

Dear Editor,

Thank you for the opportunity to revise our manuscript entitled “A 2020 permafrost distribution map over the Qinghai-Tibet Plateau” (EGUsphere-2026-345).

We are grateful to the Community Comment reviewer (CC1) and both Referees (RC1 and RC2) for their constructive and insightful comments, which have significantly improved the quality and clarity of the manuscript. We have carefully addressed all concerns raised. The major revisions are summarized below.

- The Discussion section has been substantially rewritten and restructured. The original Results subsection on mechanisms driving permafrost degradation has been moved and expanded into a new Discussion subsection (Section 5.1). Dedicated subsections have also been added on model robustness and transferability of the mapping approach (Section 5.3) and on limitations and future directions (Section 5.4), including an explicit treatment of the reliability, applicability limits, and transferability of the space-for-time substitution strategy.
- A new uncertainty analysis (Appendix C) has been added. It quantifies the robustness of both the predicted soil parameter E and the resulting permafrost map through a 16-member ensemble that cross-combines four sampling strategies with four sample sizes.
- The statistical and physical justification of several key points has been strengthened. Welch’s two-sample t -tests and Cohen’s d effect sizes have been added to the environmental-driver analysis (now Figure 10). A clear physical explanation of the $F=0.5$ permafrost-classification threshold has been inserted in Section 3.1.
- Multiple figures and their captions have been revised.
- Terminology and wording have been carefully revised throughout the manuscript for precision and rigor. In particular, “2020” and “instantaneous snapshot” has been replaced by the more accurate “2020 period”; “degradation risk” has been replaced by “degradation potential” (or “trend”); and the description of the space-for-time substitution strategy has been expanded in both the Methods and Discussion.

A detailed point-by-point response to each reviewer’s comments is provided in the attached document. All changes in the manuscript are marked in tracked-changes doc.

Thank you again for your time and consideration.

Yuhong Chen & Zhuotong Nan

On behalf of all authors

Response to CC1

We sincerely thank the reviewer for the time and effort devoted to reviewing our manuscript. We greatly appreciate the positive assessment of our approach and the practical value of the resulting dataset. The constructive comments have helped us improve the clarity, scientific rigor, and overall flow of the manuscript. In particular, we have added Welch's two sample *t*-tests and Cohen's *d* effect sizes for the environmental driver analysis, moved the original Section 4.6 (mechanisms of permafrost degradation) into a new Discussion subsection (Section 5.1) and expanded the treatment of the space-for-time substitution strategy in Sections 5.3 and 5.4, and refined the terminology and methodological descriptions (including consistent use of "2020 period" and replacement of "degradation risk" with "degradation potential"). 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**.

General comments

I have reviewed EGUSPHERE-2026-345 by Chen et al. This paper has mapped permafrost distribution in 2020 over Tibetan plateau, using new strategy of a space-for-time substitution, using neural network and machine learning approaches. The map and estimated permafrost change since 2010 well explained ecohydrological changes in high resolution. It also explained well the limitation of applying neural network scheme for permafrost mapping. Accordingly I would be very positive to publish this in TC after following minor revisions.

Response: We sincerely thank the reviewer for this encouraging and constructive assessment. In the revised manuscript, we have added Welch's two-sample *t*-tests and Cohen's *d* to quantify differences between the positive and negative groups in the environmental driver analysis, restructured the manuscript by moving the soil parameter *E* variation and permafrost degradation mechanism analysis from the Results into the Discussion section, expanded the discussion on the applicability and limitations of the space-for-time substitution strategy, polished the language throughout, and updated the relevant literature citations. Our point-by-point responses to the specific comments are provided below.

Specific comments

174-175: any reference for this threshold?

Response: Thank you for this comment. We have added a supporting reference (Hu et al., 2020) and expanded the text to give a clearer physical explanation of the $F = 0.5$ threshold for permafrost classification.

The following text has been added to Section 3.1 of the revised manuscript:

The parameter E is a dimensionless empirical factor that captures 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 $F \leq 0.5$ are classified as SFG or non-frozen ground (Hu et al., 2020). The threshold $F = 0.5$ is physically meaningful because it represents the point at which the potential freezing and thawing effects are balanced. When $F > 0.5$, the freezing effect dominates over the thawing effect, allowing the ground to remain frozen throughout the year and thus indicating permafrost. When $F \leq 0.5$, the thawing effect equals or exceeds the freezing effect, meaning the ground thaws completely during the warm season and is therefore classified as SFG (or non-frozen ground). (Lines 190-198 in the tracked revised manuscript)

Reference added:

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

181-184: As for hypothesis of E determination, soil moisture would be changed over the decadal scale, and statistical relation between E and environmental drivers might alter.

Response: We thank the reviewer for raising this fundamental point. We fully agree that soil moisture conditions can, and do, change over a decade, and that this could in principle alter the statistical relationship between E and its environmental drivers.

The space-for-time substitution strategy was adopted precisely because concurrent 2020 field surveys to recalibrate E directly are unavailable, a limitation common to many regional-scale modeling studies. Our core hypothesis is that, while the absolute value of E at any given location may shift with changing moisture and surface conditions, the underlying statistical relationship between E and a suite of environmental covariates (topography, soil texture, and dynamic surface conditions) remains sufficiently stable over a decadal interval to allow meaningful extrapolation. To implement this, we trained the Random Forest model on the high-quality 2010 E field (derived from extensive field surveys) and then drove the model with the actual observed 2016–2020 values of the dynamic predictors (precipitation as a proxy for regional soil moisture, NDVI, and fractional snow cover). In this way, the predicted 2020-period E field can respond to the decadal changes in moisture and surface conditions that actually occurred, rather than simply reproducing the 2010 pattern.

