

RC2, Anonymous Referee #2

This manuscript evaluates a Sentinel-1 repeat-pass InSAR approach for retrieving changes in snow water equivalent (Δ SWE) 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 Δ SWE with SNOTEL Δ SWE 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 Δ SWE 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 Δ SWE/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- Δ SWE 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- Δ SWE 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., $>0^{\circ}\text{C}$), 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.