To Reviewer #1

General Comments

[Comment 1] *The manuscript presented a spatiotemporally contiguous palaeovegetation cover dataset of the Tibetan Plateau since the 16ka BP, generated by the temporal and temporal random forest (RF) models using modern pollen and fossil pollen records. This is the first mapping of past vegetation on the Tibetan Plateau which should be benefit to the palae-community. I believe that the methods used in this study is feasible and the results are robust.*

[Response] Many thanks for your careful review and constructive suggestions. We greatly appreciate your insightful comments, which have significantly improved the manuscript. We hope our revisions and explanations have satisfactorily addressed your questions and comments.

[Comment 2] *However, given the complexity of the vegetation of the TP, this study only reconstructed two types of vegetation, woody and herb vegetation, plus the total vegetation, which were too coarse for further use. The pollen records can be transformed to multiple vegetation types on the plateau, at least for example, forest, coniferous forest, shrubland and tundra or various alpine vegetation (alpine meadow, steppe and desert). The reviewer suggests that the authors should reconsider the classification of vegetation types as fine as possible.*

[Response] Thank you for this insightful suggestion. In response, we refined our vegetation classification scheme better to reflect the complexity of vegetation on the

Tibetan Plateau. Specifically, we disaggregated the original “woody” category into coniferous forest and broadleaved forest, and the “herbaceous” category into alpine meadow and alpine steppe. In total, we reconstructed seven vegetation types: total vegetation, woody plants, herbaceous plants, broadleaved forest, coniferous forest, alpine meadow, and alpine steppe.

To estimate modern vegetation cover for these types, we used the Global Land Surface Satellite (GLASS) product, extracting total vegetation cover within circular buffers around modern pollen sites (5 km radius for surface soil samples, 50 km for lake sediments). We derived fractional coverage of broadleaved forest, coniferous forest, and herbaceous plants from the MODIS Land Cover Type Product (MCD12Q1), which provides an annual plant functional type (PFT) classification. Specifically, we grouped deciduous broadleaved and evergreen broadleaved forests as “broadleaved forest”, and grouped deciduous coniferous and evergreen coniferous forests as “coniferous forest,” due to their relatively small spatial extents on the Tibetan Plateau (<7% and <1%, respectively; Figure R1). Furthermore, shrubs and trees were grouped as “woody”, rather than being shrubs as a separate vegetation type, as their distribution across the Tibetan Plateau accounts for less than 1%.

We calculated the proportion of each vegetation type within pollen-site buffers and multiplied these by the GLASS total vegetation cover to estimate the fractional cover of broadleaved forest, coniferous forest, and herbaceous plants. To further distinguish alpine meadow and alpine steppe—two ecologically important grassland types in the region—we used the updated Vegetation Map of China (1:1000000) (Su et al., 2020) to determine their relative proportions within herbaceous areas, and then apportioned herbaceous cover accordingly.

Corresponding updates have been made in both the Methods and Results sections of the revised manuscript.

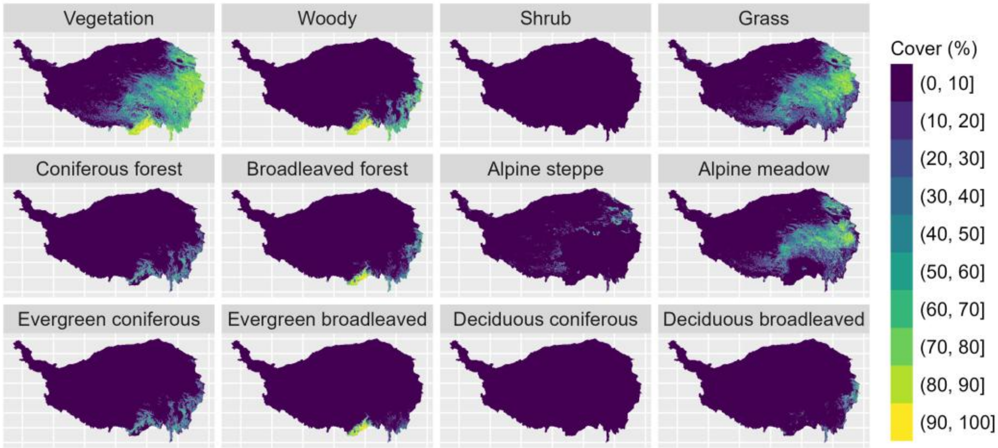

Figure R1. Spatial distribution of modern vegetation cover for different plant functional types (PFTs) on the Tibetan Plateau.