We acknowledge, however, that the assumption of a stationary driver– E relationship is an inherent limitation of the space-for-time approach. This assumption is more likely to hold over relatively short decadal timescales in relatively undisturbed alpine environments, but it may weaken or break down under abrupt disturbances (e.g., thermokarst development or intensive engineering activity) or when extrapolated over longer time periods or to strongly different climatic regimes. We have therefore taken several measures to reduce and transparently discuss this uncertainty:

- Use of representative spatial sampling (k -means clustering) across diverse environmental gradients to improve robustness, since random sampling tends to oversample the plateau's homogeneous interior at the expense of rare but distinct environmental settings where the driver- E relationship is most likely sensitive to decadal change;
- Inclusion of dynamic predictors so that E can adjust to observed decadal changes;
- Examination of the physical driver behind this shift in E : the decadal change in E tracks precipitation (used as a proxy for soil moisture), the variable most directly tied to ground thermal inertia, drier regions show a rising E and wetter regions a falling E (Fig. 10), supporting that the model is capturing a genuine moisture-driven signal rather than an extrapolation artifact;

- Testing of model generalization within the 2010 period itself, where ground-truth E values are available. The RF-based model achieved high accuracy and outperformed a more advanced MLP model, and also remained robust across multiple independent training-data sampling strategies. These results indicate that the learned relationship generalizes well;
- Explicit discussion in the revised manuscript of the conditions under which the space-for-time assumption is expected to be reliable versus where it may fail.

We believe these steps, together with the independent borehole validation and inter-comparison with existing 2020-period maps, support the reliability of the 2020-period permafrost map.

In the revised manuscript, we therefore discuss the transferability and limitations of the space-for-time substitution strategy in the Discussion section:

In Section 5.3

Despite its robustness on the QTP, the mapping framework has several limitations that constrain its transferability to other permafrost regions. The space-for-time substitution strategy relies on the assumption that the statistical relationship between E and environmental drivers remains approximately stationary over the extrapolation period. This assumption is more likely to hold in relatively stable alpine environments over decadal timescales, but may break down under rapid or non-linear environmental changes, such as widespread thermokarst development or large-scale land-use change. (Lines 679-683 in the tracked revised manuscript)

Applying the framework successfully to other regions would also require local recalibration of E . The statistical relationships derived for the QTP cannot be transferred directly, since E reflects region-specific interactions among climate, soil properties, vegetation, and topography. High-quality field survey data from the target region would therefore be needed to establish a new benchmark E field there. (Lines 684-687 in the tracked revised manuscript)

In Section 5.4:

...Second, the space-for-time substitution strategy assumes that the relationship between E and its environmental drivers remains sufficiently stable over the study period. While this assumption appears reasonable over decadal timescales on the QTP, it may introduce greater uncertainty when the method is applied over longer time periods or in areas undergoing rapid

environmental disturbances. As discussed in Section 5.3, this limitation also constrains the direct transferability of the current framework to other permafrost regions without local recalibration. (Lines 702-708 in the tracked revised manuscript)

184: a space-for-time substitution strategy, needs for more detail explanation. It is hard to understand only from Fig.2(c).

Response: We sincerely thank the reviewer for this valuable comment. We have expanded the description of the space-for-time substitution strategy in Section 3.1 and added a clear reference to the detailed explanation in Section 3.3.

In Section 3.1 of the revised manuscript, we have added the following text:

To overcome this limitation, we instead adopt a space-for-time substitution strategy, on the hypothesis that although the specific value of E at a given location may shift with changes in moisture or surface cover, its fundamental statistical relationship with environmental drivers (topography, vegetation, soil texture, soil moisture) remains applicable over a decadal scale. This allows us to predict the 2020 E field from relationships learned from the high-quality 2010 dataset (Fig. 2). More specifically, we used the high-quality 2010 E distribution map produced by Cao et al. (2023) and applied two transfer learning strategies, a neural network approach and a random forest (RF) approach, to establish the relationships between E and multiple environmental drivers. By updating the environmental drivers to their 2016-2020 values, we derived the predicted E values for the 2020 period from these learned statistical relationships. Further details on the estimation of the soil parameter E for the 2020 period are provided in Section 3.3. (Lines 204-214 in the tracked revised manuscript)

209: Any other environmental factors for multilinear regression? Any correlation between NDVI and latitude?

Response: Many thanks for this comment. Prior to finalizing the correction model, we tested additional predictors, including DEM, albedo, and longitude, and compared multiple modeling approaches, including nonlinear regression and machine-learning methods (e.g., Random Forest). The combination of LST-derived DDT, NDVI, and latitude already explained the majority of the

variance in the observed ground surface thawing indices (DDT). Adding further variables or increasing model complexity resulted in only marginal improvements.

Physically, LST-derived DDT serves as the primary thermal baseline and implicitly incorporates macro-scale temperature gradients associated with elevation. The remaining offset between satellite LST and 0 cm GST is largely governed by local surface energy balance processes. Within this framework, NDVI captures vegetation-mediated effects on albedo and latent heat partitioning, while latitude serves as a reliable proxy for regional variations in solar incidence angles and radiation budgets across the Plateau.

To evaluate potential multicollinearity between NDVI and latitude we calculated the Variance Inflation Factor ($VIF = 1.26$), which is well below the conservative threshold of 5.0, confirming that the two variables can be included simultaneously without introducing substantial collinearity. These diagnostic tests are provided here for the reviewer's information only and have not been included in the revised manuscript, as they do not change the final model formulation or the main conclusions.

