Response to RC1

**Mapping Seasonal Snow Melting in Karakoram Using SAR and Topographic Data**

Discussion: https://doi.org/10.5194/egusphere-2024-942

Comments from the reviewers are given in black.

Our responses are given in blue.

General Comments

The paper presents an innovative framework for mapping wet snow in the Karakoram region using SAR and topographic data. The integration of the Gaussian Mixture Model (GMM) to determine the Wet Snow Index (WSI) and the calculation of the Topographic Snow Index (TSI) significantly improved the accuracy of snow mapping in complex terrains. And I do appreciate the extensive data sources used by the authors, including Sentinel-1 SAR imagery, Digital Elevation Models (DEM), and Sentinel-2 Level-2A imagery, which ensure robust and detailed mapping of seasonal snow melting. The continuous analysis of Wet Snow Extent and Snow Melting Duration provides valuable insights into the temporal and spatial dynamics of snow melting in the Karakoram region, offering excellent references for water resource management and hydrological modeling, especially in glacial and high mountain areas with complex terrain.

However, I have several concerns regarding the complexity of the methodology and the unclear use of data, as well as the lack of quantified uncertainty analysis. Therefore, I recommend that the paper undergo revisions before it can be considered for publication.

Thank you so much for taking the time to read and review our manuscript, and for your positive, thoughtful and constructive comments. The structured review and detailed comments and suggestions have helped to significantly improve the manuscript.

Line 47, should be “includes”

**Response:**

Thank you for pointing this out. We have corrected the text to use “*such as*” on Line 47 to improve the sentence flow.

Line 96, Please specify the spatial resolution of the DEM.

**Response:**

We have updated the manuscript to include the spatial resolution of the DEM by adding the phrase: “*providing global coverage at a resolution of 30 meters.*”

Line 109, Please specify the Sentinel-2 spectral bands used to calculate the LIS in the paper.

**Response:**

We have clarified the Sentinel-2 spectral bands used in the LIS algorithm by adding the following description in lines 118-120: "*The LIS algorithm employed RGB spectral bands B2 (blue), B3 (green), and B4 (red), as well as infrared bands B8 (NIR) and B11 (SWIR) for snow-cloud-ice classification.*"

Line 122, It is unclear which corrections and preprocessing methods were adopted in the GAMMA software for SAR image processing. Did you apply terrain correction or geometric correction?

**Response:**

All preprocessing steps were conducted using the GAMMA software. To improve clarity, we removed the phrase "using GAMMA software" from the specific sentence and relocated it to the end of the paragraph. We confirm that both geometric and terrain corrections were applied using the GAMMA software. In the original manuscript, this process was referred to as "terrain-based radiometric normalization," a term that may be less familiar. To prevent confusion, we have revised the terminology to "terrain-based radiometric correction."

Line 131, The focus season of the paper is summer. Why was a multi-year winter value chosen for the reference $\gamma$, Does the reference $\gamma$ have a spatial pattern climatology or a single value for each basin?

**Response:**

The reference image was selected to emphasize the contrast between wet snow surfaces, which typically exhibit low backscattering intensity, and dry snow or snow-free surfaces, which have higher backscattering intensity. In other alpine regions like the Alps, summer images are often used as the reference due to snow-free conditions caused by melting. However, the Karakoram region, with its significantly higher altitude, experiences snowfall and snow cover throughout the year. Therefore, winter images were chosen as the reference to utilize the dry-snow surface to enhance the backscattering contrast. The reference image does not contain specific spatial patterns or unique values for each basin.

To clarify this in the manuscript, we added the following explanation in lines 133-136:

"*Note that this differs from other alpine regions, such as the Alps, where summer months are often used as the reference due to snow-free conditions. In the Karakoram, due to the all year long snow cover at the higher elevations, we used winter images as to leverage the dry snow conditions in winter for highlighting the contrast in backscattering intensity between summer wet snow (lower intensity due to water absorption) and winter dry snow.*"

Line 146, In the GMM model, which covariance structure did you choose? Were other parameters set by default?

**Response:**

We used the full covariance structure in the GMM model, meaning that each component has its own general covariance matrix. This choice was made after testing different covariance structures. All other parameters were set to their default values. To clarify this, we added the following details to the manuscript:

- Line 150: "We used the full covariance structure in the GMM model, i.e., each component has its own general covariance matrix, after testing different types of covariance structures."
- Line 154: "During the training, the maximum number of EM iterations was set to 100, and the convergence threshold was set to 10^−3."

