Response to Reviewer

Dear reviewer,

We sincerely thank you for taking the time and effort to review our paper. Your insightful feedback and constructive suggestions are invaluable and not only help to refine our research, but also deepen our understanding of the topic. We believe we have adequately addressed all of the major and minor comments, and the research has been substantially enhanced. Our point-by-point responses to your comments are listed below in black.

In this paper, the authors examine the snow phenology over the Northern Hemisphere (NH), based on satellite observations of snow cover by MODIS and IMS, and of snow depths by microwave instruments. The study focuses on snow onset, snow end date and snow cover days, as well as snow peak day.

They propose a dynamic threshold selection in constructing the snow phenology indicators, based on the local seasonal snow evolution, rather than a fixed threshold. The method has significant advantages, e.g., over the Tibetan Plateau (TP) where the snowpack can be shallow (Fig7) and the dynamic method allows for a considerably longer snow phenology (Fig9). These differences are strongly influenced by topography, and it is in the mountainous areas that the dynamic method offers the greatest benefits. In the mid and high latitudes, the difference between the two methods is small (Fig7) though, 4-5 days at most. Over the NH overall, the differences appear small (compare Fig8 c and d), albeit it influences the indicator trends.

The study is detailed and comprehensive, and the article is fairly clearly written. It should prove a valuable study to understand the phenology and snow variability, over the TP in particular. I recommend the paper for publication provided the main comments are addressed.

Response:

We are grateful for your comments on our work, which help us to further enhance the manuscript. Our responses to your questions and comments are listed below.

Main comments:

1. I question the choice of treating on the same footing snow cover and snow depth (SD) in the first part of the paper. Indeed, some phenology indicators (like SCED, end of snow season) exhibit large differences exceeding one month over the TP, depending if one is considering snow cover or depth. In the 2nd part of the paper, only SD is considered. One possible way to clarify this paper is to re-structure it to address snow cover phenology (comparing the 2 instruments) in a first part, and then focus SD in a

next part.

Response:

We appreciate the reviewer's valuable and thoughtful comments.

The purpose of treating snow cover fraction (SCF) and snow depth (SD) equally in the first part is to compare whether there are significant differences in the snow phenology extracted from different data and methods. Both types of snow data are commonly used for snow phenology extraction. Further, we would like to validate our idea that the extraction results of snow phenology can be significantly affected by the method. This highlights the need to improve the rationality of snow phenology extraction methods. In the second part, we select snow depth as the driving data because it shows a stable single-peak pattern in each latitudinal band, reflecting the process of snow accumulation and melting. In contrast, snow cover fraction data exhibit greater randomness and heterogeneity, especially on the Tibetan Plateau (TP). Moreover, SCF data are affected by the polar night, leading to large errors north of 60°N, and cloud cover further impacts data quality. Considering these factors, we originally choose only snow depth as the driving data for the second part. Here, we apply the same method to SCF data.

Since different data types describe different snow curves, the dynamic threshold for determining the snow season cannot be directly the same as the snow depth. We use the SCF data to calculate the first order derivative at each grid point and look for inflection points. We find that most of the grid points change rapidly not at 10%, but more within the interval from 10% to 20%. 58.51% of the Northern Hemisphere (NH) during the snow accumulation period and 62.89% during the snow melting period fall within the 10% to 20% interval (Fig. R1). The dynamic threshold is eventually set at 15%. On the Tibetan Plateau, this threshold is significantly higher than the interval 10-20%, which we hypothesize is related to the definition of the hydrological year. In this study, we define the hydrologic year as spanning from September 1 to August 31 of the following year. However, on the TP, SCF begins to grow in August. To account for this, we analyzed TP separately using a hydrologic year beginning in August and found that a

15% threshold is reasonable (Fig. R2).