The revised manuscript does, however, contain an explicit statement on the transferability of the correction (Section 5.3):

The correction applied to satellite-derived thawing indices would also need to be adapted to local conditions. The multiple linear regression model used here was optimized for the QTP, where winter snow cover is generally thin and ephemeral. In regions with thick, persistent snow cover, such as much of the Arctic or boreal Canada, an explicit snow insulation term would be needed when calculating the freezing index. (Lines 687-690 in the tracked revised manuscript)

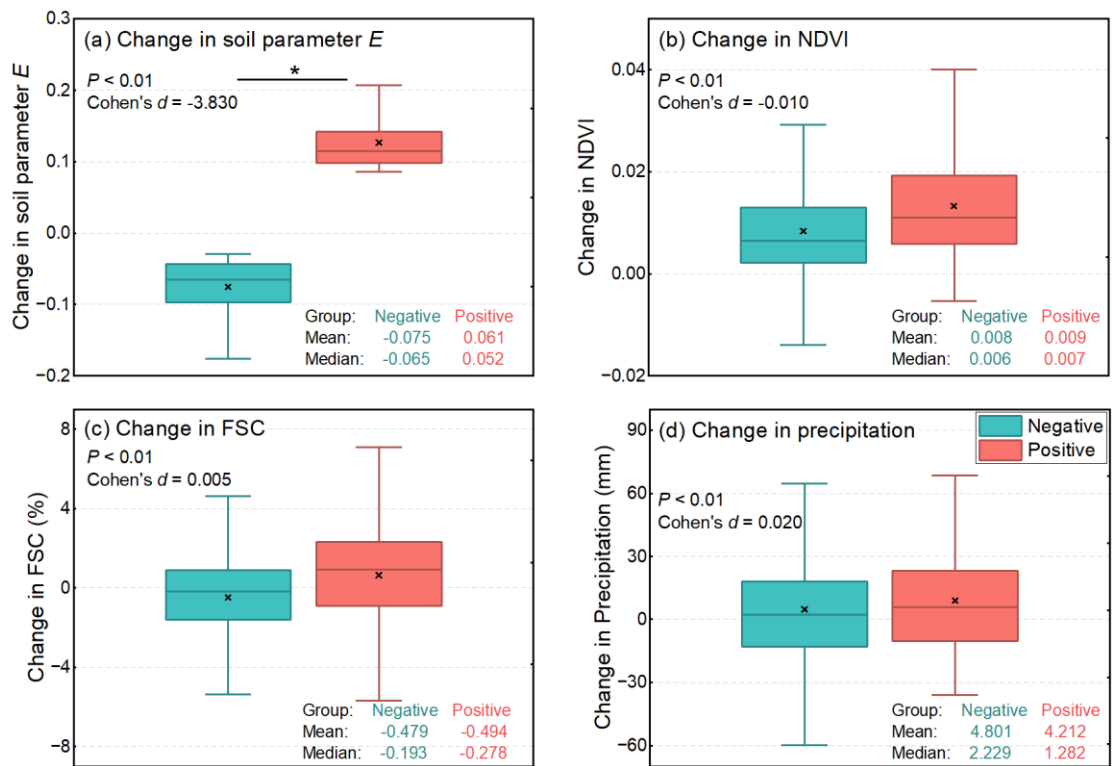
322: Is usage of the term 'risk' appropriate?

Response: We sincerely thank the reviewer for this rigorous reminder. According to the IPCC definition, risk depends not only on the hazard itself but also on exposure and vulnerability. In this study, an increase in parameter E primarily reflects enhanced conditions conducive to permafrost degradation; therefore, the term “risk” lacks scientific rigor in this context. We have replaced “degradation risk” with “degradation potential” or “degradation trend” throughout the revised

Fig.5: Can you say that these boxplots differ between Negative and Positive? Any statistics?

Response: We thank the reviewer for this valuable comment. To quantitatively assess the differences between the Negative and Positive response groups, we applied Welch’s two-sample *t*-test (suitable for large samples without assuming equal variances) together with Cohen’s *d* to evaluate effect size. The difference in parameter *E* between the two groups is highly significant ($P < 0.01$, Cohen’s $d = -3.83$), far exceeding the conventional threshold (0.2) for a large effect. For NDVI and FSC, the *P*-values are both below 0.01, but the effect sizes are negligible (Cohen’s $d = 0.010$ and 0.005 , respectively). For precipitation, the effect size (Cohen’s $d = 0.02$) is several times larger than for NDVI or FSC, indicating that precipitation has a comparatively stronger influence on the divergent trajectories of the two groups under the background of general warming and wetting.

We have updated the text and the figure caption accordingly:



Revised Figure 10 Distribution of decadal changes (2010-2020) in key environmental factors,

categorized by the trend in soil parameter E . Boxplots compare changes in (a) soil parameter E , (b) NDVI, (c) FSC, and (d) precipitation between regions of increased degradation potential (positive group) and reduced degradation potential (negative group). In each box plot, the centre line shows the median, the box represents the lower and upper quartiles (25th-75th percentiles), and the whiskers extend to the furthest data points within 1.5 times the interquartile range. The 'x' marks the mean. Asterisks indicate significant differences between the two groups (Welch Two Sample t -test, $P < 0.01$ and $|\text{Cohen's } d| \geq 0.2$).

The updated texts now read:

The decadal changes in E offer useful insight into the factors controlling permafrost stability on the QTP. Although the plateau experienced a general wetting trend during the study period (Zhang et al., 2021), regions with increasing E received substantially smaller precipitation increases (median +1.28 mm) than regions with decreasing E (median +2.23 mm) (Fig. 10). Because high-resolution soil moisture products are not available for the QTP, mean annual precipitation is used here as a proxy for regional soil moisture conditions. The effect size for precipitation (Cohen's $d = 0.020$) was notably larger than those for NDVI and FSC (Cohen's $d = 0.010$ and 0.005 , respectively). These results suggest that spatial differences in soil moisture, as indicated by precipitation, play a larger role than vegetation or snow cover in modulating ground thermal conditions. (Lines 556-562 in the tracked revised manuscript)

334-335: How we see Fig 5 to confirm this description?