Line 175, I am a little confused about the SI definition. From my understanding, is the WSI first scaled by the TSI, meaning SI = WSI/ TSI? Then you have an SI map, and to further differentiate the wet snow, a threshold of 3.5 is selected. Is this threshold basin-dependent or applied across the entire Karakoram? How do you determine that 3.5 is the optimal threshold, and how do you quantify the uncertainty resulting from this coefficient?

**Response:**

We appreciate your feedback and the opportunity to clarify the SI definition and thresholding approach. The SI (Snow Index) is not defined as SI = WSI / TSI; rather, it is generated by the pixel-wise multiplication of WSI and TSI, i.e., SI = WSI × TSI. This process scales the WSI according to terrain characteristics, creating a more precise linkage between SAR backscattering and terrain properties.

The threshold used for snow classification is not a fixed value of 3.5 but rather a dynamic value calculated specifically for each basin using the formula:

$$SI\ Threshold = 3.5 \times WSI|_{R_c=-2}$$

Here, 3.5 reflects the conditions applied to the TSI, and $WSI|_{R_c=-2}$ represents the WSI at an $R_c = -2\ dB$ for each basin. The value of 3.5 was empirically determined through extensive experiments aimed at optimizing the trade-off between precision and recall to maximize the F1-score. This value was found to provide the best performance in distinguishing wet snow from other surfaces by synergistically integrating SAR backscattering and terrain information.

The uncertainty associated with this coefficient was considered during the optimization process. While further sensitivity analysis could refine the threshold, our approach currently provides a balance between accurate classification and adaptability to diverse terrain conditions. Future work will focus on enhancing threshold calibration and further quantifying uncertainties with additional validation data.

To better explain the choice of the coefficient, and discuss the limitation of our method, we revised the section 3.4 in the manuscript as the following:

` ` `

*As illustrated in Figure 2, the final step generated an integrated Snow Index (SI) map by performing pixel-wise multiplication of WSI and TSI. This multiplication scales the WSI by incorporating terrain characteristics, thereby linking the observed SAR backscattering ratio directly with terrain properties.*

*In order to classify the integrated SI into binary snow maps, it is crucial to apply an adaptive threshold that accounts for the variation in topographic features across different basins. The variation in SAR backscatter response within a basin is inherently handled by the GMM when deriving the WSI. In contrast, the TSI is time-varying and basin-specific, requiring an optimal coefficient to condition the SI for classification. To determine this coefficient, we performed a sensitivity analysis, evaluating F1-score, precision, and recall across different values using the S2 validation snow map.*

*The results, shown in Figure 5, demonstrate that Hunza and Shyok exhibit similar responses, with optimal coefficients close to 3.5, while Shigar reaches its optimum at approximately 2.5. However, to avoid overfitting to specific basins or validation dates, we selected 3.5 as an overall coefficient to balance classification performance across all basins. This coefficient also reflects a moderate threshold applied to the TSI to determine the overall SI threshold for each basin.*

*The threshold is calculated using the following equation:*

$$SI\ threshold = 3.5 \times WSI|_{R_c=-2}$$

*where* $WSI|_{R_c=-2}$ *represented the WSI at a backscatter ratio Rc = -2dB for each basin. This value is basin-specific, allowing the threshold to adapt based on each basin's distinct characteristics. Together, these conditions form an integrated, basin-adaptive thresholding mechanism, combining SAR backscatter and topographic information into a single index to determine the SI threshold.*

*It is important to note that while the SI threshold is basin-specific, it is time-independent. The WSI is derived from a GMM trained on samples collected from multiple summer scenes over several years, ensuring that it represents an aggregated measure for each basin and is not tied to individual scenes or seasons. This design ensures robustness to seasonal variations in liquid water content and enables consistent application across different validation dates.*

` ` `

Eq. 7, should be “$\rho B$”?

**Response:**

The correct term is “$\rho_{red}$”. The text in line 194 was incorrect and we have changed from “$\rho_B$” to “$\rho_{red}$”.