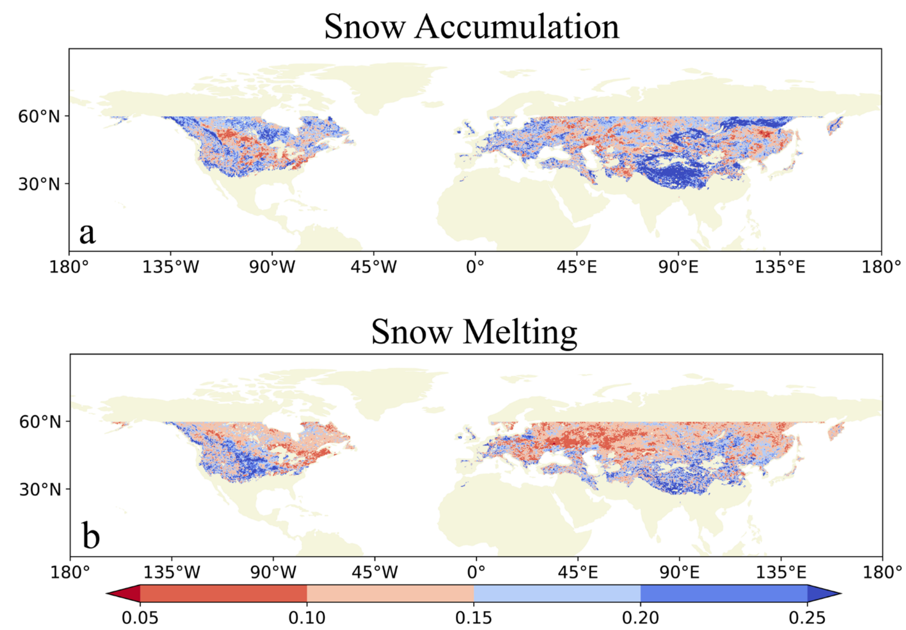

Figure R1. Spatial distribution of thresholds extracted from the first-order derivatives using MOD10C2 data in NH. (a) Percentage thresholds associated with snow accumulation (SCOD) from the first half of the first-order derivative maximum value. (b) Percentage thresholds associated with snow melting (SCED) from the last point of half of the first-order derivative minimum value.

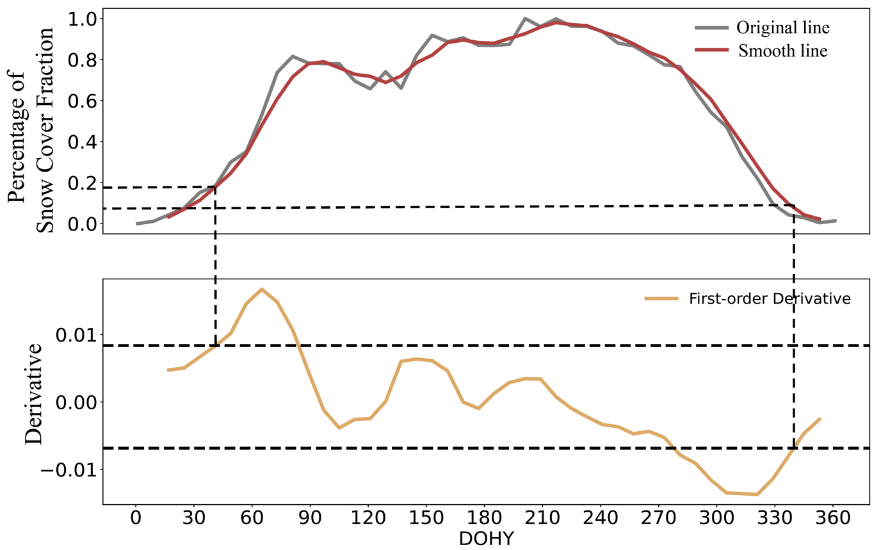

Figure R2. Intra-annual variability in the normalized snow cover fraction in the Tibetan Plateau. The gray curves represent the original SCF, the red curves represent the 30-day smoothed SD, and the yellow curves represent the first-order derivative. The unit DOHY is an abbreviation for day of the hydrological year, defined as August 1 through July 31 of the following year.

We normalize and plot the intra-annual SCF change curves for the four latitudinal bands on the same graph for comparison and find that the snow curves become wider with increasing latitude, which implies that the snow season is lengthening. In contrast to the other latitudinal bands, the snow curve of TP fluctuates a lot. The 15% position is a good match to the inflection point of the SCF curve. Employing the 15% dynamics threshold method, we extract SCF thresholds in the NH. Except for TP, the new thresholds for the entire NH are lower than the traditional fixed thresholds, implying that the traditional thresholds may result in an underestimation of the snow season.