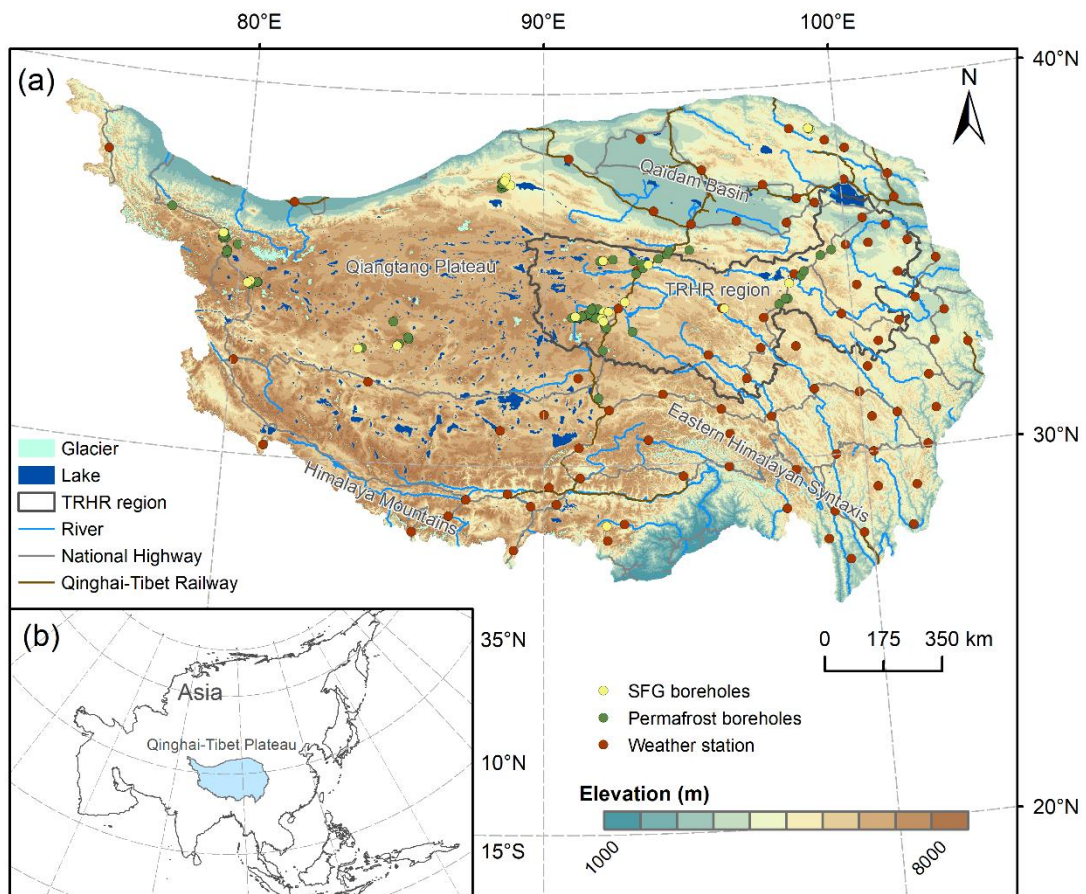
Response: We apologize for the error in the original manuscript. FSC actually exhibited a decreasing trend rather than an increase. This has been corrected in the revised text, and the figure has been renumbered as Figure 10. The corrected sentence now reads:

Although the plateau experienced a general wetting trend during the study period (Zhang et al., 2021), regions with increasing E received substantially smaller precipitation increases (median +1.28 mm) than regions with decreasing E (median +2.23 mm) (Fig. 10). Because high-resolution soil moisture products are not available for the QTP, mean annual precipitation is used here as a proxy for regional soil moisture conditions. The effect size for precipitation

(Cohen's $d = 0.020$) was notably larger than those for NDVI and FSC (Cohen's $d = 0.010$ and 0.005 , respectively). These results suggest that spatial differences in soil moisture, as indicated by precipitation, play a larger role than vegetation or snow cover in modulating ground thermal conditions. (Lines 557-563 in the tracked revised manuscript)

Table 2: Include citations for MAGT and TTOP-based maps. Can you show borehole location in any of the figures?

Response: Thank you for this helpful comment. We have added the corresponding references for the MAGT-based map (Zou et al., 2025) and the TTOP-based map (Yan et al., 2023) in Table 2 and the main text. In addition, the locations of all 109 validation boreholes (83 permafrost and 26 seasonally frozen ground sites) have been added to Figure 1, together with the boundary of the Three-River Headwater Region (TRHR) and the surrounding headwater areas.



Revised Figure 1. Topographic map and geographical context of the Qinghai-Tibet Plateau (QTP).

(a) The topography of the region. The background color gradient represents elevation derived from the Shuttle Radar Topography Mission DEM. The 87 national meteorological stations (red dots) used for ground surface temperature correction are shown, alongside major river systems and lakes. Notably, the 83 permafrost boreholes and 26 seasonally frozen ground (SFG) boreholes for validation are largely concentrated in the Three-River Headwater Region (TRHR) and surrounding headwater areas. (b) The location of the QTP in Asia.

Section 4.6: better to move to Discussions.

Response: Many thanks for this structural suggestion. We have reorganized the manuscript: the original Section 4.6 (Mechanisms driving permafrost degradation) has now been integrated and expanded within the new Discussion Section 5.1 (“Mechanisms driving changes in soil parameter *E* variation and permafrost degradation”) to improve the overall logical flow.

