

**We thank Reviewer #1 for their thoughtful comments. The reviewer's text appears in black with our responses in bolded blue.**

**Based on both reviewers' comments, we significantly revised and reorganized the introduction of the paper. We also moved the section of governing error equations from the appendix to the main body of the text (now Section 2), provided more detail in the derivations, and removed all references to the unpublished companion paper.**

### **Summary and Recommendation**

This paper quantifies and compares non-snow errors for InSAR-based retrievals of change in snow water equivalent (SWE). The analysis is conducted at 13 SNOTEL sites in the western United States. Six error sources are considered, including ionospheric effects, atmospheric humidity and pressure, soil permittivity, vegetation permittivity, and surface deformation. The paper finds that errors due to the ionosphere are large and can easily exceed the median SWE value at many sites with lower snowpack accumulation. If ionosphere effects are removed, then the remaining cumulative errors are on the order of 2 to 7 cm in SWE, with some errors offsetting each other. For 10 out of 13 sites, these errors are within 10% error for April 1 SWE, which is within the target accuracy set forth by the U.S. decadal survey.

I find this to be a straightforward and useful analysis with good potential to support SWE error assessments with InSAR-based retrievals like from NISAR. I think it would be a great contribution to the journal following attention to some comments, as elaborated below.

**Thank you for these positive comments; we are glad you found the paper to be straightforward and useful.**

1. The specific selection of these 13 SNOTEL sites could use better justification. I understand they were sampled from different regions/zones of the western United States. However, in each of those zones, there are multiple sites which could have alternatively been selected, and the low sample size leads me to wonder about the stability/robustness of the results. Would the results have changed with different sites select
2. Two of the SNOTEL sites (Disaster Peak and Quemazon) are in post-burn areas and likely have altered snow accumulation and melt dynamics relative to their pre-fire condition (see Smoot and Gleason, 2021). One might argue these two sites may not be broadly representative of their respective geographic zones due to the fire. At the same time, there are broad swaths of the western U.S. where fire has impacted snow dynamics. Similar to the previous comment, I would ask the authors to provide more justification for specific site selection and elaborate on what is being represented by the sites.

**Thank you for these points. We added more details about the site selection criteria (lines 198-205):**

**“To illustrate how non-snow errors vary across snow climates, we applied our analysis at locations across the WUS representing 13 major mountain ecoregions denoted by Trujillo and Molotch (2014). We chose a 10-year period between water years 2016–2025 (October 1,**

2015 through September 20, 2025), which includes high (2023), average (2022), and low (2016) snow seasons, to examine the temporal variability of the errors. All other datasets we use to calculate the non-snow errors (Table 1) are also available for water years 2016–2025. Within each ecoregion, stations were selected based on data availability (98% complete data records for SWE, air temperature, and 2-inch soil moisture for each of the 10 seasons). When more than one station in an ecoregion met the data availability criteria, we selected so that the final 13 stations span a range of snow classes, vegetation types and heights, and other environmental characteristics (Table 2).”

It is indeed possible that results would change if different sites had been available. However, our intent is not to draw broad conclusions about the method based on the 13 sites selected; rather, we aim to illustrate that the non-snow factors we considered here can have different magnitudes/effects in different snowpacks. We added this language explicitly in the Discussion in lines 457-467, including some new discussion about snow/fire interactions and shallow snowpacks:

“We analyzed errors at only 13 SNOTEL stations across the WUS. Although this selection was partially dictated by data availability (see Section 2.3), it is possible that our results would change if we had selected different sites. For example, two of our three shallow snowpack stations (Disaster Peak and Quemazon) are in post-burn areas and may have snow accumulation and melt dynamics that are not fully representative of their broader geographic regions (Smoot and Gleason, 2021). However, our intent is not to draw broad conclusions about the method based on the 13 sites selected; rather, we aim to illustrate that the non-snow factors we considered here can have different magnitudes and effects in different snowpacks. Especially given that wildfire area has been increasing in snow-dominated regions across the WUS since the 1980s (Kampf et al., 2022) and WUS snowpacks are projected to continue their general decline (Siirila-Woodburn et al., 2021), it is likely that the WUS will become more reliant on shallow and post-burn snowpacks for water resources in the coming years and decades. We encourage future work in snow remote sensing to continue examining shallow and marginal snowpacks (López-Moreno et al., 2024).”