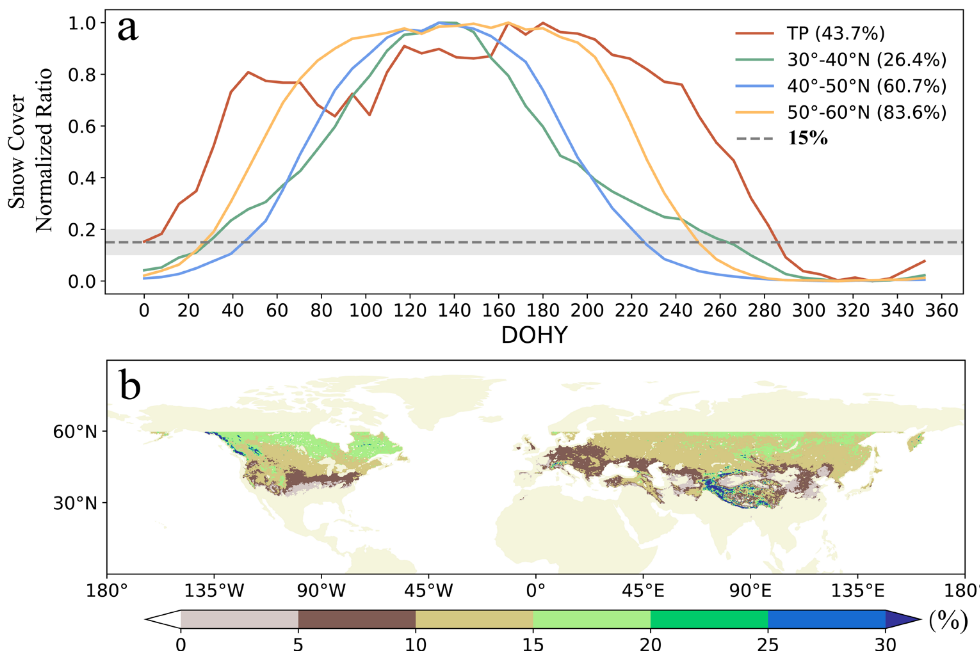

Figure R3. The dynamic threshold of the MOD10C2 dataset across the NH. (a) Intra-annual variability in the normalized MOD10C2 data for four latitudinal zones. Shading represents the interval 10%-20%, and the dashed line represents the dynamic threshold of 15%. The actual maximum snow cover fraction for each latitude band is in parentheses. The unit DOHY is an abbreviation for day of the hydrological year. (b)

Spatial distribution of multi-year average snow cover fraction thresholds in the NH extracted using the snow dynamics threshold method.

The spatial distributions of the snow phenology indicators extracted using the snow dynamics threshold method are similar to those of the original method with increasing latitude and elevation, snow cover duration (SCD) lengthens, snow cover onset day (SCOD) advances, and snow cover end day (SCED) delays. The use of the new method has resulted in a generally longer snow season due to the lowering of the thresholds. SCED demonstrated greater differences compared to SCOD.

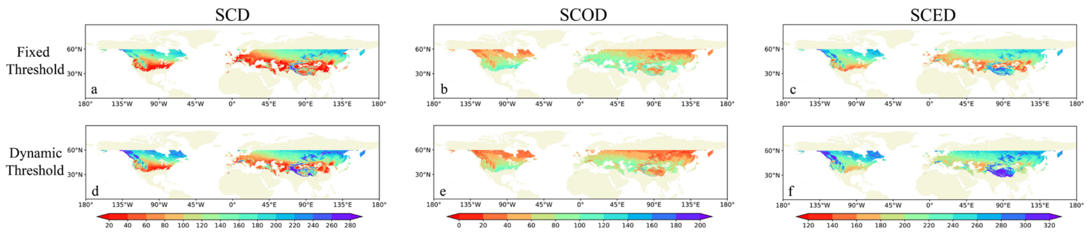

Figure R4. Spatial distribution of snow phenology extracted by MOD10C2 data over the NH for the hydrological years 2000-2018. (a) Multiyear averaged snow cover duration (SCD), (b) snow cover onset day (SCOD), (c) snow cover end day (SCED) extracted by the fixed threshold method. (d, e, f) Same as (a, b, c) but for the dynamic threshold method.

In order to show the difference between the snow phenology extracted by the dynamic threshold method and the fixed threshold method more clearly, we perform statistics in four latitudinal bands (including TP) and in the whole NH. We find that similar to the results for snow depth, the largest difference in snow phenology remains in the TP, where the SCD differs by 86 days. And the main difference in SCD is caused by SCED, with a smaller contribution from SCOD.

