Reply on RC1

General comments:

The manuscript titled "Duration of vegetation green-up response to snowmelt on the Tibetan Plateau" by Ni and colleagues investigates the complex interactions between spring phenology and snowmelt dynamics on the Tibetan Plateau (TP), a region noted for its ecological sensitivity to climate change. The authors employ satellite-derived phenological data along with various statistical analyses to explore the spatiotemporal patterns and drivers of the time difference between snowmelt and vegetation green-up. This work offers valuable insights into TP ecosystem dynamics, aligning with the research interests of EGUsphere readers. However, there are several issues that could affect the robustness of the conclusions. The following major suggestions aim to enhance the paper's scientific impact and clarity.

Response: We sincerely appreciate your feedback and suggestion. In this revision, we have introduced an additional criterion for identifying the start of snowmelt (D_{SOM}), recalculated the time differences (ΔD), and incorporated the updated D_{SOM} data into the subsequent statistical analysis. These revisions have been presented more concisely and clearly. Additionally, we have tightened the significance level to $p < 0.05$. Based on these new results, we have refined the language and restructured the results section. Our specific responses to each suggestion are as follows.

Special comments

Major concern 1: The first concern pertains to the snow coverage. While the TP experiences frequent snowfall, snow cover duration can be brief due to sublimation and wind dispersal. It is essential to verify that the study areas experience sustained snow cover throughout winter, not just isolated pixels as depicted in Figure 2. Additionally, consider streamlining the main text by moving certain figures (e.g., Figures 1 and 2) to the supplementary materials.

Response:

Indeed, while the Tibetan Plateau (TP) experiences frequent snowfall, its snow cover is typically transient, shallow, and patchy (Lei et al., 2023). For snow to have a measurable impact on vegetation, it must persist over time rather than appear only briefly. Therefore, following Chinese snow classification standards, we have introduced an additional criterion to the original two in Section 2.3.1. The duration of winter snow cover for each pixel must exceed 10 days:

Additionally, to ensure that the snow cover is not transient and can influence vegetation, we have introduced a third criterion: the snow cover duration in winter must exceed 10 days (Zhao et al., 2022).

The corresponding results and figures have been modified in the original text.

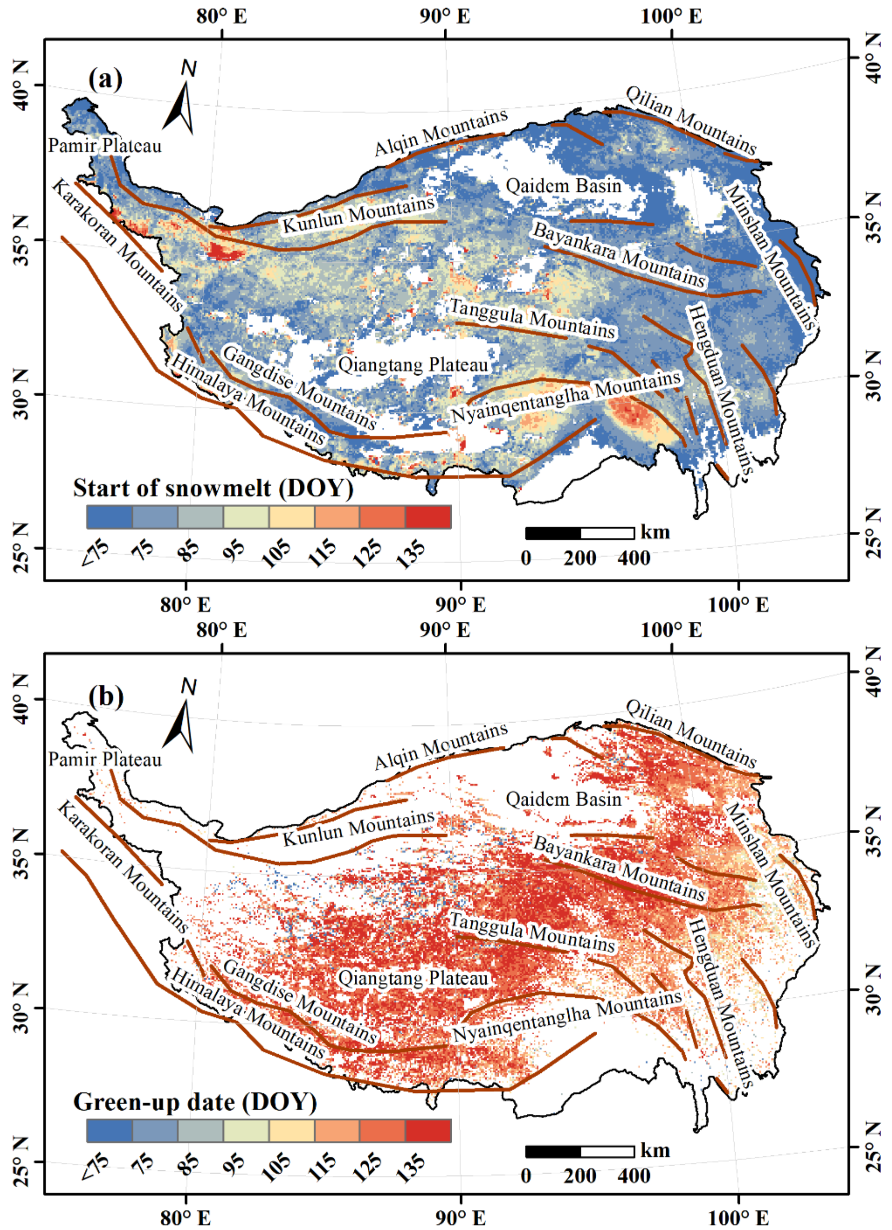


Figure 2: Spatial distribution of multiyear averaged (a)$\boldsymbol{D_{SOM}}$ and (b)$\boldsymbol{D_{GU}}$ from 2001 to 2018 on the Tibetan Plateau.

Additionally, following your valuable suggestion, original Figure 2 has been moved to the supplementary material to streamline the main text, and the figure numbering has been adjusted accordingly.

References:

Lei, Y., J. Pan, C. Xiong, L. Jiang & J. Shi.: Snow depth and snow cover over the Tibetan Plateau observed from space in against ERA5: matters of scale. Climate Dynamics, 60, 1523-1541, 2023.

Zhao, Q., Hao, X., Wang, J., Luo, S., Shao, D., Li, H., Feng, T., and Zhao, H.: Snow Cover Phenology Change and Response to Climate in China during 2000–2020, Remote Sens., 14, 3936, 2022.

Major concern 2: The second concern is about the statistical analysis. Firstly, the variables—spring mean temperature, total spring rainfall, and daily snowmelt—are likely to exhibit multicollinearity, given the interdependence of temperature/rainfall and snowmelt. A multicollinearity check is recommended to identify and potentially exclude highly correlated variables. Employing a structural equation model could provide a more nuanced understanding of these interdependencies. Secondly, the chosen significance level ($p < 0.1$) was too large. Despite this, a substantial number of pixels on the TP show non-significant trends, suggesting an absence of robust relationships between green-up and snowmelt. To address this, consider categorizing pixels by significance, explaining the underlying causes for these patterns in each category.

Response:

We agree that testing for multicollinearity is crucial in multiple linear regression to ensure no redundancy among independent variables. Therefore, we assessed multicollinearity using the Variance Inflation Factor (VIF), with the results as follows:

Table R1: Variance Inflation Factor for T_{spring}, P_{spring} and S_{StoG}

Variable	VIF
T_{spring}	1.243
P_{spring}	1.209
S_{StoG}	1.174

Since all variables have a VIF < 3, the data passes the collinearity criteria. This ensures the validity of proceeding with multiple linear regression analysis. We have further elaborated on this test in Section 2.3.4 of the methodology.

The partial correlation coefficient between each variable and ΔD was calculated to quantify their relationship. Subsequently, a multiple linear regression *model was established for each pixel (Equation 8). A prerequisite for multiple linear regression is passing the collinearity test, which requires the Variance Inflation Factor (VIF) to be less than 3, indicating no collinearity. In this study, the VIF values for* T_{spring}, P_{spring}, *and* S_{StoG} *were 1.243, 1.209, and 1.174, respectively, confirming that the collinearity test was satisfied.*

Compared to traditional regression models, the structural equation model (SEM) not only quantifies the contribution of independent variables to dependent variables but also reveals relationships among independent variables and mediating effects. This provides deeper insights into the influence mechanisms of T_{spring}, P_{spring}, and S_{StoG} on ΔD. Your suggestion is highly valuable. However, to ensure the credibility of the results, we assessed precision, model fit, and significance using R^2, SRMR, and p-values, respectively. The results are presented in Figure R1:

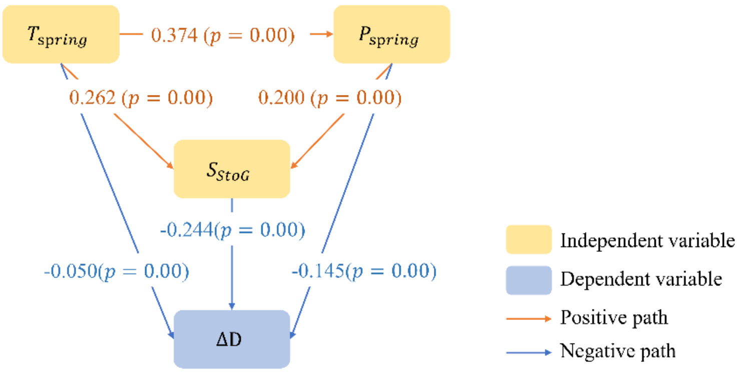


Figure R1. SEM path effect diagram for the response of D_{GU} to D_{SOM}

Although all model paths are significant (p < 0.05) and the fit is acceptable (SRMR = 0.00 < 0.08), the precision (R^2 = 0.117) is insufficient for drawing definitive conclusions. Therefore, further investigation using additional methods is needed to better elucidate the underlying mechanism.

Regarding significance, $p < 0.05$ is indeed a more appropriate criterion. Accordingly, all results have been revised based on this standard.

In this study, significance testing is conducted using the t-test, as detailed in Equation R1, which directly depends on the correlation coefficient (r) and degrees of freedom (df).

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{df} \tag{R1}$$

The t-value can be converted to a p-value by looking up to the critical value table, which depends on the sample size (df). In this study, we utilized the stats library in Python 3.8 to perform this calculation. Since each pixel has an average of 12 samples with three independent variables (x), the average degrees of freedom (df)

is 8 ($df = n - x - 1$). For df = 8, we conducted an experiment to calculate the significance indicator p-values for different correlation coefficients (r), as shown in Table R2, R3.

