Review of "Mapping seasonal glacier melt across the Hindu Kush Himalaya with time series SAR" by Scher et al.

This study uses Sentinel-1 SAR data to map the seasonal melt characteristics (melt onset, freeze onset, and melt duration) across the Hindu Kush Himalayas. Using this dataset, the study investigates spatio-temporal variations across subregions and as a function of elevation. Two sets of automatic weather stations are used to validate the interpretation of these studies and highlight where glacier ablation models would perform poorly. Given the lack of in-situ observations and seasonal remotely sensed observations in this region, this type of dataset would be very valuable to modeling glacier melt. The use of systematic C-Band SAR observations in this region also appears to be the first time this has been done now that the satellite imagery is available.

While I believe this would be a valuable contribution to the field, I unfortunately believe there to be several major issues with the interpretation of the dataset. The primary issue being the interpretations of the ablation area, which the authors note is challenging and/or were excluded; although this exclusion is unclear because analyses suggest that all area were included. Furthermore, for a study of this scale, I expected there to be some type of validation to support their findings. Instead, the two sets of automatic weather stations were partially used for validation and partially used to highlight where melt models would perform poorly. I read the author's response to reviewer's where they state that this is not possible, except for the two automatic weather stations; however, I don't agree with this assessment. For example, other relevant studies (some pointed out in the review below), in-situ measurements, and/or other sources of satellite imagery (other microwave datasets, optical datasets, etc.) would have been useful to support the interpretations and conclusions.

In my opinion, there needs to be considerably more validation. I would also suggest that unless the methods can be improved to handle the challenges in the ablation area (which would need to be shown through a rigorous validation process to provide confidence), that the ablation area be excluded entirely, and the results be limited solely to the timing of melt in the accumulation zones. Another alternative is that if debris-covered areas are the only problem, then limit the study to clean-ice glaciers only. Either way would still provide useful information to modelers, although this would greatly reduce the novelty and impact of the study overall. At that point, the novelty of the study would need to be reconsidered. I therefore recommend the paper to be reconsidered after major revisions. Please see my major comments and minor comments below.

<u>Major Comments:</u>

The methods show results (e.g., Figure 4 middle; Figure 5). If these datasets were used to develop the method or solely for validation, then this should be explicitly stated. Either way, the assessments of the methods should be in a separate results section. This performance assessment could then focus on validating the methods before the large-scale assessment of spatial trends.

The z-score metric, Figure 4 middle panel, is indicative of the magnitude of the radiometric response that we are isolating. This information is key when defining our methodological approach. Similarly, Figure 5 is used to confirm our interpretation of the radar signature in our methodology. We use the results section to communicate the major outcomes of the paper, a description of regional timings and extent of glacier surface melting.

I found the references to support various statements and relevant work needs improvement. I have highlighted many examples in the introduction, but also believe this is necessary for validating the interpretation of the SAR signals as well.

Thank you for your comment. We have addressed this in lines L34, L29-46, L47-49, L65-68, L73, and L139-141. We also extended the discussion on SAR backscatter over ablation areas with this addition to the manuscript at L118: **"Like in the accumulation zone, the surface melting response in the ablation zone will dominate the seasonal trends in backscatter because of absorption from liquid water at the surface over both bare-ice and debris-covered portions of ablation areas. Although the absolute fraction of backscatter at C-band frequencies over debris covered portions of ablation zones attributed to volume scatter is not well known, there is evidence that for low frequencies it can account for a majority of radar observations (Huang, et al. 2017)."**

The method does not appear to work over ablation areas. It is unclear from the results presented if this is an isolated issue with debris-covered areas or if it is an issue with clean ice as well because the example figures were only for debris-covered glaciers. Given debris-covered areas are highly prevalent in HKH, this is a major issue. This is clearly apparent from Figure 7 Top. The Khumbu Glacier is debris-covered below the Khumbu icefall. The interpretation of the SAR signal on this debris-covered area is that it indicates refreeze onset between DOY 250-270, which is in September. This is highly inaccurate and highlights the lack of validation within the study. For example, this part of the debris-covered ice is clearly still melting in September (Rowan et al. 2020; Journal of Glaciology). Similar issues appear to persist with the Freeze Onset in Figure 8C,D.

We find a strong seasonal loss in radar backscatter over glacier surfaces in all regions and elevation ranges of the HKH as illustrated in Figure 4B. Given our understanding of the physical radar response at these frequencies, these signals are resultant from absorption by liquid water at or near the glacier surface and are not easily explained by other physical processes. Our analysis excludes locations that do not satisfy the z-score metric ($z > 2$). Glacier surfaces in the ablation zone, including areas of debris cover, are included in our analysis because they exhibit a radar response indicative of surface melting and satisfy the z-score metric.