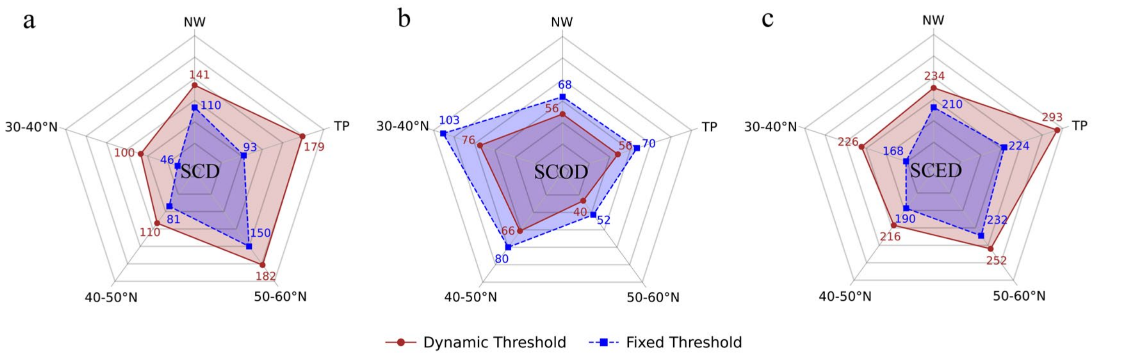

Figure R5. Comparative radar maps of (a) snow cover duration (SCD), (b) snow cover onset day (SCOD), (c) snow cover end day (SCED) extracted by the snow dynamic threshold method and the traditional fixed threshold method in five regions. The solid red line represents the dynamic threshold method, and the blue dashed line represents the fixed threshold method.

In summary, our results using SCF data for snow phenology method improvements show agreement with those based on SD. The dynamic thresholds for both datasets fall within the 10%-20% range, and both methods extend the snow season at low latitudes. The results are similar for both data, and there are flaws in the SCF data, such as the effects of polar night.

We have included the SCF results in Section 3.2 and Discussion section, and placed the corresponding graphs in the supplement. This preserves the original structure of our study—first comparing differences in existing snow phenology extraction methods and then improving them. Moreover, the results of SCF can be shown to prove the universality of the methods.

Line 348:

Given the limitations of MOD10C2 data, such as its susceptibility to polar night effects and fluctuations, we select SD as the primary driving data for this study. However, to demonstrate the robustness of the dynamic threshold method, we also apply MOD10C2 data for dynamic snow phenology extraction (see supplement). Our results indicate that for MOD10C2 data, a dynamic threshold of 15% is more appropriate. After applying the dynamic threshold method, the snow phenology results closely align with those *obtained from SD, exhibiting longer SCD, earlier SCOD, and later SCED in mid- and low-latitude regions. The most pronounced discrepancies are observed over the TP. However, since snow cover data are influenced by the polar night at high latitudes, direct comparisons cannot be made at these latitudes.*

Line 449:

Since the accuracy of passive microwave detection increases with snow depth, the passive microwave remote sensing data is more effective for analyzing snow phenology in regions with consistent snow cover (Armstrong & Brodzik, 2001; Savoie et al.,2009). In areas with shallow snow with wet snow, the accuracy of passive microwave remote sensing data is reduced, and the snow depth indicator may not accurately capture accumulation and melting processes. In addition, for the transient snow area, the snow depth curve is more volatile, which makes the assumed single-peak structure untenable. After comprehensive consideration, the snow cover fraction may be a more reliable indicator in such cases. Therefore, we perform another extraction of dynamic snow phenology using the snow cover fraction data, and the results are similar to SD, but with greater differences in TP (see supplement).

2. I am concerned about the methodology at locations where the snow curve is not monotonous across the season and has several maxima and minima linked to episodic snowfall and melt. Smoothing or climatological averaging should alleviate this potential problem. This seems be the case for the snow cover data over the TP (Fig4). This restricts the applicability of the method, and the authors expressed this concern (L445-446), even for SD at some locations over the TP where the snow layer is shallow. Please discuss this issue in the Methodology section.

Response:

Thank you for your suggestion. As you mentioned, the occurrence of sporadic snowfall events results in snow curves that are not strictly unimodal throughout the seasonal cycle, particularly in regions with unstable snow conditions. This underscores the importance of applying smoothing techniques in our approach to mitigate the influence of noise. To determine the smooth window, we analyze the dynamic threshold using percentage of snow depth extracted from the first-order derivatives under different smoothing windows. Snow depth percentage stabilizes when the smoothing window reaches around 30 days. The curve drops sharply for smaller windows but fluctuates minimally beyond this point. And snow on the lunar scale is more stable and reliable, less affected by disturbances. Thus, a 30-day smoothing window is chosen. Following your suggestions, we have extended the discussion on this issue in the Methodology section.

