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.

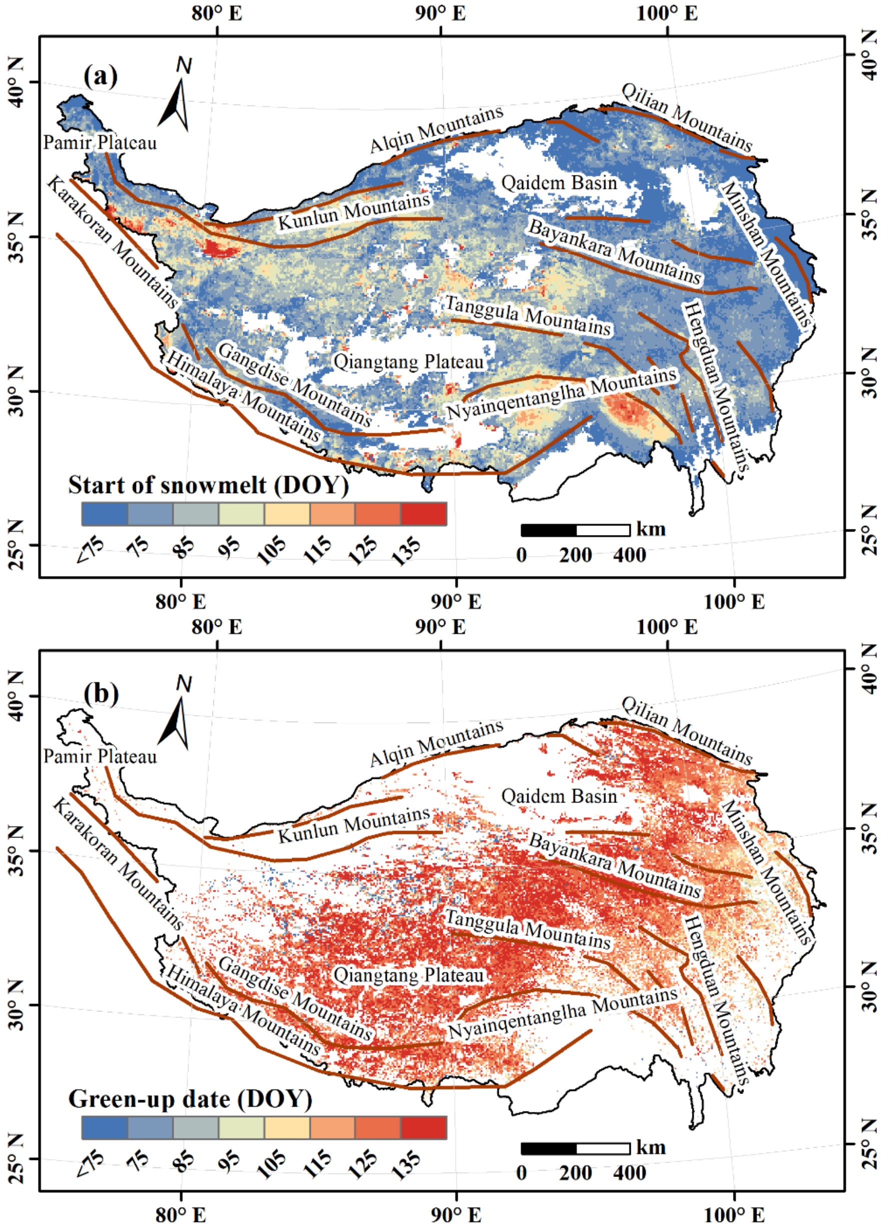

Figure 2: Spatial distribution of multiyear averaged (a)$\boldsymbol{D_{SOM}}$ and (b)$\boldsymbol{D_{GU}}$ from 2001 to 2018 on the Tibetan Plateau.