A strong reduction in surface backscatter is observed seasonally over the Khumbu Glacier, above and below the Khumbu icefall, during 2018, 2019 and 2020. These radar signals satisfy the z-score metric and we estimate the average yearly refreeze onset to be September 17the (DOY 260). Rowan et al., (2021, Journal of Glaciology) include a record of surface and below-surface temperature measurements on the Khumbu Glacier for the year 2014. It is apparent in Rowan Figure 2 (below) that mean daily air temperatures around the Khumbu icefall (KH4) are consistently below zero in late September, 2014. This is similar to the refreeze timings especially given the 12-day repeat of the SAR observational record. Although ablation is observed to persist until October 22 at KH4, there is a disconnect between surface conditions (e.g. air temp.) and melting at the debris/ice interface. It is not surprising that melting under thick layers of debris cannot be observed directly using radar. A rigorous comparison with this data is extremely difficult since these measurements are only from one year and observations are not contemporaneous to the Sentinel-1 observational record.

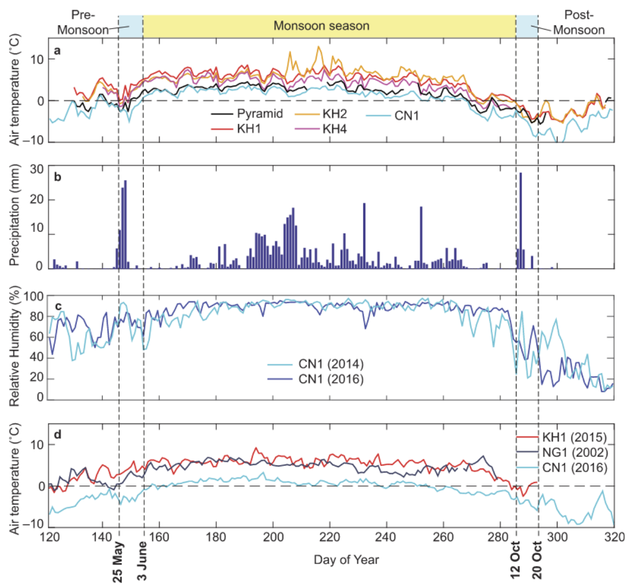

Fig. 2. Daily off-glacier and on-glacier meteorological data. (a) Mean daily air temperatures measured at Khumbu Glacier, Changri Nup Glacier and the Pyramid Observatory in 2014, (b) daily precipitation amount measured at the Pyramid Observatory in 2014 (as rain plus snow in water equivalent), (c) mean daily relative humidity measured at Changri Nup Glacier in 2014 and 2016, and (d) mean daily air temperatures measured at Khumbu Glacier in 2015, Ngozumpa Glacier in 2002 and Changri Nup Glacier in 2016.

Rowan Figure 2: Air temperature record from Rowan et al. (2021) illustrating below-zero temperatures at the Khumbu icefall (KH4) starting in late September, 2014.

Rowan, A. V., Nicholson, L. I., Quincey, D. J., Gibson, M. J., Irvine-Fynn, T. D., Watson, C. S., ... & Glasser, N. F. (2021). Seasonally stable temperature gradients through supraglacial debris in the Everest region of Nepal, Central Himalaya. Journal of Glaciology, 67(261), 170-181.

It's unclear if similar issues exist for the melt onset signals as well. The color bar in Figure 8A,B is too hard to read to discern if melt onset appears to be happening as early as March 1st, which would be unlikely. However, the fact that the color bar is shows DOY 60 suggests there are some areas where it is melting at this time; otherwise, why stretch the color bar outside the values shown in the figure? The issues with the freeze onset and melt onset and lack of any validation, in my opinion, undermine the entire study.

We have corrected the extent of the color bar in Figure 8 (Figure AR3). The color bars were stretched erroneously to the maximum extent of the data across the HKH and we thank the reviewer for pointing this out. Across the three years of retrievals, the average minimum melt onset within the Central Himalaya occurred at day of year 82 between 3200m - 3300m a.s.l.. The average minimum melt onset in the Karakoram across the record occurred at day of year 93 and at the lowest elevation bin (3000m – 3100m a.s.l.). We have scaled the color bars for Figure 8 to represent the data accordingly.

