidar and ground based measurements, including calibration as discussed in the manuscript. The method proposed by Liens et al., (2015) will still require an estimate of optimal α. Besides, InSAR retrieval using this method will only provide change in SWE and to convert it into absolute SWE, we will need prior SWE measurements or snow depth and snow density. In general, snow depth measurements are more readily available (e.g. lidar or ground measurements) and models

with data assimilation can provide close estimate of snow density, so we use InSAR retrievals of snow depth for this study.

We have added following text in the revised manuscript:

Ln 234 – "Also, InSAR retrievals only provide change in snow depth or SWE, so even for SWE changes, one would require prior SWE measurements or snow depth and snow density to obtain absolute SWE for assimilation. In general, snow depth measurements are more readily available (e.g. lidar or ground measurements) and models with data assimilation can provide close estimate of snow density, so this study also provides a framework for using InSAR snow depth change for data assimilation."

Detailed suggestions/comments

- The abstract states that the inversion was "modified", yet the common Guneriussen et al (2001) equation is used. The main difference I see is in the way the authors estimate the density and atmospheric delay, but the inversion seems the same? If not, maybe more details help here.

We wanted to emphasize the end-to-end methodology including density estimate and atmospheric delay besides the Guneriussen et al (2001) equation. In the revised manuscript, we have removed the term "modfiled".

- L76: refraction doesn't cause the phase delay. The phase is delayed because the radar wave moves slower through snow. The time of flight actually decreases due to the refraction, due to shorter path length, rather than delaying the signal.

Thank you for the correction. We rephrased the sentence.

Ln 76 – "..while slow propagation of radar waves through snowpack (depending on dielectric property) results in phase delay."

- L80: great place to mention the 2020 NASA SnowEx campaign

Added.

- Note that accurate density estimates are not required for SWE inversion, just depth retrievals as shown by Leinss.

Noted.

- L156: consider rewording. The domain was stratified into 9 classes based on tree density and snow depth. Reference SnowEx experimental plan so readers can find more details.

The sentence has been rephrased. Breen et al. (2022) includes the details about the snow pole measurements. For more details, as suggested by the reviewer, we have included the link to SnowEx experimental plan (https://snow.nasa.gov/sites/default/files/users/user354/SNEX-Campaigns/2020/NASASnowEx20_ExperimentPlan_v15.pdf).

- L175: Confusing. Are you using a single forecast or ensemble?

We use an ensemble run for data assimilation.

- L212: phase delay is an integral over phase changes in each layer, but total phase change is very similar whether you account for all layers or just use a bulk density – but again the appropriate density here is that from the upper part of the snowpack where the change in SWE/depth is occurring.

We agree with the reviewer, but in cases where snow microstrucre or snow density is also changing in deeper layers due to multiple physical processes at different temporal scales, it is important to account for this density variation also.

- L225: would be a good place to mention that the many SnowEx UAVSAR papers that have compared to lidar did the same approach here.

Added.

Ln 233 – "Earlier studies (e.g. Bonnell et al., 2024; Marshall et al., 2021; Palomaki and Sproles, 2023) have also used the same approach."

- L234-242: Calibration step here needs much more detail. Put the location of the calibration measurements (time lapse snow poles) on your maps. Did you use the average measured change for calibration? Or something more complex?

We have extended the details of calibration steps. We used average measured change for calibration. The location of snow pole measurements used for calibration are also highlighted in Fig.1. See comments above.

- L275: How did you choose magnitude of the perturbation (standard deviations in U)?

U stands for uniform distribution. The magnitudes were choosen based on the earlier study (Shrestha and Barros, 2024) that quantified precipitation variability of 40 % between two nearby SNOTEL stations in Grand Mesa and uncertainty of the daily total measurements in incoming SW and LW of ±5 % and ±10 % at the NASA SnowEx site respectively.

- L386: It is interesting that the authors report almost the exact same error as Marshall et al., 2021 with this same dataset (~5cm), but different data (timelapse snow poles) used for the evaluation – encouraging that this is consistent between studies and worth mentioning here.

We agree and mention it in the revised manuscript.

"(which is consistent with earlier findings by Marshall et al. 2021 for same InSAR retrieval but different data for evaluation)".

- L400: Forested areas showed HIGHER coherence? This probably a typo?

It was a mistaken interpretation; it has been removed from the revised manuscript.

- L430: difference in DA and DAU negligible at pits…could this be partly due to disturbed snow around pits?

Yes.

- L445-447: These results depend highly on the calibration approach, which was hard to evaluate due to a lack of detail. How different would your results be if you used the near-surface measured density? How big was the atmospheric impact? Was this applied for each UAVSAR flight? More calibration details would help reader interpretation here.

Please see reply above. Atmospheric correction impact is small in this case. Yes, it was applied for each flight. We trust details have now been added that provide a comprehensive description of the work.

- L459: how do you measure model improvement at the end of Feb/early March when there is no lidar? Is this with the pits or the time lapse snow poles?

Here, we evaluate model improvement in terms of snow depth change from the three InSAR retrievals that were not used for data assimilation.

- L493: More detail needed here – how are you doing depth-weighted averaging? And it says "or average between flights"? Did you do this differently in different situations?

