We would like to thank Jorge Jorge Ruiz for his very useful review. We appreciate the reviewer's time and effort. We believe the manuscript is now in a much better shape. We hope this manuscript help the community in future approaches for SWE retrieval missions. We would like to point out that we have changed the reference point from the average of all in situ stations to the average of two stations SWE change with reliable time series. This will address your concern about reference point.

Please find our responses below.

RC1, Referee #2, Jorge Jorge Ruiz:

The manuscript from Oveisgharan et al presents the results of the analysis of Sentinel1, C-band, InSAR timeseries over Idaho for Snow Water Equivalent (SWE). The generation of the InSAR derived SWE timeseries over the SNOTEL stations was demonstrated using them as reference for phase. Additionally, the total SWE retrieved over several months showed great agreement with SWE when compared to LIDAR snow depth. The work is of interest for the snow community focusing on InSAR SWE retrievals. The results presented in this manuscript would be useful for future satellite missions such as NISAR or ROSE-L. Furthermore, it also demonstrates the technique using Sentinel1, that will operate along ROSE-L. Congratulations to the authors for the great work. I have some concerns I think should be addressed before publication.

- Thanks for your useful review and compliment! Your feedbacks improved the paper a lot.

<u>General Comments:</u>

1. I think an interesting addition would be a figure (maybe in the annex) with the time series of in situ vs retrieved SWE for all stations. This is at your consideration. Also, in Figure 2.b add the number of the stations, at least the ones you used for Figure 9.
    - We added the station numbers in the google map too. The requested figure would be similar to figure 6(a) with fewer point in each plot and won't add much information. The correlation and RMSE of it is reported in figure 6(b). So, we decided not to include it.
2. Regarding the calibration, while I agree that the large number of stations reduce the bias, I am a bit concerned that this method can "mask" unreliable phase measurements. Take as an example the interferogram presented in Figure 5 (c), where most of the stations have very low coherence. If you do an averaging around the low coherence stations (considering homogeneity and a sufficient large number of pixels), phase is noisy and the expected value should be zero, which translates to 0 DeltaSWE. If now you add the mean DeltaSWE from the stations, you'll probably get something closer to the in situ value for that station, although the phase had no usable information. I'm not asking to change the methodology, just comment on this. Additionally, adding error bars with standard deviation in Figure 9 would be useful to assess the effect.
    - We have changed the calibration method so it would be easier for the readers. The point that you brought up is valid however, using the average of all stations remove that kind of errors as we have stations with different sort of temporal

coherence and SWE change. Note that we don't include stations with low temporal coherence for that average calculation. Having said that, we now are using 2 stations with very good correlation and temperature for the entire time series and it also should not have the problem you mentioned.
    - Regarding the error bars to figure 9, I am not sure what you mean. In this figure we are just adding the delta SWE to get total swe at each observation. What std are you referring to?
3. See comments regarding Section 6.
    - Done!
4. Perhaps it is a bit picky but be more consistent with coherence and correlation along the manuscript (e.g., lines 196, 205, F5 caption, 213… these could be coherence or temporal coherence). I am also missing some explanations why you neglect other sources of decorrelation and assume that the product is just the temporal component. Also, when discussing temperatures, I'd use 0°C over zero.
    - We thought we captured all coherences! Thanks for pointing that out. We went through the paper again. Hopefully, now it is consistent.
    - We changed all zero to 0°C.
    - We added this explanation for figure 4: "Bottom row of figure 4 shows the coherence of the images in top row of figure 4. Interferometric decorrelation has different sources, such as temporal decorrelation, volume decorrelation, signal to noise ratio decorrelation, geometric decorrelation, ... The volume decorrelation is small due to relatively small Sentinel-1 perpendicular baseline. Temporal decorrelation is the dominant source of decorrelation. For the rest of this study, we assume the observed interferometric decorrelation is approximately the temporal coherence."

<u>Specific Comments:</u>

<u>Section 1:</u>

Good introduction. I am missing perhaps some mentions to co-polar phase difference for snow height estimation.