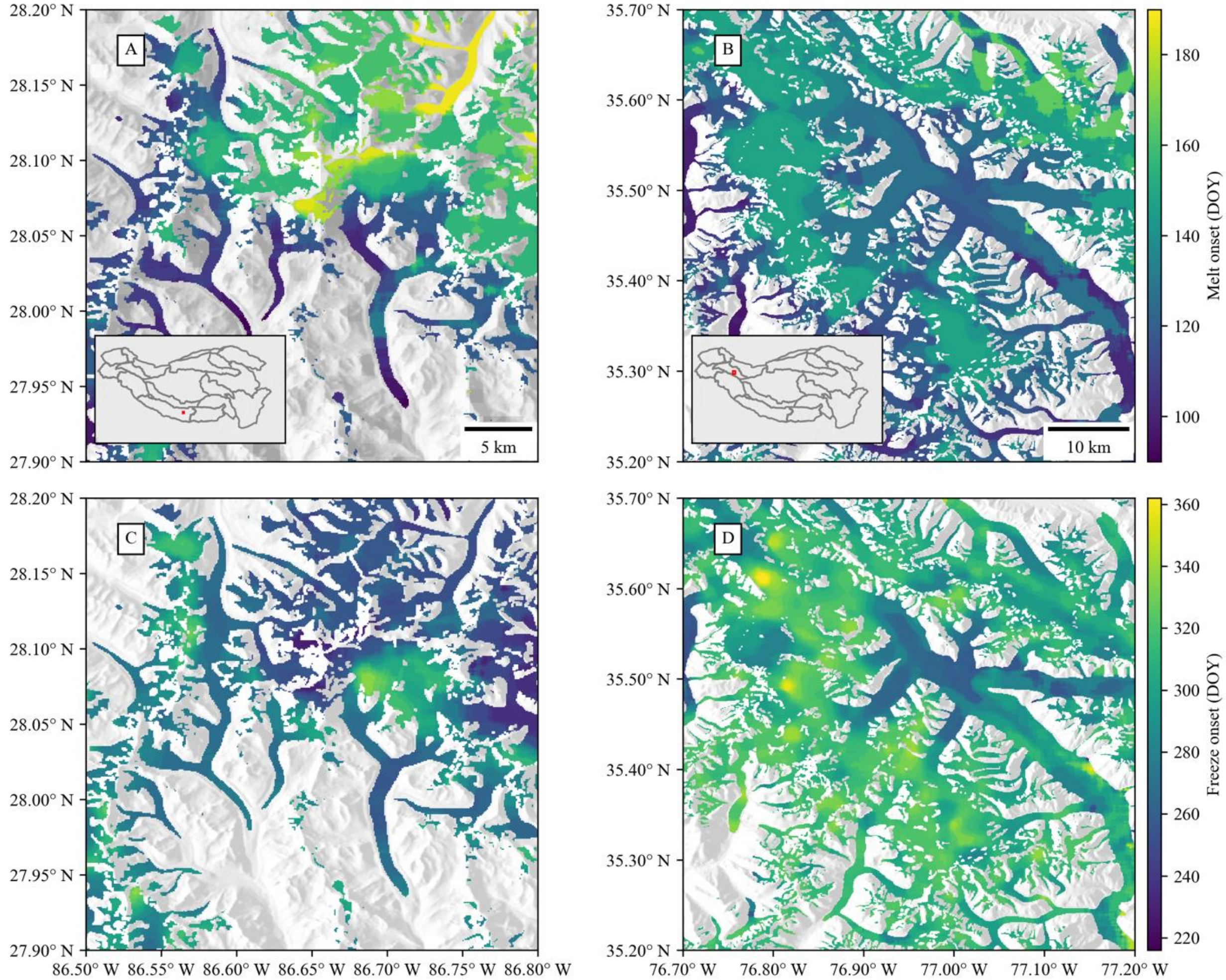

Figure AR3. Melt retrievals averaged over the calendar years 2017-2019 in the Central Himalaya and Karakoram regions. (A) Mean melt onset (DOY) in the Central Himalaya. (B) Mean melt onset (DOY) over the Siachen glacier in the Karakoram region. (C) Mean melt offset (DOY) in the Central Himalaya. (D) Mean melt offset (DOY) over the Siachen glacier in the Karakoram region. Data overlay a 30m Shuttle Radar Topography Mission (SRTM) DEM hillshade (Farr, 2007).