The depth-weighted averaging here refers to the modeled multilayer snowpack with different thicknesses. The depth weighted average density is same as the bulk density of snowpack and we use average bulk density between two repeat pass dates. We used the same method for all retrievals.

- L553 – please fix link / make repo you are trying to share here public. This will be a huge service to the community, thank you!

The MPDAF source code will be made openly accessible with the publication of Shrestha and Barros (2024, WRR) paper.

- L591: Please acknowledge SnowEx participants, especially field observers and airborne teams. A shout out to Hackweek, which started collaborations between students and resulted in the open source software packages used in this paper, would be awesome here. We really want to keep Hackweek going so more software and data can be published for others to use! Thank you.

Yes, thank you for the recommendation. Done.

- It is stated that the UAVSAR retrievals do not agree with lidar in the dense forest, but does agree with the timelapse snow poles? Why is that? Does the lidar not agree with the snow poles in the dense forest?

In the forested environment, the pattern of snow depth change between the lidar and InSAR retrieval tend to differ (Feb 1-12;Fig. 5). For the snow pole measurements, InSAR retrievals shows agreement in most cases for snow depth change from multiple retrievals in different dates (Fig. 6). RMSE of the lidar measurements at the snow pole site used in the study was around 7-8 cm for the two dates exhibiting good accuracy. However, it must be noted that the lidar measurements were not available for other retrievals. More data are required for spatial comparison to better understand the differences in retrievals over forested environments. We have added following sentence in the discussion.

Ln 558 – "RMSE of the Lidar measurements at snow pole sites used in the study on Feb 1 was also around 7-8 cm for the two Lidar flight dates exhibiting good accuracy. So, more data are required for spatial comparison in forested environments to better understand the differences between lidar and InSAR retrievals."

- Fig 1: why does land cover classification only cover part of the domain?

The model domain is only the white box.

- Fig 3: consider putting the timing of the UAVSAR flights on this figure

The timing has been added.

- Fig 4: The locations of the snowpits used is needed on the map in Fig 1 for both dates

The snow pits are shown in the map now.

- Fig 5: add locations of the timelapse poles and snowpits used on this figure

This was added in Fig 1. In Fig. 5 spatial extent, only one snow pole station is present, so not shown.

- Fig 6: Are these validation sites? I thought these were used for calibration?

W1A, W1B and W3A were only used for calibration.

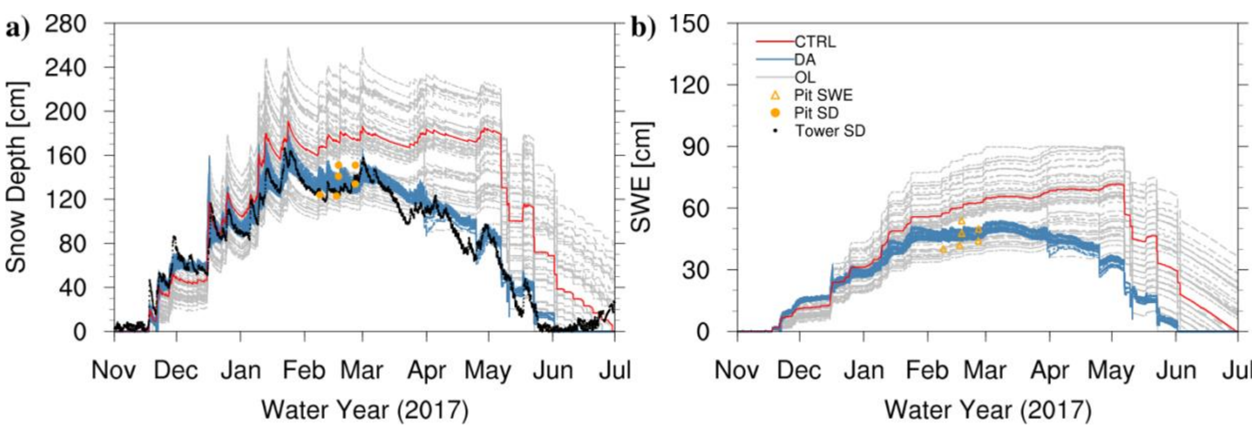
- Fig 7: see comment about interpreting snow pit depths at different locations in the context of temporal changes – some of these changes are less than expected spatial variability

These snow pits were within few meters of each other, but we have acknowledged the expected spatial variability as suggested by the reviewer in the text.

Ln 467 – "Here, we must note that the temporal variability in the snow depth is also partly contributed by the spatial variability due to change in location of pits."

- 8: Where is this comparison being done? Can you do it at the Mesa West weather station and show the measured snow depth time series? Show the location of Mesa West on your maps.

Here the comparison is being done for the spatially averaged model domain without including forested environments and lakes (white outline, Fig. 1). The idea is not to do point scale comparisons, but rather show the impact of assimilation of spatial data (lidar and UAVSAR) on modeled snowpack. We did such point comparison with tower snow depth measurements previously(source: Shrestha and Barros 2024a).



- Fig 9: Hard to see changes at this scale. Consider showing a smaller domain, and/or a full-page figure for this many panels.

With the shape of the domain along with the discussion in text (full domain), this is the optimal level for Fig. 9. The original "png" file of the image is scalable for zooming.

Excellent work here. This is a super important next step for combining L-band InSAR and modeling!

We are very thankful for the detailed review, which has helped to further improve the manuscript.