Table R2: Significance levels for correlation coefficients ($r = 0.1 \sim 0.9$ with 0.1 interval) at $df = 8$

Correlation coefficient (r)	Significance level (p)
0.1	0.783
0.2	0.580
0.3	0.400
0.4	0.252
0.5	0.141
0.6	0.067
0.7	0.024
0.8	0.005
0.9	0.000

Table R3: Significance levels for correlation coefficients ($r = 0.61 \sim 0.69$ with 0.01 interval) at $df = 8$

Correlation coefficient(r)	Significance level (p)
0.61	0.061
0.62	0.056
0.63	0.051
0.64	**0.046**
0.65	0.042
0.66	0.038
0.67	0.034
0.68	0.031
0.69	0.027

Only when r exceeds 0.64 does the test pass with $p < 0.05$. However, it is generally accepted that $r \geq 0.75$ indicates a strong correlation, $0.5 \leq r < 0.75$ indicates a moderate correlation, and $0.25 \leq r < 0.5$ indicates a weak correlation. Consequently, with small sample sizes, only pixels with strong correlations pass the test (Bonett and Wright, 2000), potentially overlooking some valid information. In our study, a valid sample requires that both the green-up date (D_{GU}) and D_{SOM} be valid, with D_{SOM} preceding D_{GU}. Therefore, the average sample size of 12 makes passing the significance test difficult. To address this, we applied the first law of geography by expanding the sample size to include the pixel itself and its eight neighboring pixels (Figure R2).

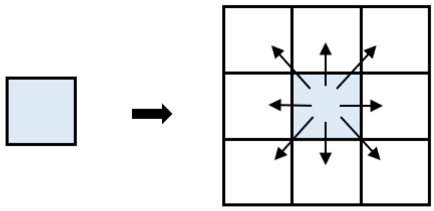


Figure R2. Diagram of sample size expansion

As a result, at the $p < 0.05$ level, the significance ratios for T_{spring}, P_{spring}, and S_{StoG} with ΔD were 23.5%, 28.8%, and 35.4%, respectively (Figure 6). Meanwhile, the dominant factor influencing ΔD was recalculated using the same methodology (Figure 7). While enhancing the significance, the original conclusions remain largely unchanged, with only minor revisions to some text and results.

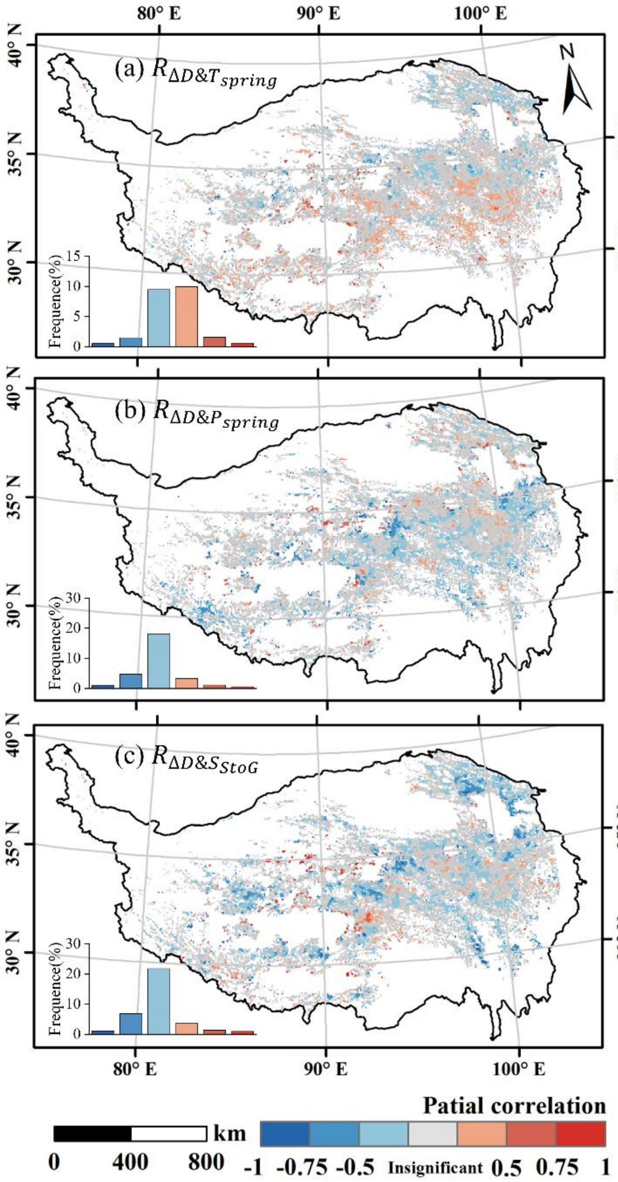


Figure 6: Spatial distribution of the partial correlation between ΔD and (a) spring mean temperature ($R_{\Delta D\&T_{spring}}$), (b) spring total rainfall ($R_{\Delta D\&P_{spring}}$), and (c) daily snowmelt from D_{SOM} to D_{GU} ($R_{\Delta D\&S_{StoG}}$).

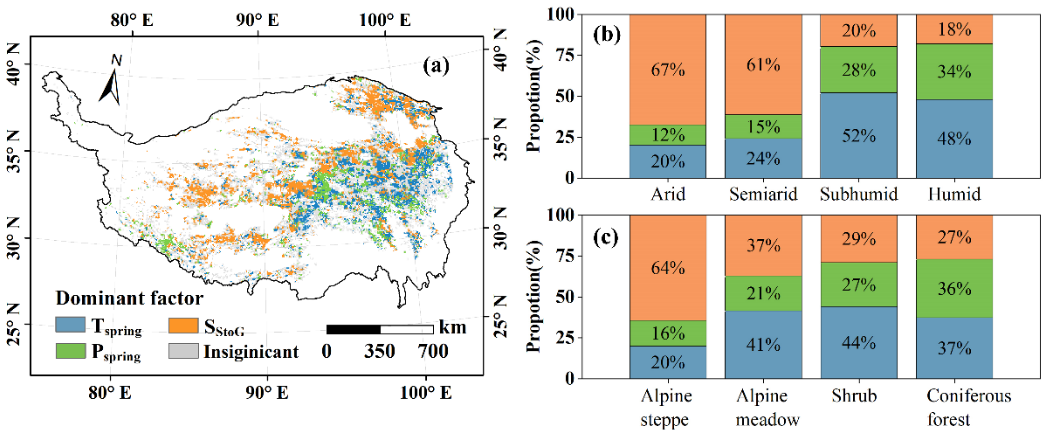


Figure 7: (a) Spatial distribution of dominant factor of ΔD and its proportion diagram among (b) different geographical zones and (c) different vegetation types.

References:

Bonett, D. G. and Wright, T. A.: Sample size requirements for estimating Pearson, Kendall and Spearman correlations, Psychometrika, 65, 23-28, https://doi.org/10.1007/bf02294183, 2000.

Major concern 3: The third concern is about the structure of results. The results section currently includes an extensive number of figures, which may hinder clarity. Consider reorganizing this section, for example, separating spatial and temporal characteristics into distinct parts and using concise figures. This restructuring could improve readability and emphasize key findings more effectively.

Response:

To improve clarity, we have restructured original Section 3.1 into three subsections: Section 3.1 (Spatial variation of D_{SOM} and D_{GU}), Section 3.2.1 (Spatial characteristics of ΔD) and Section 3.2.2 (Temporal characteristics of ΔD). Additionally, we have streamlined the language in the Results section by removing unnecessary details and focusing on key findings. For example, instead of listing specific values for each region in the spatial distribution of ΔD, we now emphasize the general downward trend. We have also removed the specific area proportions for different patterns of local Moran's I value.

Minor concern:

1. L15: Clarify whether you mean the "duration or date" of vegetation green-up.

Response: We would like to emphasize that while the impact of snowmelt on vegetation has been verified, the response of vegetation is not instantaneous and exhibits a certain degree of lag. Therefore, the focus of this study is to investigate the length of this lag, or the time difference between two dates. In our view, the term "duration" more accurately captures the concept of a time difference (a period) compared to "date."

2. L18-19: It is unnecessary to listing all methods here.

Response: As suggested, we have removed *the heatmaps and box plots*, retaining only the two primary methods.

3. L50-55: too many abbreviations make this paper hard to follow

Response: We have streamlined the text to reduce the use of abbreviations and improve the flow of the narrative. The revised text is as follow:

Snow phenology serves as a crucial indicator of changes in snow cover. Several studies have analysed the impact of snow phenology on D_{GU} *in the TP. The Snow cover end date typically exhibits a significant positive correlation with* D_{GU}*, with each 1-day advancement leading to a 0.56 days earlier* D_{GU} *(Potter, 2020; Wu et al., 2023). In contrast, the effect of snow cover duration on* D_{GU} *is more complex and region-dependent. For instance, a longer snow cover duration leads to a delayed* D_{GU} *in the western TP, while it advances* D_{GU} *in the eastern TP(Huang et al., 2019; Xiong et al., 2019). Notably,* D_{GU} *is most sensitive to the start of snowmelt (*D_{SOM}*) among various snow cover phenology metrics on the TP (Xu et al., 2022).*

4. L118: Provide more details for this treatment. For instance, if merging 10 pixels with various plant functional types (PFTs), specify which PFT the combined pixel represents. Confirm if all PFTs were included, and consider excluding bare and arable land, which lack seasonal dynamics relevant to this analysis.

Response: The original land cover types in the dataset remain unchanged, and non-seasonal and non-vegetated land were excluded from the experimental samples. Following your suggestion, we have added further clarification in Section 2.2.4:

Considering that some land covers are non-seasonal or non-vegetation, this study focuses exclusively on alpine steppe, alpine meadow, shrub, coniferous forest, and broad-leaved forest.

5. Figure 3: please make sure two figures have matched pixels and adjust the colorbar in 3b for better visibility.

Response: The discrepancy in the number of valid pixels for D_{GU} and D_{SOM} arises from differences in pixel selection criteria and identification methods. The objective here is to clarify the distinct distribution of valid pixels for the two dates. In the subsequent calculation of ΔD, only matched pixels that were valid for both dates were considered. Additionally, in response to your suggestion, the color bands in both figures have been standardized (Figure 2).

6. Figure 4: With 18 subplots, distinguishing annual differences is challenging. Move this figure to supplementary materials and replace it with a simplified version, such as a comparison between two periods (e.g., 2001-2009 vs. 2010-2018).

Response: Replacing the annual subplots with period average values is a constructive suggestion. However, the Mann-Kendall test results indicate no significant temporal changes in ΔD. Therefore, we have calculated the average over the entire period (Figure 3a) and moved the original Figure 4 to the supplementary materials.

