Dear Editor and Reviewers,

We thank the three reviewers for their careful reading of our manuscript and for their thoughtful and constructive comments. The feedback has helped us improve the clarity, organization, and scientific framing of the paper, particularly regarding the novelty of this study relative to Oveisgharan et al. (2024), the interpretation of the LIDAR and SNOTEL validation analyses, the role of temporal coherence and other environmental controls, and the limitations associated with sensing geometry, spatial scale, and uncertainty.

We have revised the manuscript accordingly and provide detailed, point-by-point responses below. All major revisions have been incorporated into the updated manuscript, and the responses refer to the revised version.

Sincerely,
The Authors

Please find our responses below.

RC1, Anonymous Referee #1

Accurate SWE estimate from satellite is critical for hydrological and climate modelling. InSAR SWE estimation technique as the potential to offer SWE at the required resolution (spatially and temporally) but further investigation is necessary. This works offer insight on the challenges of this technique using C-band data from Sentinel-1. Similar to the work in (Oveisgharan et al. 2024) Lidar and Snotel sites are used to evaluate the retrieval at several sites in the Western US. The paper would benefit from clearly stating what is the difference between this previous work. At first glance, it looks similar even though the analysis goes further in the parameter affecting the temporal coherence.

- We would like to thank the reviewer for great feedback. We address all of the concerns below. As seen below, we tried to distinguish and highlight the contribution of this work better.

I think focus on the soil background is missing on the analysis on the temporal coherence. It seems that the C-band decorrelation would highly be link to changes in the soil background (Tampuu et al. 2020, Ouaadi et al. 2024) since the soil is probably not frozen. A 6-days revisit seems to also be crucial in coherency (Ouaadi et al. 2024). This not good for NISAR since the revisit is 12 days. However, L-band seems to be less sensitive to decorrelation from the ground, and a 12-day revisit is probably fine. I suggest discussing this because someone could easily conclude that this method won’t be successful with NISAR because of the revisit time.

- Thank you for bringing up these concerns. We discuss about the 6-day vs 12-day for NISAR as you can see below in your last note. We added this to the paper to address your concern: "In contrast, longer wavelengths such as L-band are intrinsically less sensitive to small-scale scattering changes and therefore maintain higher temporal coherence over longer revisit intervals. As a result, missions such as NISAR are expected to enable robust SWE retrieval despite a 12-day repeat cycle, while also benefiting from reduced phase ambiguity."

- However, soil moisture's impact should be minimal. When snow is wet the microwave signal barely penetrates all the way to the ground. Therefore, most of the decorrelation should come from wet snow itself. We added this to the text to mention that soil permittivity change due to soil moisture also impacts temporal coherence but since there has been no indication of such in the snow accumulation season in the literature, we leave that discussion out. "In practice, however, temporal variations in scatterers—such as freeze/thaw, soil permittivity change, the movement of leaves, branches, or snow particles, reduce this correlation, leading to temporal decorrelation (Zebker and Villasenor, 1992; Kellndorfer et al., 2022; Lavalle et al., 2012; Tampuu et al., 2020; Ouaadi et al., 2024)."

Specific comments:

L11: The main driver is temporal coherence followed by temperature, wetness… Those factors affect the coherence as well. So it's still the same driver? Maybe rephrase.

- They impact coherence too but not in the same way. So, we first mentioned the parameters impacting retrieval and then the ones impacting temporal coherence. I added this sentence to make it clearer "*The influence of these parameters on temporal coherence is only partially consistent with their effect on SWE retrieval performance.*"

L22: "shifting snow accumulation toward higher". Is it really shifting the accumulation or just lowering accumulation down south?

- As stated in Klos et al. "As demonstrated through the mapped temperature-precipitation phase relationship across the western U.S., the climatic rain-snow transition zone will move up in altitude and latitude. The western U.S. is projected to see an average monthly reduction from ~53% to ~24% in the extent of the land area within a wintertime snowfall regime". Therefore, it seems that these regions are shifting.

L25: Suggest modified to "if the snowpack completely melts"

- Done

L31: I think the resolution of GlobSnow is still 25km.

- I changed it to "(tens of km)" to include the improvement reported in Kelly 2009 paper using AMSR-E data with 10km resolution.

L48: If the ratio picks up volume scattering, an incoherent process... then this would contribute to losing coherency faster at C-band than L-band due to volume scattering. I don't think it's the case, but I think it needs to be mentioned at some point.

- We think this method needs more validation. Therefore, we just mentioned some studies evaluating it. Getting into details of this method and what we think should impact it is beyond the scope of this paper.

L55: I would specify microstructure here. Stratigraphy (permittivity changes) is also "layering" which is what you want here.

