Response to RC1

We sincerely thank Reviewer #1 for the time and effort devoted to reviewing our manuscript. We appreciate the positive recognition of the novelty and practical value of the 2020-period permafrost map. The constructive comments, particularly regarding the need for explicit uncertainty quantification and a deeper discussion of the space-for-time substitution strategy, have helped us substantially strengthen the manuscript.

In response, we have added a new Appendix C presenting a multi-configuration ensemble uncertainty analysis for both the soil parameter *E* and the resulting permafrost distribution. We have also significantly expanded the Discussion section to address the applicability, potential limitations, and proper interpretation of the space-for-time approach. All comments have been carefully addressed in the revised manuscript.

Below, the reviewer's original comments are shown in **blue**, our responses are given in **black**, and the corresponding revisions added to the main text are highlighted in **red**.

The authors present in this manuscript a new high-resolution permafrost distribution map for the Qinghai–Tibet Plateau for the 2020 period. The main novelty lies in the gridded estimation of the empirical soil parameter *E* under the extended FROSTNUM framework, using a space-for-time substitution strategy in the absence of concurrent large-scale field surveys. Overall, the results of this study have good practical value and provide a useful reference for this community.

Response: We sincerely thank the reviewer for recognizing the novelty of our work and the practical value of the generated permafrost dataset. Following the constructive suggestions, we have added a new Appendix C with a comprehensive uncertainty analysis of both the predicted soil parameter *E* and the resulting permafrost map. We also have substantially expanded the Discussion to address the reliability, potential failure scenarios, and scientific interpretation of the space-for-time substitution strategy. Our point-by-point responses are detailed below.

Main points:

(1) The final results lack an explicit uncertainty analysis. At present, the manuscript selects a single optimal scheme from multiple methods and configurations to generate the final map, but does not further quantify the uncertainty of the final results. I recommend that the authors include an assessment of the robustness of both the predicted *E* parameter and the resulting permafrost distribution.

Response: Thank you for this important and constructive suggestion. The original manuscript focused on identifying the optimal configuration and validating the final map against borehole observations and existing products, but did not explicitly quantify uncertainty in the predicted soil parameter *E* or the permafrost distribution.

Following the reviewer's recommendation, we have added a new Appendix C that presents an ensemble-based uncertainty assessment. We generated 16 realizations of the *E* field by combing four sampling strategies with four sample sizes. For each realization we produced a corresponding permafrost map using the FROSTNUM model. We then calculated the ensemble mean, standard deviation (SD), and coefficient of variation (CV) of *E*, as well as the pixel-wise agreement probability across the 16 permafrost maps (Figures R1, R2).

This analysis quantifies the uncertainty arising from the choice of sampling strategy and sample size, an important but not exhaustive component of total uncertainty. We acknowledge that other sources of uncertainty, such as input data errors, the stationarity assumption of the driver–*E* relationship, and model structural choices, are not captured in this ensemble. Nevertheless, the results indicate high robustness: 98.5% of pixels show CV < 5% for *E* and only 1.15% of the plateau falls in the classification transition zone (0.2 < agreement probability < 0.8) for permafrost vs. seasonally frozen ground (SFG).

The corresponding figures and detailed description have been added as Appendix C in the revised manuscript.