One recommendation is to put the results and interpretation into better context of glacier zones. For example, the discussion/interpretation primarily focused on the liquid water in frozen snow and percolation faces, but the results that were being interpreted were aggregated from 3000-7500 m a.s.l. The lower elevations are clearly in the ablation zone, i.e., clean ice or debris-covered ice. Therefore, it's unclear how much of the discussion of the snowpack and percolation faces is warranted. Or is this partially being included to discuss seasonal snowpack covering the ablation zone? Including a rough estimate of the glacier zones, even if they are roughly estimated based on median elevation or some other metric, and their respective hypsometries may provide some useful context for interpreting the results. Otherwise, results/discussion like L406-414 come across as interpreting the entire glacier as being in the percolation zone.

We fully agree with the reviewer, it would be ideal to discuss results in context of glacier zones. However, we are not aware of a detailed mapping of glacier zones at a scale appropriate for data record. We use aggregates of elevations bins at 1000 m intervals to communicate our results in Table 3.

Line 406 refers to results over specific elevation bins where delayed refreeze is apparent (i.e. at higher elevations). We have changed the sentences beginning with L406 to clarify that this paragraph is meant specifically to discuss observations of delayed refreeze at high elevations and how these observations relate to meltwater retention specifically within percolation zones: **"Signals of delayed refreeze are observed at elevation ranges similar to greatest z-score across summary statistics of FO (*Supplementary Fig. S2*). Notably, we find specific high elevation ranges in select catchments in the western sub-regions (Eastern Hindu Kush, Western Pamir, and Karakoram) and some eastern sub-regions (Tanggula Shan, Nyainqentanglha, Eastern Tibetan Mountains, and Hengduan Shan) where there is a signal of delayed refreeze apparent in summary statistics. Although sub-regional aggregate FO statistics do not show delayed refreeze in larger sub-regions (i.e. the Central Himalaya), we observe signals of delayed refreeze on individual glaciers within the Central Himalaya indicative of meltwater retention within percolation facies (Figure 7)."**

Given the major issues with the ablation area, I would suggest either (i) limiting the study to only accumulation areas, or (ii) limiting the study to only clean-ice glaciers. Either way, there needs to be significantly more validation performed to provide confidence in the results. This validation should ideally be done for both the ablation and accumulation areas, albeit that the accumulation zones interpretation should be much stronger as they are based in theory as the authors clearly discuss in the main text and mention in response to a previous reviewer.

There is a well-known shortage of validation data (e.g. automated weather stations) throughout the HKH. Therefore, it is vital that we advance alternate means of glacier monitoring like remote sensing. We contend that the sensitivities of radar to liquid water at the glacier surface make it an ideal instrument to record glacier surface melting. The physical response of SAR to liquid surface-water is well established in various applications. We fully agree, uncertainty assessments would be ideal if validation data was available. To our knowledge, there is no remote sensing available at the spatial and temporal resolution required to define this uncertainty. For additional discussion on this topic, please see Author Responses (AR) to Reviewer #1 during the first round of reviews. Specifically, AR3, AR9, AR14, available at this link.

This validation is important as it is unclear how sensitive the results are to various aspects of the methods. For example, is there any sensitivity to the chosen dB threshold of 3? A previous reviewer asked for a sensitivity analysis, but the author's simply responded that this was *conservative based on literature values. Even if this is only done for a handful of glaciers, a sensitivity analysis would provide more confidence. Similarly, what about using ascending vs. descending orbits? The choice of orbit direction clearly affects the interpretation of the signal at high elevations (Figure 5) and yet for the full region a composite was used. If this composite is used moving forward, it'd be good to see some sensitivity/error analyses because it appears to have a strong impact on the results, but perhaps this is only for high altitudes.*

We use a universal threshold (i.e. not spatially or temporally varying) that is well established in prior studies and easy to implement on a large-scale dataset. As indicated in the text, this value was first suggested in Ashcraft and Long (2001) using a simple scattering model and is equivalent to a loss of one-half power. A spatially or temporally varying threshold requires a more complex approach but given the computational requirements and large-scale dataset (i.e. >100 Terrabytes) we believe this is beyond the scope of our current study.

Ashcraft, Ivan S., and David G. Long. "Azimuth variation in microwave backscatter over the Greenland ice sheet." IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217). Vol. 4. IEEE, 2001.

Similarly, were the high-altitude automatic weather stations used for validation? The energy balance modeling likely has its own issues (e.g., sublimation could be important here?), but the modeling appears to be used to validate the SAR signal (L320). If it is used for validation, does that mean that the SAR signal overestimates the amount of melt by 33-43% (L345)? This would appear to be a considerable amount of error.