- We changed it to: "key advantage of the interferometric approach is that SWE retrieval is largely insensitive to snowpack stratigraphy, permittivity changes, and layering (Yueh et al., 2017), and does not require prior knowledge of snow micro-structural properties."

L60: How is this different than Oveisgharan et al., 2024? Stating the difference here would be nice because it looks really similar at first.

- Very fair point, I address this in several places in the paper to make it clear the main contribution of the paper
    - In the abstract we added: "While previous work demonstrated the feasibility of SWE retrieval using Sentinel-1 interferometry over limited sites, this study provides a systematic, multi-site evaluation across diverse snow and land-cover conditions to identify the key factors controlling retrieval performance."
    - At the end of introduction "While (Oveisgharan et al., 2024) established the feasibility of SWE retrieval using Sentinel-1 interferometry and demonstrated its performance over a limited number of sites, this study extends that work by providing a systematic, multi-site evaluation across diverse snow and land-cover conditions. In particular, we investigate the environmental and geophysical factors controlling retrieval performance, including temporal coherence, temperature, vegetation, and terrain, and validate the approach using a significantly larger set of LIDAR and in situ observations. This analysis provides new insight into where and under what conditions InSAR-based SWE retrieval is reliable, moving beyond proof-of-concept toward broader applicability. The methodology developed here is also directly applicable to the recently launched L- and S-band NASA–ISRO's NISAR mission."
    - At the end of section 2 (before 2.1): "The main contributions of this study are: (1) a comprehensive multi-site validation of InSAR-based SWE retrieval using an expanded set of LIDAR and in situ observations, (2) a systematic analysis of environmental and geophysical factors controlling retrieval performance, and (3) identification of the conditions under which reliable SWE retrieval is achievable using C-band SAR data."
    - Beginning of section 4: "Unlike (Oveisgharan et al., 2024), which focused on demonstrating the feasibility of SWE retrieval using Sentinel-1 interferometry over a limited number of sites, this section evaluates retrieval performance across a broader range of snow conditions and land-cover types to identify the key factors controlling retrieval accuracy."
    - Beginning of section 5: "While previous work (Oveisgharan et al., 2024) established the capability of InSAR to retrieve SWE, it did not systematically assess the

environmental and geophysical controls on retrieval performance. Here, we extend that work by analyzing how factors such as temporal coherence, temperature, vegetation, and terrain influence retrieval accuracy across multiple sites."

  - At the beginning of section 5.1: "This analysis builds on the limited LIDAR validation presented in (Oveisgharan et al., 2024) by incorporating a larger number of LIDAR scenes across diverse environments, enabling a more comprehensive assessment of retrieval performance and its controlling factors."
  - At the beginning of section 5.2: "In contrast to the site-specific validation in (Oveisgharan et al., 2024), this section leverages multi-site SNOTEL time series to investigate how environmental variability affects SWE retrieval performance and temporal coherence."
  - End of the section 5: "This study moves beyond demonstrating the feasibility of InSAR-based SWE retrieval by providing a systematic evaluation of its performance across varying environmental conditions. The results identify the dominant controls on retrieval accuracy and define the conditions under which the method is reliable, which are critical steps toward operational SWE monitoring using current and future SAR missions."
  - At the beginning of the conclusion section: "While previous work demonstrated the feasibility of SWE retrieval using Sentinel-1 interferometry, this study provides the first comprehensive, multi-site evaluation of retrieval performance and its controlling factors. By quantifying the roles of temporal coherence, temperature, vegetation, and terrain, we establish the conditions under which InSAR-based SWE retrieval is reliable, representing a key step toward large-scale and operational applications. These results also provide a foundation for quality assurance (QA) in future SWE products by identifying the key parameters that control retrieval reliability and uncertainty."

L72: True at C-band and L-band. I would be precise here

- Done: "For terrestrial snow, the contribution of volume scattering to the interferometric phase is minimal compared to the strong ground return at high radar frequencies such as C- and L-band."

L76: Suggest changing "dual-polarization dual-frequency retrieval approaches" to "intensity based-approaches"

- Done

L124: freeze/thaw or soil permittivity change of the ground could change the backscatter response.

- I added this: "In practice, however, temporal variations in scatterers—such as freeze/thaw, soil permittivity change, the movement of leaves, branches, or snow particles, reduce this correlation, leading to temporal decorrelation (Zebker and Villasenor, 1992; Kellndorfer et al., 2022; Lavalle et al., 2012; Tampuu et al., 2020; Ouaadi et al., 2024)."

L129: Can you explain why?

- I added this to explain why: “This is because higher frequencies (shorter wavelengths) are more sensitive to small-scale variations in the scattering medium, such as snow microstructure and vegetation elements. These small scatterers introduce phase variations within a resolution cell that change between acquisitions, leading to increased decorrelation.”

L135: Would be nice to mention quickly how temporal coherence is calculated even if its just a reference.

