Initial response to reviewer comments on preprint egusphere-2026-1594, entitled: "Interferometric synthetic aperture radar (InSAR) phase data assimilation for Bayesian snow water equivalent estimation" by Margulis et al.

We greatly appreciate the constructive comments from the two reviewers and will implement the suggested changes and clarifications as indicated below in the revised manuscript. In several instances, the insightful comment is best suited to follow-on work with real phase data. In those cases that point will be made more explicitly in the revised manuscript as indicated more specifically below. In the text below, the original reviewer comment text is shown in grey font, while the response is shown in blue font.

Reviewer #1 Comments:

RC1: 'Comment on egusphere-2026-1594', Benoit Montpetit, 04 Jun 2026

This paper presents an OSSE on assimilating phase difference in repeat pass interferometric SAR data at C- L- and P-band, and its implication in estimating dSWE and SWE compared to direct retrievals. The manuscript is well written, and I really appreciated reading this paper. In my opinion, this paper is of great relevance to the community and highlights the strengths and weaknesses of the different sensors/platforms with respect to snowpack properties on the ground. The authors are very transparent in highlighting that this is not an extensive experiment covering all the possible outcomes of what can be expected in reality, but present the impact of key components such as wavelength, snow state (wet-dry), temporal correlation, signal-to-noise ratio, and phase ambiguity.

This paper also highlights the advantage of Bayesian approaches compared to more direct retrievals described by Eq. 2. This shows that even the Bayesian approaches to retrievals could be more suitable, simply by providing an uncertainty on the retrieved value.

Very minor comments:

1. One thing that is not discussed and could support the discussion is the higher sensitivity to atmospheric phase delays with lower frequencies. This gives more certainty to estimations at C-Band, since atmospheric corrections mostly rely on modelled atmospheric conditions. Maybe adding a line on this in the discussion could help guide future work.

This is a good point. In the revised manuscript we will add text to the Discussion section indicating the frequency dependence of other factors including atmospheric

phase delay. This point can be reinforced by citing the recent Palomaki et al. (2026) paper about non-snow effects as suggested by Reviewer #2.

2. I would standardize the phase units across the text to avoid confusion between degrees and radians. I would stick with the units that is used to convert dSWE to phi.

We attempted to consistently use degrees throughout, but will standardize across the revised manuscript to use phase units of degrees.

3. This might be outside the scope of this study and could be included in future studies is combining the different frequencies. It's been shown that combining C- and L-band phase information greatly improves dSWE/SWE estimates. In this data assimilation framework, I feel it would makes things very interesting showing that this method might/would benefit from the advantages of each frequency while minimizing the down sides.

This is an excellent point and meant to be the focus of future work. While this study focuses on information content of individual frequencies, the framework used could be applied to multi-frequency data assimilation. In the revised manuscript, we will include text in the Discussion section to this effect.

4. If I understood correctly, the conversion from dSWE to phi assumes a homogeneous snowpack. Given that FSM2 can provide up to three layers, it would be nice to have an understanding of the stratification of the snowpack for the three sites. If the snowpacks are very heterogeneous, this also has implications in how phi changes between two observations, which can be a source of uncertainty in this analysis. Maybe adding a line in the discussion about this.

We will clarify the stratification of the snowpack and its impact on phase prediction in the revised manuscript. We do not assume or model a homogeneous (single-layer) snowpack, but use the full 3-layer FSM2 information in Equation (2) by computing the equivalent bulk snowpack density and snow depth from the 3-layer snowpack at the pair of measurement times. For dry snow this is expected to have limited impact due to the transparency of the snowpack at the frequencies examined. An alternative could be to use the 3-layer snowpack directly in Equation (1) for predicting phase. Figure S2 shows that this results in a limited difference and so we chose to use Equation (2) to be consistent with what is typically used in retrieval algorithms where no layering information is available.

5. As for temporal correlation between wet-wet pairs, this assumes there is no major snow surface roughness change between the two acquisitions. A change in surface roughness, especially during melt/refreeze events is not uncommon. This could be

added in the discussion, but I don't feel it is possible to quantify this in the current OSSE.

We agree that it is difficult to quantify the impact of snow surface roughness changes in the current OSSE without adding more ad-hoc assumptions. So, instead we limit ourselves to simple assumptions about coherence time scales (i.e. no correlation beyond 12 days) that allow for pair-wise measurement identification when dry pairs exist. Follow-on work will look at real data where we will be able to assess the impacts of coherence/decorrelation in a more quantitative way. In the revised manuscript, we will add additional discussion on the fact that snow roughness changes could impact coherence, thus reducing the number of available pairs.