- Added: "The co-polar phase difference (CPD) between VV and HH polarization of X-band SAR acquisitions is used for estimating the depth of *fresh* snow (Leinss et al., 2014)."

Line 15: wouldn't mass be more precise than cover here?

- Changed it to "snow mass and cover". We need cover to know where it is and yes how much it is.

Line 50-56: perhaps add a few details. Phase center at the snow surface or the volume. I got a bit lost in this paragraph, when you say this method do you refer to the method for dry or wet snow? A sentence clarifying this could be useful.

- We shift the sentences to make it clearer: "The phase change of specularly reflected signals in Signals of Opportunity (SoOp) is shown to be strongly dependent on SWE changes for dry snow (Yueh et al., 2017, 2021; Shah et al., 2017). The theory behind using SoOp for SWE retrieval is similar to repeat pass interferometry that is explained in section 2. The advantage of this method is that the stratigraphy of the snow has little impact on the SWE retrieval (Leinss et al., 2015; Yueh et al., 2017) similar to SWE retrieval explained in section 2. Using the long wavelength signal at P-band in SoOp is very helpful for addressing the loss of temporal coherence and phase unwrapping challenges of this method. However, the phase sensitivity to SWE changes decreases at lower frequencies. There has been very limited data showing the success of this method at P-band. Achieving high resolution is another challenge of this method. The phase change of specularly reflected signals in SoOp is dependent on snow depth change for wet snow (Yueh et al., 2017, 2021; Shah et al., 2017)."

Line 55: remove second "this method".

- Changed the sentence now, Done!

Line 57: remove small.

- Yes, done!

<u>*Section 2:*</u>

In Section 2.1 you could add that snow is a thermal insulator and it reduces temporal changes in the underlaying layers. Is this relevant for mountainous areas?

- We add this sentence: "On the other hand snow cover has a thermal insulation effect on the ground and underlaying layers (Gu et al., 2019). The insulation increases with the snow depth. Therefore, during the snow season we assume the ground remains frozen even when snow becomes wet. Hence, temporal decoherence from the ground is negligible."

Line 68: I think the idea is clear without "very" and "small"

- Fair enough, done!

Line 72: higher than what? or relatively high frequencies.

- Changed it to "relatively high frequencies such as C-band".

Line 75: Extend the equation to include . I guess makes some sense since you are focusing on SWE in the manuscript. Also details that snow permittivity depends on its density and can be approximated from the density and that the complex part is negligible for dry snow.

- I think the sentence is incomplete here. To include what? But your argument is right. We are focusing on dry snow that imaginary part of e is small and permittivity depends on mainly density.

Line 93 to 96: is this the correct place for explaining what you have done, or should it go in the introduction?

- We explained briefly in introduction section what our method is. However, we needed a section to explain what had been done and then explain what we did different or compliment to what have been done. I include this sentence to introduction section to highlight the difference over there too: "We show for the first time that SWE estimation using repeat pass interferometry performs well by using a *long* time series of Sentinel-1 interferometric data in winter 2021."

Line 101: two different times, for generality.

- Done!

Line 102 to 103: maybe make this idea about decorrelation more general, not just for vegetation.

- We changed it to: "However, the movements of the scatterers such as leaves and branches or sea ice particles decrease the temporal coherence."

Line 105: Consider changing faster sampling for shorter temporal baseline, sounds more connected to repeat pass interferometry.

- We changed it to: “Methods such as using two frequencies or shorter revisit time are used to overcome these problems.”

Section 3:

Would be nice to explain that at the time of the study, Sentinel-1 operated a constellation of two satellites. You can mention that each of these satellites had a repeat pass of 12 days, but due to the orbit offset between them, the effective temporal baseline is 6 days.

- We add this sentence to the paper: “Sentinel-1 constellation includes Sentinel-1A and Sentinel-1B. These two satellites are in the same orbit with a 180^0 orbital phasing difference. The revisit time for each of the satellites is 12 days. However, revisit time can get to 6 days if both satellites make observations.”