Line 184, It would be better if you could clarify the exact thresholds used in the S2 snow cover detection. Currently, the flowchart for S2 is unclear. Specifically: (1) First step to meet the condition NDSI >0.4 and $\rho B$>0.2; (2) the dark cloud region only defined by $\rho B$>0.3? Can you explain what is bi-linearly down-sampled red band, from which resolution to which resolution?; (3) and then calculate the snow cover fraction (SCF) at which resolution? (4) Line 196, is this for dark cloud pixels above the snowline? Based on your description, above snow line, ‘no-snow’ is defined by NDSI >0.15 and $\rho B$>0.04 and $\rho B$<=0.1, and dark cloud is defined by NDSI >0.15 and $\rho B$>0.1? How consistent are these relaxed thresholds compared to the strict threshold $\rho B$>0.3 mentioned in 191? Generally, if you are strictly following Gascoin et al. (2019), you can briefly refer to their method, but any changes should be clearly mentioned; otherwise, restate the algorithm clearly and concisely.

**Response:**

Thank you for your detailed comments and suggestions. We appreciate the opportunity to clarify the steps and thresholds used in the S2 snow cover detection process. Below, we address each of your points:

1. **Initial Thresholds (NDSI > 0.4 and $\rho$_red > 0.2):**
   The initial step in the algorithm uses the condition NDSI > 0.4 and $\rho B$ > 0.2, as per Gascoin et al. (2019). This combination aims to identify snow-covered areas while excluding non-snow features such as bare ground and vegetation.
2. **Dark Cloud Definition ($\rho$_red > 0.3):**
   Dark clouds are defined by applying a threshold of $\rho$_red > 0.3 on the bi-linearly down-sampled red band, which reduces the spatial resolution from 20 m to 240 m by a factor of 12. This down-sampling process helps mitigate the impact of cloud shadows and high-altitude clouds, thereby refining the provisional snow masks. We followed this step strictly as recommended by Gascoin et al. (2019) to ensure consistency in cloud exclusion.
3. **Snow Cover Fraction (SCF) Calculation Resolution:**
   The SCF was calculated at 100 m elevation bands using the provisional snow masks. The DEM was resampled to the same resolution as the S2 images at 20m pixel sizes. Each elevation band’s SCF represents the proportion of snow-covered pixels within that band, allowing for a more granular understanding of snow distribution across varying elevations.
4. **Thresholds for Dark Cloud Pixels Above Snowline (Line 196):**
   For dark cloud pixels above the snowline, we applied the relaxed thresholds of

NDSI > 0.15 and $\rho$_red > 0.04 to $\rho$_red ≤ 0.1 for 'no-snow' classification and NDSI > 0.15 and $\rho B$ > 0.1 for dark cloud classification. These relaxed thresholds were introduced to differentiate between snow and dark clouds above the snowline, given that the conditions at high elevations often result in lower NDSI values and different spectral responses compared to lower elevations.

The consistency of these relaxed thresholds with the strict $\rho$_red > 0.3 threshold mentioned earlier (Line 191) lies in the specific purpose of each step: the initial strict threshold ensures accurate cloud exclusion, while the relaxed thresholds above the snowline allow for finer classification adjustments in challenging high-altitude conditions.

We have modified the description of our method to clearly distinguish these steps and their respective thresholds and have noted any deviations from the original method by Gascoin et al. (2019). We believe these clarifications will address the flowchart's ambiguity and better illustrate our approach.

The revised method description is as follows:

` ` `

*The proposed method was validated using S2 snow cover maps generated following the LIS algorithm proposed by Gascoin et al 2019. Before running the LIS algorithm, the input S2 multi-spectral bands were resampled to a pixel size of 20m x 20m to match the resolution of different bands. The COP-30 DEM was also resampled to the same pixel size as the S2 images.*