We fully agree with the reviewer's assertion that surface energy balance modeling at high elevations is complex and therefore requires considerations that are beyond the scope of this exercise. Surface energy balance outputs are used to confirm that widespread signals observed, for the first time, in SAR time series are very likely due to surface melting. We have no direct measurements of liquid water in snow and firn for full validation. Agreement between energy balance models and radar demonstrate that surface melting at high elevations is difficult to predict using only measurements of air temperature. Since the radar response to liquid water is largely unambiguous there is not a strong argument to attribute error from this modeling exercise to SAR, especially where modeling exercises are not well constrained (e.g. sublimation).

Specific Comments

Given there's no word limit, I would suggest writing out acronyms like melt onset and freeze onset to make the study more readable to the average reader.

We have limited the acronyms that we define in the paper to freeze onset (FO) and melt onset (MO). These acronyms signal to the reader that we are referring to our record of glacier surface melting. This limited use of acronyms should be accessible to the average reader, we are careful not to add any additional acronyms to the text.

L34: "Ice caps" have a particular meaning. In HKH, they are primarily (if not all) mountain glaciers or valley glaciers.

We have removed references to "ice caps" in L34 and instead referred to "mountain glaciers."

L29-46: This introduction bounces back and forth between discussing global issues (e.g., contributing 25% of sea level rise) to HKH specific issues (e.g., freshwater for 2 billion people). It's useful to show how HKH changes fit into the larger picture, but I'd be conscious of this change in scales to make it easier to read.

The implications of glacier wasting in the HKH present both regional and global hazards. Stylistically, we sought to illustrate a broad overview of the hazards associated with glacier wasting in the HKH as those hazards range in scale from local (i.e. outburst flooding) to global (i.e. sea level rise).

L47-49: Litt et al. (2019) focuses on two glaciers over a period of a several years. It does not address "projected disturbances". More appropriate reference would be one of the GlacierMIP studies (e.g., Hock et al. 2019 (Journal of Glaciology) or Marzeion et al. 2020 (Earth's Future)). The same is true for L55, which discusses projections but does not mention relevant studies where these uncertainties exist.

We have changed the sentence beginning at L48 accordingly: **"Substantial uncertainties exist in projected disturbances associated with a changing climate, environment, and hydrologic regime across the greater Himalayas due in part to a lack of observations of *in situ* hydrology and meteorology at high elevations (Hock et al., 2019; Litt et al., 2019; Marzeion et al., 2020)."**

We have changed L55 accordingly: **"Although the general trajectory of changes to the HKH cryosphere is understood (i.e. accelerated glacier mass loss on a decadal scale in the Central and Eastern Himalaya) (Fujita and Nuimura, 2011), a consensus in projecting changes to HKH hydrology is lacking largely because of missing *in situ* snow and ice monitoring data across these glaciated river basins (Marzeion et al., 2020)."**

L65-68: Statement concerns inability of studies to capture variability in patterns and magnitude of melting, but only references a study that measures the geodetic mass change (Brun et al. 2017). The more recent study by Shean et al. (2020; Frontiers in Earth Science) would be appropriate to reference here as it provides refined estimates of mass change compared to Brun et al. (2017). Furthermore, the recent study on projections in this region (Rounce et al. 2020; Frontiers in Earth Science) explicitly captures the variability reported by these measurements using a degree-day model.

A degree-day model would not capture the melt occurring for several months of the year at elevation ranges above the 0ºC summer isotherm, as presented in this study. Brun, et al. (2017) is cited to highlight sub-regional variability in the sentence beginning L66

L73: "operational monitor" – consider changing "operational monitoring system" perhaps?

We have changed this sentence accordingly in addition to all references to "operational monitor" made throughout the text at L23, L73, and L569.

L127 – this would be useful for assessing all models of glacier ablation, not just energy balance models.

We have changed the sentence accordingly: **"It is possible that intense incident solar radiation is driving these melt processes at elevations above the 0ºC summer isotherm (Matthews et al., 2019) across the entirety of the HKH, and that the sensitivity of SAR backscatter to changes in the glacier surface melt/freeze condition as seen when water transitions between solid and liquid phases provides a real alternative to temperature elevation lapse rate estimates of melting (Litt et al., 2019) for assessing models of glacier ablation."**

L137 – for The Cryosphere I would suggest using more standard glacier mass change terms (e.g., mass loss, mass balance, mass change) as opposed to the term "wasting" that is less frequently used.

We have modified language around glacier mass loss accordingly by changing references to "mass wasting" to "mass loss" at L29, L38, L532, L533, L537, and L539.