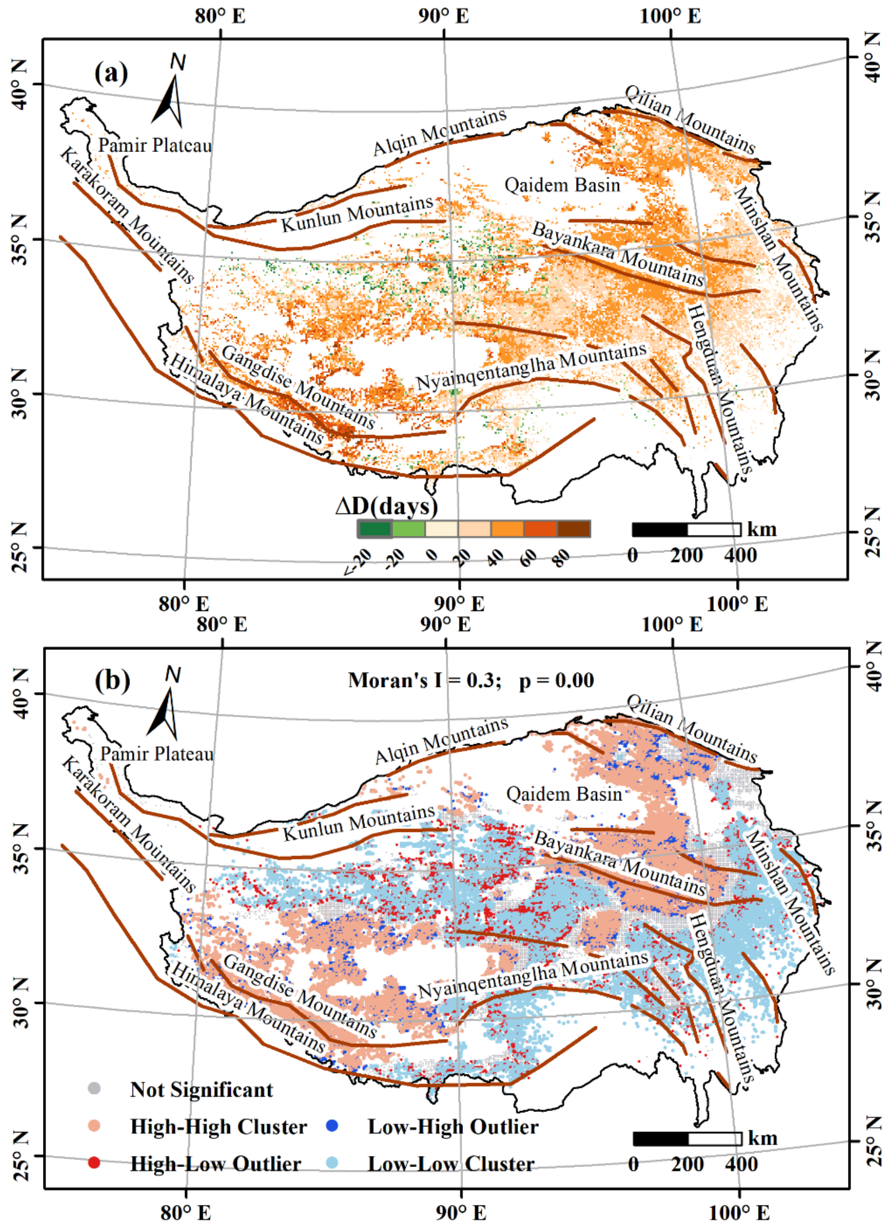


Figure 3: (a)Spatial, frequency distribution histograms and (b) local and global Moran's I values of average $\Delta \boldsymbol{D}$ *on the Tibetan Plateau over 2001–2018.*

7. Figure 5: Similar suggestion as Figure 4—consider a more concise format.

Response: We have redrawn Figure 3b using the same methodology, and the original Figure 5 has been moved to the supplementary materials.

8. The direct relationship between temperature/precipitation and green-up may be more pronounced than that of snowmelt. If so, this would suggest a lesser role for snowmelt, especially given the year-round snowfall on the Tibetan Plateau.

Response: Indeed, temperature and precipitation are the two most significant determinants of D_{GU}, and snowmelt has also been shown to influence D_{GU}. However, the duration and driving factors of vegetation response to the onset of snowmelt remain unclear. Therefore, this study focuses not only on D_{GU} itself but on the vegetation's response to D_{SOM}. The independent variable in this study is not the green-up date (DOY), but the time difference between D_{GU} and D_{SOM} (ΔD).

9. Figure 7: Why the figure 7b use a different pattern unlike 7a? it is much better to testing Tspring and Sstog, and Pspring and Sstog effects on deltaT effect on ΔT separately for clarity.

Response: Indeed, our intention was not to compare the combined effects of two independent variables. Thus, there was no need to use a dual-axis heatmap, which conveys different information compared to the box plot and is less suitable for comparisons between variables. The updated figure now uses consistent boxplots for all variables, improving clarity in the comparisons (Figure 5). Additionally, we have revised Section 3.2 based on the new results:

Figure 5 illustrates the mean value of ΔD under varying spring meteorological conditions. ΔD exhibits a clear stepwise decline from cold to warm regions, decreasing from approximately 48 to 37 days (Fig. 5a). In colder or hotter spring conditions (i.e., $T_{spring} < 270\ K$ or $T_{spring} > 275\ K$), ΔD decreased slightly. However, near the freezing point (270–275 K), ΔD shortens by 3 days with each 1K increase in T_{spring}. Under various precipitation conditions (Fig. 5b), ΔD shortens by 0.29~1.96 days for every 10 mm increase in P_{spring}. Fig. 5c reveals a strong negative correlation between ΔD and S_{StoG} when S_{StoG} exceeded 6 mm day^{-1}. For each 1 mm increase in S_{StoG}, ΔD decreases by approximately 0.615 days. The dispersion within each snowmelt category remains relatively consistent, with a standard deviation of about 16.8 days.

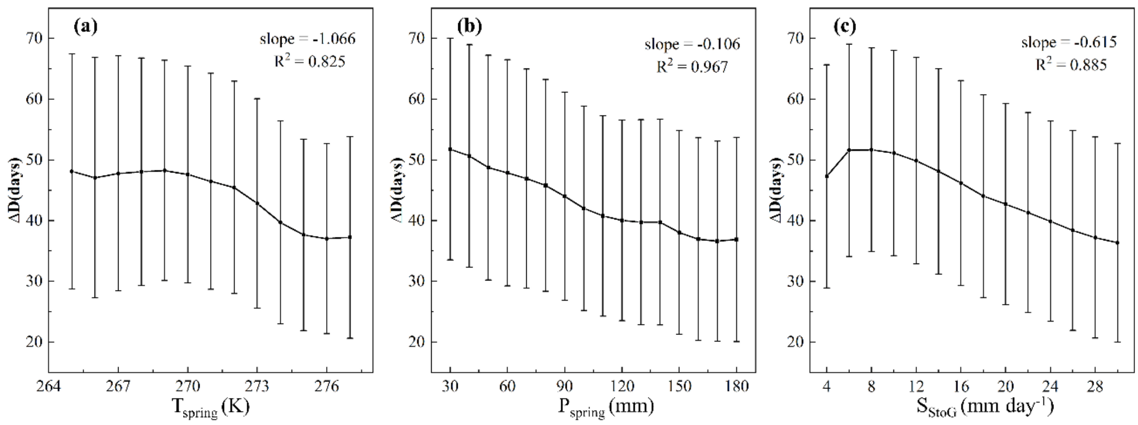


Figure 5: Variations in ΔD across regions with differing (a) spring mean temperature (T_{spring}), (b) spring total rainfall (P_{spring}), and (c) daily snowmelt from D_{SOM} to D_{GU} (S_{StoG}). Points represent the mean ΔD, while error bars denote one standard deviation. The slope and R2 value reflect the coefficient and precision of the linear regression, respectively, with a significance level of 0.01.

10. Conclusion: Condense to focus on primary findings for a stronger impact.

Response: We have streamlined the conclusion by removing some detailed result-oriented information and emphasizing the core findings and key contributions of the study.

This study investigates the dynamic response of vegetation to snowmelt on the Tibetan Plateau from 2001 to 2018. Our results reveal that the effect of snowmelt on vegetation is not immediate, with a mean response lag of 38.5 days from D_{SOM} to D_{GU}. Notably, the false spring was observed in the north-western TP, which warrants further exploration. As precipitation and snowmelt increase, the response time shortens. More complex than these factors, temperature exerts a greater influence on D_{GU} than D_{SOM} in colder regions, thus shortening the response time. Conversely, in warmer areas, increased temperatures have a stronger impact on D_{SOM}, which lengthens the response time. Furthermore, vegetation in arid regions is more dependent on water than heat, and low-vegetation areas rely more on sub-snow habitats than external climatic factors. These findings provide valuable insights into how vegetation responds to snowmelt in the context of climate change, deepening our understanding of the relationship between snowmelt onset and green-up dates. This knowledge is essential for predicting vegetation phenology and managing ecosystem services under changing climate conditions. Future research should focus on the impacts of snow cover and false spring.

Reply on RC2

General comments:

The manuscript *Duration of vegetation green-up response to snowmelt on the Tibetan Plateau* by Ni *et al.* addresses the responses of date of snowmelt and date of green-up - and the time lapsed between the two - to climate change. The study raises up questions relevant within the scope of BG, and present some novel results that highlight the importance of studying snowmelt patterns in alpine regions. The scientific methods are clearly outlines, but the statistical analysis is described in too little detail, which does not allow to reproduce the results. The title reflects the content of the article, and the abstract is good, though with too many details, it can be written in a more concise way. Some more references could be added, but in general they provide good and appropriate references.

Response:

We sincerely appreciate your constructive suggestions and have revised our manuscript accordingly. First, we have added more details to the methods and results sections. Following your recommendations, we increased the font size in all figures and refined their details. Additionally, we polished the text to enhance clarity and eliminate ambiguity. We then integrated the latest results and clarified our objectives. Based on these revisions, we rewrote the introduction to better align with the study's structure and adjusted the article's organization accordingly. The specific responses to each suggestion are detailed below.

Major concerns:

There are two major concerns I have about the article.

1. The statistical analyses are not actually described. In the Material and methods there are no details describing what program was used, what functions, how they dealt with problems in the different models... And in the results they do not provide any details either, they only indicate whether something is significant or not. They should at least mention the number of replicates, p-values, confidence intervals... They could include all these details in one table, for ease of reference. And they use different significant thresholds for different tests. They should provide the reasoning behind that, otherwise it looks like they were chosen *a posteriori*, which is not scientifical. As the other referee points out, in one of the tests they use a significant threshold of 0.1, which seems quite high.