- We added this to the beginning of section 2.2: “The correlation between two signals is computed as the normalized cross-correlation of two complex SAR images (Oveisgharan and Zebker, 2007).”

L163: Are comparing the depth for lidar and InSAR at every 5m for each site?

- Good point, we added this: “To enable direct comparison with Sentinel-1–derived SWE, the LIDAR data were resampled to 80 m resolution.”

Figure 10: It seems like soil moisture content would highly impact the temporal coherence at C-band. This effect seems to be less of a problem at L-band.

- We added this to the paper to address your concern: “While soil moisture can influence temporal coherence, its impact is expected to be limited during the snow season, as the ground is typically frozen under dry-snow conditions. In situations where the ground is not frozen, the snowpack is likely to be wet, violating the assumptions required for InSAR-based SWE retrieval. Therefore, the effect of soil moisture on temporal coherence is considered negligible for the dry-snow conditions analyzed in this study.”

L357: “increase particle motion” Do you mean water percolation?

- We meant snow particle motions. When the snow becomes wet, snow particles start to move or slide. We add “increase snow particle motion”.

L374: NISAR is a 12-day repeat cycle, and you just said its bad for coherence…

- We added this to the paper to address your concern: “In contrast, longer wavelengths such as L-band are intrinsically less sensitive to small-scale scattering changes and therefore maintain higher temporal coherence over longer revisit intervals. As a result, missions such as NISAR are expected to enable robust SWE retrieval despite a 12-day repeat cycle, while also benefiting from reduced phase ambiguity.”

Reference

Oveisgharan, S., Zinke., R., Hoppinen, Z. and, Marshall, H. P. Snow water equivalent retrieval over Idaho– Part 1: Using Sentinel-1 repeat-pass interferometry. The Cryosphere, https://doi.org/10.5194/tc-18-559-2024

Tampuu, T., Praks, J., Uiboupin, R., & Kull, A. (2020). Long Term Interferometric Temporal Coherence and DInSAR Phase in Northern Peatlands. Remote Sensing, 12(10), 1566. https://doi.org/10.3390/rs12101566

Nadia Ouaadi, Lionel Jarlan, Ludovic Villard, Adnane Chakir, Saïd Khabba, Pascal Fanise, Mohamed Kasbani, Zoubair Rafi, Valerie Le Dantec, Jamal Ezzahar, Pierre-Louis Frison. Temporal decorrelation of C-band radar data over wheat in a semi-arid area using sub-daily tower-based observations, Remote Sensing of Environment, Volume 304, 2024, https://doi.org/10.1016/j.rse.2024.114059

RC2, Anonymous Referee #2

This manuscript evaluates a Sentinel-1 repeat-pass InSAR approach for retrieving changes in snow water equivalent (ΔSWE) and then integrating those changes to estimate "total SWE" over several western U.S. SnowEx sites. The authors compare Sentinel-1 retrievals to two independent data sources: airborne LiDAR snow depth maps (SnowEx QSI and ASO) and in situ SNOTEL observations (SWE, snow depth, and near-surface air temperature). They report that 6-day Sentinel-1 acquisitions show moderate agreement with LiDAR snow depth patterns (site-level correlations ~0.42–0.66), while 12-day acquisitions perform poorly due to temporal decorrelation and phase ambiguity. Comparisons of Sentinel-1 ΔSWE with SNOTEL ΔSWE over selected frames yield correlations up to ~0.81 with sub-cm RMSE, and the paper further analyzes how coherence, temperature, snow conditions, and terrain/vegetation characteristics influence retrieval performance.

[Major comments]

Major concern: comparability and representativeness of the validation comparisons.

The paper’s validation strategy would benefit from a more thorough treatment of what each measurement/product represents and how differences in sensing geometry and scale affect "apples-to-apples" comparison. The study combines measurements from different sensors/platforms (Sentinel-1 InSAR, airborne LiDAR, and point-scale SNOTEL), but it does not sufficiently describe their measurement characteristics (support/footprint, geometry, and physical quantity observed) nor how these differences are accounted for in the evaluation.

- We expanded on this with couple of paragraphs in section 4.1 (marked in red) to address the resolution, angle, swe vs snow depth, … .

1. Geometry and slope effects (nadir vs slant path):

Over flat terrain, geometric discrepancies are limited, but in mountainous terrain the differences can be systematic. Airborne LiDAR (nadir-looking) measures snow depth approximately along the z-axis (distance between a snow-free reference surface and snow surface). In contrast, the InSAR retrieval used here is based on interferometric phase delay and is interpreted as ΔSWE under dry-snow assumptions (Eq. 1). Because radar propagation is along the *slant path* (satellite line of sight) and interacts with terrain slopes and viewing geometry, it is uncertain whether LiDAR depth patterns and InSAR-derived ΔSWE/SWE patterns should be expected to match without explicitly treating slope/local incidence effects and the definition of "effective path length" through snow. As slope increases, the observed quantities may diverge even if the underlying snowpack is the same.