*The LIS algorithm started with generating provisional snow masks by applying thresholds on the Normalized Difference Snow Index (NDSI) and the red band reflectance ($\rho_{red}$) with the condition:*

$$(NDSI > n_i)\ And\ (\rho_{red} > r_i)$$

*where $n_i = 0.4$ and $r_i = 0.2$ (Gascoin et al 2019). This step was designed to identify snow-covered areas while excluding non-snow surfaces such as vegetation and bare ground. However, this approach could sometimes exclude snow-covered pixels due to errors in cloud masking. To correct the errors, a refinement step was introduced to reassign dark cloud pixels that were initially misclassified. Following Gascoin et al. (2019), dark clouds were identified by applying a threshold of 0.3 on the bi-linearly down-sampled red band, which reduced the resolution of the red band from 20m to 240m by a factor of 12. This process helped to exclude cloud shadows and high-altitude cirrus clouds from the snow classification. Afterwards, the provisional snow masks were further refined using the basin snowline calculated from the COP-30 DEM. We calculated the total snow cover fraction (SCF) within every 100m elevation band using the provisional snow mask, and defined the snowline using the lowest elevation band where the SCF exceeded 30\%. For pixels identified as dark clouds above the determined snowline, the conditions used in Equation 7 were reapplied with adjusted thresholds to account for the unique conditions at high altitudes. Specifically, the relaxed thresholds of $n_i = 0.15$ and $r_i = 0.04$ were used to classify snow pixels, and dark cloud pixels with $\rho_{red} > 0.1$were reassigned as cloud, while* *other pixels were categorized as "no-snow." These adjusted thresholds help to differentiate snow from dark clouds in challenging high-altitude environments, ensuring a more accurate classification. Following the adjustment of the snow mask, we extended the LIS algorithm by further applying a threshold on the NIR band with $\rho_{NIR} > 0.4$ to classify glacier ice and water bodies from snow (Paul et al 2016).*

```
```

Line 203, When you mention summer, which months are included?

**Response:**

The term "summer dates" refers specifically to the dates listed in Table 2. To avoid confusion, we have revised the sentence to read: "*As the S2 snow cover maps classify both dry and wet snow rather than only wet snow, we selected only the summer dates listed in Table 2 to ensure that the S2 snow cover maps predominantly reflect wet snow conditions*."

Line 204, I am curious about the accuracy in S2 classification, and is it possible to estimate its uncertainty? Generally, in which confidence (and at what elevation) can S2 snow mostly be considered melt snow during summer, given that the accuracy matrix is calculated by comparing snow-free or snow from S2 with no-or-dry snow or melt snow from SI?

**Response:**

The uncertainty of the S2 classification can be estimated using several approaches, such as cross-validation with ground observations, spectral analysis combined with temperature data, or probabilistic uncertainty quantification methods. However, a rigorous validation scheme is beyond the current study's scope, as the S2 results were generated using a method established in previous research, where detailed uncertainty analysis can be found.

Regarding elevation, as mentioned in the manuscript (line 225), areas above 5500 m a.s.l. are confidently considered non-melting zones due to consistently low air temperatures throughout the year. Therefore, we focused our comparison between SAR and S2 snow maps on areas below 5500 m a.s.l. to avoid errors associated with dry snow misclassification.

Figure 6, How do you explain the remaining mismatch in the ice and melt categories between SI/Rc and S2 detection in the Hunza region?

**Response:**

The observed mismatches in the ice/water category primarily occur on glacier ice surfaces. In the S2 results, this category is determined using the NIR band, whereas the SAR method does not specifically distinguish ice from snow. On the observed date, glaciers might have experienced some melting, which could lead to a decrease in the SAR backscattering ratio, resulting in glacier ice being misclassified as wet snow in the SAR-based methods.

To provide more context for comparison, we added an RGB image column to the figure (Figure 7 in the revised manuscript). We also included the following text in the manuscript to address this concern: "*However, a consistent mismatch in the ice/water category can be observed between the SAR (both SI and ratio methods) and S2 results, particularly over glacier surfaces. This discrepancy arises from the differing detection principles of the two approaches: the S2 results classify glacier ice using thresholds on the NIR band, while the SAR-based methods do not explicitly resolve glacier ice. On the observed date, glacier ice in the ablation zone may have partially melted, reducing the SAR backscatter ratio and leading to its misclassification as wet snow in the SAR results. As discussed earlier, glacier surfaces present unique challenges for SAR-based methods due to their complex scattering mechanisms. While the inclusion of TSI improves the robustness of our method by integrating topographic controls, it does not explicitly account for the heterogeneity of glacier surfaces. Consequently, glacier-specific conditions, such as localized melting or scattering from mixed ice-snow surfaces, may lead to underestimation or misclassification.*"

Figure 7, How do you explain the steeper slope of the snow coverage profile in the SI method compared to the flatter slope in the Rc method?