Specific Comments

[Comment 3] *Line 106-108, we used the MODIS Land Cover Type Product (MCD12Q1), which provides an annual Plant Functional Type (PFT) classification (DiMiceli et al., 2022). Trees were classified as "woody", while shrubs and grasses were grouped as "herbaceous." Why not keep the tree, shrub and grass PFTs? These three PFTs should be the appropriate vegetation types of the Tibetan Plateau. Otherwise, trees and shrubs should be classified as "woody", while grasses were grouped as "herbaceous."*

[Response] Thank you for your comment. In our revised analysis, we redefined “woody” as including both trees and shrubs, and “herbaceous” as including only grasses, following the reviewer's recommendation.

Regarding the classification of shrub as a separate PFT, the reason it was not initially treated independently is that the product of MODIS (MCD12Q1) and GLASS data resulted in a very low average shrub cover (<1%) across the Tibetan Plateau (Figure R1), making it difficult to reconstruct reliable spatial and temporal patterns for this category alone. We have refined our classification scheme and now reconstruct seven vegetation types separately: total vegetation, woody plants, herbaceous plants, broadleaved forest, coniferous forest, alpine meadow, and alpine steppe (see detailed responses to **Comment 2** by ***Reviewer #1***). These categories better reflect the ecological diversity of the Tibetan Plateau and are supported by both pollen data and remote sensing products.

[Comment 4] *Line 117-118, for the palaeovegetation models, we standardized all forest vegetation types in the models as woody and all grass and shrub vegetation types as herbaceous. This is not acceptable too.*

[Response] Thank you for your comment. To ensure consistency between pollen-based reconstructions and model-simulated paleo-vegetation cover, we revised the classification scheme used in the vegetation models. Specifically, trees and shrubs were grouped as “woody”, and grasses were classified as “herbaceous”, consistent with common plant functional type (PFT) conventions.

Unlike the pollen-based reconstructions, which included seven PFTs (e.g., total vegetation, woody plants, herbaceous plants, broadleaved forest, coniferous forest, alpine meadow, and alpine steppe), the vegetation outputs from the vegetation models did not include alpine-specific types such as steppe and meadow. As such, only the five shared PFT categories were used for comparison between reconstructed and simulated vegetation cover.

Corresponding changes have been made to both the Methods and Results sections of the revised manuscript to reflect this harmonized classification approach.

[Comment 5] *How to select the past vegetation type from a fossil pollen record at 400-year interval?*

[Response] Thank you for your constructive comment. To account for vegetation adaptation to climate changes over the past 20,000 years, we reconstructed palaeovegetation cover for each 400-year interval by building separate RF models to predict spatial vegetation patterns for that time slice. Therefore, the quality of fossil pollen records in each time interval is critical for model reliability.

We first evaluated the temporal resolution of 61 filtered fossil pollen sequences, which showed a median resolution of 220 years and a 75th percentile of 360 years (Figure R2). We selected a 400-year time interval as a compromise between two key considerations: (i) avoiding overly fine intervals that would require interpolation in RF-temporal reconstructions and introduce additional uncertainties into RF-spatial models, and (ii) ensuring sufficient sample size for each time bin. Our analysis revealed that fossil pollen sample counts increase almost linearly with coarser bin widths between 100 and 400 years, but the gain becomes marginal beyond 400 years; for 400-600 year intervals, the sample sizes per bin are similar (Figure R3).

To further assess the robustness of this choice, we compared RF-spatial model performance at resolutions of 200, 300, 400, 500, and 600 years (Figure R4). The results showed that 400-year bins yielded predictive accuracy comparable to 500–600-year bins but better captured vegetation responses to centennial-scale climatic events, such as the Younger Dryas and Bølling–Allerød periods (Figure R5).

In summary, the 400-year interval represents an optimal balance between model accuracy and the temporal resolution needed to resolve rapid vegetation changes. Nevertheless, future research would greatly benefit from more in *situ* fossil pollen records with higher temporal resolution, as current datasets remain limited in both time and space. All of this information has been added to the revised manuscript.

