

Author's Responses to RC1's comments on "*Ensemble numerical simulation of permafrost over the Tibetan Plateau from Flexible Permafrost Model: 1950–2023*"

Wen Sun and Bin Cao

State Key Laboratory of Tibetan Plateau Earth System Environment and Resources (TPESER), National Tibetan Plateau Data Center (TPDC), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

Correspondence: Bin Cao (bin.cao@itpcas.ac.cn)

The authors would like to thank the reviewer for their constructive feedback, and the thorough assessment of the manuscript. Below, we provide a point-by-point response to each comment. Reviewer comments are given in black and responses in blue. Additionally, we have included details of how we intend to address these changes in a revised submission.

Permafrost underlies roughly 15% of the Northern Hemisphere's exposed land. Its thaw is reshaping hydrology and ecosystems and undermining the stability of infrastructure. Understanding the trajectory of permafrost dynamics under continued warming is therefore essential. At regional-to-hemispheric scales, numerical models are essential for reconstructing past states, attributing observed trends, and projecting future permafrost dynamics. Here, Sun and Cao introduce the Flexible Permafrost Model (FPM), a standalone land-surface scheme configured in one-dimensional heat conduction, and apply it to a long (1950–2023) ensemble simulation over the Tibetan Plateau (TP). The experiments are forced by ERA5-Land re-analysis data and use a deep soil column (150 m) to reconstruct the permafrost thermal regime. A 45-member ensemble to represent the broad uncertainty of hydrological parameters, and yields spatially consistent estimates of active-layer thickness (ALT), mean annual ground temperature (MAGT), and permafrost areas with observed and previous studies. Evaluation against site observations shows skill of the correct order of magnitude, and the experiments clarify how shallow-column diagnostics can bias permafrost area and trend estimates relative to deep-column simulations.

Overall, this work should be considered by The Cryosphere, provided that the authors address the comments below and supply the requested clarifications.

Comments

1. The author stated that the geothermal numerical model lacks the link with the atmosphere, and the land surface models are not good at representing the permafrost processes. FPM coupled the advantages of these two models to deal with the land-atmosphere interactions and extend the soil column more deeply. Please articulate the specific advantages of FPM relative to existing models: e.g., demonstrably higher accuracy, computational efficiency, or novel parameterizations that capture landscape dynamics.

Responses: RC#2 also suggested to better "*explain how this new model positions itself relative to the existing ones in order to justify its relevance*". In the revision, we will revise as below to clarify.

Significant efforts have been made to understand the permafrost changes over the TP based on simulations. A significant portion of these contributions comes from the hydrological community, employing models originally designed to simulate hydrological processes in permafrost-affected regions. However, many of the models implemented detailed representations of hydrological processes (e.g., water mass balance) while simplifying the surface energy balance and soil thermal processes. For instance, the DHTC model parameterizes ground heat conduction as a linear function of net radiation (Linmao et al., 2024), and the FLEXTopo-FS model uses the Stefan equation rather than a numerical solution for heat conduction (Gao et al., 2022). Beyond such hydrological models, the process-based models used for recent transient permafrost simulation over the TP can be generally divided into geothermal numerical models (i.e., GIPL model) and the common land surface models (i.e., CLM and Noah-MP). The geothermal numerical models typically have rich permafrost-specific processes, such as suitable numerical solver in heat transfer with soil phase changes (Nicolsky et al., 2007; Tubini et al., 2021), deep soil column (tens to hundreds of meters), and well-defined lower boundary, but lack representation of land-atmosphere interactions (i.e., Qin et al., 2017, Sun et al., 2023). On the other hand, the land surface models benefit from the consideration of land-atmosphere processes, and therefore outperform in describing the responses and influences of permafrost to climate warming (i.e., Guo et al., 2018, Wu et al., 2018, Zhang et al., 2021, Cao et al., 2022). Recently, a few

permafrost-specific land surface scheme models—combining the advantages of these two types of models—were proposed. The stand-alone models yield promising potential for application to cross-scale permafrost processes (Fiddes et al., 2015; Westermann et al., 2016). However, dedicated stand-alone permafrost models remain scarce for the TP. Most existing simulations rely on distributed hydrological models that have been enhanced with permafrost process representations (e.g., Gao et al., 2018; Song et al., 2020). Although these models generally offer more realistic and detailed simulations of permafrost-influenced hydrological processes, they are typically confined to site or regional scales and short time periods due to their demand for extensive spatial data and high computational cost (e.g., Pan et al., 2016; Zhang et al., 2017; Zheng et al., 2020).

