Response to Reviewer

Dear Dr. Xiao,

We sincerely appreciate your time and effort in discussing our study and sharing your insightful suggestions. Your feedback has been incredibly valuable in deepening my understanding of the topic and identifying areas for improvement. Our point-by-point responses to your comments are listed below in black.

Main comments:

1.According to numerous previous works and my analysis (https://doi.org/10.1016/j.isprsjprs.2024.07.018), the uncertainty in snow phenology is mainly derived from the accuracy of daily snow cover products. You always mentioned there are potential bias when using the fixed threshold to determine snow phenology. Through your whole text, you didn't specifically illustrate this type of uncertainties. May I ask you to give some examples to clarify this uncertainty what you mentioned?

Response:

We agree with your perspective that uncertainty in data will pass to snow phenology calculation. In this study, we did not explicitly quantify the bias because there are no direct observational data for snow phenology. Facing snow data uncertainty and a lack of snow phenology references, we try to optimize the methods for snow phenology extraction.

We do not mean that the traditional fixed threshold approach is incorrect, but we identify the limitations of the traditional fixed threshold method and explore possible improvements. Spatially, snow conditions vary across regions. Applying a uniform threshold implies treating all snow conditions as identical, which is unreasonable. Snow conditions also differ significantly over time, such as from the Industrial Revolution to the present.

We think that when snowfall can be steadily converted into snow on the ground, it means the onset of the snow season. In this work, we define the inflection point of snow coverage or depth curve, serving as a dynamic threshold for identifying the start of the snow season. In contrast, a fixed 2 cm or 50% threshold will appear as different phases of the snow season in different regions and time periods. Similarly, snow cover extent data are obtained via pre-processing, such as the sub-grid snow information extraction, which may introduce several thresholds or criteria.

As mentioned above, we have recognized the impact of data uncertainty on snow phenology identification, but we believe that the spatial and temporal heterogeneity of snow cover needs to be considered. This work is a preliminary exploration of dynamic snow phenology, and we hope to further refine it in future work, including your important suggestions. According to your comment, we revise our manuscript to further clarify the above statement.

---In the Abstract, "Previous studies commonly employed … leading to potential biases of snow phenology." This description is too general.

Response:

Thank you for the suggestion. In this sentence, we want to express the limitation of fixed thresholds. We have revised this sentence for the better understanding of the readers.

Line 7:

Previous studies commonly employed fixed threshold methods to extract snow phenology, which cannot represent the differences in the beginning/end of the snow period under different snow conditions in the Northern Hemisphere, leading to potential uncertainties of snow phenology.

--- Lines 65-66, "The fixed threshold for snow phenology fails to account for the variations in snow cover across the NH". Please give us more explanation on this. Do you have any related references to support your statement? Additionally, the following sentence "In fact, snow cover increase …. Especially on the TP" cannot support your statement on the uncertainties due to using fixed threshold.

Response:

As in the previous responses, we believe that a fixed threshold fails to represent the actual beginning/end of snow cover under different snow conditions, particularly when considering long-term climatic shifts and regional differences in snow.

We propose that the initiation of the snow season should be determined by identifying the point at which snowfall consistently leads to stable accumulation on the ground. In our approach, we define this using a dynamic threshold based on the inflection point of the snow coverage or depth curve. This inspiration comes from the methodology used in the extraction of vegetation phenology. It is well known that snow and vegetation in the Northern Hemisphere exhibit similar large-scale spatial patterns, both varying with latitude. However, while vegetation coverage decreases with increasing latitude, snow cover follows the opposite trend, expanding at higher latitudes. Additionally, both snow and vegetation are influenced by climate change, leading to temporal variations in their distribution and dynamics. The word "phenology" was originally derived from vegetation and was subsequently extended to include snow phenology. Consequently, methods for extracting vegetation phenology are more advanced and well-developed compared to those for snow phenology. At present, it has been proven in the methods of extracting vegetation phenology that fixed thresholds cannot accommodate spatio-temporal heterogeneity, ignore inter-annual variations, and are not applicable to diverse vegetation types, among a series of other problems (White et al., 1997; Mo et al., 2012). However, this issue regarding snow phenology extraction methods has not yet received attention and resolution. Building on insights from vegetation studies, we hypothesize that similar principles apply to snow. The fixed threshold method, however, fails to account for the temporal and spatial variability of snow, limiting its effectiveness in capturing dynamic snow processes.