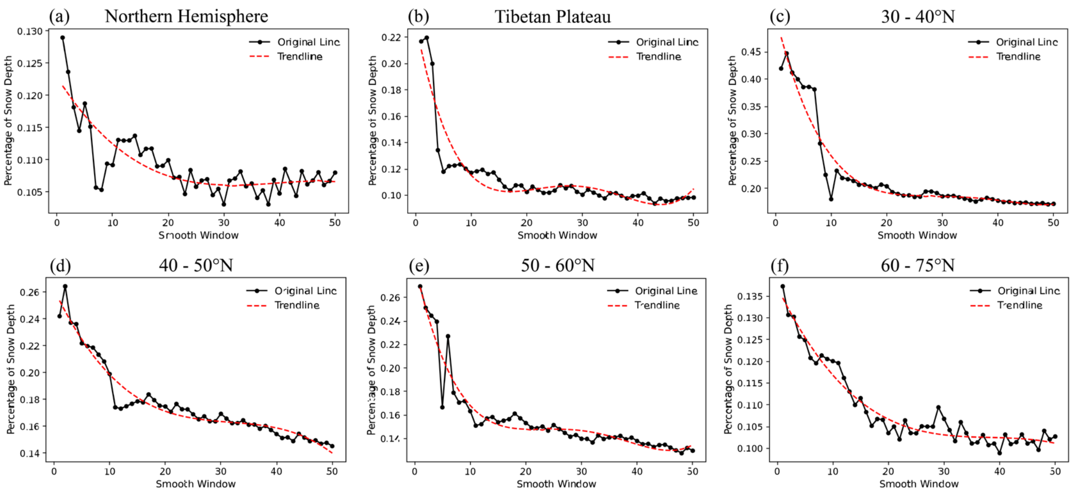

Figure R6. The relationship between the percentage of snow depth (dynamic threshold) and smooth window in (a) the Northern Hemisphere, (b) the Tibetan Plateau, (c) 30°N–40°N, (d) 40°N–50°N, (e) 50°N–60°N, and (f) 60°N–75°N. The black line is the original line, the black dot is the specific value for each year, and the red line is the trend line.

Line 143:

Occasional snowfall events, such as short-duration or localized snowfall, can introduce anomalous fluctuations in the snow curve, leading to multiple small peaks or atypical maxima. These fluctuations represent short-term meteorological noise rather than the long-term seasonal evolution of the snow cover. This is particularly common, especially *in unstable snow areas (e.g., the Tibetan Plateau). Smoothing can reduce these instabilities and make the snow curve more reflective of seasonal snowpack changes.*

Minor comments:

1. L 36: Concerning the decreasing length of the snow season in Notarcola (2022): wasn't this paper focused only on the mountainous regions?

Response:

Thank you for the insightful comment. The study by Notarnicola primarily focuses on mountainous regions, and its conclusions regarding the shortening of the snow season are also specific to these areas. This was not explicitly clear in our original text. We have now revised the manuscript to clarify this point.

Line 35:

Moreover, the length of the snow season and the number of snow days are shortened in global mountain regions (Notarnicola et al., 2022).

2. L98: It is a bit unclear what the authors mean by "replacing the dataset"?

Response:

The global snow depth data set of Che et al. is affected by satellite orbits, leading to substantial missing measurements at low and middle latitudes, especially on the Tibetan Plateau. The location of the missing data varies from day to day with no regularity. In the other data set of Che et al. on snow depth in China, an algorithm is used to fill in the missing data. Given that the mid- and low-latitude regions are representative of unstable snow conditions, they are the primary focus for improving snow phenology using our dynamic threshold approach. However, the presence of significant data gaps introduces uncertainties in snow phenology extraction that need to be addressed. To mitigate this, we replace the snow depth data for China in the global dataset with Chinese snow depth data, ensuring more accurate phenology representation in this region.

In order to give the reader a better understanding, we have changed the sentence in the manuscript.

Line 97:

The long-term series of daily global SD is affected by satellite orbits, leading to substantial missing measurements at low and middle latitudes. To minimize the negative effects of data gaps, we substitute the China region in the global dataset with another set of snow depth data for the China region (Che et al., 2015).

3.L106: What is SCE and where is it defined ? Or is it a Typo and should it be SCA as it refers to IMS ?