Reviewer #2 Comments:

RC2: 'Comment on egusphere-2026-1594', Anonymous Referee #2, 10 Sep 2026

Margulis et al. present an OSSE-based analysis that compares deterministic and Bayesian approaches for SWE estimation using synthetic phase delay time series at three radar frequencies (C-band, L-band, P-band) with 6-day repeat for two experimental cases (perfectly unwrapped vs. wrapped). The experiments are carried out with the FSM2 snow model (for generating “truth” and for assimilation experiments with the particle filter) at three stations in the well-studied Tuolumne River Basin (California, USA) in water year 2017. The analysis demonstrates the efficacy of Bayesian assimilation for SWE estimation for all three radar frequencies, while the deterministic approach yields larger errors, especially when phase change is ambiguous (wrapped between -pi and +pi). Additional sensitivity experiments are conducted to assess the impact of phase error and repeat frequency (6-day vs 12-day).

Overall, this is generally a well-designed and executed study, and provides compelling evidence of the use of Bayesian methods for InSAR SWE estimation. These results are timely and should be of high interest to the snow science community, given the recent launch of NISAR and other ongoing efforts to develop InSAR SWE estimates from existing and proposed SAR missions. While the study is not comprehensive, I think it clearly demonstrates the advantages of the data assimilation approach over the deterministic dSWE retrieval approach. I have several suggestions and questions for the authors to consider as they revise the manuscript.

Comments

1. Fundamental to InSAR is the problem of phase ambiguity, but I do not think the paper clearly explains the concepts of perfectly unwrapped vs. wrapped phase in Section 2. I think the paper would be more readily understood by those who

are not InSAR experts if a conceptual figure was added in Section 2 to explain these central concepts and the problem of phase wrapping to dSWE retrievals and SWE at different SAR frequencies. This new figure might be similar to Fig. S5, but could be simplified to focus on dSWE rather than the individual SWE components (depth and density).

We agree that the issue of phase wrapping and its associated ambiguity on dSWE retrievals is worthy of additional discussion. In the revised manuscript we will use the suggestion to include a new figure in Section 2 that is a simplified version of Figure S5, but with a focus on SWE rather than depth and density separately.

2. On the subject of conceptual figures, I find it odd that the paper's only conceptual figures (Fig. 1 and 2) are somewhat redundant of each other (and with Fig. S3). These figures present the differences in viewing geometry and snow scattering for monostatic and bistatic SAR configurations. Those two configurations do not appear to be relevant to the rest of the paper, which raises the question: why include two conceptual figures to distinguish these two configurations? I think it would be more effective to merge those into a single figure to emphasize the key information or move them to the supplement. Such a change would accommodate the addition of a conceptual figure about phase wrapping (see previous comment).

   The original intent to include Figure 1 and 2 was to communicate that despite the differences in transmitter/receiver geometry the physical basis of SWE retrieval is similar for monostatic and bistatic configurations. Figure 1 was included to show the large-scale viewing differences, and Figure 2 was included to show the specific impact of snow on refraction and phase delay. (Figure S3 was meant to be the wet-snow equivalent of Figure 2, which shows the change in primary reflection boundary.) We agree that there is some redundancy between Figures 1 and 2 and can combine them to make room for a phase wrapping figure (as discussed above in Comment 1) in the revised manuscript.

3. The paper's focus on C-band, L-band, and P-band is logical given existing, new, and proposed SAR missions. However, a common frequency that the paper does not consider is X-band SAR, which is already available from commercial providers (e.g., ICEYE, Capella) at temporal repeat as fine as sub-daily to 3-day. I think it could help to establish (in Section 2.1) why X-band was not considered in the analysis. Like C-band, I assume X-band would have high SNR but even greater phase ambiguity and more limited penetration into snow. However, the phase wrapping issue may be mitigated somewhat by the higher temporal repeat (1-3 days) of current X-band commercial systems.

This is a good point and the method described in this paper is meant to be applicable to any frequency suitable to InSAR, including X-band, S-band or others. In the revised manuscript, in both Section 2.1 and the Discussion we will make clear that other frequencies could be used (including X-band) and what the implications of those frequencies are relative to those examined in more detail in this paper.