L139-141 – appears to imply that Karakoram, Kunlun Shan, etc. are not affected by increases in global average temperature. I'd suggest references Karakoram anomaly and recent studies finding that they are starting to lose mass (e.g., Farinotti et al. 2020 (Nature Geoscience)).

This sentence was meant to highlight sub-regional variability in melt patterns. We have changed this sentence accordingly: **"Glacier wasting in the HKH is heterogeneous and the increase in global average temperature has caused wasting of mountain glaciers across all HKH sub-regions (Farinotti et al., 2020; Gardelle et al., 2012)."**

We have also changed the sentence beginning at L535 accordingly: **"Sub-regions with slower mass loss (Eastern Hindu Kush, Western Pamir, Karakoram, Tibetan Interior) show on average one month less of melt duration relative to regions with accelerated mass loss."**

L148 – is there a reference for this ICIMOD dataset?

We have added a citation for this shapefile at L148.

L158-171 – this is a lot of detail on the glacier outlines used in a study that I don't see being relevant. If debris-covered areas are an issue, there's no mention of datasets that explicitly delineate it (Scherler et al. 2018 (Geophyiscal Research Letters); Herreid and Pellicciotti 2020 (Nature Geoscience)). This should be discussed.

We thank the reviewer for this suggstion. We agree that too much discussion of the process for producing glacier outlines is not appropriate for this paper. We mention the issue of debris covered glaciers and the production of glacier outlines in a single sentence and as an aside (L166-169). We mention the use of SAR interferometry to delineate debris-covered glaciers.

L189 – "this" should be "the" or "the timeframe of this study"

We have incorporated this suggestion into the text at L189.

L204 – suggest stating the name of the cloud-computing platform here, i.e., Google Earth Engine, to make this explicit for readers as opposed to only mentioning it at the end of the paragraph.

We have restructured L204: **"A cloud-computing platform and application programming interface (Google Earth Engine) with pre-processed radiometrically terrain corrected Sentinel-1 A/B data was used to detect melt characteristics across the region (Gorelick et al., 2017)."**

L216 – It is my understanding that these are both on Khumbu Glacier. Therefore, this should be "over a high elevation glacier", not glaciers.

We have restructured L216 accordingly: **"Measurements from two automated weather stations (AWS) are used to estimate surface energy balance (SEB) and evaluate surface melting conditions over the Khumbu Glacier."**

L245 – Why is this weather station not stated in Section 2.4?

We thank the reviewer for pointing this out and are happy to have made the following changes to better organize presentation of AWS data in the paper. We have moved Table 2 to Section 2.4 while also adding the Everest AWS data to Table 2.

We have rewritten the sentence at L215 as follows: **"Measurements from two automated weather stations (AWS) are used to estimate surface energy balance (SEB) and evaluate surface melting conditions over the Khumbu Glacier and measurements from two additional AWS are used to calculate temperature-elevation lapse rates for comparison with melt retrievals (Table 1)"**

We have appended the following sentence at L222: **"AWS data collected within the Langtang Valley are used to estimate temperature elevation lapse rates following prior studies and serve as data for comparison with Sentinel-1 backscatter values (Table 1) (Shea, 2016). "**

L247 – Introduction stated that assuming 0 degC threshold is poor assumption and was part of the motivation for this study. May want to preface that this generally works well, except for at *extreme altitudes where other processes (e.g., sublimation) may be important (if that is the point that is trying to be stated)?*

We thank this reviewer for this suggestion and strongly agree that temperature elevation lapse rates do well to capture the processes of glacier melt within the HKH at elevations where the maximum daily average temperatures exceed 0ºC, as stated in L567. We however contend that our statements regarding temperature-elevation lapse in fact working well at elevations where mean daily air temperatures exceed 0ºC (e.g. through agreement presented in L247, statements in L494, L567) will suffice.

L296 – It's unclear how debris-covered areas were identified (see previous comment). Nonetheless, not including debris meant between 30-48% of the ice in the ablation area is not being monitored (Kraaijenbrink et al. 2017; Nature). It would be good to state how much over the actual glacier area was able to be monitored, and if statements are being made about the ablation area (like that being mentioned here), then stating the percentage of the area in the ablation area being monitored is relevant as well.

In Table 3, we tabulate the area over which melt signals were retrieved within each sub-region following median window filtering. In Figure 4, Z-score data by elevation shows that there are only two sub-regions (Eastern Tibetan Mountains and Eastern Hindu Kush) where the mean z-score is well below the threshold of 2, suggesting that these low elevation areas (~3000m - 34000m a.s.l.) were the only portions of the study region where Sentinel-1 did not record a strong enough signal related to seasonal melt processes.