Response: The statistical methods used in this study include Moran's I index, the Mann-Kendall test, partial correlation analysis, and multiple linear regression analysis.

Global and local Moran's I values were computed using the Spatial Autocorrelation, Cluster and Outlier Analysis tools in ArcGIS 10.8, with all settings except the input layer kept at their default parameters (detailed in Section 2.3.2). Information about the software and tools has been incorporated into the manuscript.

The calculations were performed using the Spatial Autocorrelation, Cluster and Outlier Analysis tools in ArcGIS 10.8.

The Mann-Kendall test was implemented using MATLAB R2021a, where a custom function was developed to calculate the slope and significance level based on the formula in Section 2.3.3. Since MATLAB serves merely as an execution tool, further details are unnecessary.

For partial correlation analysis, we calculated the second-order partial correlation coefficient between the time difference (ΔD) and T_{spring}, P_{spring}, and S_{StoG} to mitigate indirect effects. These coefficients were derived from simple correlation coefficients, calculated using the *corr* function in Python 3.8. Additional methodological clarifications and function usage details have been incorporated into the manuscript.

The partial correlation coefficient between each variable and ΔD was calculated to quantify their relationship. To isolate the direct effects of T_{spring}, P_{spring}, and S_{StoG} on ΔD while minimizing indirect influences, the second-order partial correlation coefficient was employed (Equation 8 & 9).

$$r_{ij\cdot n} = \frac{r_{ij} - r_{in} \times r_{jn}}{\sqrt{(1 - r_{in}^2)(1 - r_{jn}^2)}} \tag{8}$$

$$r_{ij\cdot mn} = \frac{r_{ij\cdot n} - r_{im\cdot n} \times r_{jm\cdot n}}{\sqrt{(1 - r_{im\cdot n}^2)(1 - r_{jm\cdot n}^2)}} \tag{9}$$

where r_{ij} represents the simple correlation coefficient between variables i and j, calculated using the corr function in Python3.8. $r_{ij\cdot n}$ is the first-order partial correlation coefficient between i and j, which accounts for their correlation after removing the linear effects of control variable n. Similarly, $r_{ij\cdot mn}$ is the second-order partial correlation coefficient between i and j, controlling for both m and n. Accordingly, we calculated $r_{\Delta DT_{spring}\cdot P_{spring}S_{StoG}}$, $r_{\Delta DP_{spring}\cdot T_{spring}S_{StoG}}$, and $r_{\Delta DS_{StoG}\cdot T_{spring}P_{spring}}$.

The linear regression model was implemented in Python 3.8 using the *OLS* function for model estimation and the *variance_inflation_factor* function for the collinearity test.

The OLS and variance_inflation_factor functions were used to construct the linear regression model and perform the collinearity test in Python 3.8.

To provide a more detailed presentation of the results, we included the corresponding p-values for each global Moran’s I index (Figure 3b, S4). Additionally, for the global Moran’s I index over the period 2001–2018, we have included the full report generated by ArcGIS 10.8 in the supplementary materials (Figure S3). Regarding the linear regression analysis, the model's fitting accuracy (R^2) for each pixel has been added to Figure S5.

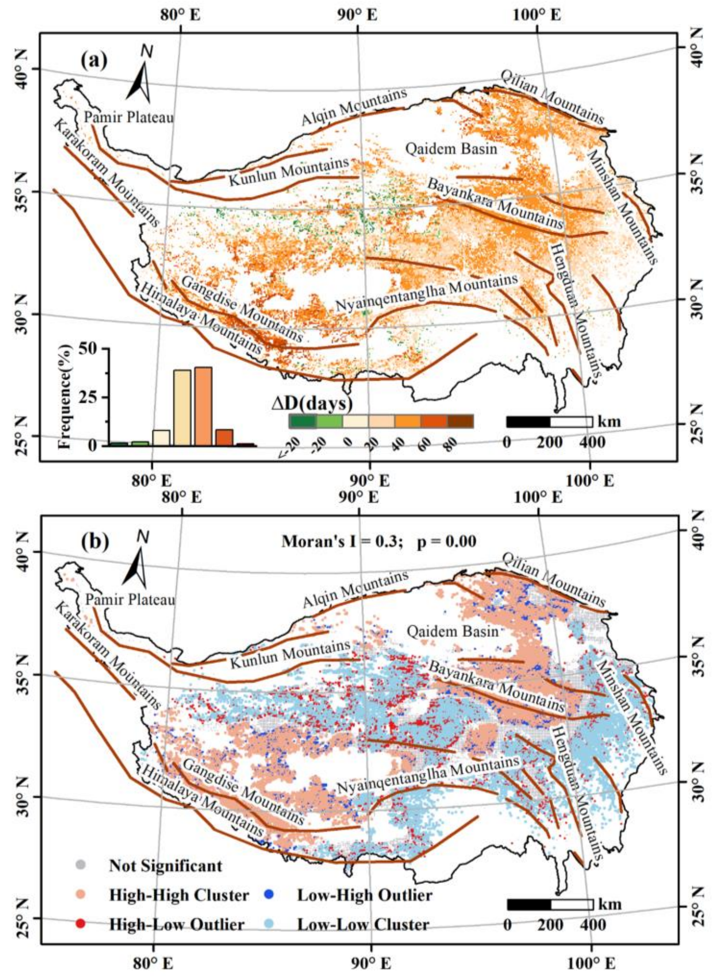
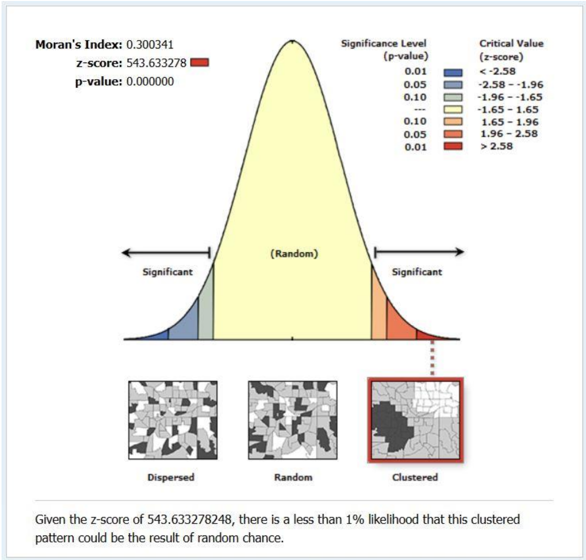


Figure 3: (a)Spatial, frequency distribution histogram and (b) local and global Moran’s I value of average ΔD *on the Tibetan Plateau over 2001–2018.*



Global Moran's I Summary

Moran's Index:	0.300341
Expected Index:	-0.000032
Variance:	0.000000
z-score:	543.633278
p-value:	0.000000

Figure S3: Report of the global Moran's I over 2001-2018 on the Tibetan Plateau

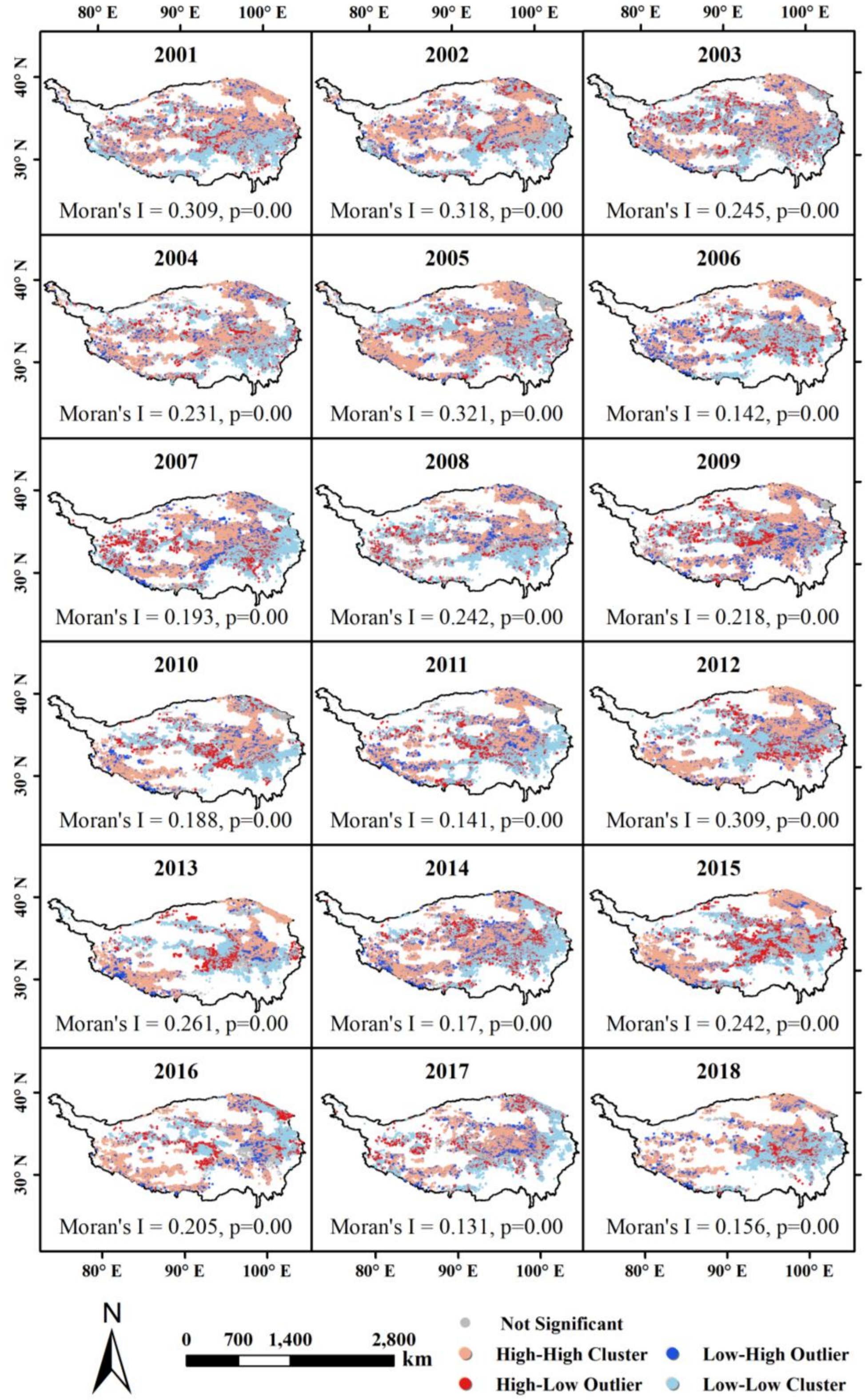


Figure S4: Global and local Moran's I values of ΔD from 2001-2018 on the Tibetan Plateau

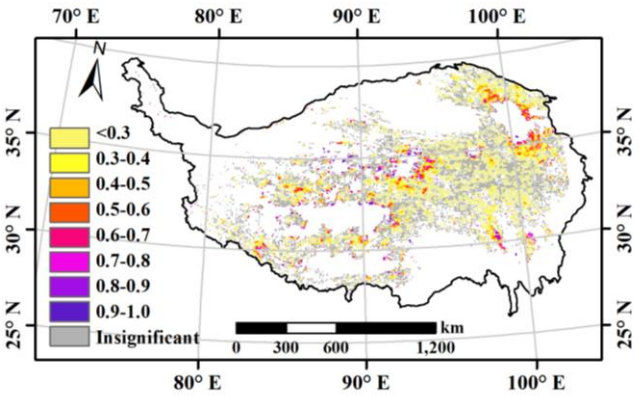


Figure S5: The fitting accuracy R^2 of linear regression model of ΔD for each pixel.