I think Figure 2 deserved better visualization. The caption should mention that the green frame is the S1 image (as you did in Line 160), and what are red and green diamonds.

- We changed the caption to: “© Google Earth View (a) Google Earth View of Sentinel-1 path:71, frame:444 in Idaho. (b) zoomed to the Sentinel-1 path:71, frame:444, shown by big green rectangle. Red boxes show the location of LIDAR data acquisition. The green diamonds show SNOTEL stations with ΔSWE error less than 2cm in the entire time series. The red diamonds show SNOTEL stations with ΔSWE more than 2cm in at least one observation in the time series. Yellow squares are SNOTEL stations 1 and 11 used for reference point.”

Line 143: … at a central frequency of 5.405 GHz

- Done!

Line 145: comment that you can get S1 data at the Copernicus Open Access Hub :)

- Added: “The data are free and available through Alaska SAR Facility (ASF) or The Copernicus Data Hub distribution service.”

Line 167: … for each of the SNOTEL stations.

- Done!

Line 209: change squares to diamonds?

- Good catch! Done!

Figure 3, (a): Change x label to “Day from 12/01”.

- Done!

Section 4:

Line 185: are these sources of error general for all applications or just particular to snow?

- These are main sources of error for any application using Sentinel-1 data. We modified it to: “The main sources of error in the science and applications using Sentinel-1 repeat-pass interferometry are”

Line 195: be more specific specifying that it is data from SNOTEL stations and that this data is snow measurements.

- We changed it so it better captures your concern: “The temperature is also an important factor. Equation 1 is valid for dry snow ((Leinss et al., 2015), and we use near surface air temperature above 0^0 C as a metric that indicates wet snow in snow season. Any SNOTEL data with in situ near surface air temperature more than 0^0 C is unreliable in our study."

Line 213: this idea about temporal coherence being a limitation of InSAR is repeated a couple of times in the paper. Consider removing this one.

- Agree, done!

Line 214-215: I don't understand the point of this sentence. Isn't it as short as possible? Or would you discuss limitations related to satellite capabilities?

- We removed that sentence as it is confusing. You are right. Our point is to use different data set to find the needed revisit time for future snow missions. We cannot afford as short as possible (every minute?) for a mission ☺

Line 221: water vapor content between passes.

- Done!

Section 5:

Line 239: is results the correct word here? Or LIDAR data/scan?

- Changed to LIDAR data.

Somewhere in this section you should state what is the window (number of pixels) used to extract the values of coherence and phase for the stations.

- We add this to this section: "As mentioned in section 3.1, the resolution of the Sentinel-1 InSAR data from Hyp3 is 80mx80m. We used a 10x10 multi-looks window of retrieved SWE and temporal coherence around the SNOTEL locations to reduce the speckle noise. Therefore, we compared the SNOTEL SWE with the 800mx800m retrieved SWE around the SNOTEL site. The heterogeneity of the environment such as vegetation cover, vegetation fraction, land type, and SWE distribution in the 800mx800m around the SNOTEL station affects our accuracy. We will analyze the effect of the heterogeneity of the environment on the SWE retrieval for SNOTEL stations in the future work of this study."

Line 255: change discovered to observed?

- Good suggestion, done!

Section 5.1.1: I would like to read some discussion about the points with error=0

- Some of the points were out of Sentinel-1 frame but in the square. Therefore, there was not data over there. We removed those stations since they were not actually in the frame. So, instead of 43 stations we have 31 stations in the new figures.
- There is no more zeros in figure 6(b) except for station 4 that we added this explanation: " Note that station 4 has just one observation with temporal coherence more than 0.35. That observation is the first observation with zero SWE change. Therefore, there is not enough points to calculate ρ_i. Hence, RMSE and correlation are zero."

Figure 7,c): are the dots placed on the reference acquisition?

- We add this sentence to clarify: "Note that the labels on x-axis show the first date of each interferometric observation."