3. To what extent might error/uncertainty in satellite location/position induce a change in InSAR phase and thus an error in retrieved SWE? I assume this might manifest as a bias over the scene and could be detected/corrected. However, if much of the scene included a change in SWE along with other non-snow errors, then I wonder to what extent orbit error might be detected or corrected.

Other than during periods of very high solar activity, the final estimates of satellite orbits tend to be highly accurate, especially in comparison to the airborne UAVSAR platform referenced in several studies in the introduction. Baseline estimation errors appear as consistent linear phase ramps in interferograms that are quite recognizable and easily removed using fast-Fourier-transform techniques. These effects should be not a significant barrier to InSAR-derived SWE measurements.

## Line Comments

- L. 43: Delete “a”.

Done.

- L. 51: Delete “will”.

Done.

- L. 66-67: Why is it  $2\pi$  for equation 1 but  $4\pi$  for the appendix equations? I assume the  $4\pi$  is due to the two-way distance from satellite to surface, so it is a little odd that equation 1 is developed as  $2\pi$ .

**This comment has been addressed by the expanded equation derivations (now Section 2). The  $4\pi$  is indeed due to the two-way travel distance, but the Phase-to-SWE equation has a factor of 2 in the denominator. We made this explicit in Section 2.1 (specifically lines 80-81).**

- L. 162: Delete “season”.

Done.

- L. 210: Remove first “early”.

Done.

- L. 220: Add “the” before “largest error”.

Done.

- L. 236-243 and Fig. 6: What about cases when  $\Delta SWE = 0$ ? This may not be an uncommon occurrence (e.g., cold regions/periods with no new snow and no ablation), and would be important to document. I understand the challenge with including this case (i.e., cannot divide by zero for Fig 6b) but perhaps there is another way to summarize and report expected errors for the  $\Delta SWE=0$  case?

**Thank you for this suggestion. We extended this exceedance analysis by separating the daily observations into days where  $\Delta SWE = 0$  and days where  $\Delta SWE \neq 0$ . As noted by the comment, it is not feasible to add a curve to Figure 6b when  $\Delta SWE = 0$ , and the separate zero/nonzero curves were quite similar to the ones already in Figure 6a. We did not change the error curves in the figure but added a new paragraph in lines 363-369:**

**“We used a similar analysis to investigate whether non-snow errors behave differently on days with vs. without a snow event. We split the October 1 through April 1 dataset into days where  $\Delta SWE = 0$  (1784 total measurements) and days where  $\Delta SWE \neq 0$  (21,698 total**

measurements), representing both accumulation and ablation events. When  $\Delta\text{SWE} = 0$ , the 50% exceedance thresholds were 0.210~m for total non-snow error and 0.031 m for non-ionospheric error. When  $\Delta\text{SWE} \neq 0$  the 50% exceedance thresholds were 0.230~m for total non-snow error and 0.030 m for non-ionospheric error. The error curves for the two cases (not shown) were quite similar to those in the left panel of Figure 6. This result indicates that non-snow factors affect the phase relatively consistently, regardless if a snow event occurs or not.”

During this analysis we also caught a typo in the title of Figure 6a. Previously we had n=23,397 but we revised this to n=23,482 to reflect the true number of observations in the dataset.

- L. 338: Add “of” after “effects”.

Done.

- L. 380: add “(i.e., surface deformation)” at the end of this line because that is what is causing the change in distance.

**This comment has been addressed by the expanded equation derivations (specifically Section 2.2.5).**

- L. 410: This canopy height is from LandFIRE, right? Please clarify.

**This comment has been addressed by the expanded equation derivations (specifically line 178 in Section 2.2.4 and Table 1 in Section 2.1).**

## Figures and Tables

- Table 1 – how is surface deformation mapped to the SNOTEL locations, given that these two stations may be separated by some distance?

