Scale patterns of the Sentinel-1 SAR-based snow depth product compared to station measurements and airborne LiDAR observations

General Comments:

This paper addresses an important topic by evaluating the scale-dependent performance of the C-SNOW Sentinel-1 snow depth product against both in-situ measurements and airborne LiDAR observations. The multi-scale analysis and inclusion of geographic and land cover effects are valuable contributions to the remote sensing and snow hydrology communities. However, the manuscript would benefit from greater clarity in its introduction and methodological explanations, Additionally, some of the comparisons between datasets are not well-aligned in terms of spatial or temporal scale, which limits the interpretability of the findings. The Discussion section leans heavily on restating results rather than offering critical insight into the causes and implications of observed discrepancies.

Overall, the paper has the potential to contribute knowledge on the accuracy of Sentinel-1 snow depth at different spatial scales, but revisions are needed to improve its structure, clarity, precision, scientific depth, and accuracy.

Major Comments:

1. The rationale behind the use of Sentinel-1 data at 10 km and 25 km resolutions is unclear. If the primary aim of the paper is to evaluate the spatial and temporal performance of Sentinel-1 for snow monitoring, then standard higher-resolution products (e.g., 0.5 km and 1 km) would suffice. The inclusion of coarser resolutions needs to be better motivated and should be clearly stated and supported by appropriate literature and methodological context.
2. Since NRCS SNOTEL provides direct snow depth measurements, please clarify why a conversion from SWE to SD was performed. A fixed density to go from SWE to depth will impact the S1 evaluation, which is another consideration. Also, explicitly explain the choice of using a fixed snow density value of 0.24 g/cm³ for the Russia-SWE dataset. Is this based on regional averages or prior literature? Including a brief rationale would improve clarity.

3. In Figure 10, the comparison of the C-SNOW time series with spatially distributed ASO LiDAR data appears unrelated comparison. Since C-SNOW is also available as a spatial product over much of the Northern Hemisphere, it would be more appropriate to extract the C-SNOW spatial data closest to date to the ASO flight and perform a spatially explicit comparison. This would enable more meaningful evaluation and avoid misleading conclusions from mixed-scale comparison. Furthermore, the inclusion of air temperature in this figure is puzzling. The manuscript does not provide a rationale in the methodology for using temperature as a covariate or validation proxy. Since it is not used quantitatively in the analysis, I recommend removing it unless a clear scientific justification is provided. Overall, please revise this section to focus on meaningful spatial comparisons (e.g., using representative basins), and consider summarizing the evaluation using average statistics or appropriate spatial accuracy plots (scatter, bias maps, etc.).
4. There are several issues in section 3.3. First, the phrase "other types" of land cover should be clarified—please specify which land cover categories are included beyond forest and permanent ice. Additionally, while the inclusion of permanent ice regions is noted, it's important to question the relevance of evaluating C-snow accuracy in such areas. Does the Lievens et al. algorithm or Sentinel-1 backscatter perform reliably in permanent ice environments and dense forests? Please justify why this analysis was included and consider whether they should be treated as known limitations of the remote sensing platform and dataset. In continuation, some findings are counterintuitive and warrant further explanation. For instance, significant overestimation in areas with low forest fraction (0–0.2, figure 12c) is unexpected, as reduced vegetation cover typically enhances radar retrieval accuracy. Similarly, while elevation and elevation variability appear to influence performance, the paragraph lacks synthesis on why certain ranges (e.g., moderate forest and elevation variability) yield better results. Please also clarify how elevation differences between stations and grid cells were calculated, and how these mismatches propagate error across scales (this can go in methodology).
5. The current Discussion section is written more in the style of a results narrative. While the detailed reporting of accuracy metrics, grid-level behavior, and ice-related overestimates is important, most of the text focuses on what was found rather than interpreting what it means. To strengthen this section, I recommend separating the descriptive content into the

results section and expanding the discussion with deeper analysis. For example, lay error analysis background, consider critically why ASO and station-based validation trends differ, what the implications of spatial representativeness are for coarse-scale validation, and how the known limitations of Sentinel-1 in forested or glaciated areas affect the broader applicability of C-snow. Comparisons with previous studies and a more explicit articulation of limitations and future directions would also help to better contextualize the results.

Minor Comments

1. The introduction is currently in a general tone. Several statements are made without citing relevant studies or offering clear justification. Please consider grounding key claims with references and clarifying the motivation and novelty of the study more explicitly.
2. Land cover classifications such as forest and permanent ice are introduced in the Results section without prior explanation. These should be defined and justified in the Methodology section, including the source of the land cover data and how the categories were used in the analysis.
3. Sections 3.3 and 3.4 both focus on geographic and environmental influences on snow depth retrieval. Merging them into a single cohesive section would improve readability and thematic consistency by presenting all location-based findings together.
4. Reorganize Figures and reconsider adding all figures for Improved Flow
    a. Figure 14 could be merged with Figure 2 to streamline the presentation of the study area and grid setup. Label the nested grid structure directly in the map and refer to it when introducing the experimental design.
    b. Figure 16 is conceptually related to Figure 12. Placing them closer together would enhance the narrative flow and allow readers to better understand the progression of results.
    c. Please consider reducing the number of figures and balancing the proportion of figures and text. Refer to comments that suggest either removing or combining figures.