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:

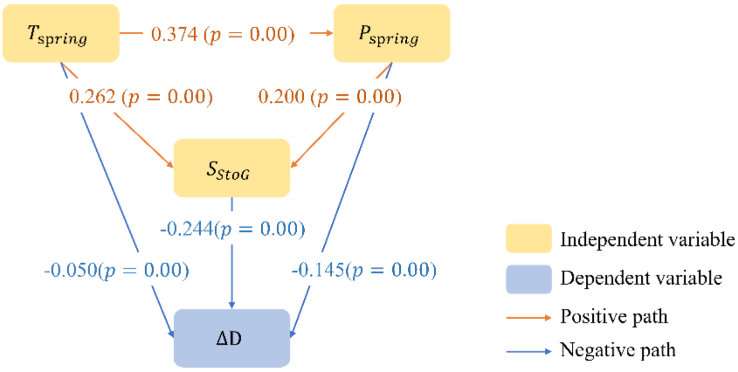

Figure R1. SEM path effect diagram for the response of D_{GU} to D_{SOM}

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).

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

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 ($r = 0.1$~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).

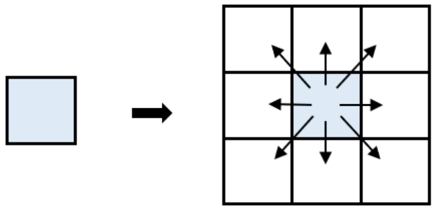

Figure R2. Diagram of sample size expansion

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.

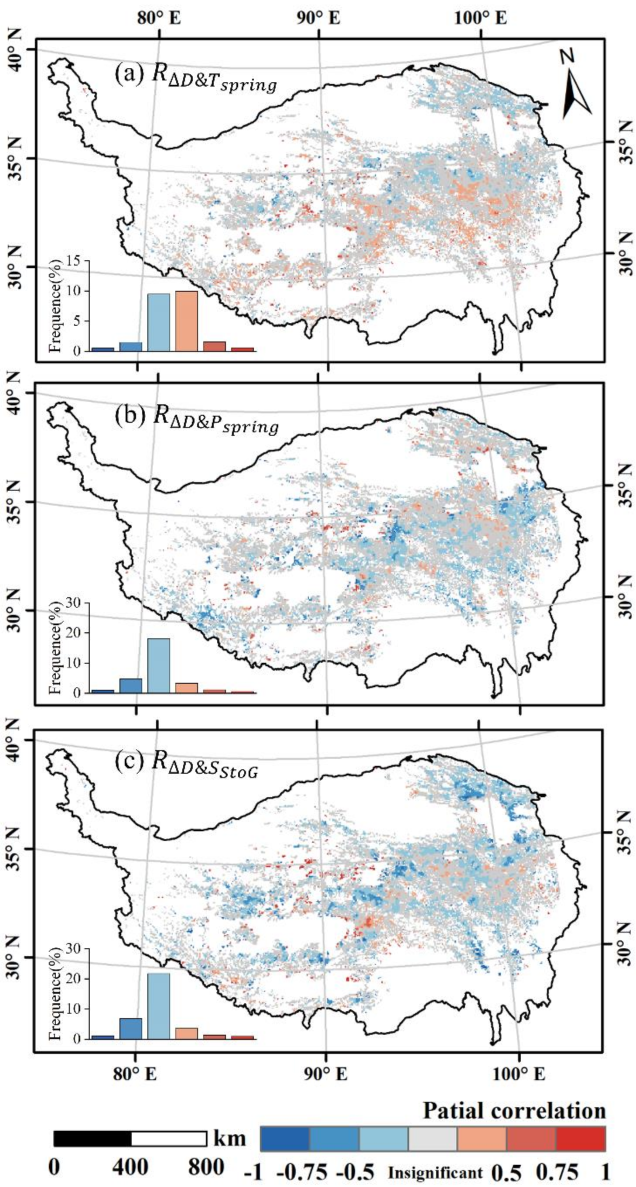

Figure 6: Spatial distribution of the partial correlation between ΔD and (a) spring mean temperature ($R_{\Delta D\&T_{spring}}$), (b) spring total rainfall ($R_{\Delta D\&P_{spring}}$), and (c) daily snowmelt from D_{SOM} to D_{GU} ($R_{\Delta D\&S_{StoG}}$).