2. Lines 5–6: The author states, “The FPM accounts for both vertical and lateral heat flow ...”. Yet the present application appears strictly 1-D. Please make this distinction explicit in the Abstract/Introduction/Methods to avoid implying that lateral heat-flux parameterizations are active in this study, or provide details if they are.

We recognize that references to 2D capabilities could be potentially misleading (Referee #2 raised a similar point), as a full assessment of 2D-model suitability requires further applications and evaluation. In the revised manuscript, we will remove the description of lateral heat transfer in FPM from the model description section and relocate it to the outlook section.

3. Section 3.3: I was wondering do the ensemble parameters come from both Table 1 and Table 2 (Lines 177–178), or only from the hydrological parameters in Table 2? In addition, please justify the choice of 45 members per grid cell.

Response: We agree this is misleading. Only the hydrological parameters in Table 2 are used to produce the ensemble member. In the revision, this part will be revised as:

“In this study, the ensemble simulation is produced using reasonable ranges of parameters (Table 2).”

4. Line 190: Could you clarify the spin-up convergence criterion? For example, which variables were evaluated, and what thresholds or tolerances were applied to judge convergence? I also suggest considering a dynamic spin-up for regional runs, which may be more efficient than a fixed 1000-year spin-up per grid cell.

Response: The soil temperature (difference for annual mean soil temperature $< 0.01\text{ }^{\circ}\text{C}$) was used as indicator. We used the 1000 years as spin-up after a large amount experiments. To reduce possible uncertainty, the soil profile was set as 150 m, while only the upper most 100 m is used for further analyses. We agree a dynamic spin-up can be more efficient. In the revision, this part will be changed to:

“To ensure the convergence of soil temperature profile, the model was initialized through a 1000-year spin-up process. This was achieved by cyclically applying the climate forcing data from the first decade (July 1950 to June 1960) one hundred times.”

5. Figure 2:

First, please add a legend. It is hard for me to recognize the meaning of the different lines.

Response: The figure will be revised as below with a legend and revised caption.

Second, reanalysis forced simulations exhibit a larger seasonal amplitude (colder winters, warmer summers) than observations; however, it seems cannot be explained by the cold bias of reanalysis forcing.

Response: The seasonal amplitude is primarily attributed to biases in soil moisture. In general, a wet-biased soil column leads to a colder soil temperature. For example, in subplot (a), the simulated summer soil temperature was cold-biased due to an approximately 25% overestimation in input soil moisture. This clarification will be added to Section 5.1 (Model Evaluations) in the revised manuscript.

“FPM showed relatively worse performance in areas with alpine swamp meadow ($\text{RMSE} = 3.0\text{ }^{\circ}\text{C}$), with warm bias in summer and cold bias in winter. This is primarily attributed to poorly prescribed soil information, i.e., the absence of peat layer in alpine swamp meadow and soil moisture. At the sites with alpine desert, the overestimated soil moisture (by about 25%) at 0.5 m depth leads to a colder simulated soil temperature in summer (Fig. 2a).”

Third, are simulation-observation comparisons shown for the same calendar year for all subfigures? If so, which year?

Response: The daily soil temperature were first averaged to a mean climatological series for each day of the year (DOY) based on all available sites and years for each vegetation type and soil depth. The model simulations were subsequently compared against observations using these DOY time series. This clarification will be added to the figure caption.

Fourth, could the author explain the meaning of the red and blue numbers? I assume they report RMSE and BIAS for reanalysis forced versus observation forced runs.

Response: Yes, red numbers were the RMSE and BIAS of ERA5-Land forced simulations, while the blue ones are for the observation forced runs. The color legend will be added to the revised figure (see below).

Fifth, Table 3 lists the vegetation type at the four sites as alpine marsh meadow. Why is the number 3 in Figure 2?

Response: This is because among the four alpine swamp meadow sites, only three provided measurements at 0.1 m depth, and only two at 1.2 m depth.

In the caption, we will add *"The soil depth and numbers of sites (N) are given in parentheses. The sites used for each vegetation type and depth differ based on data availability."* to clarify.