Line-to-Line Comments

23 and elsewhere: The errors statistics are reported to the hundredth of a cm, but this level or precision is not realistic or warranted. Please consider the significant units here and elsewhere when reporting errors.

26: Remove "Especially an".

35: To clarify how snow depth (SD) information contributes to water availability, state "…estimated from snow depth (SD) and snow density.".

44: Add "one or several" before "meters".

45: These density values are too high for most seasonal snow and in the range of when snow transitions to firn. Recommend revising to the typical range of density values (~100 to 550 kg/m3) for seasonal snow (see Sturm et al., 2010, J. Hydrometeorology).

38- 50: The paragraph introduces the use of microwave remote sensing for SWE retrieval and suggests that SAR offers advantages over passive microwave techniques. However, it would benefit from greater specificity and clarity. For instance, clearly states that while passive microwave remote sensing is widely used, its coarse spatial resolution (~25 km) limits its ability to capture the fine-scale spatiotemporal variability of snowpacks in complex mountainous terrain. This will help establish a stronger context for the discussion of SAR advantages.

Additionally, when discussing mountain snowpack complexities (**lines 45–50**), it would be helpful to explicitly connect these challenges—such as variable snow density, grain size, wind, and gravity-driven compaction, elevation, and aspect—to the limitations of passive microwave sensing. You may also consider briefly noting the difficulty of ground-based observations in remote, high-elevation regions, which further highlights the value of satellite-based approaches. These additions will create a smoother transition to the next paragraph (**lines 51–61**), which focuses on SAR. Lastly, in line 43, replace "snow cover" with "snowpack" for technical accuracy.

51: Delete "the" before "monitoring".

51-52: While this paragraph highlights recent advances using C-band SAR for SWE monitoring, it would be helpful to first acknowledge earlier foundational studies that have explored SAR for retrieving snow characteristics such as (Ulaby and Stiles, 1980; Bernier et al., 1999; Shi and Dozier, 2000; Chang et al., 2014; Lievens et al., 2019), there are many more. Including a broader set of references would better reflect the extensive work by the remote sensing community and provide context for the transition to C-band applications.

52-54: The sentence discussing snow volume scattering being stronger in Ku-band compared to "other bands" could be made clearer. Please specify which bands are being referenced (e.g., X-, C-, or L-band) and provide clearer citations. If this theoretical point is derived from Rott et al. (2010) or others, please state this explicitly and consider rephrasing for clarity and precision.

55-56: Instead of using “VH/VV,” which may be unclear to some readers, consider using the more descriptive phrase “cross- to co-polarization ratio.”

56-59: To avoid redundancy, consider replacing the second use of “notably” with an alternative phrase. Additionally, rephrase “due to the nonspherical properties of snowpack” to “due to the anisotropic nature of snow grains,” which is a more accurate physical description. You may also cite relevant literature on this point—Lievens et al. (2022) references several useful studies that could support this claim.

62-64: It would be helpful to clarify that the original C-snow product developed by Lievens et al. (2019) was produced at 1 km resolution without wet snow masking. In later studies, higher-resolution (e.g., 500 m) products were developed with wet snow flagging. Mention explicitly that there are no C-Band products at 10 km and 25 km resolution. Additionally, before introducing subsequent evaluation studies, please report the original study’s metrics (e.g., RMSE, bias) from the original Lievens et al. study. This will help establish a clear baseline and better contextualize the results from later evaluations, thereby strengthening the motivation for your analysis.

67-68: Please clarify that Hoppinen et al. (2024) evaluated the performance of the C-snow retrieval algorithm using airborne LiDAR data across two separate water years: 2020 and 2021. As currently written, “2020–2021” may be misinterpreted as a single water year.

69-71: The statement about Broxton et al. (2024) and Yang et al. (2024) using C-snow products at 10 and 25-km resolutions appears to be inaccurate. Broxton et al. used machine learning to enhance C-snow snow depth estimates at 0.5 and 1 km resolutions and compared these to University of Arizona SWE and airborne LiDAR data. They did not use C-snow to improve passive microwave SWE. Additionally, the studies you cited used either the 1 km or 500m C-snow product—not the 10 or 25 km. Please check for accuracy and revise.

72-73: Please rephrase the sentence about Lievens et al. (2022) for clarity and precision. They used Sentinel-1 backscatter observations to retrieve snow depth across multiple resolutions in the European Alps and evaluated retrieval performance. This distinction is important to avoid confusion between backscatter versus the derived SD observation also in the monitoring and evaluating SD at different resolutions.

84-85: ASO data are spatially extensive than stations, but do not have wider global coverage (i.e., they are only available in the western U.S.). Please revise phrasing for accuracy.