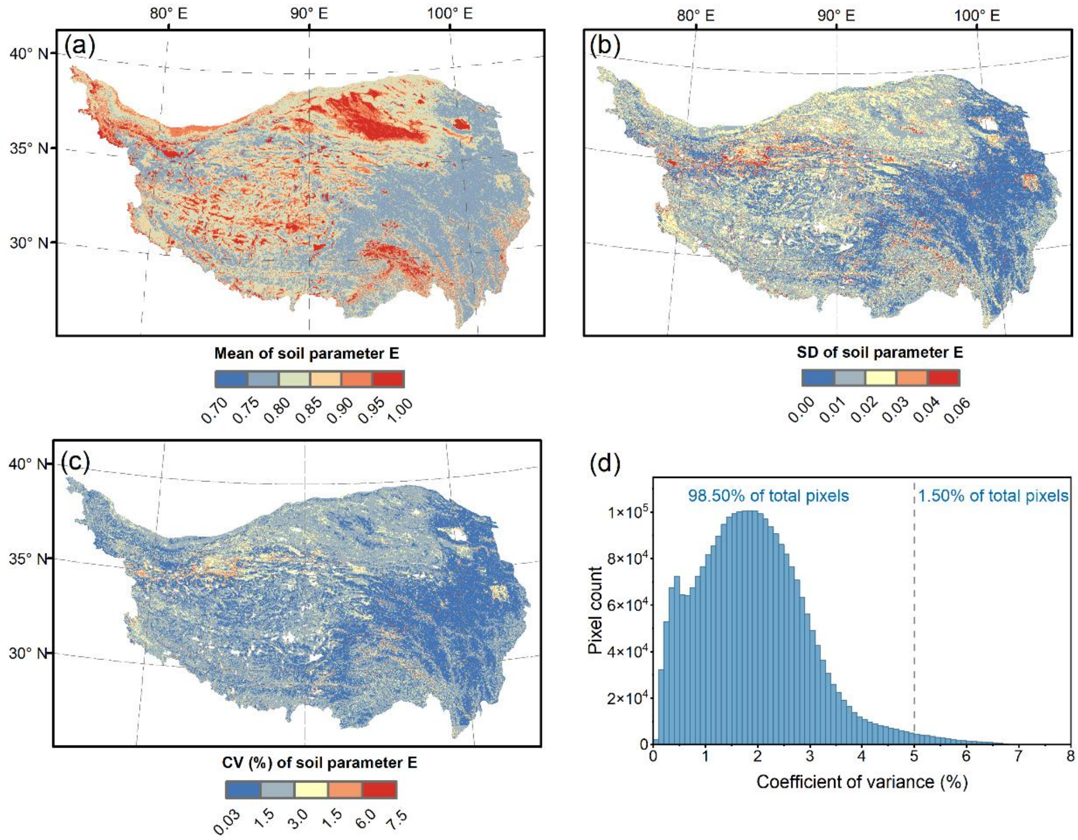

Figure R1 Uncertainty in the soil parameter *E* arising from different training sample selection strategies and sample sizes. Sixteen realizations of the *E* field were generated by combining four sample strategies with four sample sizes. (a) Ensemble mean of *E*; (b) Standard (SD) of *E*; (c) Coefficient of variance (CV, %) of *E*; (d) Frequency distribution of pixel-level CV values across the QTP. The dashed line in panel (d) marks a CV threshold of 5%. This ensemble quantifies uncertainty associated with the choice of sampling strategy and sample size. It does not capture other sources of uncertainty, such as errors in input datasets or potential non-stationarity of the driver–*E* relationship between 2010 and the 2020 period.