4. The study focuses on a single study basin (3 stations) and a single year (wet 2017 in California), and as a result only assesses the case of very deep snowpack (1300-2000 mm peak SWE). I appreciate that the authors acknowledge these deep conditions throughout the paper. However, I think it would be useful if they expand the discussion about the potential/limitations of the methodology in more moderate or shallow snowpacks, and how/whether the three SAR bands might perform differently.

   This is a good point. In the revised manuscript we will more clearly describe how the results would be expected to change in more moderate or shallow snowpacks and across the three SAR bands. We expect that the primary difference would be linked to fewer wrapped phase events in cases where the dSWE values between pairwise measurements are smaller.

5. The OSSE framework leverages the FSM2 snow model with a maximum of 3 snow layers. What is the impact of the number of snow layers on the model results? The supplement document indicates that too few snow layers may result in an underestimation of wet snow effects, which is important to the InSAR experiments. At a minimum, it would help to include some review and discussion on the influence of snow model vertical complexity (e.g., Webb et al., 2022; Cristea et al., 2022; Decharme et al., 2016). More ideally, a sensitivity experiment of LWC to number of snow layers could be useful (note: FSM2 allows the user to change the maximum number of layers).

   In the revised manuscript, we will include some discussion of the impact of snow model vertical complexity on bulk snow state estimation, including the references to the suggested work. While the layering is likely to have an impact on the simulated snow states and their accuracy, in the OSSEs constructed in this paper the focus is on whether the retrieval and Bayesian estimates can reproduce the true (synthetic SWE). Hence we think the impact of layering on the true estimation error might be better addressed in the context of future work with real data where we can compare estimates (using different layering schemes) to a ground-truth value. In the revised manuscript we will include discussion on the need for examining this in future work with real data where

different layering schemes could be used to estimate dSWE and SWE that could be compared to ground-truth estimates.

6. A consistent result from the OSSE experiments was that the Bayesian approach typically yielded estimates that had little or no bias. To what extent is this an artifact of the OSSE design? Namely, the "truth" in the OSSE setup was an FSM2 simulation with a precipitation bias correction factor identified for each site, while time-varying errors (e.g., random errors) in precipitation were not considered (L. 533). In the current setup, missing observations (e.g. wet snow in the early season) may therefore have similar precipitation errors as other times when observations are available. One could argue that temporally varying precipitation errors might be included as a side analysis (similar to Section 4.3.1, analysis on phase error magnitudes).

   For the case of unwrapped phase (Figure 10), both the retrieval and Bayesian estimates have negligible bias. This is likely optimistic for both estimates, where systematic errors are likely to enter into the estimation procedure in various ways. For the OSSE, the baseline case is useful to see how bias may change as a result of degrading key aspects of the measurement process (larger measurement error, longer repeat frequency). With respect to the modeling of truth and prior, the choice of a constant precipitation bias correction factor was motivated by previous work with a different methodology (Fang et al., 2022) whereby a constant precipitation bias correction factor model was assumed and shown to successfully estimates SWE. That said, it is invariably a simplifying assumption. Including such time varying errors (which are not well known) was thought to be too ad-hoc for inclusion in the "truth" and therefore not examined in the paper. We think this too would be best investigated with real data in future work. In the revised manuscript we will highlight this specific line of research to be investigated with real data and mention how the OSSE setup is a simplified version of reality that likely has smaller biases than would exist with real data.

   Fang, Y., Y. Liu, S.A. Margulis, 2022: A western United States snow reanalysis dataset over the Landsat era from water years 1985 to 2021, Scientific Data. https://10.1038/s41597-022-01768-7

7. The experiment is conducted independently (i.e., point or pixel mode) at the three sites, but does not consider that data availability/uncertainty may vary in space and how spatial relationships may be leveraged when only certain parts of a basin have data gaps. A prime example is Experiment #2 with C-band, where the Bayesian approach yields large overestimates of SWE during an early season gap in measurements at PDS (Fig. 12 top) when there is no such measurement

gap at DAN and SWE estimates have low error (Fig. 11 top, L 590-598). I think it would be helpful to include some discussion on the possibility of adapting the Bayesian SWE estimation approach in a spatial framework in future work.

This is a good point and one we are targeting for future work. In the revised manuscript we will make this more explicit that the spatial correlations in measurements and their errors have the potential to gap-fill in space and time.