We have restructured L309 in order to articulate more clearly that there are two elevation ranges and regions where the z-score test show an inverse response to seasonal melt: **"Mean seasonal melt magnitude averaged over 100m elevation bins over all three calendar years of data shows strong (z > 2) melt signals across glacio-climatic sub-regions and across all elevation ranges of significant glaciation except below ~3,400m a.s.l. in the Eastern Tibetan Mountains and Eastern Hindu Kush sub-regions."**

L307 – if the debris-covered area is excluded, then how are statements being made concerning those areas? Is the assumption at the regional scale that when evaluating variability, the debris-free areas that were measured at lower elevations are representative of the debris- covered areas? If so, this should be stated explicitly, as it's confusing within its present form. Note that the Khumbu Glacier example clearly shows debris-covered areas being included.

Exclusion within any area due to the presence of liquid water not dominating the signal will be conducted on a per-pixel basis and interpolated over at the end of product generation between areas where a melt signal was retrieved.   Areas are excluded that do not show a seasonal sensitivity to the backscatter signal representative of the presence of liquid water at or near the surface.

L322 – "are" should be "is"

There are multiple weather stations installed at the Khumbu glacier.

L324 – If the energy balance modeling is an important part of this study, which it appears to be, then additional detail should be provided in the main text. At a minimum, this should include the values used to force the model (i.e., those in Table SI) as well as relevant information pertaining to the timestep and the terms considered (i.e., Equation 1 in the supplement).

We contend that inclusion of model parameters is not central to this paper as the model was parameterized identical to Matthews, et al. (2019). All relevant model parameters are included in the supplementary section.

L330-350 – these are results, not methods.

This is a confirmation that our methodology is correctly interpreting the melting signal. This information is not included in the resulting data record that we derive.

L337-339 – I was thinking this same point, so I'm glad the authors brought this up. However, the timing of satellite overpasses is known (Figure 5 caption), so why was this not performed? This appears to be the primary validation of how well SAR performs, and one of the primary conclusions of the study that SAR can detect melting where models otherwise wouldn't, so it should be presented in a rigorous analysis to provide confidence.

The SEB model was run at half temporal resolution at two hour time steps. When matching the overpass timing of the descending and ascending passes (6:00am and 6:00pm respectively) we find that the SEB determined melting duration is largely the same and does not change the interpretation of our data. In order to avoid this notable confusion in the text, we thank the reviewer for pointing this out and have removed the following sentence at L337: **"A more robust comparison would match the timings of satellite overpasses and meteorological observations and acknowledge that some disagreement between melting estimates is resultant from this difference."**

L351-355 – High degree of confidence from two weather stations, where the days of melting appears to be overestimated by 33-43% (L345) seems to oversell the results. I agree that this is a challenging topic, especially since there are uncertainties with the SAR data and uncertainties with the energy balance model (unfortunately, there's no validation data for the energy balance model, so are these differences due to the SAR data or the energy balance model – likely both?).

We have a high degree of confidence melt is occurring where SEB results and Sentinel-1 retrievals show that the glacier surface is melting despite air temperatures below 0ºC.

L355 – "HMA" (High Mountain Asia) is not defined in text. It's used later in text as well. I'd suggest using HKH throughout.

The "HMA" abbreviation was mistakenly included. We have changed this to HKH at lines L355, L531, and L533.

L386 – except for debris-covered areas, no (L297)?

There are only two regions where the mean z-score is less than 2, which we highlighted in the response to the specific comment on L296.

Table 3 – caption and table are opposite directions making it very difficult to read. Also, 1 km elevation bins are very coarse considering that some regions (if not most regions) have most of their glacier area spanning 1-2 km (Figure 4 Bottom). This means that the statistics shown for the other elevation bins is for a tiny fraction of the glacier area, which in my opinion detracts from the overall value of this table. The 100 m bins in Figure 4 are much more meaningful.

The data presented in Figure 4 and the Supplementary Figures on sub-regional melt timing and z-score sufficiently summarize the data for the purpose of this communication. Those interested in working with the data will be able to retrieve it and calculate statistics relevant to their study regions at 90m maximum spatial resolution.