We have added to the manuscript in order to make it more comprehensible to the reader.

Line 65:

The fixed threshold for snow phenology fails to account for the variations in snow cover across the NH. In fact, snow cover increases with latitude, with thick and stable snow cover at high latitudes and shallow and short-lived snow cover at middle and low latitudes, especially in the TP (Orsolini et al., 2019). In addition, the snow changes from year to year due to many aspects of the climate, and the regional snow cover trends exhibit a heterogeneous and non-linear response to its regional warming rate (Blau et al., 2024). Snow conditions are variable, but thresholds are always fixed, which can lead to inaccuracies of snow phenology. Therefore, employing different snow data and a fixed threshold will lead to uncertainties in extracted snow phenology. At present, it has been proven in the methods of extracting vegetation phenology that fixed thresholds *cannot accommodate spatio-temporal heterogeneity, ignore inter-annual variations, and are not applicable to diverse vegetation types, among a series of other problems (White et al., 1997; Mo et al., 2012). However, this issue regarding snow phenology extraction methods has not yet received attention and resolution. In fact, the onset of the snow season is marked by the sustained accumulation of snowfall as ground snow, rather than being determined by a fixed threshold. We aim to propose a novel method that incorporates both spatial heterogeneity in snow cover and temporal variability to extract snow phenology, reducing the uncertainty associated with the fixed threshold method from a physical meaning perspective.*

References:

Mo, J., Zhu, W., Wang, L., Xu, Y., and Liu J.: Evaluation of remote sensing extraction methods for vegetation phenology based on flux tower net ecosystem carbon exchange data. Chinese Journal of Applied Ecology, 23(2), 319-327. https://doi.org/10.13287/j.1001-9332.2012.0072, 2012.

White, M. A., Thornton, P. E., and Running, S. W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles, 11, 217–234. https://doi.org/10.1029/97GB00330, 1997.

--- Lines 68-69: "Snow conditions are variable, … underestimation or overestimation of snow phenology." Please add more explanations on this underestimation and overestimation. Is there any published works to support this statement?

Response:

It is a misrepresentation here. As mentioned in your first question, there are no reference data to support our assessment of whether snow phenology is underestimated and overestimated.

Therefore, we have revised the description to highlight our consideration of the impact of spatial and temporal heterogeneity of snow in the Northern Hemisphere on snow phenology. Please see the answer to the previous question for more details.

2.Line 14-15: “At low and middle latitudes, the snow cover duration (SCD) extends, the snow cover onset day (SCOD) advances, and the snow cover end day (SCED) delays, …” This conclusion is quite different from what we got up to now. Does your analysis conclusion tell us we have more snow in this region in a warming world?

Response:

We apologize for any ambiguity in my previous explanation, which may have led to a misunderstanding. The comparison of extending and advancing here specifically refers to the traditional fixed threshold method. In particular, compared to the conventional 2 cm threshold, the dynamic threshold method results in a longer snow cover duration (SCD), an earlier snow cover onset date (SCOD), and a later snow cover end date (SCED) at low and mid-latitudes, especially over the Tibetan Plateau, where the SCD difference can reach up to 28 days. Conversely, at higher latitudes, the changes follow the opposite pattern. Therefore, the reference here does not pertain to absolute changes in time scales but rather to the relative differences in snow phenology results derived from the two methods. To prevent misinterpretation, clarifications and additional labels have been incorporated into the manuscript.

Line 13:

Using the dynamic threshold method, there is an earlier snow cover onset day (SCOD), a later snow cover end day (SCED), and a longer snow cover duration (SCD) at low and middle latitudes, especially on the Tibetan Plateau, where the SCD differences can reach 28 days. The differences in snow phenology at higher latitudes are reversed.

3.MOD10C2 products provide a maximum snow cover extent during this 8-days period. That means you would potentially overestimate the snow cover phenology metrics (snow cover days, snow onset date and snow end date) when you used this data. “For the SCF dataset, … after the last identified snow cover.” (Line 112-113) Please cite references to support this processing’s reasonability.