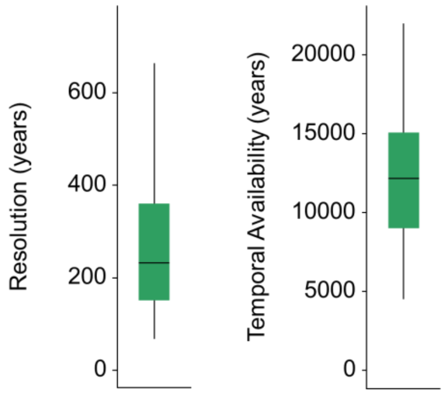

Figure R2. The statistical distribution of the resolution and temporal coverage of the fossil pollen records. The horizontal black line indicates the mean values, bar ends represent the 25th and 75th percentiles, and horizontal lines represent the 5th and 95th percentiles.

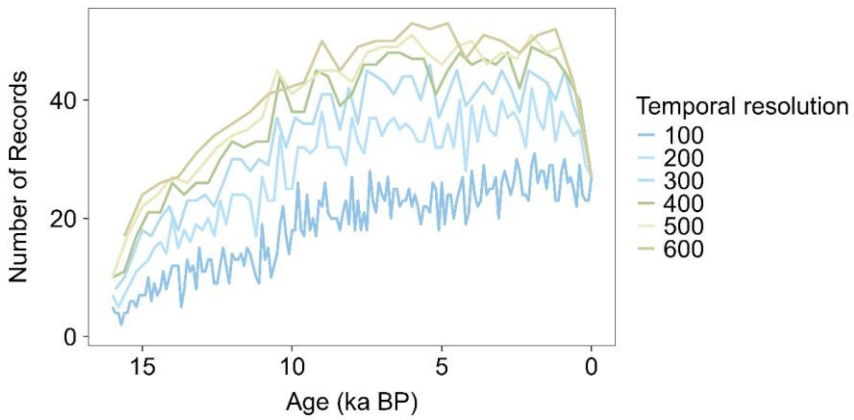

Figure R3. Changes in the number of fossil pollen records from 16 ka to the present at different temporal resolutions.

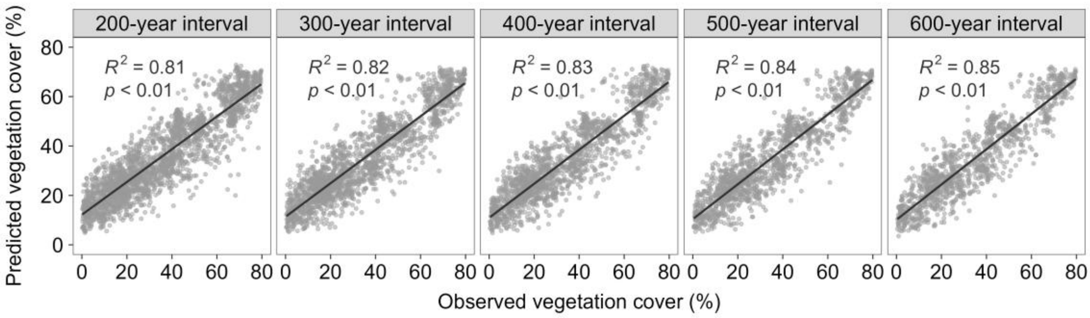

Figure R4. Model performance of Random Forest reconstructions across different temporal resolutions based on 10-fold cross-validation.

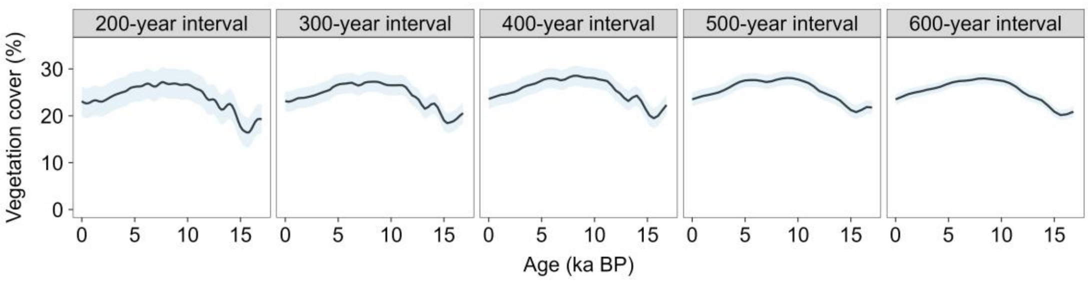

Figure R5. Temporal trends of reconstructed vegetation cover under different temporal resolutions. Solid lines represent regional means estimated using a Generalized Additive Model, and shaded areas indicate 95% confidence intervals.

