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 reanalysis 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.

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.

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.

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.

5. Figure 2: First, please add a legend. It is hard for me to recognize the meaning of the different lines. 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. Third, are simulation–observation comparisons shown for the same calendar year for all subfigures? If so, which year? 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. Fifth, Table 3 lists the vegetation type at the four sites as alpine marsh meadow. Why is the number 3 in Figure 2?

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 $250\,\mathrm{kg\,m^{-3}}$ may be high for the TP, several studies (e.g., Dai et al., 2018, Yin et al., 2021) report values closer to $150\,\mathrm{kg\,m^{-3}}$.

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.

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

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

Reference:

Dai, L., Che, T., Xie, H., and Wu, X. 2018. Estimation of Snow Depth over the Qinghai-Tibetan 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 km^2 scale. Advances in Climate Change Research, 12, 814–827.