**Response:**

The steeper slope observed in the SI method can be attributed to the non-linearly enhanced response of the TSI to WSI in the transition zone between 4500 m and 5500 m a.s.l. The mixed snow conditions within this zone increased uncertainty in WSI, resulting in exaggerated TSI values, particularly at the lower and upper ranges of the region. This highlights the need for a more carefully calibrated TSI model, such as one using topographic bins of higher resolution.

We added the following explanation to the manuscript: "*In contrast, the SI method provided snow classification results that were closer to the S2 profiles at lower elevations. Between 4500 m and 5500 m a.s.l., the SI curve displayed a significantly steeper slope compared to the S2 curve, with snow coverage being underestimated between 4000 m and 5000 m and overestimated between 5000 m and 5500 m. This pattern suggests that uncertainties in the SI method, particularly in mixed snow conditions within transition zones, may have led to a nonlinear exaggeration of the TSI response to snow cover. A more precisely calibrated TSI model could help align snow coverage profiles more closely between the SI method and S2 results*."

Additionally, we rephrased this paragraph to enhance readability and ensure clarity.

Line 243, when calculating the SWE and SMD, do you also include the area where a.s.l. over 5500m? Then how to make sure the accuracy above 5500m, given that these areas were excluded in the S2 validation?

**Response:**

In calculating the Wet Snow Extent (WSE) and Snow Melting Duration (SMD), we included areas above 5500 m a.s.l. This inclusion does not conflict with the validation using Sentinel-2 (S2) data, as the S2 validation was performed on specific dates and regions to assess the overall effectiveness of our proposed method. The validation results demonstrated that our approach successfully maps wet snow distribution across the study area, supporting the application of this method to the entire region and for all image series.

Ensuring accuracy above 5500 m a.s.l. remains challenging due to limited data availability for direct validation in these high-altitude areas. Quantifying accuracy in such regions would ideally require additional validation sources, such as snow distribution modeling, in-situ measurements, or cloud-free multi-spectral time-series observations. However, these additional validation steps are beyond the current study's scope. We recognize the need for accuracy assessments at higher altitudes and plan to enhance the quality of our products in future work through improved validation strategies tailored to these challenging environments.

Line 256, what level of precipitation and temperature are adopted here, surface or pressure level?

**Response:**

The temperature data used in our study refers to the air temperature at 2 meters above the land surface, and the precipitation data represents the accumulated liquid and frozen water that falls to the Earth's surface. We have added these details to the manuscript with the following rephrased sentences:

"The temperature data, averaged weekly, includes mean, maximum, and minimum air temperatures at 2 meters above the land surface across the Karakoram region, providing insights into the thermal conditions influencing snow melting. The precipitation data, representing the total accumulation of liquid and frozen water on the surface, was compiled and averaged monthly to complement the temperature analysis, highlighting precipitation trends and their impact on the snowpack."

Line 269, does it mean that each pixel you have the wet snow days in total in terms of the whole summer season, then I don't know the purpose of rescale it into 365 days since I guess most of the melting is in summer? It is not appropriate to extend the summer research into the annual. You could just use real wet duration instead.

**Response:**

Each year, we had a varying number of observed days, creating a time series with uneven intervals. For a given year, suppose we have N observed days (where $N < 365$), and a pixel is covered by wet snow on M of those days. To estimate the Snow Melting Duration (SMD), we calculate it using the formula (M/N)×365(M/N) \times 365(M/N)×365. This operation does not extend summer conditions to the entire year but rather standardizes the observation period to an annual basis, allowing for consistent comparisons between years.

We revised the paragraph as follows:

"The SMD reflects the temporal persistence of wet snow cover within a given year, enabling consistent comparisons across years with varying numbers of observation days. To compute the SMD for each year, we first determined the ratio of days with wet snow cover (M) to the total number of observed days (N) for each pixel. Since the number of observed days (N) varied each year and was typically less than 365, we rescaled this ratio to a 365-day basis using the formula $(M/N)\times365$ . This standardization allows us to calculate the annual average of wet snow cover days, facilitating consistent comparisons between years."

Line 290, It may not be suitable to use words like "greatly" to describe the improvement if the new classification's improvement is around 5%.

**Response:**

We have adjusted the language to avoid using "greatly" and have rephrased the sentence as follows: "The proposed method has effectively improved the mapping accuracy in the validation."