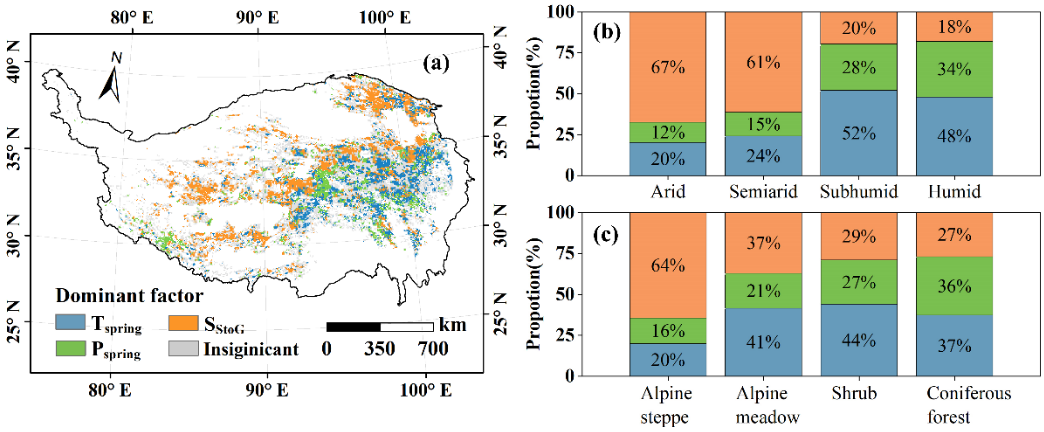

Figure 7: (a) Spatial distribution of dominant factor of ΔD and its proportion diagram among (b) different geographical zones and (c) different vegetation types.

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.