As this study includes only one regression model, it is not feasible to provide information such as confidence intervals, sample size, and goodness of fit as suggested. Each model has its own evaluation metrics for significance, but some require reference to distribution tables for determination. To enhance readability, we standardized the significance assessment by converting all indicators to p-values using unified formulas and degrees of freedom. For instance, partial correlation significance is typically evaluated using a t-test (Equation R1).

$$t = \frac{r}{\sqrt{1-r^2}}\sqrt{df} \qquad (R1)$$

where r and df represent the partial correlation coefficient and degree of freedom, respectively. And we use the *stats.t.sf* function in Python 3.8 to convert the t-value to a p-value, retaining only samples where $p < 0.05$. In the figures, significant areas are shown in color, while non-significant areas are rendered in gray. The results section discusses only the significant pixels.

Regarding the significance threshold, in the revision, we adjusted the significance level to 0.05. Since each pixel has an average of 12 samples with three independent variables (x), the average degrees of freedom (df) is 8 ($df = n - x - 1$). For df = 8, we conducted an experiment to calculate the significance indicator p-values for different correlation coefficients (r), as shown in Table R1, R2.

Table R1: Significance levels for correlation coefficients ($r = 0.1\sim0.9$ with 0.1 interval) at $df = 8$

Correlation coefficient (r)	Significance level(p)
0.1	0.783
0.2	0.580
0.3	0.400
0.4	0.252
0.5	0.141
0.6	0.067

0.7	0.024
0.8	0.005
0.9	0.000

Table R2: Significance levels for correlation coefficients ($r = 0.61 \sim 0.69$ with 0.01 interval) at $df = 8$

Correlation coefficient(r)	Significance level(p)
0.61	0.061
0.62	0.056
0.63	0.051
0.64	**0.046**
0.65	0.042
0.66	0.038
0.67	0.034
0.68	0.031
0.69	0.027

Only when r exceeds 0.64 does the test pass with $p < 0.05$. However, it is generally accepted that $r \geq 0.75$ indicates a strong correlation, $0.5 \leq r < 0.75$ indicates a moderate correlation, and $0.25 \leq r < 0.5$ indicates a weak correlation. Consequently, with small sample sizes, only pixels with strong correlations pass the test (Bonett and Wright, 2000), potentially overlooking some valid information. In our study, a valid sample requires that both the green-up date (D_{GU}) and D_{SOM} be valid, with D_{SOM} preceding D_{GU}. Therefore, the average sample size of 12 makes passing the significance test difficult. To address this, we applied the first law of geography by expanding the sample size to include the pixel itself and its eight neighboring pixels (Figure R1).

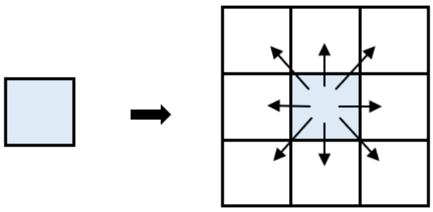


Figure R1. Diagram of sample size expansion

As a result, at the $p < 0.05$ level, the significance ratios for T_{spring}, P_{spring}, and S_{StoG} with ΔD were 23.5%, 28.8%, and 35.4%, respectively (Figure 6). Meanwhile, the dominant factor influencing ΔD was recalculated using the same methodology (Figure 7). While enhancing the significance, the original conclusions remain largely unchanged, with only minor revisions to some text and results.

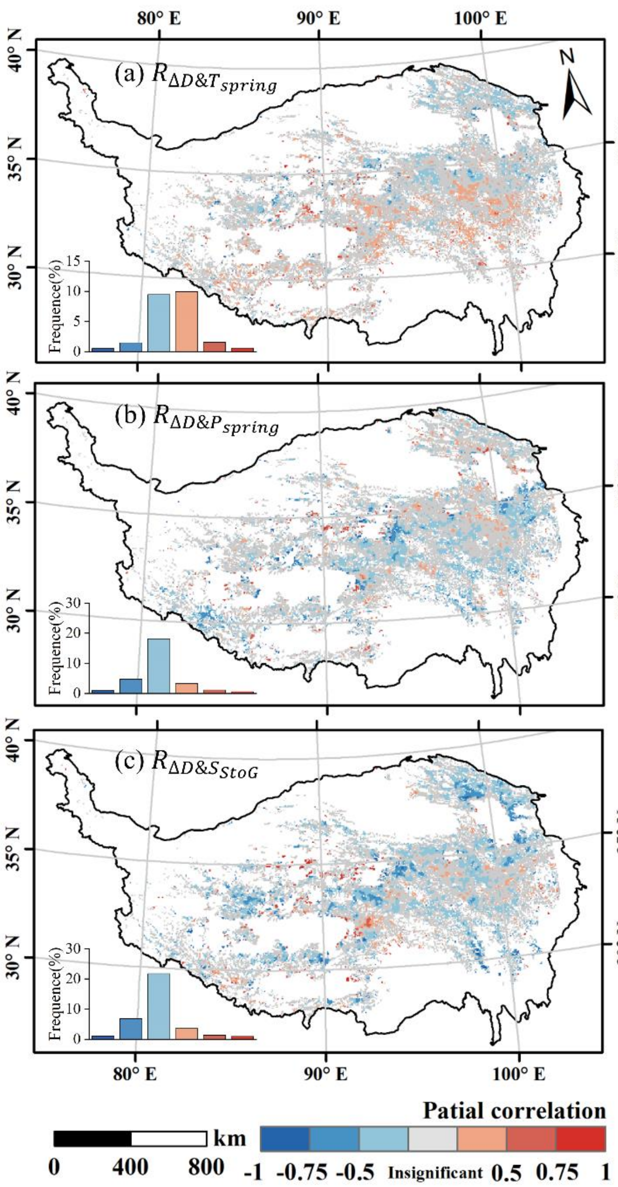


Figure 6: Spatial distribution of the partial correlation between ΔD and (a) spring mean temperature ($R_{\Delta D\&T_{spring}}$), (b) spring total rainfall ($R_{\Delta D\&P_{spring}}$), and (c) daily snowmelt from D_{SOM} to D_{GU} ($R_{\Delta D\&S_{StoG}}$).

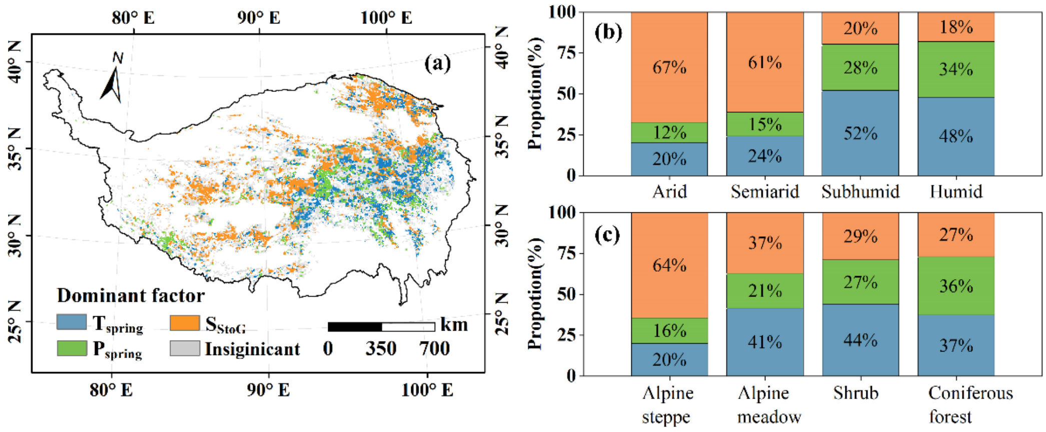


Figure 7: (a) Spatial distribution of dominant factor of ΔD and its proportion diagram among (b) different geographical zones and (c) different vegetation types.

References:

Bonett, D. G. and Wright, T. A.: Sample size requirements for estimating Pearson, Kendall and Spearman correlations, Psychometrika, 65, 23-28, https://doi.org/10.1007/bf02294183, 2000.

2. They do not write anything about what the objectives of the study are, and what their hypothesis are either. I think they need to state them clearly in the introduction, and the rewrite the structure of the Results and Discussion to answer the questions they pose. At the moment the structure is not completely cohesive. Having some clear hypothesis would help to streamline the flow of ideas. For example, the first subsection in the discussion is about a point that has not been raised before, and is not mentioned either in the Introduction or in the Results. I reckon this would also help the authors decide which results and figures to show in the Results section, something the other referee also mentions.

I think these points should be addressed before working any further in the manuscript. I consider that these points, though they require quite some work, do not change the content of the manuscript in a major way, so I think just a minor revision (though extensive) is needed.

Response: We have integrated the latest results and clarified the objectives and hypotheses of the study. While previous research has reported a positive correlation between the start of snowmelt (D_{SOM}) and the green-up date (D_{GU}), they did not address whether D_{GU} responds to D_{SOM} instantaneously or with a time lag. The first objective of this study was to explore this temporal relationship. Additionally, given geographical variations, we hypothesized that response times might differ regionally, which formed the second objective of the study—to verify regional differences in ΔD and further explain the spatiotemporal variation and underlying response mechanisms. Accurate extraction of D_{SOM} and D_{GU} is fundamental to addressing these issues. Therefore, our third objective was to develop a more precise method for extracting D_{SOM} at a higher resolution. We have incorporated these purposes into the introduction to streamline the flow of ideas.