**We make the assumption that the 12-day surface deformation at a given SNOTEL site is equivalent to that of the nearest GNSS station, even though the SNOTEL and GNSS points are not colocated. We do not attempt to spatially interpolate or otherwise adjust the GNSS data to account for the separation distance. We added these details explicitly in lines 218-221.**

- Figure 1 – I think it is better stack the 4 panels vertically to align their common axis (time).

**We agree and have made this change (now Figure 2).**

- Figure 2 – Similar to previous comment. I think this would display more effectively if it was flipped with 2 rows by 3 columns. That way, the SWE errors can be compared more readily to SWE on the common axis (time).

**We agree and have made this change (now Figure 3).**

- Figure 2 – remove “6” at the start of the caption.

**Done (now Figure 3).**

- Figure 3 – I realize these are ordered by site number, but think that this could be more effective if arranged by bias (from most negative to most positive, on a median basis).

**During revisions we remade a draft of the figure as suggested. While it is perhaps more visually appealing, sorting this figure in a different order than all others in the paper suggests that we are trying to draw some conclusion about the sites based on the median bias. As the rest of our analysis shows, the errors interact in complex ways in both space and time and we think it would be inappropriate to draw general conclusions based on the median error values at these 13 sites. Ultimately we decided not to implement the change in the revised manuscript. This figure (now Figure 4) is unchanged.**

- Figure 7 – the longitude and latitude labels on the horizontal axes of all 3 panels should be swapped (e.g., longitude should be latitude, and vice versa).

**Done.**

- Figure 7 – would it help to denote the intersection points between the transects on each of the 3 panels?

**We attempted to add this to the figures during revision but could not find a clean way to do it, since the intersection points are not common between all subpanels. Ultimately, we decided that the additional text/labeling required to denote the intersection points distracted from the data and ultimately did not improve the figure.**

## **References**

- Smoot, E. E. and Gleason, K. E.: Forest Fires Reduce Snow-Water Storage and Advance the Timing of Snowmelt across the Western U.S., *Water*, 13, 3533, <https://doi.org/10.3390/w13243533>, 2021.

**We thank Reviewer #2 for their thoughtful comments. The reviewer’s text appears in black with our responses in bolded blue.**

**Based on both reviewers’ comments, we significantly revised and reorganized the introduction of the paper. We also moved the section of governing error equations from the appendix to the main body of the text (now Section 2), provided more detail in the derivations, and removed all references to the unpublished companion paper.**

### **General Comments**

The authors assess six potential error sources that may influence NISAR L-band snow water equivalent retrievals using the phase-change approach. To do so, they apply equations describing the theoretical influence on signal phase for each component, using a combination of observed and modeled data to derive the expected influence. These phase shifts are used as input to the SWE retrieval algorithm to determine the expected magnitude of SWE change errors from each factor. The study is a necessary and well-conceived sensitivity analysis, not an observationally focused effort. While the authors do not assess all potential factors that influence SWE, the authors offer a thoughtful discussion of their study limitations and the recommendations that can realistically be drawn from this study. They provide well-written, clear, and concise methods, with effective and visually appealing figures and results. It is already a strong paper, and it was enjoyable to read and will be a good addition to the literature on InSAR-based SWE retrievals – especially as NISAR data begins to become available for SWE mapping. I have only minor comments that I believe would further strengthen the manuscript.

**Thank you for these positive comments; we are glad to hear that you enjoyed the paper.**

### **Specific Comments**

1. The concept of the 12-day baseline may warrant a clearer description in the introduction or methods. How it differs from a 12-day moving-window approach should be clarified.