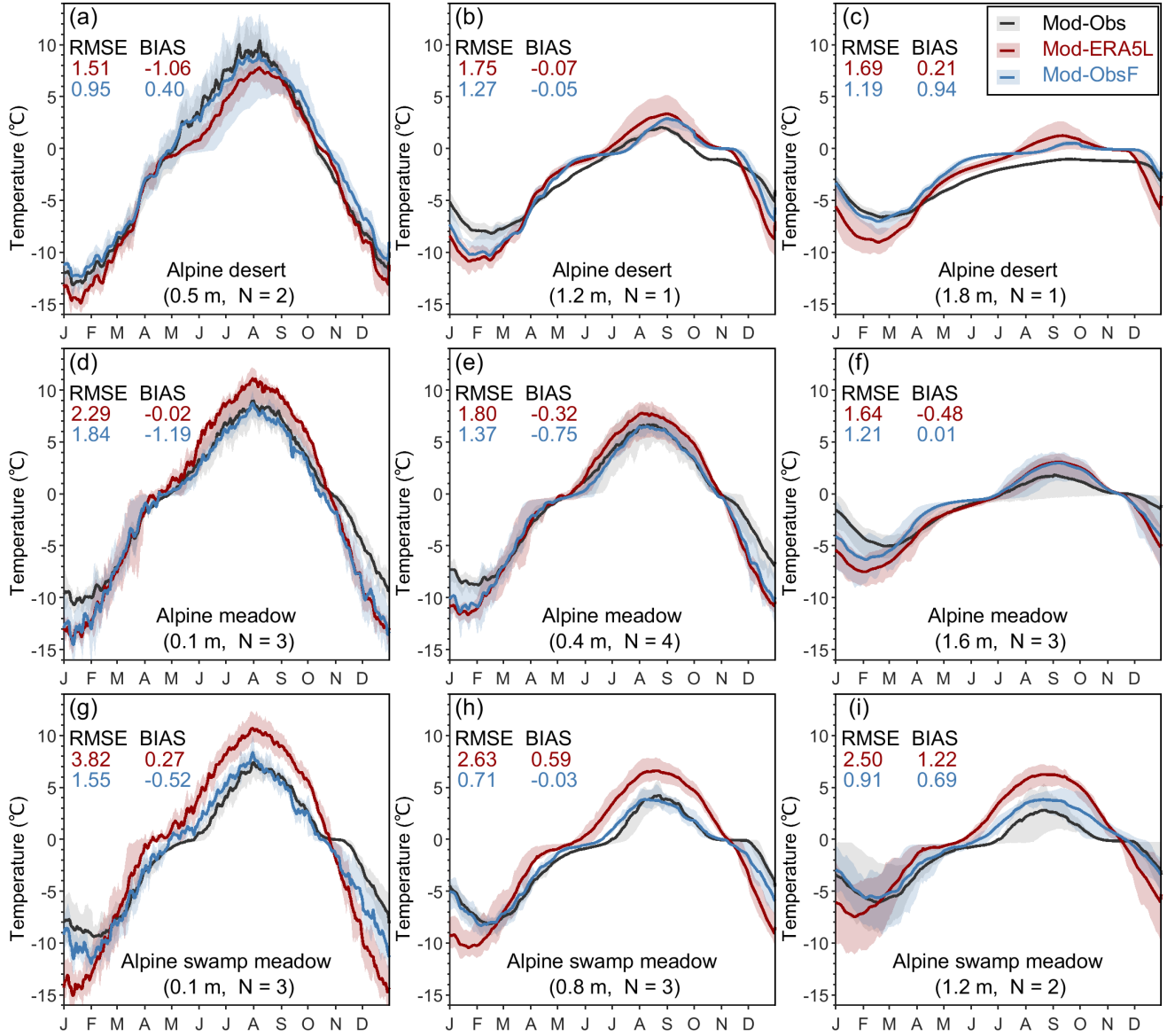


Figure 2: Comparison of simulated and observed day-of-year soil temperature in the active layer across the synthesis sites. The daily soil temperature present is averaged for each vegetation type and soil depth based on all available sites and years. The soil depth and numbers of sites (N) are given in parentheses. The sites used for each vegetation type and depth differ based on data availability. Observations are in black, red lines show the simulation forced by reanalyses, and the blue lines represent that forced by observed atmospheric forcing and *in situ* soil information (if available). The shaded areas depict the ensemble range from the 25th to 75th. The ensemble of observation forced simulation are produced using results from different sites and additional ranges of soil moisture (see Table 2).