8. The determination of wet vs. dry vs. mixed snow from LWC is unclear. The supplement includes some of the background theory on how the complex permittivity is calculated while considering LWC (equation S4). I assume that dry snow is LWC=0%, but I am not sure how the paper classified the wet and mixed states and how this classification is different for each SAR band (Fig. 6). This is important because it determines the availability of phase observations in the experiments. Is it necessary to distinguish between mixed and wet, or could these be merged into a single wet snow class? It seems that for both of those cases, there are no phase observations available.

   We will clarify this in the revised manuscript. To provide more context, in Figure 6, the "mixed" and "wet-wet" cases could be considered a single class that does not yield a usable pairwise measurement. The distinction is made primarily to emphasize that there is often a transition, where one member of a pair is dry and one is wet before fully transitioning to wet snow. Under those conditions it would be expected that there would be limited coherence and therefore no useable measurement pair. The other point is to show that there is some frequency dependence of this effect. The layered phase equation (represented by Eq. (1), Text S2), which is a more general form of Eq. (2), is used as a diagnostic to identify whether the phase response behaves like the expected dry-snow case of Eq. (2). As LWC increases in the snowpack, there will be certain LWC profiles where the phase response instead behaves like the wet-snow case of Eq. (S9). In the manuscript we label such conditions as "sufficiently wet" snow. If either member of a pair exhibits sufficiently wet conditions it is assumed there will be no measurement (i.e. coherence would be too small). However, longer wavelengths (i.e., P-band) can penetrate through slightly wet snow more easily than shorter wavelengths (i.e., C-band). So, at low LWC values, there are cases where the phase response still behaves like that described by Eq. (2). In all the pairs examined in the OSSE, this never happened at C-band and only happened once at L- and P-band (see Figure 6 for the DAN site), where around April 1st there is a pairwise measurement available at L- and P-band (that is not at C-band) due to limited LWC at that time. For real data, the coherence would be used to make this identification of a useful pair and so the method used in the OSSE (which cannot predict coherence) is a simplification.

9. In general, I think many figures might be better presented if the panels were rearranged to align common variables (e.g., so errors can be more easily compared horizontally). Currently, some adjacent panels have SWE errors and dSWE errors, which are not directly comparable. I identified a few figures (see below) that may particularly benefit from this type of redesign. Where possible, it would help to match the limits of these common axes so their relative magnitudes can be easily understood by the reader.

   This is a good point. We will include these changes in the revised manuscript as described below in response to specific Figure suggestions.

Line Comments

- 148-150: Could consider citing the recent Palomaki et al. (2026) paper that attempted to quantify these non-snow effects on phase change for InSAR dSWE measurements.

  This is a good suggestion. We will cite this as part of the discussion as to how non-snow effects influence phase change. This also supports the suggestion by Reviewer #1 above (their comment #1).

- 246: Why was this nominal error value selected for phase measurement error? What types of effects might induce a 20 deg error? More generally, is it reasonable to assume that all three radar bands would have a similar uncertainty? I understand the need for the same error in the context of the experiment, but it would help to elaborate whether a similar phase error would be expected in practice for each of the three bands.

  In practice the measurement errors are expected to vary in time, space, and across frequencies (due to non-snow effects, coherence, and phase wrapping) as acknowledged on lines 157-158 and Text S3. We would argue that attempting to synthesize these complicated error structures in the context of an OSSE would lead to many ad-hoc assumptions that would be hard to justify. The OSSE study in this paper is best thought of as a parametric study meant to illustrate how the information content in phase measurement can yield dSWE and SWE estimates at different frequencies and under different measurement errors. The measurement error ranges examined (i.e., from 10 – 60º) were chosen to span from higher to lower signal-to-noise ratios (i.e., relative to a +/-180º phase wrapping cycle). The 20º nominal case was chosen as an illustrative baseline case within this range where the retrieval and estimation would be expected to perform reasonably well. The frequency dependence of the signal-to-noise (especially in the more straightforward unwrapped case) was discussed when presenting the results (e.g., lines 440, 458, 607, 644, 679, 690, etc.) . In the

revised manuscript, we will more clearly discuss this issue and point to the fact that more realistic error structures will be best examined in the context of real data (already mentioned on line 762).

- 292-294: Please provide quantitative justification for the assumption of decorrelation after 12 days.