333-340: Better to move to Discussions.

Response: Accepted. This material has been moved to the Discussion section as part of the broader reorganization described above.

Response to RC1

We sincerely thank Referee #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, and 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 provided 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 combining four sampling strategies with four sample sizes. For each realization, we produced a corresponding permafrost map using the FROSTNUM model, and 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 C1 and C2 in the revised manuscript).

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 within the classification transition zone ($0.2 < \text{agreement probability} < 0.8$) for permafrost versus seasonally frozen ground (SFG).

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

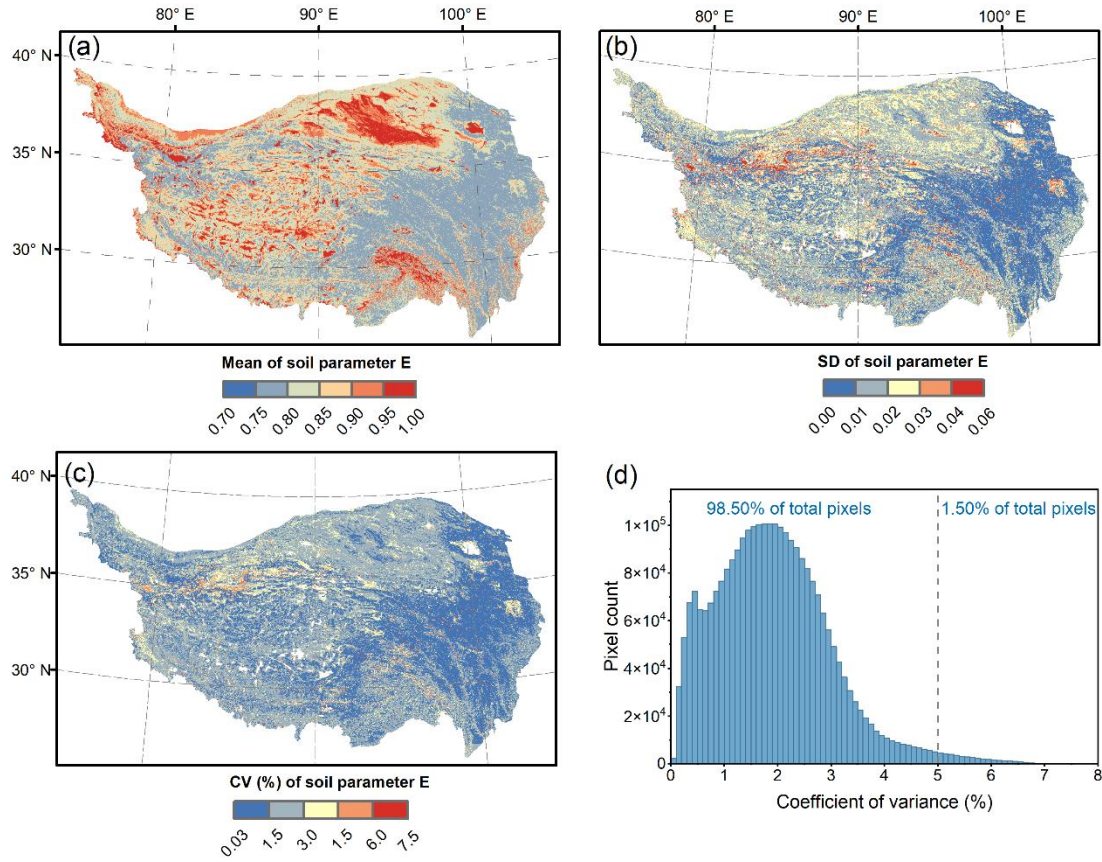


Figure C1. Uncertainty assessment of the soil parameter E derived from 16 model configurations combining different sampling strategies and sample sizes. (a) Ensemble mean of E ; (b) Standard deviation (SD) of E ; (c) Coefficient of variance (CV, %) of E ; and (d) frequency distribution of pixel-level CV values across the QTP. The dashed line in panel (d) indicates a CV threshold of 5%, showing that 98.5% of pixels exhibit CV values below this threshold. 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.