Wang et al. (2015) found that 39.9% of meadows and 36.7% of steppes on the TP showed a significant correlation between D_{GU} *and* D_{SOM}*. Additionally, Wang et al. (2018) reported positive Pearson correlation coefficients between* D_{SOM} *and* D_{GU} *for most regions of the TP, with exceptions in warmer and drier areas. Although these studies have confirmed that snowmelt affects vegetation green-up, the response time of vegetation to snowmelt remains unknown. Recent studies have investigated the delayed effect of meteorological factors on vegetation (Xu et al., 2023; Wu et al., 2015). We also propose a hypothesis: the response of vegetation green-up to snowmelt is delayed, and this delay exhibits regional heterogeneity.*

To test our hypothesis, this study calculates and analysed the time difference between the D_{GU} *and* D_{SOM} *on the TP. Accurate extraction of* D_{SOM} *and* D_{GU} *is essential for addressing these issues. For* D_{SOM}*, optical remote sensing primarily detects the presence or absence of snow, limiting its ability to identify the melting state. Although microwave remote sensing can more accurately detect snowmelt, its spatial resolution is lower. In this study, we used a daily snow depth dataset with a spatial resolution of 0.05° to identify* D_{SOM} *with higher resolution for the TP from 2001 to 2018. For* D_{GU}*, we used the* D_{GU} *dataset for the TP from 2001 to 2018, which was generated by Xu et al. (2022b) through the optimal combination of six vegetation indices and four extraction methods, resulting in the highest accuracy dataset. We then calculated the time difference between* D_{GU} *and* D_{SOM}*, denoted as* ΔD*, to explore the response of vegetation green-up to snowmelt. In addition, we employed exploratory spatial data analysis and the Mann-Kendall test to examine the spatiotemporal variation of* ΔD*. To further explore the spatiotemporal variation and response mechanisms, we applied partial correlation analysis and multiple linear regression to examine the relationships between* ΔD *and various influencing factors.*

Based on these objectives and additional feedback, we have reorganized the structure of the Results and Discussion sections. First, Section 3.1 presents the higher-resolution $\mathrm{D_{SOM}}$ data. Section 3.2 has been further divided into 3.2.1 and 3.2.2, focusing on spatial distribution and temporal dynamics, respectively, to address Purpose 1. Section 3.3 fulfills Purpose 2 through statistical analysis. In the first part of the Discussion, we evaluate the accuracy of the $\mathrm{D_{SOM}}$ data to achieve Purpose 3. In the second part, we further explore the mechanisms behind vegetation's response to snowmelt.

Specific comments:

1. L30-31. You say that the TP plays a role in maintaining global biodiversity and ecological security. Maybe add short sentence explaining what that role is?

Response: We have expanded this statement to further elaborate on the diversity and ecological security of the region.

Moreover, its unique natural environment and diverse habitats facilitate the interaction and integration of various biota (Chu et al., 2024; Yu et al., 2021). These features help mitigate global warming, protect biodiversity, and provide carbon sequestration, thereby playing a crucial role in ecological security (Liang and Song, 2022).

2. L50. Regarding snow phenology in the TP, I would appreciate a short description of when the first snow tends to arrive, for how long it lasts and when it starts melting. I think it would make it easier for the reader to picture it. Maybe include a table showing the different values for the different regions? Seeing those values would help appreciate the changes experienced by the different regions

Response: As your advice, we have added an overview of TP snow phenology.

In high-altitude areas (around 17.69%), snow cover typically begins in October and ends in April, whereas in low-elevation areas (around 56.69%), the snow cover duration is usually less than 20 days, concentrated in December and January (Xu et al., 2024).

3. L61-65. This is part of the materials and methods. Here you should indicate what your objectives and hypotheses are

Response: We have reorganized the introduction and highlighted the purposes of this study in this paragraph.

To test our hypothesis, this study calculates and analysed the time difference between the D_{GU} and D_{SOM} on the TP. Accurate extraction of D_{SOM} and D_{GU} is essential for addressing these issues. For D_{SOM}, optical remote sensing primarily detects the presence or absence of snow, limiting its ability to identify the melting state. Although microwave remote sensing can more accurately detect snowmelt, its spatial resolution is lower. In this study, we used a daily snow depth dataset with a spatial resolution of 0.05° to identify D_{SOM} with higher resolution for the TP from 2001 to 2018. For D_{GU}, we used the D_{GU} dataset for the TP from 2001 to 2018, which was generated by Xu et al. (2022b) through the optimal combination of six vegetation indices and four extraction methods, resulting in the highest accuracy dataset. We then calculated the time difference between D_{GU} and D_{SOM}, denoted as ΔD, to explore the response of vegetation green-up to snowmelt. In addition, we employed exploratory spatial data analysis and the Mann-Kendall test to examine the spatiotemporal variation of ΔD. To further explore the spatiotemporal variation and response mechanisms, we applied partial correlation analysis and multiple linear regression to examine the relationships between ΔD and various influencing factors.

4. L89-90. Where was the dataset obtained from? Where did the data come from?

Response: The D_{GU} dataset was produced by Xu et al. (2022), who are also the collaborators of this study, and they directly provided this data. The provenance of this dataset has been added to the manuscript.

Since Xu et al. are also the collaborators of this study, they directly provided this data.

Reference:

Xu, J. Y., Tang, Y., Xu, J. H., Chen, J., Bai, K. X., Shu, S., Yu, B. L., Wu, J. P., and Huang, Y.: Evaluation of Vegetation Indexes and Green-Up Date Extraction Methods on the Tibetan Plateau, Remote Sens., 14, https://doi.org/10.3390/rs14133160, 2022b.

5. L143, Figure 2 caption. I write it here, but it applies to other captions as well. Could you write a more detailed caption?

It is explained in the text, but looking only at the figure I do not know what "Filtered" means, or what the 47 days, 3 days and 6 days indicate, for example. It would be easier for the reader if these values are explained, shortly, in the caption was well.

Response: Following your suggestion, we have further clarified the information in the caption. Additionally, as another reviewer recommended placing this figure in the supplementary material to simplify the text, the revised version is now available in Figure S1.

Figure S1: Diagram illustrating the identification of the start of snowmelt (D_{SOM}). The snow depth is smoothed using Sacitzky-Golay filtering, represented by the orange line. Each continuous snowfall event is identified based on set criteria and highlighted with a blue background. The duration of each snow event, indicated by the number of days, is used to determine the longest snowfall process. The final identified D_{SOM} is marked by the orange point.

6. L152. When calculating I_I, how do you determine what the neighboring regions are? Is it just the adjacent pixels, or is it more complex?

Response: In our study, we assessed the degree of spatial autocorrelation between all pixels and region i. Different pixels are assigned weights based on inverse distance, meaning that the closer a pixel is to region i, the higher its weight. We have added a more detailed explanation of this method in the manuscript.

The local Moran's I (I_i) measures the degree of clustering or spatial autocorrelation of ΔT within a specific region i, relative to all other regions, as calculated using Equation 3.

7. L186-187. Do you mean that, at temperatures below 0 degrees, the correlation between SCED and Tgu was low, while at temperatures above 0 the correlation was high? I think you could rewrite this sentence to make the message clearer

Response: In this revision, we have rewritten this sentence for clarity.

Specifically, when the temperature is below 0℃, the correlation between the snow cover end date and D_{GU} *weakens as temperature increases. In contrast, when the temperature exceeds 0℃, the correlation strengthens with rising temperature (Wu et al., 2023).*

8. L188. "a strong negative correlation prevailed". Between T_{SOM} and T_{GU}? So, in areas of high temperature, a late T_{SOM} was related to an early T_{GU}, and an early T_{SOM} to a late T_{GU}? Or am I misunderstanding it?

Response: Yes, that is what I intended to express. The correlation between D_{SOM} and D_{GU} varies under different hydrothermal conditions. I have added the full context to clarify any potential misunderstandings.

Increased humidity from precipitation strengthened the positive correlation between D_{SOM} *and* D_{GU}*, whereas in high temperature areas, a strong negative correlation between* D_{SOM} *and* D_{GU} *prevailed (Xu et al., 2022a).*

9. L189. What environmental conditions? The factors mentioned afterwards are either temperature-based or precipitation-based.

Response: The references do not explicitly indicate this. Here, we aim to highlight that, in addition to temperature and precipitation, it is essential to consider the amount of snowmelt.

10. L190. Regarding "influencing factors". Why did you end up choosing these factors? Was it based on the literature, did you perform any tests to see what factors had the largest coefficients...?

Response: We refer to the conclusion of Shen et al., "The temporal changes in phenology from long-term observations on the QTP have been mainly associated with three climatic drivers: changes in temperature, precipitation, and snowmelt date." We also reviewed these three factors prior to this sentence, which led us to focus on them. Following your suggestion, we have cited their article accordingly.

Reference:

Shen, M., Wang, S., Jiang, N., Sun, J., Cao, R., Ling, X., Fang, B., Zhang, L., Zhang, L., Xu, X., Lv, W., Li, B., Sun, Q., Meng, F., Jiang, Y., Dorji, T., Fu, Y., Iler, A., Vitasse, Y., Steltzer, H., Ji, Z., Zhao, W., Piao, S., and Fu, B.: Plant phenology changes and drivers on the Qinghai-Tibetan Plateau, Nat. Rev. Earth Environ., 3, 633-651, https://doi.org/10.1038/s43017-022-00317-5, 2022.

11. L195. Any reasons for not considering interactions between the factors?

Response: Indeed, interactive effects between the factors may better explain the heterogeneity of ΔD. Following Du et al. (2019), we included the product terms of each pair of the three original variables in the linear regression model to account for their interactive effects (Equation R2). However, before performing the linear regression, we conducted a collinearity test among the variables (i.e., $VIF < 3$). Upon including the product terms, severe multicollinearity issues emerged (Table R3, R4). Consequently, we retained these three factors as the independent variables in the original analysis.

$$\Delta D = aT_{spring} + bP_{spring} + cS_{StoG} + dT_{spring} \times P_{spring} + \mathrm{e}T_{spring} \times S_{StoG} + \mathrm{f}P_{spring} \times S_{StoG} + \mathrm{g} \quad (R2)$$

Table R3: Variance Inflation Factor for T_{spring}, P_{spring} and S_{StoG}

Variable	VIF
T_{spring}	1.243
P_{spring}	1.209
S_{StoG}	1.174

Table R4: Variance Inflation Factor for T_{spring}, P_{spring}, S_{StoG} and their product terms

Variable	VIF
T_{spring}	32.416
P_{spring}	12584.8
S_{StoG}	1.185
$T_{spring} \times P_{spring}$	20254.1
$T_{spring} \times S_{StoG}$	30.826
$P_{spring} \times S_{StoG}$	21424.3

Reference:

Du, J., He, Z. B., Piatek, K. B., Chen, L. F., Lin, P. F., and Zhu, X.: Interacting effects of temperature and precipitation on climatic sensitivity of spring vegetation green-up in arid mountains of China, Agric. For. Meteorol., 269, 71-77, https://doi.org/10.1016/j.agrformet.2019.02.008, 2019.