One thing that comes to mind is if you can say something before the 12/25/20 acquisition (as it is the common one between the two interferograms you show) that could explain the noisy phase. E.g. wet snow, a storm…

- Thanks for the suggestion! We added this argument to the paper: “We observe that 4 out of 6 days between 12/19/20 and 12/25/20 (observation 4) are relatively warm including day 12/19/20. All 31 stations have temperature between -7^0C to 6^0C at 6 am in those four days. The warm days cause a lot of melting and refreezing in those 4 days. Hence, we expect to have small temporal coherence and consequently noisier fringes. On the other hand, the temperature is relatively warm only on 12/26/20. The rest of the 5 days between 12/25/20 and 12/31/20 (observation 5) are mostly colder than -7^0C for all 31 stations. Therefore, higher temporal coherence and less noisier fringes.

Line 261: Do you mean observations 5 and 4? Otherwise, I don’t think we know where stations 5 or 4 are…

- Yes, we meant observation not station, thanks for catching it, it is fixed now.

Section 5.1.2:

Line 269: As mentioned in Section 5.1.1? or in Section 4? But check this sentence if I understood correctly you use DeltaSWE values when temporal coherence is below 0.35 and temperature higher than 0°. So, does this mean all values?

- It was explained better in section 4 but also in section 5.1.1. On the other hand, there was a typo here. We only used data with high temporal coherence and temperature lower than 0^0C. Here is the new sentence: “However, as mentioned in section 4 and 5.1.1, we only used ΔSWE values with temporal coherence more than 0.35 and temperature less than 0°C.” We discarded high temperature or low coherence data.

Please indicate that you set SWE_t1 as 0. Some stations have snow already for the first acquisition. I think you need a comment on this.

- We add this sentence to clarify: “Note that the SWE(t_i+1) is measured compared to SWE(t_1). For simplicity, we assume the SWE at time t_1 is equal to zero.”

Figure 9: Do the timeseries expand the same time span. I think they do but the x axis is different for the three subfigures… Also, drop some of the decimal precision? See comment about std for each acquisition.

- We changed the time to observation number with viable retrieval (coherence >0.35 and T<) to avoid confusion. We explained it in the text: “We had 18 observations for the entire time series. Discarding some observation due to low temporal coherence or high temperature, changes the time series length. As seen in figure 8, we keep all 18 observations for station 20 but only 15 observations for station 12.”

Line 295: could you be a bit more precise about how a station is labeled as red or green? What is the threshold for it? I see this is explained in the caption of Figure 10, but shouldn’t it be in the text also?

- We add this threshold to the text and corrected in the caption: “The SNOTEL sites are shown by small diamonds in figure 2(b). The green small diamonds have a total SWE error less than 2 cm in the entire time series, similar to stations 12 and 30. The red diamonds have a total SWE error more than 2cm, similar to station 20. However, the retrieved SWE has a similar pattern as in situ SWE.”

Section 5.2:

I think the analysis here could be extended. Is there anything that is increasing the error like steep slopes or vegetation (although from figure 2 seems there is not much vegetation)? Could you add some comments about coherence for those images? Was it generally conserved? Another point just out of curiosity, are the results similar if you compare the LIDAR map to S1 SWE map from 03/13/21? Looking to Figure 3.b) does not seem there were much SWE accumulated between the pass dates.

- Thanks for the suggestions! The section 5.2 is expanded now. We are working on a second paper to better discuss the effect of different parameters such as vegetation, slope, temperature, … . It is beyond the scope of this work. However, we added two figures to show the mean of temporal coherence for these two sites and how it affects the correlation.
- We also added a figure showing the correlation between LIDAR snow depth and retrieved SWE for different observation dates.

Line 303: Add the complete date of the LIDAR scan.

- We add the flight date: "The terrain DEM is measured by LIDAR sensor during September 2021. The DEM is used to measure the snow depth using the LIDAR data collected on 3/15/2021."

Line 305: isn't the closest date the 03/13/21?