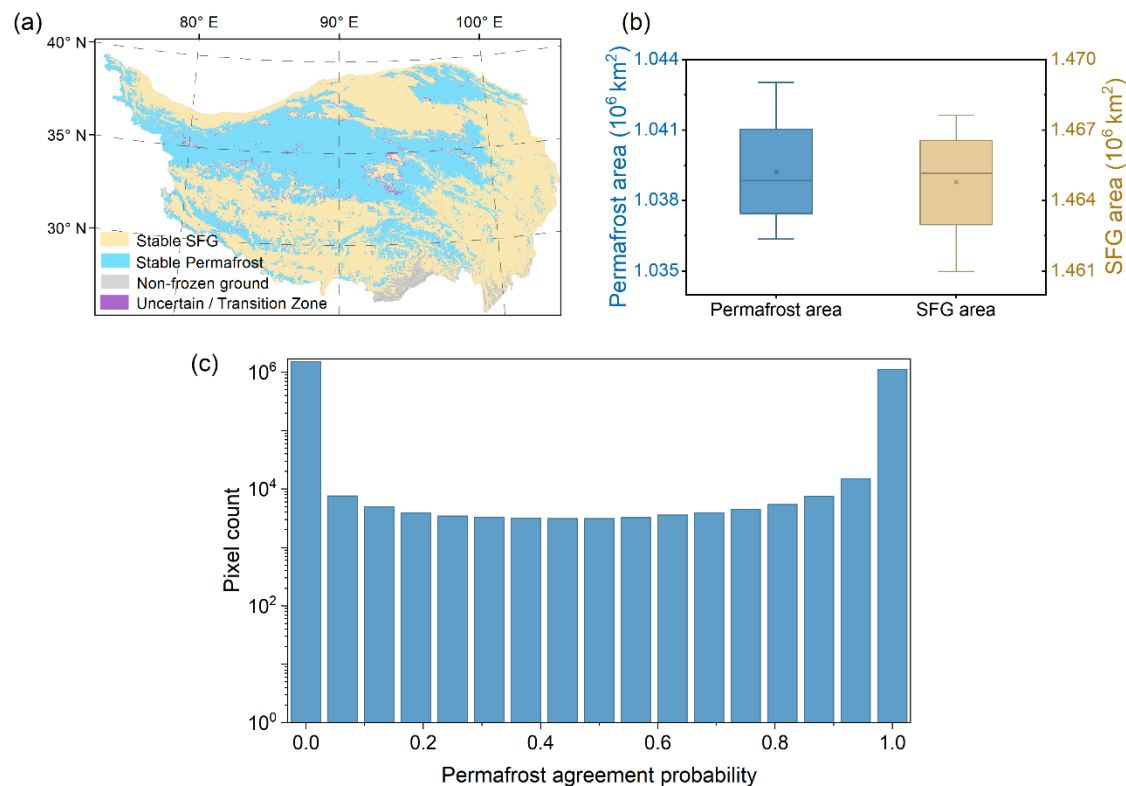


Figure C2. Uncertainty assessment of simulated permafrost distributions derived from 16 parameter E maps. (a) Spatial classification consistency of the 16 ensemble members. Blue pixels indicate areas consistently classified as permafrost by all ensemble members (16/16), yellow pixels indicate areas consistently classified as SFG, and purple pixels denote pixels with inconsistent classifications among ensemble members. (b) Uncertainty ranges of the total permafrost area (left axis) and SFG area (right axis) estimated from these ensemble members. (c) Frequency distribution of permafrost agreement probability (P_{agree}) across the QTP shown on a logarithmic scale (log₁₀). Stable SFG ($P_{agree} \leq 0.2$) and stable permafrost ($P_{agree} \geq 0.8$) occupy 56.69% and 42.16% of the total pixels, respectively, whereas 1.15% of pixels fall within the classification transition zone ($0.2 < P_{agree} < 0.8$). 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 new Discussion Section 5.3 (“Model robustness and transferability of the mapping approach”):

Despite its robustness on the QTP, the mapping framework has several limitations that constrain its transferability to other permafrost regions. The space-for-time substitution strategy relies on the assumption that the statistical relationship between *E* and environmental drivers remains approximately stationary over the extrapolation period. This assumption is more likely to hold in relatively stable alpine environments over decadal timescales, but may break down under rapid or non-linear environmental changes, such as widespread thermokarst development or large-scale land-use change. (Lines 679-683 in the tracked revised manuscript)

Applying the framework successfully to other regions would also require local recalibration of *E*. The statistical relationships derived for the QTP cannot be transferred directly, since *E* reflects region-specific interactions among climate, soil properties, vegetation, and topography. High-quality field survey data from the target region would therefore be needed to establish a new benchmark *E* field there. (Lines 684-687 in the tracked revised manuscript)

We also added the following text to section 5.4 (“Limitations and future directions”):

...Second, the space-for-time substitution strategy assumes that the relationship between *E* and its environmental drivers remains sufficiently stable over the study period. While this assumption appears reasonable over decadal timescales on the QTP, it may introduce greater uncertainty when the method is applied over longer time periods or in areas undergoing rapid environmental disturbances. As discussed in Section 5.3, this limitation also constrains the direct transferability of the current framework to other permafrost regions without local recalibration. (Lines 702-708 in the tracked revised manuscript)

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. The following text has been added to the revised manuscript:

The parameter E is a dimensionless empirical factor that captures 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 $F \leq 0.5$ are classified as SFG or non-frozen ground (Hu et al., 2020). The threshold $F = 0.5$ is physically meaningful because it represents the point at which the potential freezing and thawing effects are balanced. When $F > 0.5$, the freezing effect dominates over the thawing effect, allowing the ground to remain frozen throughout the year and thus indicating permafrost. When $F \leq 0.5$, the thawing effect equals or exceeds the freezing effect, meaning the ground thaws completely during the warm season and is therefore classified as SFG (or non-frozen ground). (Lines 190–198 in the tracked revised manuscript)

Reference added:

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

(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. The revised text now reads:

The MLP-predicted E field showed unrealistic spatial artifacts and failed to reproduce the benchmark 2010 permafrost distribution of Cao et al. (2023) (see Appendix B for details). This suggests that unconstrained neural networks may lack sufficient physical constraints for robust spatial extrapolation in heterogeneous environments such as the QTP. (Lines 368-374 in the tracked revised manuscript)

Response to RC2

We sincerely thank Referee #2 for the constructive and detailed review of our manuscript. We appreciate the positive overall assessment and the valuable suggestions that have helped us improve the clarity, precision, and scientific rigor of the work. In particular, we have strengthened the description of the historical validation for the 2010 period, adopted more precise temporal terminology throughout, expanded the discussion on the physical interpretation and limitations of the map with respect to permafrost depth, and clarified the transferability of the methodology. 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**.

Chen et al. present a 1x1 km map of permafrost conditions on the Qinghai–Tibet Plateau for the 2020 period, along with a novel methodology for deriving said map without access to extensive field data for this period. The authors also perform extensive validation of their methodology and output against available observational data and permafrost maps. I applaud Chen et al. for including the MLP model despite its deficiencies, which can help guide similar applications in the future. Overall, my judgement is that this study is well suited for publication in *The Cryosphere*.