12. L224. When you say that a test or model was significant (here and later on) can you provide the statistical values connected to the test (p-value, confidence interval...), in order to give a better idea of the strength of the effect? And can you say anything about the goodness of fit of the model in the supplementary material?

Response: We have included the p-value for significance in the manuscript, as well as in Figure 3 and Figure S4. Besides, we added a full report of global Moran's I (Figure

S3) by ArcGIS 10.8. However, Moran's I is an indicator for evaluating the degree of spatial clustering, not a fitting model. Typically, p-values or z-scores are used to test significance, rather than goodness of fit (Anselin, 1995; Li et al., 2024).

Reference:

Anselin, L.: LOCAL INDICATORS OF SPATIAL ASSOCIATION - LISA, Geog. Anal., 27, 93-115, https://doi.org/10.1111/j.1538-4632.1995.tb00338.x, 1995.

Li, S. J., Sun, Z. C., Guo, H. D., Ouyang, X. Y., Liu, Z. Q., Jiang, H. P., and Li, H. W.: Localizing urban SDGs indicators for an integrated assessment of urban sustainability: a case study of Hainan province, Int. J. Digital Earth, 17, https://doi.org/10.1080/17538947.2024.2336059, 2024.

13. L249. You indicate how much ΔT advances near the freezing point. But do you mean that it advances 3 days if we consider an increase in temperature of 1 K? Or what do you mean?

Response: Yes, an increase of 1K in temperature can shorten ΔD by 3 days. We have modified figure 4 and the corresponding text based on your suggestion and feedback from another reviewer.

Figure 4 illustrates the mean value of ΔD under varying spring meteorological conditions. ΔD exhibits a clear stepwise decline from cold to warm regions, decreasing from approximately 48 to 37 days (Fig. 4a). In colder or hotter spring conditions (i.e., $T_{spring} < 270\ K$ or $T_{spring} > 275\ K$), ΔD decreased slightly. However, near the freezing point (270–275 K), ΔD shortens by 3 days with each 1°C increase in T_{spring}. Under varying precipitation conditions (Fig. 4b), ΔD shortens by 0.29~1.96 days for every 10 mm increase in P_{spring}. Fig. 4c reveals a strong negative correlation between ΔD and S_{StoG} when S_{StoG} exceeded 6 mm day^{-1}. For each 1 mm increase in S_{StoG}, ΔD decreases by approximately 0.615 days. The dispersion within each snowmelt category remains relatively consistent, with a standard deviation of about 16.8 days.

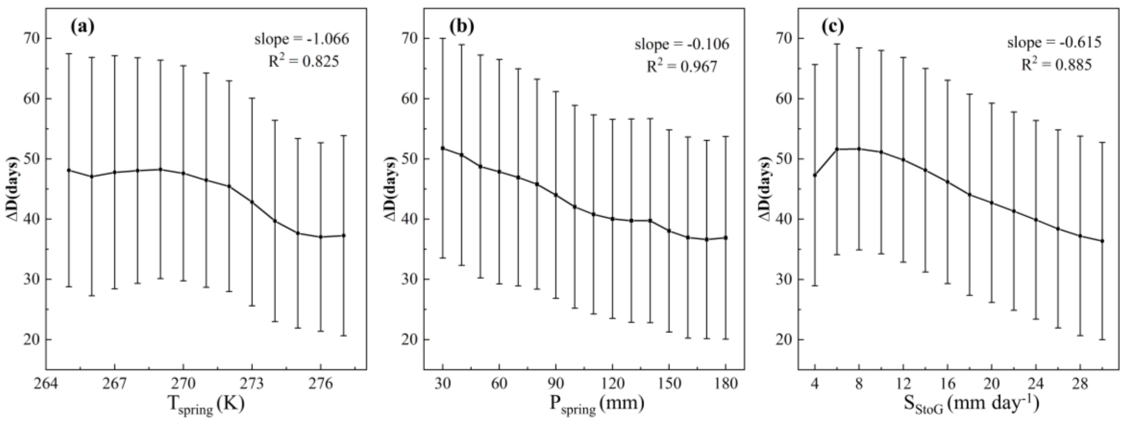


Figure 4: Variations in ΔD across regions with differing (a) spring mean temperature (T_{spring}), (b) spring total rainfall (P_{spring}), and (c) daily snowmelt from D_{SOM} to D_{GU} (S_{StoG}). Points represent the mean ΔD, while error bars denote one standard deviation.

The slope and R^2 value reflect the coefficient and precision of the linear regression, respectively, with a significance level of 0.01.

14. L250. "ΔT fluctates with temperature". Do you mean that there is not a consistent trend, i.e., that higher temperatures do not always have shorter ΔT? Or something else? Could you make it clearer?

Response: Yes, while there is a general trend that higher temperatures lead to shorter ΔD, it is not a strict linear relationship. The revision is consistent with specific comment 13.

15. L264. Just to make sure I understand your point. Even though higher T_{spring} was connected to lower ΔT values, in more than 50% of the plots that had a significant correlation a greater average temperature in spring led to a delayed green-up after the start of the snowmelt, right?

Response: What we intend to convey is that 15.7% of the samples passed the significance test. Among these significant pixels, 59.4% exhibited a positive correlation, while 40.6% showed a negative correlation between T_{spring} and ΔD. We have revised the expression following suggestions from another reviewer.

At a significance level of 0.05, ΔD was significantly correlated with temperature in 23.5% of the samples. Among these significant pixels, 51.3% exhibited a positive correlation between ΔD and T_{spring}.

16. L286. [This comment refers to the whole section 4.1]

This is the first time you mention that this study will look at different ways of identifying T_{SOM}. You should include a paragraph, or at least a couple of sentences talking about it in the introduction, describe in the material and methods how the values are obtained, and present the results in the Results section

This section is good, but comes up abruptly. I was not expecting this after reading the rest of the article.

Response: Indeed, proper foreshadowing enhances the coherence of the article and improves readability. We have incorporated the need for high-resolution D_{SOM} data into the introduction. Additionally, the extraction method has been described and revised based on another feedback (Section 2.3.1).

Accurate extraction of D_{SOM} and D_{GU} is essential for addressing these issues. Xu et al. (2022) compared six vegetation indices with four extraction methods and used the optimal combination to generate the D_{GU} dataset for the TP. For D_{SOM}, optical remote sensing primarily detects the presence or absence of snow, limiting its ability to identify the melting state. Although microwave remote sensing can more accurately detect snowmelt, its spatial resolution is lower. Therefore, accurately extracting D_{SOM} at higher resolutions remains a challenge.

17. L311. It would have been interesting if you could have had at least one sample area per land cover type (10 in total), as I imagine they differ in the snow cover pattern and properties. But maybe it was not possible?

Response: In the revision, we selected one sample area for each type of land cover. Corresponding adjustments were also made. First, permanent snow/glacier, lake, and arable land were excluded due to the lack of seasonal snow cover and the small number of samples. Second, since most land cover types are not distributed contiguously, the original 10 × 10 km area was almost composed of mixed land cover type. We reduced the sample area to 5 × 5 km. The recalculated results are presented in Figure 8:

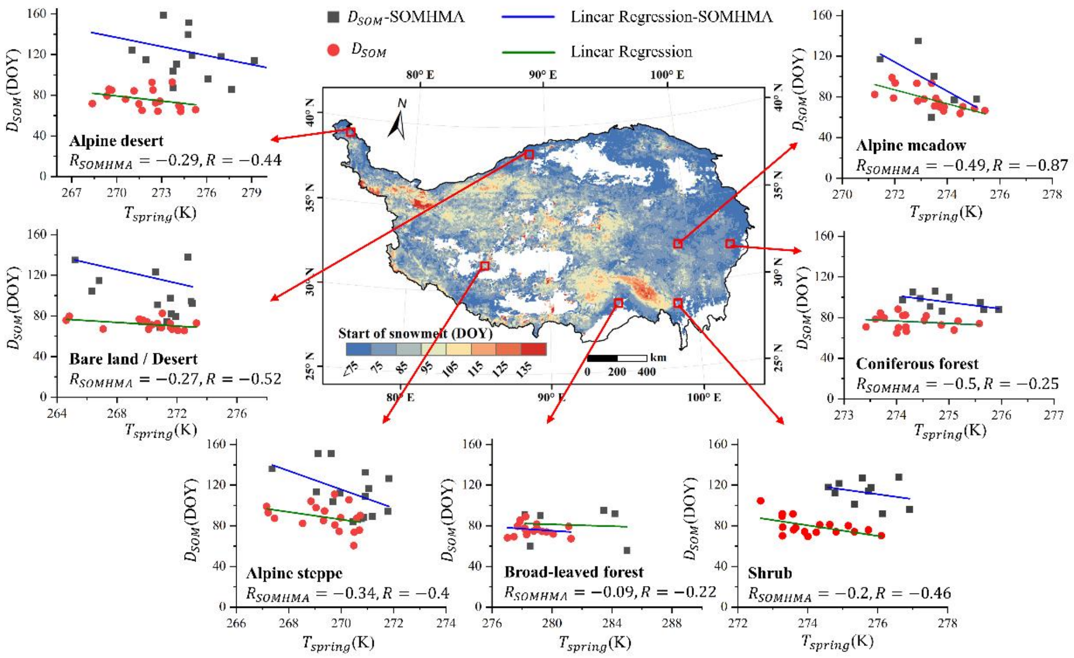


Figure 8: Relevance between the average D_{SOM} and mean April temperature in a sample area of each land cover.

Overall, the identification of alpine meadow yielded the best result, while the identification of forests was the least effective. When comparing the two D_{SOM} data, our result was still superior to SOMHMA in all types except for coniferous forests. We modified the original text accordingly:

Thus, following Grippa et al. (2005), we randomly selected one sample area for each land cover on the TP and calculated the correlation coefficient between average D_{SOM} and mean April temperature (Fig. 8). Our results showed stronger consistency with temperature trends compared to SOMHMA.

18. L318. I reckon you could first have a section where you discuss T_{SOM} and T_{GU} by themselves (section 3.1 in the Results): their spatial distribution, connecting it to the different land cover types (and maybe other factors); and their temporal distribution and how you find very few areas with a significant trend.

I think it would be best to cover these points before discussing the ΔT (which is section 3.2 in the Results).

Response: It is essential to first discuss D_{SOM} and D_{GU}. We have split the original section 3.1 into 3.1 and 3.2 to analyze their spatial distribution. While temporal variation analysis should have been included, the lack of clear patterns (Figure R2) hindered the interpretation of subsequent ΔD trends.