Response:

Thank you for your question. As you mentioned, a major limitation of optical remote sensing is cloud contamination, which obscures snow data on most days during the snow season. To mitigate this issue, 8-day and 16-day composite snow cover products are generated to reduce cloud interference (Hall et al., 2002). The MOD10C2 data is one of the products of optical remote sensing, which provides the maximum snow cover extent over an 8-day period. Consequently, using it for snow phenology extraction may cause an overestimation of the snow season. In order to compare snow phenology results from different data, we chose this data set with reference to Chen et al. (2015). However, the bias that this data set brings to snow phenology is not negligible. Therefore, we will follow up by choosing daily-scale, more accurate SCF data for the extraction of phenology. The method used in L112 is adopted from the study by Chen et al. (2015). The citations have now been added to the manuscript.

Line 112:

For the MOD10C2 dataset, considering the 8-day temporal resolution, the SCOD is the date four days before the first identified snow cover, and the SCED is four days after the last identified snow cover. SCD is determined by multiplying the number of snow occurrences by eight (Notarnicola, 2020; Yue et al., 2022; Guo et al., 2022; Chen et al., 2015).

References:

Chen, X., Liang, S., Cao, Y., He, T., and Wang, D.: Observed contrast changes in snow cover phenology in northern middle and high latitudes from 2001–2014. Scientific Reports, 5. https://doi.org/10.1038/srep16820, 2015.

4.1: which part of data do you use to plot this figure? 1989-2018 or 2000-2018? Single data and which data? Additionally, this curve is like for vegetation growth instead of snow cover evolution. Please check it again. The fact is that Summer is no snow (DOY=180 days). Finally, how can I understand the term "percentage of snow depth"? For your reference, here is a in-situ observation of snow evolution within a snow cover year (10/1 – 9/31) https://www.climatehubs.usda.gov/hubs/northwest/topic/30-year-normals. Regarding only snow depth variable used in your method, I am confused how you use snow cover products (IMS and SCF) to calculate snow phenology metrics.

Response:

Thank you for your question. We will address each point individually.

1. The snow depth data used to generate this figure covers the period from 2000 to 2018. I apologize for any ambiguity in my previous explanation that may have led to a misunderstanding. In fact, the figure presents results based on the hydrological year, which spans from September 1 to August 31 of the following year. As a result, Day 180 falls within the winter season, leading to a high SD value. To avoid such confusion, I have revised "DOY" in the text to "DOHY" (day of the hydrological year), which I hope will improve clarity.

2. The label "percentage of snow depth" represents the value after normalizing for snow depth. The schematic equation is as follows.

$$percentage\ of\ snow\ depth = \frac{Snow\ depth - Snow\ depth_{min}}{Snow\ depth_{max} - Snow\ depth_{min}}\ , \quad (1)$$

3. In Section 3.1, Snow depth is selected as the driving data because it consistently follows a stable single-peak pattern across different latitudinal zones, effectively representing the accumulation and melting processes of snow. In contrast, snow cover data are more irregular and spatially diverse, especially over the Tibetan Plateau (TP). Furthermore, snow cover measurements are significantly impacted by the polar night, causing considerable inaccuracies north of 60°N, while cloud interference further compromises data reliability. Therefore, we use the snow depth dataset to improve the snow phenology extraction method. In fact, the dynamic threshold method can also be applied to SCF data. The results of the new method using a dynamic threshold for SCF are nearly similar to those of SD, and we show the results for SCF in the Appendix. However, IMS data have already undergone preprocessing. Since their thresholds cannot be adjusted, it is not possible to apply a dynamic threshold to IMS data.

5.Through your manuscript, you used a 5% or 10% to determine snow phenology metrics. Is this another type of "fixed threshold" method. Additionally, you didn't include some comparisons, validations, and evaluations. How do you convince our readers of the responsibility (not accuracy) of your snow phenology results? Because your method quite largely different the method we usually used.