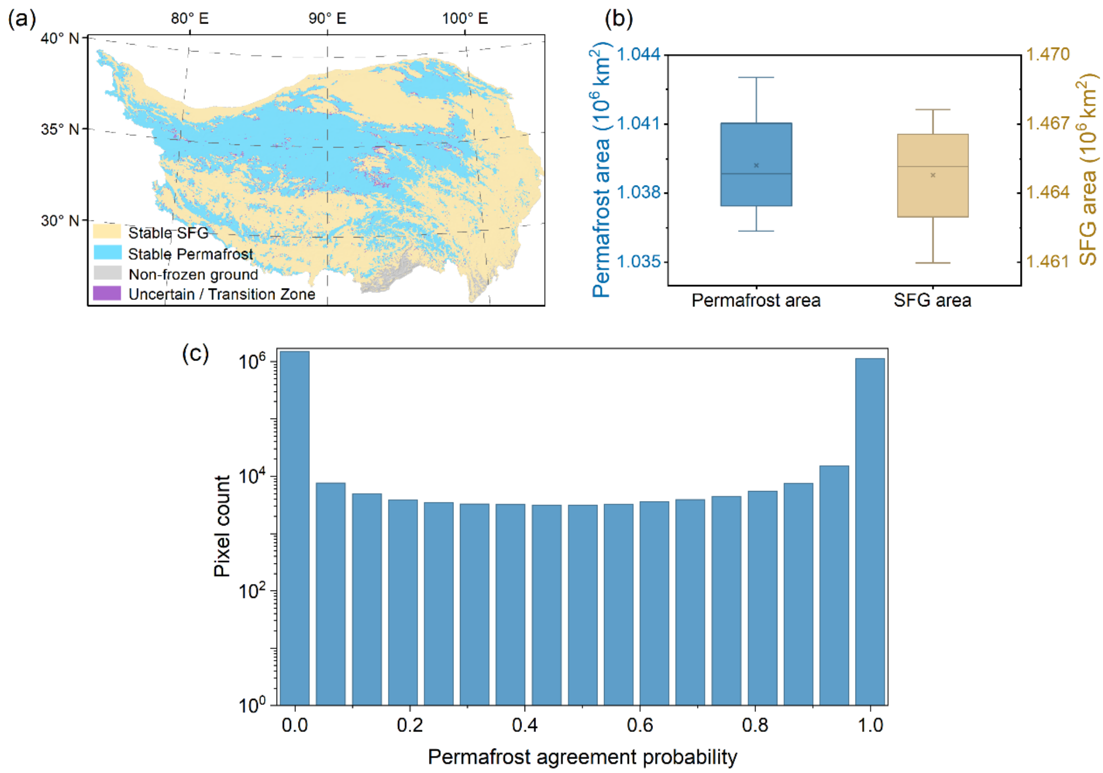

Figure R2 Uncertainty in simulated permafrost distributions arising from different training sample selection strategies and sample sizes. Sixteen permafrost maps were generated using the 16 realizations of the soil parameter *E* shown in Figure R1. (a) Spatial classification consistency across the 16 ensemble members. Blue pixels indicate areas classified as permafrost by all 16 members, yellow pixels indicate areas classified as SFG by all 16 members, and purple pixels indicate areas with inconsistent classifications among ensemble members. (b) Range of total permafrost area (left axis) and SFG area (right axis) estimated from the 16 ensemble members. (c) Frequency distribution of permafrost agreement probability (P_{agree}) across the QTP, shown on a logarithmic scale (log10). Pixels with $P_{agree} \leqslant 0.2$ (stable SFG) and $P_{agree} \geqslant 0.8$ (stable permafrost) occupy 56.69% and 42.16% of the total area, respectively, while 1.15% of pixels fall within the transition zone ($0.2 < P_{\mathrm{agree}} < 0.8$). This analysis quantifies the robustness of the permafrost classification to variations in training sample selection. It does not represent the full uncertainty of the 2020-period permafrost map, as other important sources of uncertainty (e.g., input data errors and the space-for-time stationarity assumption) are not included in the ensemble.

(2) The methodological implications of the space-for-time substitution strategy are not discussed in sufficient depth. In particular, the manuscript would benefit from a clearer discussion of where this strategy is likely to be most reliable, where it may break down, and how this affects the interpretation of the final map.

Response: We thank the reviewer for this valuable comment. We have thoroughly revised and expanded the Discussion section to explicitly address the reliability, potential breakdown conditions, and implications for map interpretation of the space-for-time substitution strategy.

The following text has been added to the Discussion:

"The space-for-time strategy is reliable over decadal timescales in relatively undisturbed alpine ecosystems, where vegetation–soil–permafrost feedbacks track atmospheric warming in a quasi-linear manner. However, this assumption is less applicable under abrupt, non-linear environmental shifts. In localized hotspots undergoing rapid disturbances (e.g., thermokarst collapse or intensive infrastructure development), subsurface hydrological changes can cause the deep soil thermal regime to decouple from surface indicators.

We fully acknowledge that, while the absolute value of E at any location may change with evolving moisture and surface conditions, the space-for-time approach assumes that the underlying statistical relationship between E and its environmental drivers remains sufficiently stationary. This assumption constitutes an inherent methodological limitation. To reduce associated uncertainty we employed representative spatial sampling across environmental gradients and incorporated dynamic predictors (precipitation, NDVI, and fractional snow cover) so that the predicted 2020-period E can respond to observed decadal changes. Nevertheless, the 2020-period map should be interpreted as a quasi-equilibrium representation of near-surface permafrost conditions."

Minor points:

(1) P7 L172–175. A brief explanation of why F > 0.5 is used as the threshold for permafrost classification would improve the clarity of the method description.

Response： We have added the reference and a clear physical explanation of the threshold. In the revised manuscript we have added the following text:

"The parameter E is a dimensionless empirical factor that accounts for the local environmental modification of the ground thermal regime. Theoretically, E reflects the combined effects of soil thermal properties and moisture conditions in both frozen and thawed states. Based on this formulation, the threshold for permafrost occurrence is defined at $F > 0.5$. Pixels with $F \leqslant 0.5$ are classified as SFG or non-frozen ground (Hu et al., 2020). The threshold $F = 0.5$ is physically meaningful because it corresponds to the case where the maximum thawing depth equals the maximum freezing depth. When $F > 0.5$, the thaw depth is smaller than the freeze depth, indicating permafrost; When $F \leqslant 0.5$, the thaw depth exceeds the freeze depth, indicating SFG."

Reference added:

Hu, J., Zhao, S., Nan, Z., Wu, X., Sun, X., & Cheng, G. (2020). An effective approach for mapping permafrost in a large area using subregion maps and satellite data. Permafrost and Periglacial Processes, 31(4), 548-560. https://doi.org/10.1002/ppp.2068

(2) P2 L63–68; P4 L110–115. The manuscript refers to the map as being “for 2020” or an “instantaneous snapshot”, while the forcing data actually cover 2016–2020. A more cautious and consistent expression, such as “for the 2020 period,” is recommended.

Response： We thank the reviewer for this important clarification. We have revised the manuscript throughout to use the more precise term “2020 period” and have removed expressions such as “instantaneous snapshot”.

(3) P12, L308–315. For the discussion of the MLP failure, it would be better to avoid attributing the issue simply to the “black box” nature of deep learning, and instead refer more specifically to the lack of physical constraints and poor generalization.

Response： We have rewritten the relevant paragraph to focus on the lack of physical constraints and limited generalization capability. In the revised manuscript the text now reads:

“The MLP-predicted *E* field was characterized by unrealistic spatial artifacts, with clustering narrowly near zero or scattering across implausible magnitudes (see Appendix B). Crucially, when we utilized this MLP-derived *E* field to reconstruct the permafrost distribution for 2010, the resulting map deviated significantly from the rigorously validated benchmark established by Cao et al. (2023). This failure likely reflects the limited generalization capability of the unconstrained MLP model, which was unable to maintain physically realistic relationships when extrapolated beyond the training domain. The absence of effective physical constraints allowed the model to fit local patterns within the training data but reduced its robustness when applied across the heterogeneous environmental conditions of the entire plateau.”