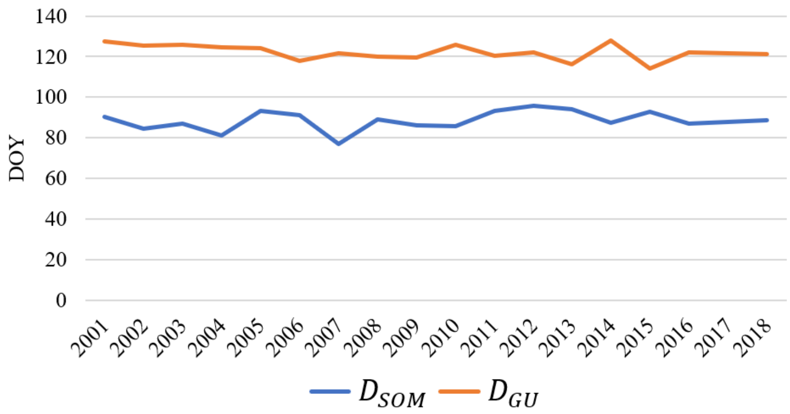


Figure R2: Mean annual D_{SOM} and D_{GU} on the Tibetan Plateau from 2001 to 2018.

19. L324. Maybe you could add a short sentence here, after "respectively", explaining that the larger deltaT in the TP is due to its characteristics - Which would link this paragraph to the following ones

Response: We have added the explanation to the manuscript.

They were shorter than those on the TP, likely due to differences in snowpack characteristics. The blocking effect of the Himalayas and Karakoram Mountains results in transient, shallow, and patchy snow on the TP (Lei et al., 2023). Consequently, the impact of snow on response times is particularly significant on the TP.

20. L360-363. I think you should write shortly what this actually means for the different vegetation types. What would happen to the vegetation (both higher and lower) if temperature increases and/or if snowmelt happens earlier?

Response: The influence of temperature and snowmelt on ΔD has been analyzed in detail in the previous paragraph. This section focuses on determining which factor dominates under different conditions. In accordance with your suggestion, we have examined the differences in response across vegetation types under temperature rise and snowmelt scenarios:

When analyzing by vegetation type, T_{spring} was the dominant factor for higher vegetation (e.g., shrubs and forests) while lower vegetation (e.g., alpine steppes and meadows) is more affected by S_{StoG} (Fig. 7c). Wind speed is high on the TP due to its special topography and elevation. This limited the snow accumulation on the branches *for higher vegetation, thus some parts of the vegetation are exposed to the outside environment. In contrast, the height of the lower vegetation is sometimes equal to the depth of the snow, which can partially or completely cover the vegetation (Tang et al., 2024). If the temperature rises, it will directly act on higher vegetation while heat will transfer slower to the low vegetation buffered by snow. If the snow melts, the melting water can directly affect the lower vegetation under the snow while the tall vegetation can only be used after absorption through the soil.*

21. L373. You could name what areas in the TP experience a negative ΔT, and explain shortly how they might be affected by climate change.
Response: Your suggestion was very constructive and we have incorporated these points accordingly.

Zhu et al. (2019b) used the SI-x model to simulate the probability of false spring in China from 1950 to 2005 and projected it until 2100. Their findings revealed that the central TP exhibited the highest probability of false spring in China. Rising temperature due to global warming are expected to further advance the D_{GU}*, thereby increasing the risk of false spring.*

22. L378. The conclusion includes to many details. You do not need all the specific numbers here, they appear in the rest of the manuscript. Here you should only include those details and number that more strongly support your conclusions.
Response: In this revision, we have modified as follows:

This study investigates the dynamic response of vegetation to snowmelt on the Tibetan Plateau from 2001 to 2018. Our results reveal that the effect of snowmelt on vegetation is not immediate, with a mean response lag of 38.5 days from D_{SOM} *to* D_{GU}*. Notably, the false spring was observed in the north-western TP, which warrants further exploration. As precipitation and snowmelt increase, the response time shortens. More complex than these factors, temperature exerts a greater influence on* D_{GU} *than* D_{SOM} *in colder regions, thus shortening the response time. Conversely, in warmer areas, the increased temperature has a stronger impact on* D_{SOM}*, which lengthens the response time. Furthermore, vegetation in arid regions is more dependent on water than heat, and low-vegetation areas rely more on sub-snow habitats than external climatic factors. These findings provide valuable insights into how vegetation responds to snowmelt in the context of climate change, deepening our understanding of the relationship between snowmelt onset and green-up dates. This knowledge is essential for predicting vegetation phenology and managing ecosystem services under changing climate conditions. Future research should focus on the impacts of snow cover and false spring.*

TECHNICAL CORRECTIONS

1. L73, Figure 1. This applies also to some of the other Figures. They are a bit difficult to read, especially the text. Could you make the font slightly larger, or the figure?

Response: We have made the font larger in all figures.

2. L73. It is difficul to see the difference in colour between "Shrub" and "Coniferous forest" in the legend, and impossible (at least for me) in the Figure. Would it be possible to use colours with more contrast?

Response: We have used a different color (R:247, G:157, B:56) to represent the coniferous forest.

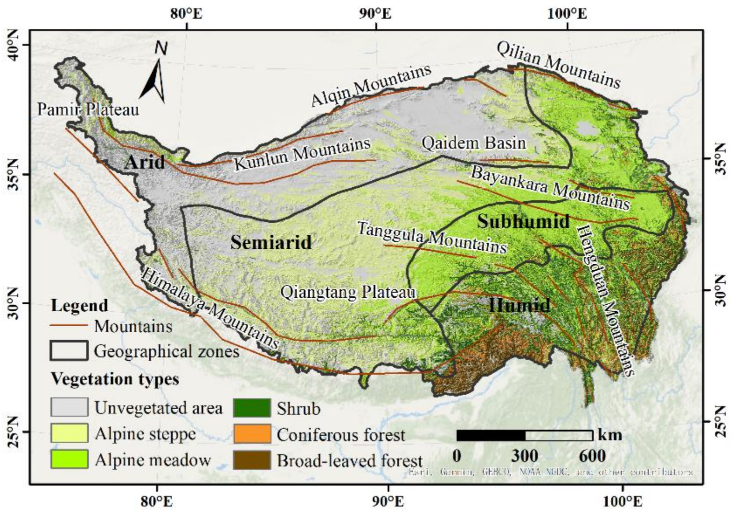


Figure 1: Map of the study area with vegetation types, distribution of mountains and geographical zones (based map from ESRI)

3. L120. It says "Subsection (as Heading 3)". I assume this is a typo, and you wanted to write only "Calculation of ΔT"

Response: We have delete "Subsection (as Heading 3)".

4. L191. T_{spring} (here and in the whole manuscript). It might be best to use a different letter for temperature. In most cases you use T to denote time (T_{SOM}: time of start of snowmelt; T_{GU}: time of green-up). Using then T for T_{spring} could be confusing for the reader.

Or use a different letter for denoting time. Maybe D, as you are always talking about dates?

Response: We have replaced $\Delta T, T_{GU}, T_{SOM}$ with $\Delta D, D_{GU}, D_{SOM}$ throughout the manuscript.

5. L211. "more concentrated"? The range is almost the same (30 days)

Response: Initially, we wanted to highlight that the overall range of D_{SOM} is broader (i.e., the disparity between the maximum and minimum values is more pronounced).

While the phrase may not have been the most appropriate, we chose to delete "and was more concentrated" to maintain contextual coherence.

6. L212. Is the progress not from East (earlier T_{GU}) to West (later T_{GU})?
Response: We have rewritten it for more accurate expression.

Vegetation green-up earlier in the southeast and later in the centre and west.

7. L223. "to gain more about spatial distribution". To gain what? More insight? More information?
Response: We have added "gain more insight" in the manuscript.

8. L233, Figure 5. Have you tried using the darkest colours to represent the high-high and low-low clusters (since they are more common), and paler colours for the other ones? Pale colours (at least in my case) are more difficult to distinguish from each other in a small image
Response: This color scheme is the default used after Cluster and Outlier Analysis tool (calculation of local Moran's I values) in ArcGIS 10.8 and is commonly employed in numerous papers (Peng et al., 2024). Therefore, we believe it is best to retain this classic color scheme for better legibility.
Reference:
Peng, R., Tang, J. H. C. G., Yang, X., Meng, M., Zhang, J., and Zhuge, C.: Investigating the factors influencing the electric vehicle market share: A comparative study of the European Union and United States, Appl. Energy, 355, https://doi.org/10.1016/j.apenergy.2023.122327, 2024.

9. L248. Do you mean "constant" instead of "consistent"?
Response: The original sentence has been removed.

10. L249. "rises" instead of "rise".
Response: Revised.

11. L253. Instead of "amplify" I think "strengthened" would be a better word. It implies that the reduction is greater, but also more consistent, which seems to be the case in your data
Response: The original sentence has been removed.

12. L280. "taller" instead of "higher".
Response: Revised.

13. L330-331. This is actually not that easy to see in the figure, maybe this figure should be larger.

Response: We have enlarged figure 7 in the manuscript.

14. L351-352. I suggest to rewrite this sentence. Something like this:
"Conversely, in colder regions increased temperatures can reduce cold stress on vegetation, resulting in a larger effect on T_{GU} (-0.27), similar to that on T_{SOM} (-0.28)."
Response: We have revised this sentence as follows:

In warmer regions with mean annual temperatures above freezing, spring temperature correlates negatively with D_{SOM} (The correlation coefficient is -0.46) and D_{GU} (The correlation coefficient is -0.07), indicating that temperature primarily influences snowmelt rather than vegetation growth, thus extending response times. In colder regions increased temperatures can reduce cold stress on vegetation, resulting in a larger effect on D_{GU} (The correlation coefficient is -0.27). However, consistent sub-freezing temperatures do not significantly lead to later D_{SOM} (The correlation coefficient is -0.28, which is similar with D_{GU}). In summary, the relationship between temperature and response time is modulated by the magnitudes of their respective influences at the local scale.

15. L353. "enhance" might be a bit of a misleading word. The reader might think that a better Time of start of snowmelt is an earlier Time of start of snowmelt. You could say "do not significantly lead to later T_{SOM}", for example.
Response: We have changed "enhance" to "lead to later".

However, consistent sub-freezing temperatures do not significantly lead to later D_{SOM} (The correlation coefficient is -0.28, which is similar with D_{GU}).

16. L364. "the growth rate of soil temperatures"? What do you mean?
Response: We have revised the expression.

This finding aligns with the observations of Zheng et al. (2022), who noted that soil temperature increased more rapidly in lower vegetation than in higher vegetation following snowmelt in Alaska, suggesting that snowmelt has a greater impact on lower vegetation.