Response:

We agree with your point of view. We acknowledge that the dynamic threshold method for snow phenology is not fully dynamic; however, it is relatively more dynamic compared to the traditional fixed threshold method. Specifically, the fixed threshold approach fails to account for interannual variations and spatial differences in snow, treating snow uniformly across all locations and time periods without distinction, which is evidently unreasonable. The fixed threshold has no physical meaning and no consideration of the different states of the snow season. In fact, the snow season is signaled when snowfall can be stabilized into snow on the ground. The currently implemented 10% dynamic threshold method partially addresses some of the limitations of the fixed threshold method, though it is not yet fully perfect. Moreover, the 10% threshold is primarily derived from mathematical calculations. In future work, we plan to integrate atmospheric variables to further enhance the physical significance of the 10% dynamic threshold.

For the second question, the rationalization of dynamic snow phenology methods should start with the problems of existing fixed methods. The limitations of the fixed threshold method are described in detail in Main Comment 1 and will not be repeated here. Please refer to the answer in Main Comment 1 for details.

6.Based on your analysis, I wonder if your approach means that snow phenology metrics will vary depending on the study area used. For example, the snow depth in the Northern-Xinjiang is generally higher, while it is lower in the TP. ---Case1: Only TP data is used to analyze snow phenology metrics for TP. ---Case2: The data both in the TP and Northern-Xinjiang are used to analyze snow phenology metrics for TP. Are the snow phenology metrics in the TP region different between Case 1 and Case 2?

Response:

We think the snow phenology of the TP obtained from Case 1 and Case 2 in your question is the same. Since the threshold is calculated individually for each grid point without applying a regional average, each grid point has its own specific threshold based on its snow curve. Given the differences in snow conditions between the TP and Xinjiang, the extracted thresholds are inherently distinct. The thresholds for grid points on the TP are derived solely from their own snow characteristics and are independent of the snow conditions in Xinjiang. Therefore, including or excluding Xinjiang's snow data does not affect the threshold values for TP grid points, nor does it influence the resulting snow phenology metrics.

Minor comments:

1. Change "on" to "in" in line 67.

Response:

We have revised the sentence.

Line 67:

In fact, snow cover increases with latitude, with thick and stable snow cover at high latitudes and shallow and short-lived snow cover at middle and low latitudes, especially in the TP.

2. Change "24 hours" to "daily" in line 95

Response:

The sentence has been modified.

Line 95:

The dataset is a daily product from 1980-2018 with a spatial resolution of 25067.53 meters, and shows a relative deviation within 30%.

3. Change "the hydrological year" to "a hydrological year" in line 111

Response:

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

Line 111:

SCOD is defined as the first day with the first continuous snow cover exceeding five days in a hydrological year, whereas SCED is the last day with the last continuous snow cover exceeding five days.

4. How do I understand the "snow index" in Line 132? Snow depth? Snow cover days? NDSI?

Response:

"Snow index" is a pronoun that can denote any of the elements of snow. For better understanding, we've changed it to snow element and given an example in parentheses afterward.

Line 129:

To investigate the snow phenology in different areas across the globe, we propose a dynamic threshold for snow phenology.

$$Snow_{ratio} = \frac{Snow - Snow_{min}}{Snow_{max} - Snow_{min}} , \quad (2)$$

where $Snow_{max}$ *is the annual maximum snow element (e.g., snow cover fraction, snow depth) and* $Snow_{min}$ *is the annual minimum snow element.*

5. Line 101-102: "This data has been validated against … less than 5 cm account for approximately 65% of all the data" Please add the reference for this statement to support this "5cm - 65%" number pair.

Response:

Thank you for your advice. We have added the reference for this statement to support this "5cm - 65%" number pair.

Line 101:

This data has been validated against meteorological observations, and absolute errors of less than 5 cm account for approximately 65% of all the data (Che et al., 2008).

References:

Che, T., Li, X., Jin, R., Armstrong, R., and Zhang, T.: Snow depth derived from passive microwave remote-sensing data in China. Annals of Glaciology, 49, 145–154. https://doi.org/10.3189/172756408787814690, 2008.

6. Line 141: Why do you use “30-day” moving window? Any specific reasons?

Response:

Thanks for the comment. The purpose of smoothing is simply to eliminate noise and avoid the effect of chance snowfall timing on the extraction of snow phenology. Snow changes are more stable and reliable on a monthly scale and are not affected by random snowfall.