- We added this to section 4.1 to address your concern: "Terrain slope in complex topography can affect both LIDAR measurements and InSAR-derived SWE. For LIDAR, steep slopes may introduce sampling biases and noise in snow depth estimation, while for InSAR, slope influences the local incidence angle and the effective radar propagation path through the snowpack, thereby affecting the interferometric phase and the retrieved SWE. In addition, steep terrain can introduce radar-specific effects such as layover and shadow, which further degrade signal quality and coherence and can impact retrieval accuracy. Because the radar footprint is significantly larger than that of LIDAR, these effects can be more pronounced in the InSAR observations. A more rigorous treatment of slope effects, including explicit modeling of radar–snow interaction geometry, is beyond the scope of this study. However, the impact of slope is partially addressed by including terrain slope as one of the parameters in the retrieval performance analysis (Section 5.1). Its influence is also evident in the comparison between ascending and descending acquisitions discussed in Section 4.1.2. These factors should be considered when interpreting validation results, as some observed differences or dependencies may arise from mismatches in measurement geometry and resolution rather than from retrieval performance alone."

2. SWE vs snow depth (density variability):

A central issue is that the manuscript compares retrieved SWE (or integrated SWE proxy) to LiDAR snow depth and reports correlations (Section 4.1). Depth and SWE are not directly comparable without accounting for snow density variability. This needs explicit discussion and, ideally, an uncertainty treatment or conversion approach (e.g., using density estimates, or restricting validation to periods/areas with representative density constraints). Since SNOTEL provides both SWE and snow depth, the authors could leverage co-located time series to at least characterize plausible density ranges/variability and how that affects depth–SWE comparisons.

- We added section 4.1.4, "Evaluating the Retrieved Density", to address the discussion of density derived from comparing SWE and snow depth. Figure 9 addresses the reviewer concern.
- We also added 2 sections, (4.1.5) and (4.2.1), on uncertainty analysis for LIDAR and SNOTEL comparisons.

3. Spatial representativeness and scaling:

The paper compares (i) pixel-scale satellite retrievals, (ii) very high-resolution LiDAR maps, and (iii) point-scale SNOTEL measurements. The manuscript should state clearly how LiDAR is resampled/aggregated to Sentinel-1 resolution and how point measurements are mapped to radar pixels (or whether a neighborhood averaging strategy is used). Without this, the reported correlations may partly reflect representativeness mismatch rather than retrieval skill.

Given these comparability issues (geometry/slope, quantity mismatch, density variability, and representativeness), I recommend major revision: clearly define what each dataset measures, how it is made comparable (geometry and scale), and how these factors influence the interpretation of agreement/disagreement.

- We added detailed explanation to sections 4.1 and 4.2 to address the differences in the three data sets (all marked red).

[Specific comments]

1. Equation (1): definition of incidence angle θ in mountainous terrain.

How is θ defined for each pixel? Is it the sensor incidence angle relative to the ellipsoid, or a local incidence angle accounting for terrain slope and aspect? If slope is not included, then the effective angle between the local surface normal (or relevant propagation direction through the snowpack) and the satellite viewing vector changes with terrain, and this could affect both the phase-to-ΔSWE scaling and differences between ascending/descending geometries. Please clarify and, if needed, quantify sensitivity.

- We have clarified the definition of θ in the manuscript. Specifically, θ represents the **local incidence angle**, which accounts for terrain slope and aspect, rather than the incidence angle relative to a reference ellipsoid. This angle is computed using the radar viewing geometry and the local surface normal derived from topography.
- By using the local incidence angle, we account for variations in terrain that affect the effective interaction between the radar signal and the snowpack. This reduces potential biases in the phase-to-ΔSWE scaling and helps explain differences observed between ascending and descending geometries. We have added this clarification to the manuscript to avoid ambiguity.

2. L156 (6 a.m. vs 6 p.m. performance)

The manuscript states that retrieval works better for 6 a.m. acquisitions than 6 p.m. and therefore focuses on descending data. Please provide a physical explanation (e.g., observation geometry, diurnal melt/freeze state, wet snow occurrence, near-surface temperature timing, changes in coherence) and show supporting evidence (e.g., coherence/temperature statistics split by acquisition time).