Response:

Thanks for your suggestion. This should indeed be changed to SCA as the IMS data can represent changes in SCA, which have been changed in the manuscript.

Line 106:

For the daily IMS binary SCE dataset, no additional processing is required to determine if the grid is covered with snow.

4.IMS also measures snow cover fraction: hence the labels (e.g. in Table 2, figures 3-4) should be more consistent: using IMS, SCF, SD mixes instruments names and the variables.

Response:

Thank you for your valuable suggestion. As you pointed out, IMS and MOD10C2 represent different forms of SCF, and using SCF to refer to MOD10C2 may cause confusion. To clarify this, we have made the following adjustments: the IMS label remains unchanged in all figures, while SCF has been replaced with MOD10C2 to explicitly refer to the dataset. This revision ensures a clearer distinction between the two data sources.

5.L145: The first derivative also goes to zero at the maximum of the snow curve.

Response:

Thank you for your reminder. Indeed, the first derivative corresponding to the maximum of the snow curve is also zero. To avoid confusion, the following annotations have been added to the manuscript.

Line 151:

When the first-order derivative equals zero (at the beginning and end of the curve, not the maximum), it shows that snow has not started accumulating or has completely melted.

6.When adapting the formula from the vegetation index, is it true that SnowMin is actually zero throughout this study?

Response:

$Snow_{min}$ is not universally zero; rather, it is determined by local snow conditions. In high-altitude regions with perennial snow cover, $Snow_{min}$ remains nonzero, whereas in mid- and low-latitude regions with relatively sparse snowfall, most $Snow_{min}$ values approach zero. This spatial variability is analogous to vegetation indices, which may not be zero depending on the specific local environment.

7.A couple of points are not clear in the Methodology: on one hand, "The above process is carried out for each grid in the NH (hence locally), yet above, it seems that the ratio is defined in "latitudinal zones" (implying a zonal average). The authors also mention multi-annual averages of the threshold, which implies that the threshold is defined for each year and then averaged, as opposed to using a climatological evolution to define a threshold. Please clarify.

Response:

For the first question, to assess the generalizability of the ratio, we conduct calculations not only for the entire NH and for individual latitude bands but also at each grid point (Figure 5). The results for the entire NH are presented in Figure 1, while those for the different latitudinal analyses are shown in Figure R6. The ratios are found to converge towards approximately 10%. When performing the same calculation at each grid point, we find that 73.05% and 82.65% of the two ratios fell within the 5%-15% range, respectively. Consequently, the final ratio is set at 10%, ensuring its validity both at the hemispheric scale and latitudinal zones, as well as at the local level.

For the second question, as you mentioned, the thresholds are defined annually, and the snow phenology for each year is extracted based on that year's specific threshold. This approach allows both the interannual variations in the thresholds and snow phenology to reflect climate evolution. For example, as the climate warms, changes occur in the maximum snow accumulation, the timing of snowfall, and snow duration, leading to shifts in the snow curve pattern. Consequently, the thresholds derived from the snow curves also vary, resulting in corresponding interannual changes in extracted snow phenology. In the manuscript, the thresholds are averaged over multiple years solely for spatial comparison with the traditional 2 cm threshold. The results indicate substantial variations in thresholds across different latitudes, suggesting that a uniform threshold is not appropriate. However, in the subsequent process of snow phenology extraction, the multi-year averaged threshold was not used. Instead, year specific thresholds are applied to ensure that the influence of climate evolution is not masked.