- Yes, but we wanted to cover the entire time series. We can go with 3/13 or 3/19. The results doesn't change significantly.

Line 311: total SWE accumulated rather than change?

- Corrected!

Perhaps indicate in the text that in Figures 11 and 12, (c) what does color represent?

- We added: "The 2D-histograms of these two images are shown in figure 10(c) and 11(c) where x- and y-axis show the LIDAR snow depth and Sentinel-1 retrieved SWE, respectively. The colors in part (c) shows the $10^{\text{number of cells}}$ with LIDAR snow depth x in part(a) and InSAR SWE y in part (b).

<u>*Section 6:*</u>

I think this section is a bit short and does not discuss important aspects of the work. You could emphasize that the total SWE vs LIDAR was done accumulating the contribution from many interferograms, and the great spatial match. This to me is a truly impressive result. On the other hand, you should also need to discuss the calibration strategy, in my opinion this plays a crucial role in your results you cannot overlook commenting on it.

- We expand this section more to cover referencing problem and LIDAR results better: "We showed that retrieved ΔSWE using Sentinel-1 is highly correlated (0.8) with in situ values, with an RMSE of 0.93cm. For reference point of interferometric phase, we used two in situ stations with temporal coherence more than 0.35 and temperature less than 0°C for the entire time series. We subtracted the difference between the average of in situ and retrieved ΔSWE of these two stations from retrieved values to calibrate the retrieved ΔSWE. The ΔSWE RMSE error is less than 2 cm for all stations and less than 1 cm for most stations. The correlation between retrieved and in situ ΔSWE is more than 0.6 for most stations. The ΔSWE retrieval performance degrades for days with small ΔSWE. We demonstrated that low temporal coherence not only degrades the SWE retrieval performance, but also the unwrapping algorithm performance. We showed that big

melting events between two Sentinel-1 acquisitions make the interferometric fringes noisy and unwrapping algorithm challenging.
The retrieved total SWE has less than 2cm RMSE compared with in situ values for 9 stations and more than 2 cm for 14 stations.
The highlight of the results of this study is the similarity between two independent measurements, retrieved SWE using Sentinel-1 data and LIDAR snow depth data. We used Sentinel-1 data between 12/01/20 to 03/19/21 to retrieve ΔSWE time series. By adding the entire time series of ΔSWE, we calculated the total SWE on 03/19/21. Total retrieved SWE using Sentinel-1 interferometric data and LIDAR snow depth images over two regions in Idaho show similar patterns and are correlated by more than 0.47. We showed that the correlation is higher for regions with higher temporal coherence in Banner Summit. … It is also shown in this study that melting due to warm temperature reduces the temporal coherence and the performance of unwrapping algorithm."

Line 324: Not completely sure about this. Definitely short temporal baselines will help with coherence but if between passes there is a snowstorm most likely coherence will drop significantly.

- Analyzing the temporal coherence in this study shows that temporal coherence decay exponentially with swe change. So, you are right that big snow storm decrease the temporal coherence but with smaller temporal baseline the swe change also decreases. As time is correlated to swe change in the snow season.

Technical Comments:

Line 191: Section 5.12, I think

- We don't have section 5.12. I am not sure what you mean here. We used all the data in section 5.2 even the ones with low coherence and high temperature.

Line 196: Section 5.1.2, I think

- I am not sure. We meant in section 5.2. Note we removed section 5.1.2 in the new version of the paper. "We used all the data in section 5.2, including the data with low coherence and high temperature."

Line 214: InSAR (capitalization)

- We removed this sentence from the new version of the paper.

Figure 7. (a): Path:71. Why 2021? Should not be from 2020 to 2021?

- Corrected: "Retrieved ΔSWE using Sentinel-1 interferometric phase versus in situ ΔSWE for all the stations with temporal coherence more than 0.35 for the entire Sentinel-1 time series from December 2020 to March 2021."

Line 256: InSAR(capitalization)

- Corrected!

Line 333: NISAR acronym already introduced.

- Removed, thanks!