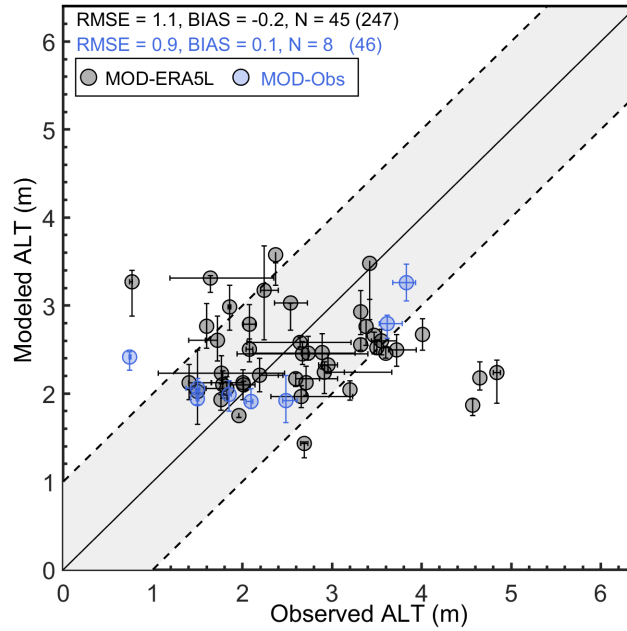


Figure 3: Evaluation of modeled active layer thickness (ALT). The ensemble mean from FPM simulations (MOD-ERA5L) are given in black dot, with the whiskers representing the range between the 25th and 75th percentiles. The observed mean was aggregated from multiple measurements at a single site or from multiple sites within the same grid. N indicates the number of grids used for evaluation after aggregating sites within the same grid, and the number of measurements was given in parentheses. The additional simulation driven by observed meteorological forcing (MOD-Obs) are given in blue. Dashed lines indicate ± 1 m.

6. Figures 3 and 4: First, what is the meaning of the horizontal error bar for each point? Second, the author attributes the cold bias of reanalysis to lead to the colder simulated MAGT and shallower ALT. However, I was wondering does the snow density gives any influence? Because 250 kg m^{-3} may be high for the TP, several studies (e.g., Dai et al., 2018, Yin et al., 2021) report values closer to 150 kg m^{-3} .

Response: The horizontal error bar is the range between the 25th and 75th percentiles for measured grids. This could be either from single site with multi-years' measurements or several sites in the same grid. In the revision, the figure and caption will be revised as above to clarify.

Regarding snow density, some studies (e.g., Dai et al., 2018; Yin et al., 2021) have adopted a bulk snow density of approximately 150 kg m^{-3} for the TP, based on observations from the China Meteorological Administration (CMA). However, recent investigations (e.g., Zhong et al., 2021; Cao et al., 2023; Che et al., 2025) suggest that the snow density values from the CMA network are significantly underestimated when compared with stand-alone measurements. This discrepancy is likely attributed to the CMA's measurement methodology, which employs a heavy snow gauge consisting of a steelyard balance and a 5000 cm^3 tube-cutter, particularly problematic given the generally shallow snowpack over the TP – with a mean snow depth of only 0.01 m across 87 CMA stations. In our preprint, we used a value of 250 kg m^{-3} , which represents a typical constant snow density.

Both RC#2 and RC#3 raised concern regarding the potential uncertainties arising from the use of a static snow density. To address this, we will incorporate the empirical snow compaction parameterization from Versegny (1991) into the FPM. In this scheme, the fresh snow density is set to 100 kg m^{-3} , and snow compaction is calculated as follows:

$$\rho_{sn}^{t+\Delta t} = (\rho_{sn}^t - \rho_{sn}^{max}) \cdot \exp(-0.24\Delta t) + \rho_{sn}^{max} \quad (1)$$

where ρ_{sn}^{max} is assumed to be $300 \text{ (kg m}^{-3}\text{)}$, and Δt is the simulation time step in day.

The updated simulations are very close to those using the static snow density of 250 kg m^{-3} as snow is very minor in most permafrost region over the TP.

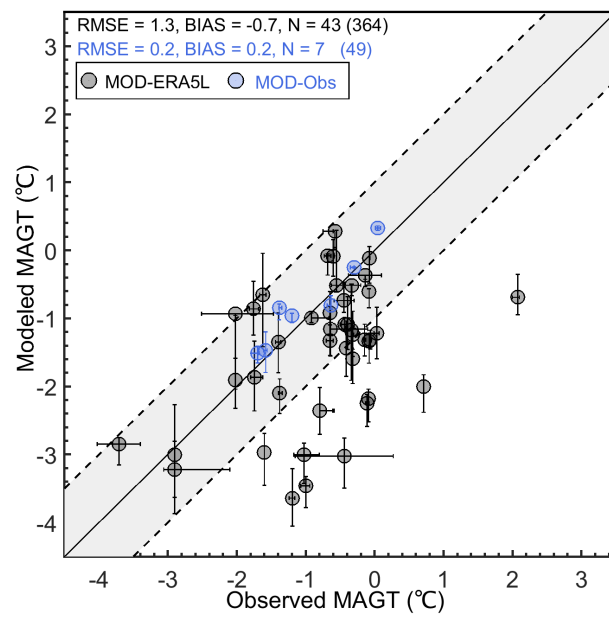


Figure 4: Same as Figure 3, but for the mean annual ground temperature (MAGT). Dashed lines indicate ± 1 °C.