**“Temporal baseline” is a common phrase in the InSAR literature but this comment helped us realize that it may be less familiar to our intended snow science audience. In our revised introduction we define the phrase the first time it appears (line 37). In other instances we changed the wording from “baseline” to “period” or something similar (e.g. Abstract line 17). We also revised our methods to include the phrase “moving window” which we agree is a more widely understood term (lines 206-215).**

2. The last sentence of the abstract ‘For individual 12-day baselines, removing as many errors as possible....’ comes across as very general. The study has several key takeaways and suggestions that are presented at the end of the conclusion – I’d suggest drawing upon those to provide a more impactful and informative concluding sentence to the abstract

**Done. The end of the abstract (lines 22-27) now reads: “Removal of other error sources requires careful consideration of the SWE monitoring application: for tracking seasonal**

**SWE accumulation in areas with deeper snowpacks, correcting some errors but not others may actually decrease accuracy by removing offsetting cumulative effects. For individual 12-day periods, wet and dry tropospheric effects (due to changes in water vapor and pressure) should be removed for accurate interpretation of spatial patterns of snow accumulation at basin to range scales, and site-specific factors should be considered to assess the relative influence of vegetation, soils, and surface deformation.”**

3. The introduction is lacking a summary of the InSAR SWE retrieval literature, particularly in terms of observed errors from field testing. I recommend that the authors draw on that literature to provide more insights into expected retrieval errors based on observational data.

**While revising the introduction, we made an effort to describe the previous literature in more detail instead of inserting long lists of references. Lines 35-38 report error metrics from previous studies, and we also mention challenging environments for the technique in lines 41-42.**

4. While I am sure each of the 6 assessed components are described in the unpublished companion paper in detail, it would be helpful to the reader if the authors could provide brief (but more detailed than is currently provided) descriptions of each component. For example, some background as to what drives ionospheric variability or surface deformations would be useful to mention early on in the introduction.

**We moved the section of governing error equations from the appendix to the main body of the text (now Section 2), provided more detail in the derivations, and removed all references to the unpublished companion paper.**

5. In Figure 6, the exceedance probability of the ionospheric error contribution looks like a step function with error only occurring in  $\sim 0.025\text{m}$  bands. This is not reflected in the red line which summarizes all error components as well. – why is this?

**Good eye – the IGX IONEX data have a precision of only one decimal point, and applying Equation (4) when  $\Delta\text{TEC} = 0.1$  is equivalent to a SWE error of  $0.02549\text{ m}$ . So this is the minimum step size in ionospheric error that we can calculate using this dataset. When all errors are incorporated (i.e. the red error curve in Figure 6) the other datasets have higher precisions and the error curve becomes essentially continuous. We added a note about the coarse precision of the ionospheric data and resulting stepwise behavior to the caption of this figure (now Figure 7).**

6. There are opportunities in the discussion and conclusion to add more statistical metrics, which would strengthen the analysis and provide key insights for readers. See Line 279 (‘largest error values’). Paragraph 1 of the conclusions attempts to clarify the expected errors using the exceedance statistics, but its presentation could be clearer in my opinion. Presenting the 50% exceedance for each component would provide a good indication of the relative influence of error across all 6 assessed error sources. The delta SWE threshold should also be consistent and clearly described.

We added the 50% exceedance error values for all curves shown in the figure in a new table (Table 4), which we agree helps clarify the text with references in the Discussion (lines 415-417) and Conclusions (lines 479-481). Based a comment from the other reviewer and to address the point in this comment about the consistent  $\Delta\text{SWE}$  threshold, we also added a new paragraph in the results comparing the exceedance errors with zero/nonzero  $\Delta\text{SWE} = 0$  (lines 363-369):

“We used a similar analysis to investigate whether non-snow errors behave differently on days with vs. without a snow event. We split the October 1 through April 1 dataset into days where  $\Delta\text{SWE} = 0$  (1784 total measurements) and days where  $\Delta\text{SWE} \neq 0$  (21,698 total measurements), representing both accumulation and ablation events. When  $\Delta\text{SWE} = 0$ , the 50% exceedance thresholds were 0.210~m for total non-snow error and 0.031 m for non-ionospheric error. When  $\Delta\text{SWE} \neq 0$  the 50% exceedance thresholds were 0.230~m for total non-snow error and 0.030 m for non-ionospheric error. The error curves for the two cases (not shown) were quite similar to those in the left panel of Figure 6. This result indicates that non-snow factors affect the phase relatively consistently, regardless if a snow event occurs or not.”