- We thank the reviewer for this valuable suggestion. In response, we added a new section comparing SWE retrievals using 6 am (descending) and 6 pm (ascending) Sentinel-1 acquisitions over Banner Summit and Mores Creek, including comparisons with LIDAR data. We also provide a physical interpretation of the observed differences. Specifically, we show that improved performance in the morning is primarily associated with colder and drier snow conditions, leading to higher temporal coherence, while differences in viewing geometry further influence spatial variations in coherence and retrieval accuracy. To support this, we included temporal coherence comparisons between 6 am and 6 pm acquisitions and discussed their relationship with environmental conditions and imaging geometry. We believe this addition strengthens the manuscript and clarifies the underlying mechanisms controlling retrieval performance.

3. Figure 1 color choice (accessibility).

Using green and red together is not ideal for color-vision deficiency. Please consider an alternative color palette or add different line styles/patterns for frames and LiDAR extents.

- We used blue instead of green. Blue and red make it color-vision efficient

4. Figure 6: clarify red lines in panels d1 and d2.

In Figure 6, what do the red line segments represent in (d1) and (d2)? If they mark a subset of stations, special periods, thresholds (e.g., >0°C), or data gaps, please add to caption/legend. As-is it is ambiguous.

- We explained in the text and the caption that "In panels (d1) and (d2), the red and blue curves represent two different SNOTEL stations within the scene, whereas only a single station (blue curve) is available in the other panels."

5. L293: define variables used in parameter analysis.

Please define "vegetation height," "ground topography," "slope," and "aspect" explicitly: data source, how computed (e.g., DEM), spatial resolution, and whether they are averaged/binned at Sentinel-1 pixel scale. Without definitions, the discussion is difficult to follow.

- We added this to the text to clarify: "Vegetation height is derived from the QSI LIDAR dataset, which provides co-registered measurements of vegetation structure and snow depth. Elevation is obtained from the Copernicus DEM (GLO-30), used within the ASF HyP3 processing workflow, with a spatial resolution of 30 m. Terrain slope and aspect are calculated from this DEM. All parameters are subsequently aggregated to match the ~80 m spatial resolution of the multilooked Sentinel-1 data used in this analysis."

6. L311: definition of temporal coherence.

Temporal coherence is introduced earlier, but please provide a concise definition where it becomes central to the analysis (e.g., complex correlation magnitude between acquisitions; whether it's VV/VH; whether it's averaged over a time window; and any masking). This will help readers interpret Figures 8-13.

- I added some explanation to section 2.2 in red to address the reviewer's concern.

7. References formatting.

The reference list is comprehensive and relevant, but there are numerous formatting errors (incorrect journal titles, missing article numbers, duplicate entries, corrupted author names). Please revise the bibliography carefully and standardize citation format.

- We thank the reviewer for this careful observation. We have thoroughly revised the reference list to correct formatting issues, including journal names, author names, capitalization, and missing publication details. Several entries contained typographical errors introduced during reference management export, which have now been corrected. We also standardized journal naming conventions and formatting across all references to ensure consistency.

RC3, Anonymous Referee #3

This study evaluates Sentinel-1 repeat-pass InSAR retrieval of snow water equivalent (SWE) using comparisons with airborne LIDAR snow depth and SNOTEL observations across several SnowEx sites. The results suggest that 6-day repeat acquisitions can recover spatial SWE patterns reasonably well at some sites, while performance degrades substantially for 12-day revisits. The discussion on temporal coherence, temperature, and SWE variability is also potentially valuable for future operational applications.

However, I have several concerns regarding the novelty, the clarity of the methodology, and the interpretation of the validation results. In particular, I am not fully convinced that the manuscript sufficiently distinguishes its contribution from the authors' previous work, and several of the methodological assumptions require stronger justification. I therefore recommend that the paper undergo major revisions before it can be considered for publication.

1. I am not sure the manuscript presents sufficient methodological novelty relative to the authors' earlier work. The SWE retrieval framework, phase-to-ΔSWE conversion, and even the reference-point calibration strategy appear to follow Oveisgharan et al. (2024) quite closely. As it stands, it is not very clear what is fundamentally new here beyond applying the same framework to more sites and comparing with LIDAR and SNOTEL datasets. Is the novelty primarily the broader evaluation across multiple SnowEx sites? Is it the comparison against LIDAR snow depth and SNOTEL? Or is there a methodological improvement that is not sufficiently highlighted? At present, the paper reads more like an application and extension of an existing framework than a clearly new methodological contribution. The authors need to state much more clearly what is fundamentally new in this paper and why this new step is important enough to warrant a separate publication.