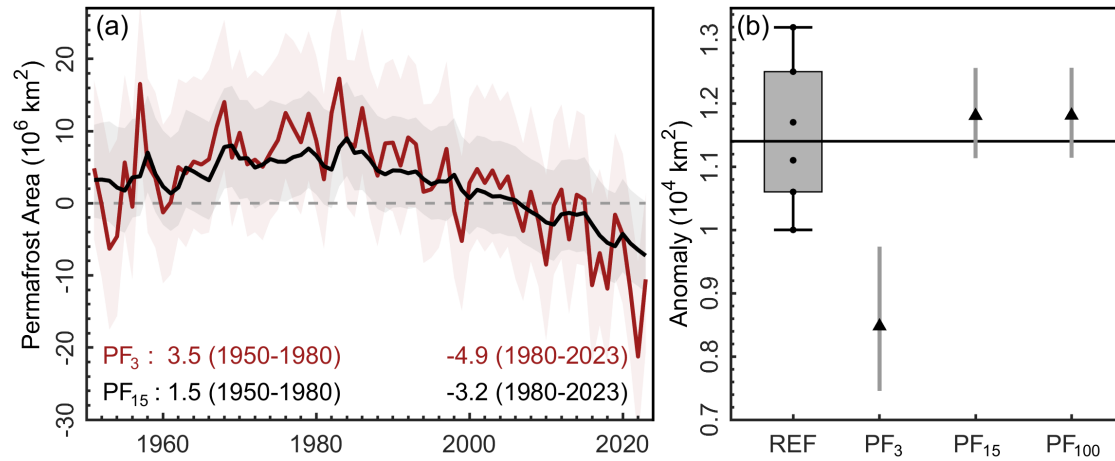


Figure 9: (a) Comparison of estimated current (2010–2023) permafrost area from FPM simulations with referenced estimations, and (b) anomaly of permafrost area since 1950. The subscript means the depth above which permafrost is diagnosed. The permafrost area trend ($10^4 \text{ km}^2 \text{ dec}^{-1}$) is estimated for the periods before and after 1980, separately.

7. Section 5.4: Because the text first discusses the time series of permafrost-extent anomalies, consider swapping the order of Figs. 9(b) and 9(a). Also, I could not locate the source of the 5.2% figure cited on line 289; please clarify.

Response: We will swap the order of Figs 9(b) and 9 (a) as above. The 5.2% indicated increased permafrost area between 1950–1980. In the revision, it will be updated to 2.6% based on the improved simulations with snow compaction scheme.

8. A residual water content of $5 \text{ m}^3 \text{ m}^{-3}$ seems implausible in line 391. Please check the value (units/decimal).

Response: It should be $0.05 \text{ m}^3 \text{ m}^{-3}$, will be revised.

9. Please ensure consistent verb tense throughout the manuscript, e.g., line 306: "introduce and demonstrated".

Response: The verb tense will be revised throughout the manuscript in the revision.

References:

- Cao, B., Wang, S., Hao, J., Sun, W., and Zhang, K.: Inconsistency and correction of manually observed ground surface temperatures over snow-covered regions, *Agricultural and Forest Meteorology*, 338, 109518, 2023.
- Che, T., Dai, L., and Li, X.: Spatiotemporal distribution of seasonal snow density in the Northern Hemisphere based on in situ observation, *Research in Cold and Arid Regions*, 17, 137–144, <https://doi.org/10.1016/j.rcar.2025.02.004>, 2025.
- Dai, L., Che, T., Xie, H., and Wu, X. 2018. Estimation of Snow Depth over the QinghaiTibetan Plateau Based on AMSR-E and MODIS Data. *Remote Sensing*, 10, 1989.
- Yin G., Niu, F., Lin, Z., Luo, J., and Liu, M. 2021. Data-driven spatiotemporal projections of shallow permafrost based on CMIP6 across the Qinghai-Tibet Plateau at 1 km² scale. *Advances in Climate Change Research*, 12, 814–827.