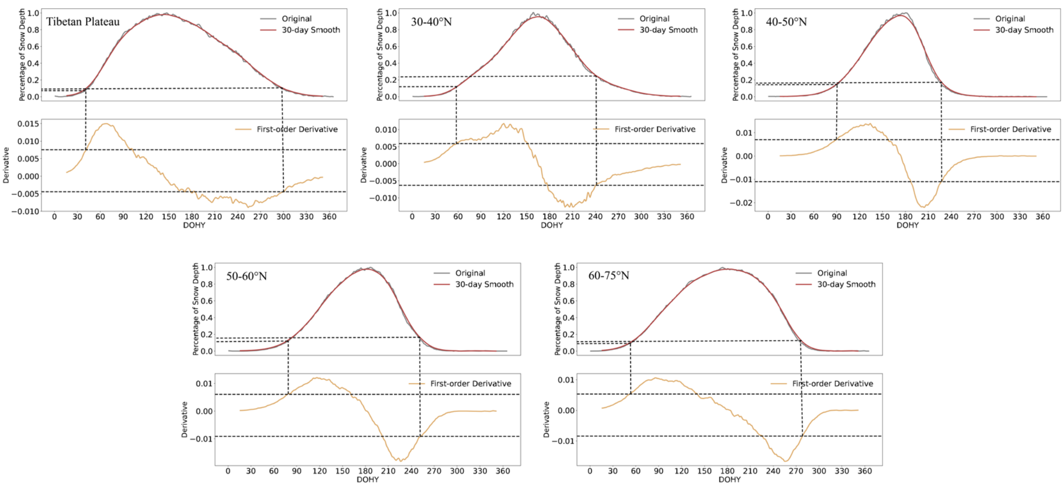

Figure R7. Schematic diagram of ratio results calculated at different latitudinal zones, including the Tibetan Plateau, 30°N–40°N, 40°N–50°N, 50°N–60°N, and 60°N–75°N. Similar to Figure 1.

8.Caption of Fig 4 over which years is this intra-annual (climatological?) variation established?

Response:

This represents the intra-annual variation curve averaged over the period from 2000 to 2018. The original figure caption do not specify this information, now it has been added.

Line 222:

Intra-annual variations of IMS, MOD10C2, and SD in five latitudinal zones (including the Tibetan Plateau) from 2000 to 2018.

9.L242-265: there are lot of repeats with the Methodology section 2.3

Response:

Thank you for your advice. We have condensed this section to try to avoid repeating sentences.

Line 140:

To identify the optimal $Snow_{radio}$*, we normalize and smooth the interannual SD variation curves for the entire NH and each latitudinal zone (including the Tibetan Plateau) with a 30-day moving window and then calculate their first-order derivatives. Occasional snowfall events, such as short-duration or localized snowfall, can introduce anomalous fluctuations in the snow curve, leading to multiple small peaks or atypical maxima. These fluctuations represent short-term meteorological noise rather than the long-term seasonal evolution of the snow cover. This is particularly common, especially in unstable snow areas. Smoothing can reduce these instabilities and make the snow curve more reflective of seasonal snowpack changes. We assume the curve has a single peak structure. The first-order derivative shows the rate of snow accumulation or melting, with its extreme points indicating the steepest changes. However, when the first-order derivative equals zero (beginning and end of the curve, not the maximum), it shows that snow has not started accumulating or has completely melted. The intermediate state between the maximum rate and no change represents when snow starts to accumulate or melting is nearly complete, which is what we are looking for in*

SCOD and SCED. So, here we simply choose the extreme midpoint of the first-order derivative as the snow curve turning point. The percentage of tthe urning point is the threshold we need. The above process is carried out for each grid in the NH, and Figure 1 shows a schematic of the entire NH. The percentages for SCOD and SCED fall between 5% and 15% (marked by red circles). This threshold range (i.e. $Snow_{ratio}$*) can therefore serve as an indicator for the snow season.*

10.The early onset of snow annual cycle for the TP is interesting; is it governed by the high-altitude areas? It might be worthed to split this Fig 6 into mountainous and non-mountainous areas, like done in the later part of the paper.

Response:

This is an interesting and complex issue. The early onset of the snow season on the Tibetan Plateau (TP) is a complex phenomenon influenced by multiple factors. The high-altitude environment of TP, with an average elevation exceeding 4,000 m, plays a crucial role in this process. At such elevations, lower temperatures prevail throughout the year, facilitating an earlier transition into winter and creating favorable conditions for snowfall. Additionally, the region's complex topography contributes to the retention of cold air, further enhancing local cooling effects (Yang et al., 2014). Moreover, TP is subject to the combined influence of the East Asian winter monsoon and the South Asian monsoon. During autumn, the intensification of the East Asian winter monsoon enhances cold air advection, leading to frequent southward intrusions of cold air masses. This process results in a rapid decline in temperature, promoting an earlier onset of snowfall (Wu et al., 2012). In short, the early snow season in TP is the result of multiple factors, and its high altitude makes a considerable contribution.

Taking your advice, I have divided Figure 6 into mountain ranges and non-mountain ranges and added it to Figure 6 in the manuscript with the appropriate captions. Regardless of the latitudinal belt, the snow curve in the non-mountainous region would be narrower than that in the mountainous region, implying a shorter snow season in the non-mountainous region. An interesting phenomenon is that the peak of non-mountain range snow arrives sooner. An interesting phenomenon is that the peak of non-mountain range snow arrives earlier in TP. Snow on TP mountains exhibits greater stability and shares characteristics with high-latitude snow, leading to a snow curve that closely resembles that of high-latitude mountainous regions. However, the snow season in the TP non-mountains is significantly earlier than in other latitudinal belts. This may be related to the fact that TP has more occasional random snowfalls, and the snow is stored due to colder temperatures. This is a very interesting phenomenon, and we will follow up by paying more attention to this phenomenon and conducting a more detailed study.