Response: We sincerely thank the reviewer for the positive assessment and for the thoughtful comments that helped us improve the manuscript. We have revised the text to more clearly present the historical validation of the methodology on the 2010 data, adopted consistent and precise temporal language (“2020 period”), strengthened the discussion of map limitations with respect to permafrost depth, and expanded the section on methodological transferability. Our point-by-point responses are provided below.

Specific comments

1. How well does the presented methodology for estimating the soil parameter E perform for the 2010 period? If I understand correctly, the authors compare estimated E for 2020 period to the E determined from the field surveys in 2010. It would be interesting to see how well the two methods are in estimating E for 2010, and how the estimated E has changed from 2010 period to the 2020

period.

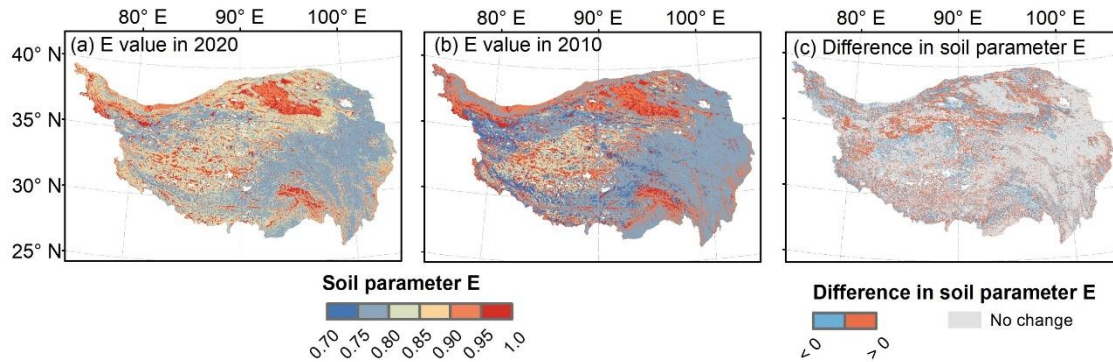
Response: We thank the reviewer for this important question. This validation is already presented in the manuscript. Table 1 reports the accuracy of four sampling strategies when the Random Forest model (trained on environmental covariates) is used to reproduce the benchmark E field derived from field surveys in Cao et al. (2023) for the 2010 period. The optimal configuration (k -means sampling with 5000 points) achieves a correlation coefficient of $r = 0.862$, $RMSE = 0.048$, and explains 74.83% of the spatial variance in the 2010 E values. These metrics demonstrate that the methodology can reliably reconstruct the 2010 E field from environmental drivers.

To make this historical validation more immediately visible to readers, the caption of Table 1 has been revised as follows:

Table 1. Comparison of the predictive accuracy of four sampling strategies for estimating the local soil parameter E for 2010. The baseline parameter E used for validation was derived from Cao et al. (2023). Accuracy metrics (RMSE, MAE, r , and % variance explained) are derived from 5-fold cross-validation using an optimal sample size of 5000 points.

Sampling method	RMSE	MAE	r	% variance explained
k -means sampling	0.048	0.029	0.862	74.83
Equal Range (ER) sampling	0.049	0.030	0.831	69.18
PCA-based sampling	0.049	0.030	0.831	69.90
Latin Hypercube Sampling (LHS)	0.054	0.036	0.806	65.74

We have also clarified the description in Section 3.3 and revised the caption of Figure 4 to explicitly distinguish the 2010 E field (from Cao et al., 2023) from the 2020-period E field predicted using the space-for-time approach. The decadal changes in E (Fig. 4c) are therefore derived from two independently generated fields, rather than from applying the new method to both periods.



Revised Figure 4 Spatial distribution of the soil parameter E in the 2020 period (a) and 2010 (b). (c) illustrates the difference (2020 period minus 2010), where positive values (red) indicate an increase in E (conditions conducive to degradation) and negative values (blue) indicate a decrease. The 2010 E map was derived from Cao et al. (2023), whereas the 2020-period E map was predicted using a space-for-time substitution strategy.

2. The authors consistently refer to the permafrost distribution being for 2020, while in they are using data aggregated over 5 years to predict this distribution. Moreover, regions are classified as seasonally frozen ground if conditions are unfavourable for permafrost during this temporal period, while it can take substantially longer for deep permafrost to completely disappear. I would thus encourage a more precise language regarding the time of validity of the map, and how the classifications relate to the transient thermal state of the ground.

Response: We thank the reviewer for this rigorous and important comment. We have revised the manuscript throughout to consistently use the term “2020 period” instead of “2020” or “instantaneous snapshot.” This better reflects that the thermal forcing indices and environmental covariates are averaged over the 2016–2020 period.

Regarding the deeper concern about permafrost depth and transience, we have added the following clarification to the Discussion (Section 5.4):

Despite its overall robustness, the methodology presented in this study is subject to several limitations. First, the permafrost distribution map primarily reflects the thermal state of the shallow subsurface, typically at depths of around 10–15 m. It does not resolve deep or relict permafrost that may persist below the depth of zero annual amplitude. Consequently, areas classified as SFG in the 2020-period map may still contain permafrost at greater depths that

has not yet fully responded to recent warming. Therefore, SFG areas identified in the map should be interpreted as regions where permafrost was not detected within the resolved depth range, rather than as areas where permafrost has completely disappeared. (Lines 695-701 in the tracked revised manuscript)

3. I would welcome a discussion on the transferability of the presented mapping approach to other climates/environments. It would be interesting to hear the authors thoughts on which aspects of the method, including assumptions, are generalizable, and which are specific to the context of the Qinghai–Tibet Plateau.