90-91: Recommend using a citation such as "The first 1 km SD product based on C-band SAR, covering all mountain ranges in the Northern Hemisphere, was developed by Lievens et al. (2019). The dataset is publicly available through the C-SNOW project (C-SNOW, 2024)." Instead of using a link in the text.

104-105: The Zenodo link should be replaced with a proper DOI citation. Also, revise the sentence for typos and formatting—for instance, the link includes a fragment ("#YdYE...") that should be removed.

106-112: Remove links and include proper citations.

127-133: ASO is not a LiDAR mission, ASO is a company that conducts LiDAR flight surveys using an airborne laser scanner (ALS). The dataset is available from 2013-2019 and 2022 to present. Please check the 2.2.2 section for accuracy. Refer to NSIDC, ASO website, and Painter et al 2016 paper. Also, clarify the reasoning behind using 3m instead of 50m.

130-134: Please clarify that only California and Colorado ASO surveys were used.

141-142: Replace the word "corresponding datasets" with a direct reference to Table 1 for clarity. Also, in line with earlier comments, avoid including direct links (e.g., to Google Earth Engine) in the main text.

164: Clarify whether this is the Pearson correlation coefficient or other.

169-170: Add a line that explains what this section is about. It starts suddenly without making any coherence.

178-182: This text (and Figure 7) does not add much to the study, as it is all obvious and expected behavior for considering different spatial scales of a dataset. I recommend removing this.

189-192: The term "wet snow season" should be replaced with "ablation" or "melt season" to better reflect the physical processes and limitations of Sentinel-1 during this period. The phrase "mismatch of snow season length" is vague. It's not clear whether the authors refer to differences in snow onset and melt timing, the overall duration of snow cover, or specific discrepancies, C-Band data is only available till the end of April anyway as Sentinle-1 is not reliable in ablation season.

195-196: This line should be part of the limitation section

223-225: What is the other type of land cover you are talking about here? Please explain how you expect Sentinel-1 to work in permanent ice regions, what are those regions?

271-276: Here and elsewhere – what is the purpose of comparing a 10-km or 25-km estimate of snow depth with a station? One would not expect the station to match those very different scales.

282-283: What evidence do you have for this statement? I do not think that spatial representativeness is guaranteed for a station in a flat area. Additionally, I would not conflate low elevation with flat, as there can be topographic complexity/variability even at lower elevations.

306-307: It should be part of the data section

Figures and Table Comments

Figure 1: Use distinct colors for SNOTEL and GHCN since some of the stations are close to each other and it is hard to distinguish.

Figures 1 and 2: Consider combining these into a single figure with 4 panels (a-d).

Figure 2: Continent labels are hiding in some places. Make it consistent—no labels or labels everywhere.

Figure 3: This figure may not be necessary, especially given the already high number of figures (19). However, if you choose to include it, consider showing ASO data coverage across the entire US and highlighting (e.g., with a box) the region used in your analysis. This would provide helpful context without redundancy.

Figure 4: Consider using station codes instead of full names to reduce clutter. Add borders to markers for better visibility and clarify the meaning of each color in the figure description.

Figure 5: Please distinguish two boxes by panels for example say panel a and panel and state it in the figure description. Also, the arrow between accuracy analysis and uncertainty analysis is misleading. Are you trying to say uncertainty analysis comes after accuracy analysis? Regardless arrow is not needed.

Figure 7: It appears that Figure 7 is only showing C-snow data at different spatial scales (1, 10, and 25 km) across three mountain regions, without any comparison to reference or evaluation datasets. If that's the case, the purpose of this figure needs to be clearly stated—whether it is to demonstrate spatial variability or resolution effects. As currently placed, the figure does not align well with the section, which focuses on comparing C-snow with observed SD. Consider either relocating the figure to a more appropriate section or removing it if it doesn't directly contribute to the evaluation narrative.

Figure 8: Please clarify whether Figure 8 shows the average time series across all stations or if it represents individual station values. If it's an average, this should be explicitly stated in the caption and main text. However, averaging across diverse stations may mask site-specific dynamics and variability. Consider instead selecting one or a few representative stations to illustrate the temporal mismatch at different scales more clearly. Additionally, using a line plot instead of scatter points would improve readability and better highlight trends over time.

Figure 10: Check major comment section

Figure 12: It's not immediately obvious that the markers show actual SD values, and the bars show relative error. Adding a brief formula or explanation for Rbias in the caption or methods would also help.

Figure 13: The left plots may be better conveyed as pie charts rather than bar charts because they add up to 100%.

Figure 14: Refer to minor comments. Additionally, the layout of this figure is not ideal, as the three selected regions overlap the global map but instead could be shown separately with more detail.

Figure 15: How figure 15 is different than Figure 8? Figure 15 makes more sense in section 3.1 in replacement of Figure 8.

Figure 16: This figure does not contribute much to the study and can be considered for removal.

Figure 17: It should be a part of the results.