- Very fair point, I address this in several places in the paper to make it clear the main contribution of the paper
    - In the abstract we added: "While previous work demonstrated the feasibility of SWE retrieval using Sentinel-1 interferometry over limited sites, this study provides a systematic, multi-site evaluation across diverse snow and land-cover conditions to identify the key factors controlling retrieval performance."
    - At the end of introduction "While (Oveisgharan et al., 2024) established the feasibility of SWE retrieval using Sentinel-1 interferometry and demonstrated its performance over a limited number of sites, this study extends that work by providing a systematic, multi-site evaluation across diverse snow and land-cover conditions. In particular, we investigate the environmental and geophysical factors controlling

retrieval performance, including temporal coherence, temperature, vegetation, and terrain, and validate the approach using a significantly larger set of LIDAR and in situ observations. This analysis provides new insight into where and under what conditions InSAR-based SWE retrieval is reliable, moving beyond proof-of-concept toward broader applicability. The methodology developed here is also directly applicable to the recently launched L- and S-band NASA–ISRO's NISAR mission."

- At the end of section 2 (before 2.1): "The main contributions of this study are: (1) a comprehensive multi-site validation of InSAR-based SWE retrieval using an expanded set of LIDAR and in situ observations, (2) a systematic analysis of environmental and geophysical factors controlling retrieval performance, and (3) identification of the conditions under which reliable SWE retrieval is achievable using C-band SAR data."
- Beginning of section 4: "Unlike (Oveisgharan et al., 2024), which focused on demonstrating the feasibility of SWE retrieval using Sentinel-1 interferometry over a limited number of sites, this section evaluates retrieval performance across a broader range of snow conditions and land-cover types to identify the key factors controlling retrieval accuracy."
- Beginning of section 5: "While previous work (Oveisgharan et al., 2024) established the capability of InSAR to retrieve SWE, it did not systematically assess the environmental and geophysical controls on retrieval performance. Here, we extend that work by analyzing how factors such as temporal coherence, temperature, vegetation, and terrain influence retrieval accuracy across multiple sites."
- At the beginning of section 5.1: "This analysis builds on the limited LIDAR validation presented in (Oveisgharan et al., 2024) by incorporating a larger number of LIDAR scenes across diverse environments, enabling a more comprehensive assessment of retrieval performance and its controlling factors."
- At the beginning of section 5.2: "In contrast to the site-specific validation in (Oveisgharan et al., 2024), this section leverages multi-site SNOTEL time series to investigate how environmental variability affects SWE retrieval performance and temporal coherence."
- End of the section 5: "This study moves beyond demonstrating the feasibility of InSAR-based SWE retrieval by providing a systematic evaluation of its performance across varying environmental conditions. The results identify the dominant controls on retrieval accuracy and define the conditions under which the method is reliable, which are critical steps toward operational SWE monitoring using current and future SAR missions."
- At the beginning of the conclusion section: "While previous work demonstrated the feasibility of SWE retrieval using Sentinel-1 interferometry, this study provides the first comprehensive, multi-site evaluation of retrieval performance and its controlling factors. By quantifying the roles of temporal coherence, temperature, vegetation, and terrain, we establish the conditions under which InSAR-based SWE retrieval is reliable, representing a key step toward large-scale and operational applications. These results also provide a foundation for quality assurance (QA) in future SWE products by identifying the key parameters that control retrieval reliability and uncertainty."

2. I do not quite follow the reconstruction of total SWE from interferometric SWE change. Equation (1) estimates ΔSWE between two acquisitions, but Equation (2) appears to reconstruct

total SWE by summing these increments from the first Sentinel-1 date after 1 December, while assuming zero SWE at that initial date. This assumption needs more explanation. In some of these mountain regions, there may already be substantial early-season snow accumulation by early December. If so, the retrieval would miss a non-negligible baseline SWE and the reconstructed total SWE would be systematically biased low. The manuscript currently treats this as a simplifying assumption, but I think it could directly affect the interpretation of the comparison with LIDAR snow depth. How sensitive are the reported correlations to this assumed zero baseline? Does this matter only for absolute magnitude, or could it also affect the spatial comparisons if early-season accumulation is not spatially uniform? I would like to see either a sensitivity analysis or at least a more rigorous discussion of the implications of this assumption.

- We added this to address this concern: "The assumption of zero SWE at t1 may introduce a bias in the reconstructed total SWE in regions where early-season accumulation is present. However, this assumption primarily affects the absolute magnitude of SWE rather than its spatial variability. For example, based on in situ measurements distributed across the Sentinel-1 frame (p71, f444), the spatial standard deviation of SWE is approximately 5 cm at the beginning of December and increases to about 23 cm by the end of March. This indicates that early-season SWE exhibits relatively low spatial variability compared to peak winter conditions. Given that the LIDAR scenes cover a much smaller spatial extent (~ 16 km × 15 km) within the larger Sentinel-1 frame (~240 km × 240 km), the spatial variability of SWE at the beginning of December is expected to be even smaller at the LIDAR scale. Therefore, neglecting the initial SWE mainly introduces an approximately constant offset in the reconstructed SWE and is not expected to significantly affect the spatial correlation with LIDAR snow depth. However, the reconstructed SWE values may be systematically biased low in absolute terms."