### Technical Suggestions

Line 40: I recommend mentioning the target +/- 10% goal when introducing the decadal survey

**Done (now lines 48-51).**

Line 46 (& referenced elsewhere): The submitted (unpublished?) companion paper (Hoppinen et al., ) should be published before this work is published, since this relies directly on those presented methods

**This comment has been addressed by the expanded equation derivations (now Section 2).**

Line 53: Briefly noting other potential error sources here would be relevant

**This comment has been addressed by the expanded equation derivations (now Section 2).**

Consider more descriptive section titles (sections 3 and 4) and adding an additional sub-section of results after the error exceedance analysis (before the gridded analysis), as this came across as a separate analysis sub-section

**Done. Based on this comment we also took a more liberal approach with subsections when revising the equation derivations (Section 2).**

Line 128-129: Providing an explicit example of such interactions would be helpful. The author's use of examples elsewhere in the paper really helps its clarity in my opinion.

We added an example from our data as well as a more hypothetical discussion in lines 242-251:

**“However, the interannual variability of the total cumulative error (i.e., the width of the boxplot) changes between stations and most stations also show positive errors (overestimated SWE) in at least one year. This indicates that it may be difficult to predict the general behavior of the total error at any site based on errors measured in previous seasons. For example, at Sunset the largest positive error (0.096 m) occurred in 2016 and the largest negative error (-0.191 m) occurred in 2025, but peak SWE measurements in the two years were very similar (0.452 m in 2016 and 0.444 m in 2025). Wide interannual variability at a site is likely driven by the complex interactions between the error factors and their physical drivers at seasonal scales. For example, a series of several early season snow accumulation/melt events would likely result in fluctuating tropospheric effects (i.e., changes in air pressure and precipitable water associated with frontal passage) but the secondary soil permittivity effects from melting snow are partially controlled by antecedent soil moisture conditions, which we do not assess in our analysis.”**

Line 222: consider adding the relative difference in error (numerically) between the morning and evening overpasses – it is clear in Figure 5, but the specific magnitude of the difference is less so, especially because of the different temporal variation patterns

**We added more detail about the AM/PM differences in ionosphere error in lines 344-352:**

**“For example, at Trial Lake, the seasonal average AM ionosphere error (mean of the middle AM orbit line) is 0.010 m with an average interquartile range (IQR) of 0.377 m. The seasonal average PM ionosphere error is 0.013 m with an average IQR of 1.031 m. The largest average difference between AM and PM errors is 0.312~m on November 27, a critical early-season period when total SWE and  $\Delta$ SWE are typically small. Ionosphere errors of 0.312 m potentially represent more than 100% of the snowpack at many WUS locations in late November. The largest difference between AM and PM IQR is 1.206 m on April 12, another critical monitoring period around the timing of peak SWE in the WUS. AM and PM results are similar at Paradise and Disaster Peak. For these reasons, we recommend using ascending NISAR overpasses to generate NISAR interferograms for  $\Delta$ SWE measurements, which will acquire data at approximately 06:00 local time.”**

Line 236: Consider adding further justification as to why the ionospheric component was excluded from Figure 6b

**This was simply to improve clarity of the figure; we added this detail explicitly in lines 374-375.**

Line 344: Where did 670% come from? The highest error was around 500% in the presented tables

**The highest total cumulative error in Table 1 is indeed 581%, but this sentence refers to the specific ionospheric cumulative error which is 670% at Quemazon (also shown in Table 1).**

Line 352: Is there a relevant manuscript/source that the authors would refer readers to review to better understand how the ionic corrections are made? Consider adding a source within the text that details ionospheric correction effects – in the discussion, if possible

**We added two relevant citations along with a few sentences describing the primary drivers, spatial and temporal variation cycles of the ionsphere, and the primary impacts of it on InSAR pairs in lines 101-106:**

**“The varying electron density of the ionosphere is affected by solar UV radiation, Earth’s magnetic field, and atmospheric gas concentrations; interactions between these factors cause electron density concentrations to vary over multiple spatial (sub-kilometer to tens of kilometers) and temporal (diurnal, seasonal, and interannual) scales (Lean et al., 2016). The resulting impacts on InSAR phase are frequency-dependent and can introduce larger errors at lower frequencies, like NISAR’s L-band measurements (Meyer and Agram, 2017).”**