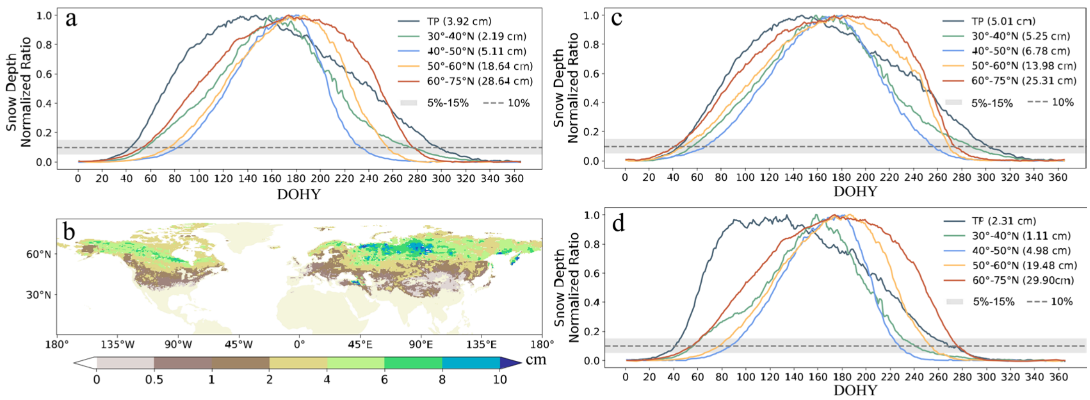

Figure R8. The dynamic SD threshold across the Northern Hemisphere. Intra-annual variability in the normalized SD for five latitudinal zones in (a) the whole Northern Hemisphere, (c) the Northern Hemisphere mountain ranges, (d) the Northern Hemisphere non-mountain ranges. Shading represents the interval 5%-15%, and the dashed line represents the dynamic threshold of 10%. Actual maximum snow depths for each latitude band are in parentheses. The unit DOHY is an abbreviation for day of the hydrological year, defined as September 1 through August 31 of the following year. (b) Spatial distribution of multi-year average snow depth thresholds in the Northern Hemisphere extracted using the snow dynamics threshold method.

Line 284:

Defining SDtopo greater than 200 as a mountain range divides the Northern Hemisphere into mountain ranges and non-mountain ranges. Regardless of the *latitudinal belt, the snow curve in the non-mountainous region would be narrower than that in the mountainous region, implying a shorter snow season in the non-mountainous region. Snow curves in mountainous regions are more stable and show the same pattern in the five latitudinal zones. Therefore, the location of the 10% threshold is appropriate in both mountain range and non-mountain range areas where the turnover change occurs.*

References:

Wu, G., Liu, Y., He, B., Bao, Q., Duan, A., & Jin, F.-F: Thermal Controls on the Asian Summer Monsoon. Scientific Reports, 2(1), 404. https://doi.org/10.1038/srep00404, 2012.

Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., & Chen, Y: Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. Global and Planetary Change, 112, 79–91. https://doi.org/10.1016/j.gloplacha.2013.12.001, 2014.

Typos & English.

1. L125 (and many other places) ratio not radio

Response:

Thank you for your correction, all statement errors in the manuscript have been corrected.

Line 125:

When the $NDVI_{ratio}$ exceeds is below a certain threshold, the corresponding day of the year is determined as the EOS.

2. Caption of Fig 1 : “has not started … has ended”

Response:

The sentence has been modified.

Line 137:

Gray shading indicates the snow season has not started or has ended, blue shading indicates the snow accumulation period, and red shading indicates the snow melting *period.*

3. L217: snow conditions at given grid points

Response:

We have revised the sentence.

Line 217:

Specifically, when the threshold is reduced, snow conditions at given grid points are more easily reached, resulting in a longer SCD, earlier SCOD, and later SCED.

4.Caption of Fig 4 : replace …"snow elements" by the annual snow maxima over the respective areas.

Response:

The sentence has been modified.

Line 224:

The values in the graphs characterize the annual snow maxima over the respective areas.