[Comment 6] *For the RF-temporal models, why have only pollen and topographic* *factors been used? Why not add climate data? But for the RF-spatial models, all the data f climate and topography were applied.*

[Response] Thank you for this important comment. The RF-temporal models were designed to reconstruct temporal changes in vegetation cover at individual site scales. Importantly, this approach is independent of climatic inputs or climate model assumptions, ensuring that our palaeo-reconstructions directly reflect past floristic changes preserved in the sedimentary pollen record. Climate simulations from Earth system models often exhibit biases in temporal trends when compared with proxy-based reconstructions, such as the well-known Holocene temperature conundrum(Liu et al., 2014; Kaufman & Broadman, 2023). Incorporating such climate data into RF-temporal models could therefore distort the true temporal patterns of vegetation change inferred from pollen.

In contrast, climate data are essential for RF-spatial models, which aim to reconstruct the spatial distribution of vegetation at specific time slices. Climate variables provide continuous spatial fields that help capture the regional-scale environmental gradients influencing vegetation patterns. Furthermore, our sensitivity analysis confirmed that the use of gridded climate data in RF-spatial models does not significantly affect the temporal variability of our spatio-temporal dataset. Specifically, we developed 20 sets of RF-spatial models using palaeovegetation cover as the response variable while randomly assigning palaeoclimate data from non-corresponding time bins as predictors. The reconstructed vegetation cover was generally consistent with that from the original models (Figure R6).

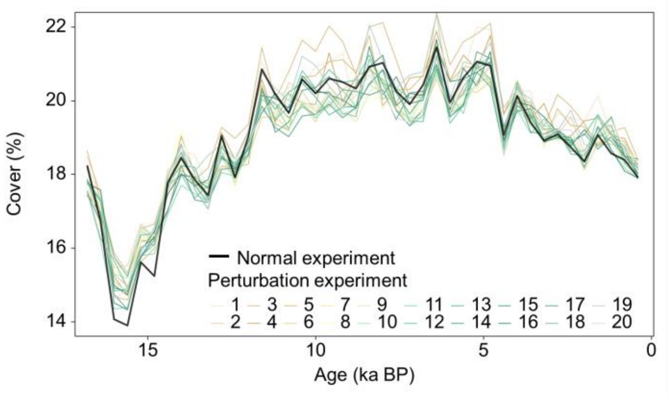

Figure R6. The comparison between the perturbation experiment and the normal experiment. In the perturbation experiment, the temporal sequence of input data used for RF-spatial is randomly scrambled.

[Comment 7] *Line 298, zhang 2024 to Zhang 2024.*

[Response] Thank you for pointing this out. We have corrected "zhang 2024" to "Zhang 2024" in the revised manuscript.

Reference

Kaufman, D. S., & Broadman, E. (2023). Revisiting the Holocene global temperature conundrum. *Nature*, *614*(7948), 425–435. https://doi.org/10.1038/s41586-022-05536-w

Liu, Z., Zhu, J., Rosenthal, Y., Zhang, X., Otto-Bliesner, B. L., Timmermann, A., Smith, R. S., Lohmann, G., Zheng, W., & Elison Timm, O. (2014). The Holocene temperature conundrum. *Proceedings of the National Academy of Sciences*, *111*(34), E3501–E3505. https://doi.org/10.1073/pnas.1407229111

Su, Y., Guo, Q., Hu, T., Guan, H., Jin, S., An, S., Chen, X., Guo, K., Hao, Z., Hu, Y., Huang, Y., Jiang, M., Li, J., Li, Z., Li, X., Li, X., Liang, C., Liu, R., Liu, Q., … Ma, K. (2020). An updated Vegetation Map of China (1:1000000). *Science Bulletin*, *65*(13), 1125–1136. https://doi.org/10.1016/j.scib.2020.04.004