3. I am not fully convinced by the comparison between retrieved SWE and LIDAR snow depth without a clearer discussion of snow density. The manuscript reports correlations between Sentinel-1-derived SWE and LIDAR snow depth and interprets these as retrieval skill. However, SWE and snow depth are not equivalent quantities,4. 5. 6. and the relationship depends on snow density, which can vary substantially in space and time. It would help if the authors clarify whether this comparison is intended only to assess relative spatial patterns rather than physical equivalence. At present, this point is underexplained, yet it underpins one of the main validation results of the paper

- We dedicated section 4.1.4, "Evaluating the Retrieved Density", to address the discussion of density derived from comparing SWE and snow depth. Figure 9 addresses the reviewer concern.

4. The uncertainty analysis is still too limited. The manuscript mentions several sources of uncertainty, including atmospheric phase delay correction, phase ambiguity, wet-snow effects, and low temporal coherence. However, these are discussed somewhat separately, and I think the

paper would be much stronger with a more systematic uncertainty assessment. For example, how large are the expected errors associated with residual atmospheric correction over complex terrain? How often might the dry-snow assumption be violated even when near-surface air temperature remains below 0∘C? How much of the retrieval degradation during larger storms is due to phase ambiguity versus decorrelation? At present, the paper concludes that 12-day repeat data perform poorly and 6-day repeat data perform better, but the contribution of each error source to that difference is not quantified clearly enough. This needs to be discussed in a more integrated way.

- We added 2 sections, (4.1.5) and (4.2.1), on uncertainty analysis for LIDAR and SNOTEL comparisons.

5. I am not fully convinced by some of the broader interpretations regarding the dominant controls on retrieval performance. The manuscript argues that temporal coherence is the dominant factor controlling retrieval quality, with additional roles for temperature, SWE variability, vegetation, topography, and other environmental parameters. While this is plausible, some of these conclusions seem largely descriptive and based on a limited number of sites or frames. There are also exceptions discussed in the text, such as RC20, TU21, and the weaker consistency between the LIDAR-based and SNOTEL-based analyses. The manuscript excludes TU21 from the later parameter-impact analysis because of low overall coherence, and RC20 is explained post hoc by frequent melt events and low SWE. So I am not sure the manuscript yet provides enough evidence to conclude that temporal coherence is the dominant control in a general sense, rather than one important control among several interacting factors.

- We soften the claim and acknowledge interactions and variability in abstract: “Analysis of retrieval drivers indicates that temporal coherence is a primary control on performance, with additional contributions from temperature, snow wetness, and vegetation cover. The influence of these parameters on temporal coherence is only partially consistent with their effect on SWE retrieval performance, highlighting the complex interplay among environmental and observational factors. Temporal coherence generally declines with increasing snow depth, slope, and temperature, but improves under dry, cold conditions and gentle terrain.”
- We also added this to section 5.1 to show that we acknowledge the exception, and we cannot generalize the temporal coherence impact: “At the same time, the presence of exceptions indicates that temporal coherence alone does not fully explain retrieval performance, and that additional environmental and observational factors must also be considered. Therefore, temporal coherence is used here as a primary metric for evaluating retrieval quality, while its interaction with other controlling parameters is examined in Section 5.1.1.

6. I do not quite follow the discrepancy between the LIDAR-based and SNOTEL-based performance analysis. In the LIDAR comparison, the message is that higher temporal coherence

generally leads to better performance, and this is used to motivate the parameter analysis in Section 5.1.1. However, in the SNOTEL-based analysis, the interpretation appears less straightforward, with performance being discussed in relation to temperature, SWE magnitude, and coherence in a less consistent way. The manuscript also uses SNOTEL measurements both to establish reference points and to evaluate retrieval behavior, which makes me wonder whether there is some circularity or at least a dependence between calibration and evaluation that should be discussed more explicitly. I think the authors need to explain much more clearly why the LIDAR-based and SNOTEL-based analyses do not lead to a cleaner and more unified picture if temporal coherence is indeed the dominant retrieval control. At the moment, the two validation approaches seem to support overlapping but not fully consistent interpretations.