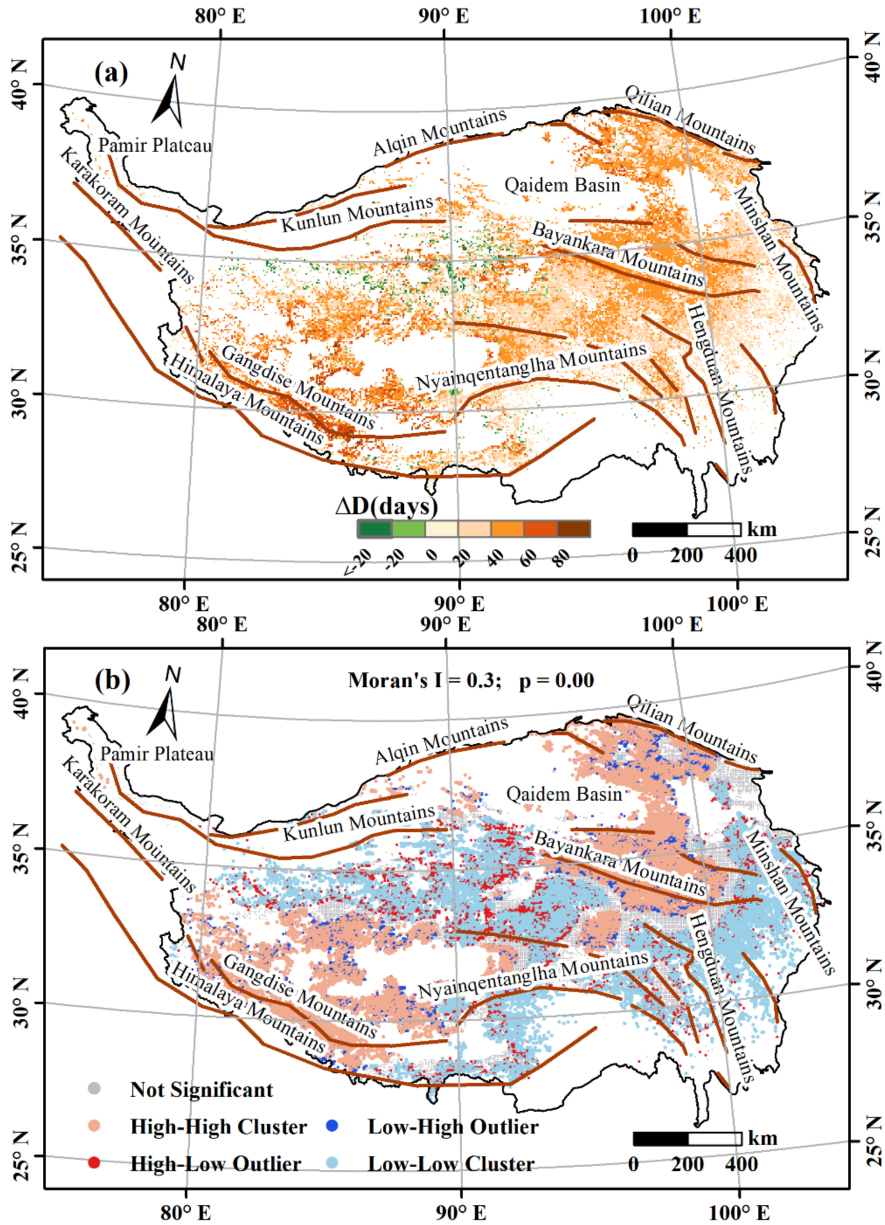

Figure 3: (a)Spatial, frequency distribution histograms and (b) local and global Moran's I values of average $\Delta\boldsymbol{D}$ *on the Tibetan Plateau over 2001–2018.*

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.

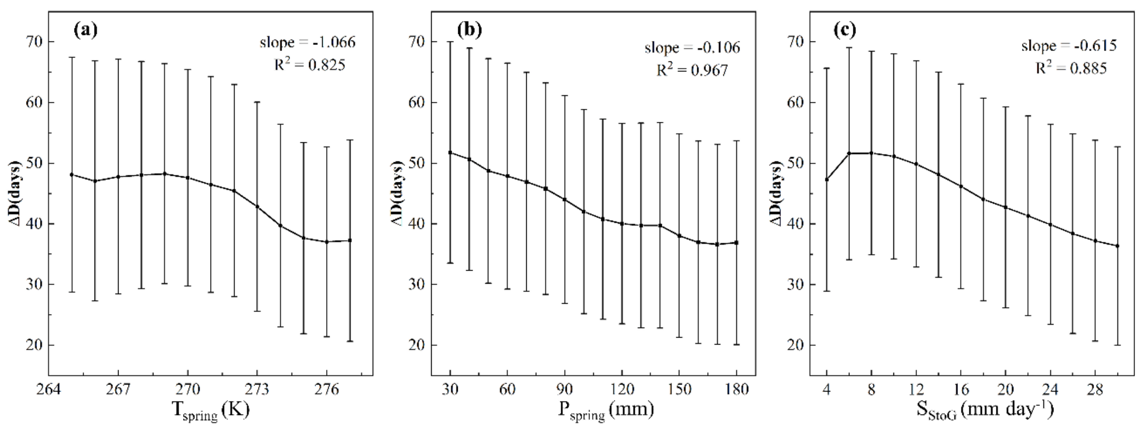

Figure 5: Variations in ΔD across regions with differing (a) spring mean temperature (T_{spring}), (b) spring total rainfall (P_{spring}), and (c) daily snowmelt from D_{SOM} to D_{GU} (S_{StoG}). Points represent the mean ΔD, while error bars denote one standard deviation. The slope and

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.*