To argue this point, we analyze the dynamic threshold using percentage of snow depth extracted from the first-order derivatives under different smoothing windows (see Methods) and present the results in a line graph (Fig. R1). The percentage of snow depth threshold stabilizes when the smoothing window reaches approximately 30 days. The curve exhibits a sharp decline for smoothing windows smaller than 30 days, while for values greater than 30 days, it shows minimal fluctuations. Based on this analysis, we select a smoothing window of 30 days. Although minor adjustments to the smoothing window have little impact on the results, it should not be too small, as this increases sensitivity to noise, nor too large, as it may erase some valid information.

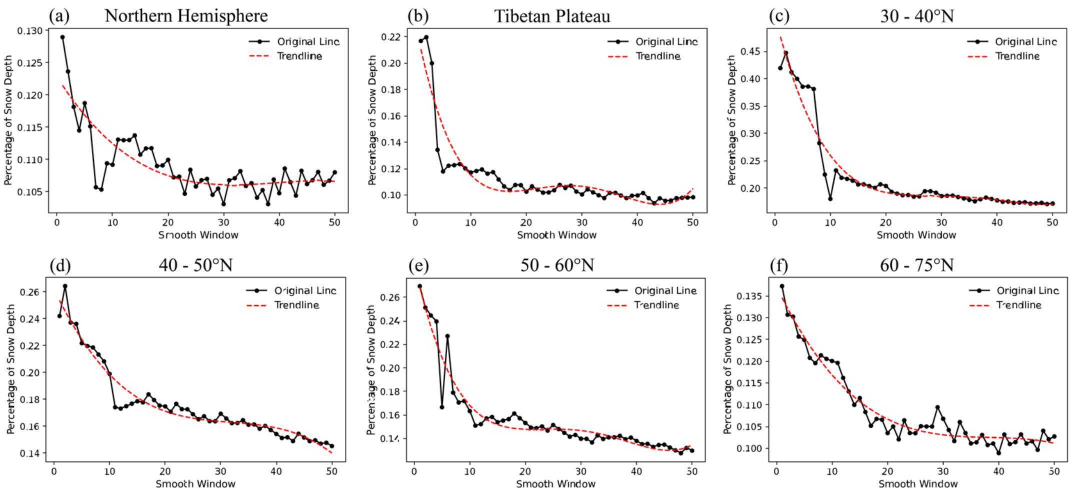

Figure R1. The relationship between the percentage of snow depth (dynamic threshold) and smooth window in (a) the Northern Hemisphere, (b) the Tibetan

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.

7. Section 2.1: how did you handle the different spatial resolution of these snow products?

Response:

For snow products with different resolutions, we have standardized them to a resolution of 0.25°. We have now commented in the manuscript.

Line 102:

For snow products with different resolutions, we have standardized them to a resolution of 0.25°.

8. Line 246-247: "Given that the zonal variations … compared to a fixed threshold." It is not clear.

Response:

We apologize for the lack of clarity in our previous explanation. Our intended meaning is that the spatial distribution characteristics of snow are similar to those of vegetation, both exhibiting a latitudinal zonation. However, while vegetation decreases with increasing latitude, snow increases. The key point is that there are clear similarities between the two. In vegetation phenology extraction methods, both fixed threshold and dynamic threshold methods exist. Numerous studies comparing these approaches have consistently shown that the dynamic threshold method is more reasonable. Given the relationship between snow and vegetation, we hypothesize that a dynamic threshold method should also be more accurate than a fixed threshold method for snow phenology extraction.

This point was not clearly conveyed in the original manuscript, and we have now revised it accordingly.

Line 251:

Given this similarity in zonal variation, we propose a dynamic threshold method for extracting snow phenology, inspired by a commonly used approach in vegetation phenology.

9. Are the days in Fig 2 and Table 2 DOYs (Day of Year)? If it is, 66 of SCOD in table 2 means the snow starts in March. Please check the whole paper’s figures again.

Response:

Similar to Figure 1, the hydrological year is still represented here. Annotations have been added to the figure notes and tables.