  This is an acknowledged simplification of the decorrelation process (as discussed on lines 160-167) and used for all three frequencies to keep interpretation simple by keeping the total number of observations of similar order across frequencies at a given site. In practice the decorrelation time will be a complicated factor of space, time, and frequency, but one that can be diagnosed from the coherence of the measurements themselves. In the case of the OSSE, where we cannot predict coherence, we must make a simplification aimed at capturing the first-order impact of decorrelation on limiting the number of available measurement pairs. The 12-day upper limit to correlation is based on literature values for C- and L-band showing sufficient correlation up to 14 days (Ruiz et al., 2022 as cited on lines 163-164). In theory, P-band would have longer decorrelation timescales. The specific choice of 12 days was based on the experimental revisit times of 6- and 12-day repeats. Choosing an integer multiple of these values leads to measurements pairs for every 1 or 2 sampling intervals and 12 days was in the range of values seen in Ruiz et al. (2022). It is expected that in future work with real data, variations in decorrelation between different frequencies could be more definitively examined. In the revised manuscript we will clarify this.

- 321: Suggest adding "(i.e., equal weights)" after "1/N".

  This edit will be made in the revised manuscript.

- 399: What is "significant LWC"? This wording seems subjective / inexact.

  There is no single threshold to LWC that makes it "significant", but it is used here as described in response to Comment 8 above to identify whether the snowpack is "sufficiently wet". We will clarify this in the revised manuscript.

- 444-446: Similar to previous comments, it is not clear that the LWC sensitivity is fully established.

  This is also related to Comment 8 above. We will clarify this in the revised manuscript.

- 467-469 and Fig. 8: This result may not be obvious as it relies on the reader to notice that the limits of the x-axis vary by panel. I would suggest revising Fig. 8

to flip the panels to align the common dimension (b value), and match the axis limits. This will make it more obvious that the posterior b-value distributions are quite narrow for C-band and progressively wider at L-band and P-band.

This is a good suggestion that will be implemented in the revised manuscript.

Figures and Tables

- Table 1 – suggest adding a row specifying the wavelength, as some readers may think more in wavelength space rather than frequency space.

This is a good suggestion that will be implemented in the revised manuscript.

- Fig. 10 – suggest stacking these panels vertically to align the common variable (retrievals errors vs. Bayesian errors).

We will implement this in the revised manuscript.

- Fig 10 and Fig. 13 – suggest merging into a single figure, where the left panels are the unwrapped case and right panels are the wrapped case (and top is dSWE errors and bottom is SWE errors). This would allow a more direct comparison and make it obvious how the unwrapped and wrapped cases yield different error magnitudes. The current configuration requires the reader to flip back and forth between Fig. 10 and 13 and remember the values.

The primary reason the two figures were separated was to separate the narrative to focus on Experiment #1 results first and Experiment #2 second. In the original manuscript, we included the horizontal green line in Figure 13 as a reference value to Figure 10 so that the reader didn't have to flip back and forth. But perhaps the comment is indicative that the reference line is difficult to see. Our preference is to keep the figures separate for narrative flow, but in the revised manuscript we will use a more noticeable line color to indicate the reference values from Figure 10.

- Fig. 14 – Suggest flipping to 2 rows (dSWE errors on top, SWE errors on bottom) by 3 columns (bands). This would align their common information (y-axis errors) for ease of comparison.

We will implement this in the revised manuscript.

- Fig. 16 – suggest flipping so dSWE errors are the top two panels and SWE errors are the bottom two panels.

We will implement this in the revised manuscript.

References

- Cristea, N. C., Bennett, A., Nijssen, B., and Lundquist, J. D.: When and Where Are Multiple Snow Layers Important for Simulations of Snow Accumulation and Melt?, Water Resources Research, 58, e2020WR028993, https://doi.org/10.1029/2020WR028993, 2022.
- Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P., and Morin, S.: Impacts of snow and organic soils parameterization on northern Eurasian soil temperature profiles simulated by the ISBA land surface model, The Cryosphere, 10, 853–877, https://doi.org/10.5194/tc-10-853-2016, 2016.
- Palomaki, R., Hoppinen, Z., and Marshall, H.-P.: A spatiotemporal analysis of errors in InSAR SWE measurements caused by non-snow phase changes, The Cryosphere, 20, 2703–2721, https://doi.org/10.5194/tc-20-2703-2026, 2026.
- Webb, R. W., Musselman, K. N., Ciafone, S., Hale, K. E., and Molotch, N. P.: Extending the vadose zone: Characterizing the role of snow for liquid water storage and transmission in streamflow generation, Hydrological Processes, 36, e14541, https://doi.org/10.1002/hyp.14541, 2022.