Figure 6 – this figure is difficult to read/interpret due to the lines. I would suggest changing both the circles and the squares to lines. The circles and squares are clearly separated, so you could simply add a note that the melt onset is on the left and freeze onset on the right, which is intuitive anyways. This would enable the reader to follow the lines and determine how the trends vary. It will also make it easier to see the 12 subregions, which currently are on top of one another and hard to discern any information from. This will also make it easier to see statements like L400 where it does not follow a linear trend, when the current form of Figure 6 looks like roughly speaking higher elevations are around DOY 250-270 and lower elevations are around 280-330, so while the linear trend is not as strong, there still appears to be a trend.

We have reformatted the figure to include lines instead of shapes (Figure AR4) and inserted it into the manuscript at L415. For more detailed illustrations on non-linear trends in melt patterns, please see Supplementary Figure 2.

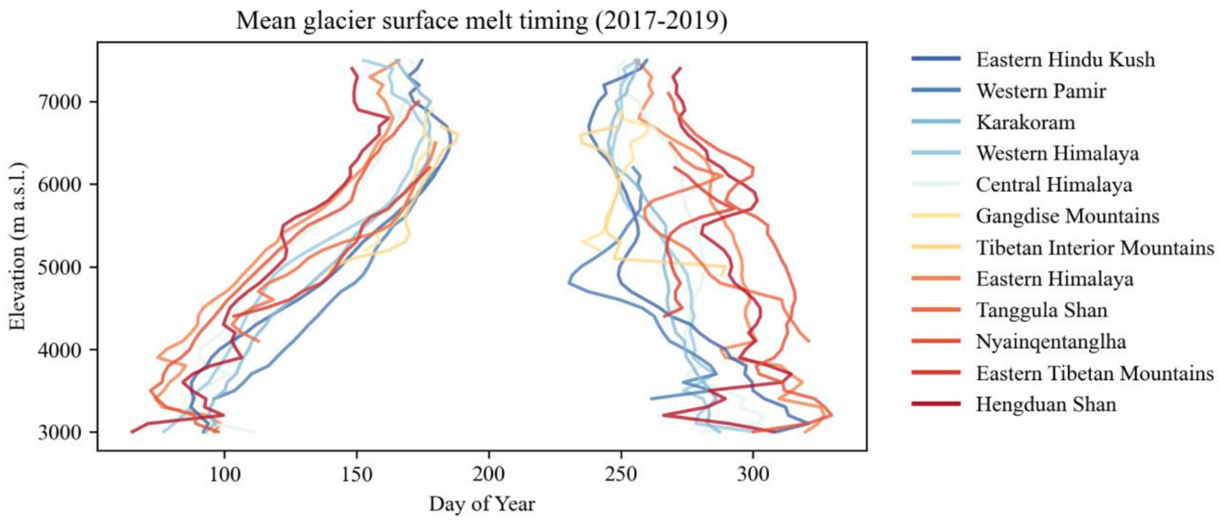

Figure AR4. Mean melt onset (MO) and freeze onset (FO) summarized in 100m elevation bins using the 30m SRTM digital elevation model (Farr, 2007) and 12 HiMAP sub-regions (Shean, 2020). The blue to red color scale indicates the longitude of the HiMAP region centroid, where the westernmost regions are shown in dark blue and eastern most shown in dark red.

L393-396 – is this supported by the energy balance modeling?

SEB models are at point-locations and therefore will not capture melt/freeze characteristics as they vary across elevations. Additionally, C-band SAR will be sensitive to the presence of liquid water across a depth on the order of meters in the percolation zone, which may not be captured in SEB models.

L429-433 – Khumbu Glacier is highly debris-covered below the Khumbu icefall, which the authors appear to interpret as a signal of refreeze onset. This is highly inaccurate. Figure 7 (Top) suggests the Khumbu Glacier would be refreezing on DOY 250-270 (sometime in September) when the debris-covered ice is clearly still melting (Rowan et al. 2020; Journal of Glaciology). Figure 8C,D show similar issues. See major comment.

Please see response to the major comment above.

L533 – Shean et al. (2019a) does not exist. Unclear therefore where the trends for 2000-2010 are coming from and why trends in later decades covering the time period observed (i.e., 2000- 2018 from Shean et al. 2020) are not used.

We have corrected this reference to refer to Shean, et al. (2020) at L533.

L557 – surface energy balance models "are constrained" by SAR data indicates that the SAR data is used to calibrate the energy balance models. This is in direct contrast to earlier where the energy balance models were being used as validation of the SAR observations (L320).

We have changed the sentence beginning at L557 accordingly: **“Melt conditions in surface energy balance models of glacier melt driven by *in situ* meteorological data from Mount Everest fall within the date ranges of melt retrievals recorded in Sentinel-1 SAR data.”**