- Fair point. We expand our discussion in section 5.2 (marked red). We hope this address your valid concern. “This behavior differs from the more consistent positive relationship observed between temporal coherence and performance in the LIDAR-based analysis (Figure 13(a)), where higher coherence generally leads to improved agreement. It is important to distinguish between the two validation approaches used in this study, as they assess different aspects of retrieval performance. The SNOTEL-based analysis evaluates SWE change (ΔSW E) between individual interferometric acquisitions, and therefore reflects the temporal performance of the retrieval at specific time intervals. In contrast, the LIDAR-based analysis compares total retrieved SWE with total LIDAR-derived snow depth at a given date, which effectively integrates SWE changes over time and represents an aggregate measure of performance. As a result, the two approaches differ in their sensitivity to errors. In the LIDAR comparison, total SWE is obtained by summing ΔSWE over time. Any error in the reference point used for phase calibration propagates as a constant offset in the reconstructed total SWE. Such an offset does not affect spatial correlation with LIDAR snow depth, which primarily reflects the similarity of spatial patterns. In contrast, the SNOTEL-based analysis compares ΔSW E across multiple dates, and is therefore more sensitive to reference point errors and temporal noise. For example, estimating a small ΔSW E for large ΔSW E can significantly affect correlation with SNOTEL measurements, whereas in the LIDAR comparison the same error would contribute only a constant bias with no on spatial correlation. More broadly, the LIDAR-based analysis evaluates "spatial" correlation of "total SWE", while the SNOTEL-based analysis evaluates "spatio-temporal" correlation of "SWE change". Consequently, the relationship between temporal coherence and retrieval performance differs between the two approaches. In the LIDAR comparison, higher average temporal coherence over the time series generally leads to improved performance, as it reflects more reliable phase measurements accumulated over time. In the SNOTEL analysis, however, the relationship is more complex. High temporal coherence often occurs during periods of small ΔSW E, when retrieval performance is inherently more sensitive to noise and may degrade despite favorable coherence conditions. Therefore, the LIDAR-based analysis captures the integrated, average behavior of the retrieval, while the SNOTEL-based analysis reveals the instantaneous performance and its dependence on temporal variability. Together, these complementary perspectives provide a more complete understanding of the factors controlling SWE retrieval performance. If a temporally consistent LIDAR time series were available, co-registered with the InSAR acquisitions, the comparison would become more analogous to the SNOTEL-based analysis, as both would involve evaluating SWE and snow depth changes between acquisition dates and their dependence on temporal coherence. In that case, the relationship between coherence and retrieval performance would be expected to follow similar behavior to that observed in the SNOTEL comparison."

7. The manuscript separates the LIDAR analysis into 6-day and 12-day revisit groups and states that retrieval performance is very poor for the 12-day repeat case. It also states that temporal coherence is generally low at C-band and that the 6-day repeat significantly improves coherence compared to the nominal 12-day repeat. However, I am still not clear what the main limiting factor is for the 12-day case. Is it mainly decorrelation over longer repeat intervals? Is it because larger accumulated. ΔSWE between visits makes phase ambiguity more likely? Is it because longer intervals make wet-snow episodes and transient melt more likely to contaminate the signal? Or is it all of these together?

- We expanded this in section 4.1.3 (marked by red). We hope it addresses the reviewer's concern. "The results (not shown here) indicate that temporal coherence degrades significantly at the 12-day revisit interval, rendering the data unreliable for SWE retrieval. This degradation arises from a combination of factors rather than a single dominant limitation. First, the longer temporal baseline increases decorrelation due to snowpack evolution, wind redistribution, and melt–freeze processes, which reduces the reliability of the interferometric phase. As discussed in Section 4.1.5, low temporal coherence leads to a substantial increase in SWE retrieval uncertainty and degrades the performance of phase unwrapping algorithms. Second, the magnitude of SWE change between acquisitions is significantly larger for 12-day intervals compared to 6-day intervals (e.g., Figure 8 versus Figure 4). For Sentinel-1, the SWE ambiguity corresponding to a 2π phase cycle ranges from approximately 1.5 to 3.5 cm depending on incidence angle. In many cases, the accumulated ΔSW E over 12 days exceeds this ambiguity threshold, making phase unwrapping more challenging and increasing the likelihood of phase errors. While such ambiguities can sometimes be resolved using high-coherence reference areas combined with in situ constraints, the simultaneous presence of low coherence limits the effectiveness of these approaches. Finally, longer revisit intervals increase the probability of transient melt events and wet-snow conditions occurring between acquisitions, which violate the dry-snow assumption required for the interferometric SWE retrieval. In addition, the reduced number of acquisitions in a 12-day time series increases the relative weight of each interferogram in reconstructing total SWE, making the retrieval more sensitive to individual errors. Overall, the combination of increased temporal decorrelation, larger SWE changes relative to phase ambiguity, reduced phase unwrapping reliability, and a higher likelihood of wet-snow conditions makes SWE retrieval from 12-day Sentinel-1 data significantly more challenging. It is worth noting that upcoming missions such as NISAR, despite having a similar nominal 12-day repeat cycle, operate at longer wavelengths (L- and S-band), which are expected to maintain higher temporal coherence and provide larger SWE ambiguity thresholds (approximately 6–14 cm). These characteristics are anticipated to mitigate some of the limitations observed for C-band Sentinel-1 data."

Minor comments:

1. Line 265: "Retrieavl" should be "Retrieval"

- Done!