The Tibetan Plateau hosts the world's most extensive high-altitude permafrost areas, covering approximately 40% of the region. Over the past few decades, widespread permafrost degradation on the Plateau has been primarily driven by climate warming. However, the harsh environmental conditions and logistical challenges have severely limited the establishment of observational sites.

To better understand the evolution of the permafrost thermal regime over recent decades, Zhao et al. present a modeling approach—the Moving-Grid Permafrost Model. This model is forced by remote sensing data that have been corrected using a machine learning technique, thereby enabling the reconstruction of the historical permafrost thermal regime in West Kunlun. Furthermore, to enhance the spatial resolution of the simulations, the study employs clustering techniques and parallel computing methods to accelerate model runs. Consequently, the model demonstrated improved performance relative to available observational data and provided valuable insights into the spatiotemporal evolution of permafrost thermal regimes in West Kunlun.

Overall, this work merits attention from The Cryosphere, provided that the authors adequately address the review comments and incorporate additional information.

General Comments

1. In terms of the transient numerical model, First, I would like to understand the distinct advantages of the Moving-Grid Permafrost Model (MVPM) compared to existing models, like GIPL, Noah-MP, CLM, and CryoGrid. The authors state that the MVPM accounts for the thermal properties between frozen and thawed soil, unfrozen water content in frozen soil, ground ice distribution, thaw settlement of the ground surface, and geothermal heat flux to address model deficiencies. However, these physical processes and parameterization schemes are also implemented in other land surface models. Could the authors clarify what specific improvements or innovations MVPM provides over these existing models? Second, how does the MVPM model deal with the water balance in the soil domain? Which scheme does it use for the dynamics of soil water contents? No flow (constant water plus ice contents)? Bucket scheme? Richards equation? Third, although the snow cover on the Tibetan Plateau is relatively thin, it can significantly affect the hydrothermal state of the permafrost beneath it. Does the MVPM model consider the insulation and cooling effect of snow cover? Fourth, does this study activate the ground subsidence module of MVPM? This means that does this study consider the existence of excess ice?

2. Aiming at model forcing, as the only model forcing variable, this study adopted three statistical and machine-learning approaches to extend the land surface temperature from Zou et al (2014, 2017). I was wondering why the authors selected these eight specific input variables—surface air temperature, precipitation, skin temperature, soil temperature, fractional cloud cover, surface radiation budget, leaf area index, and digital elevation model—for the statistical and machine learning approaches. Could the authors clarify whether including more (or fewer) variables might help to avoid issues of model underfitting or overfitting? Besides, Furthermore, what is the basis for selecting the particular datasets used for these variables? For example, the study utilizes the uppermost soil temperature from CFSR, while skin temperature is taken from ERA5Land. Given that ERA5Land also provides uppermost soil temperature data at a higher spatiotemporal resolution compared to CFSR.

Specific Comments

1. Line 40: Simth et al 2022? maybe it is a wrong reference?

2. Line 49: The reference (Zhao et al., 2019) seemed missing.

3. Line 51: Should the Qinghai-Tibet Highway and Railway be abbreviated as QTH? Not sure.

4. Line 84: "ALT" should be given its full name, this is the first time it has been abbreviated. And "refreezing" of what?

5. Line 161: Does this study activate the settlement module?

6. Line 183: What is the surface radiation budget? net radiation? net shortwave radiation? or net longwave radiation? It is not clear.

7. Line 189: The resolution of all input data is not daily.

8. Line 271: How to deal with initial water/ice content?

9. Line 274: "approximately"? It should be an exact number for the grid cell to be simulated.

10. Line 300: "Filed investigation and borehole monitoring datasets"? I guess it is "Field".

11. Line 430: "Grey shading"? only saw the grey line.

12. Section 4.2.4: The author states, " Permafrost area in the West Kunlun kept stable from 1980 to 1999, decreased in the 2000s, while increased between 2010 and 2022." However, the MVPM model is just forced by land surface temperature, which showed an increasing trend between 1980 and 2022 (Figure 4). Could the author explain why the permafrost area increased between 2010 and 2022?

13. Line 601–609: So how about the reanalysis data (like Chinese meteorological forcing datasets, ERA5 Land)? Compared with the forcing data from ESMs, the spatiotemporal resolution of them is better.

14. Could the author explain how this study can be extended to the future projections? Due to its inability to obtain the land surface temperature with higher resolution from remote sensing data in the future, how to diagnose the future condition of permafrost in West Kunlun.