Response: We thank the reviewer for this valuable suggestion. We have added a dedicated subsection in the Discussion (Section 5.3, “Model robustness and transferability of the mapping approach”) addressing the transferability of the methodology.

The added text reads:

Despite its robustness on the QTP, the mapping framework has several limitations that constrain its transferability to other permafrost regions. The space-for-time substitution strategy relies on the assumption that the statistical relationship between E and environmental drivers remains approximately stationary over the extrapolation period. This assumption is more likely to hold in relatively stable alpine environments over decadal timescales, but may break down under rapid or non-linear environmental changes, such as widespread thermokarst development or large-scale land-use change. (Lines 679-683 in the tracked revised manuscript)

Applying the framework successfully to other regions would also require local recalibration of E . The statistical relationships derived for the QTP cannot be transferred directly, since E reflects region-specific interactions among climate, soil properties, vegetation, and topography. High-quality field survey data from the target region would therefore be needed to establish a new benchmark E field there. The correction applied to satellite-derived thawing indices would also need to be adapted to local conditions. The multiple linear regression model used here was optimized for the QTP, where winter snow cover is generally thin and ephemeral. In regions with thick, persistent snow cover, such as much of the Arctic or boreal Canada, an explicit

snow insulation term would be needed when calculating the freezing index. Overall, while the core components of the methodology, including the extended FROSTNUM framework, space-for-time substitution, and machine learning-based estimation of E , are conceptually transferable, applying them successfully elsewhere depends on the availability of high-quality local calibration data and appropriate adaptation of the thermal-index correction procedures. (Lines 684-693 in the tracked revised manuscript)

Technical comments:

Line 56-62: This paragraph warrants some reference(s), e.g. for the mentioned “previous mapping efforts” or for the applications requiring “year-specific benchmarks”

Response: We have added appropriate references (Obu et al. 2019, Zou et al. 2017; Ran et al., 2012; Cao et al. 2023) to support the statements regarding previous mapping efforts and the need for year-specific maps.

Beyond these data limitations, a fundamental temporal deficiency persists across most existing products (Obu et al., 2019; Zou et al., 2017): they largely reflect long-term historical averages rather than the current distribution of frozen ground. Because ground observations on the QTP are sparse and unevenly distributed, concentrated mostly along transportation corridors, previous mapping efforts often pooled data across multiple decades to increase sample size (Ran et al., 2012). As a result, these maps depict a composite state that smooths over the rapid climatic shifts of past decades. This reliance on climatological averages creates a temporal mismatch that hinders the accurate validation of land surface models and ecological assessments that require precise, period-specific benchmarks (Cao et al., 2023). (Lines 62-69 in the tracked revised manuscript)

Line 97: please use “per decade” rather than “/10a”

Response: Accepted.

Line 162-163: This does not fit to the subsection “Existing permafrost datasets for comparison”, please move or revise header.

Response: We thank the reviewer for catching this inconsistency. We have revised the heading, which now reads “Datasets for comparison and land cover consistency”.

Section 4.1: It would be interesting to know if there is also a temporal trend in DDT_{LST}, or if the presented cooling pattern is only present in the corrected DDT_{GST} values.

Response: We thank the reviewer for this suggestion. We examined both the raw LST-derived DDT and the vegetation- and latitude-corrected DDT (错误!未找到引用源。). The eastern cooling signal is present in both, but becomes clearer and more spatially coherent after the correction. In the Three-River Headwater Region the central area shows relative differences mostly within $\pm 5\%$, while the cooling reaches 5–15% in the eastern margin of the plateau. Local meteorological-station observations further support the corrected signal: 6 of 8 stations in the eastern margin of the TRHR recorded a decreasing trend in DDT during 2016–2019.

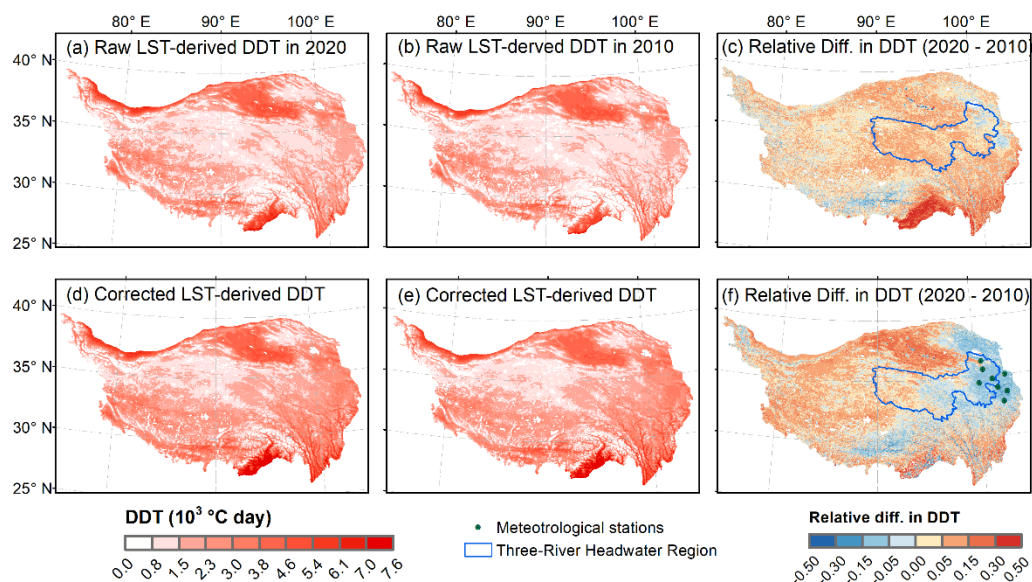
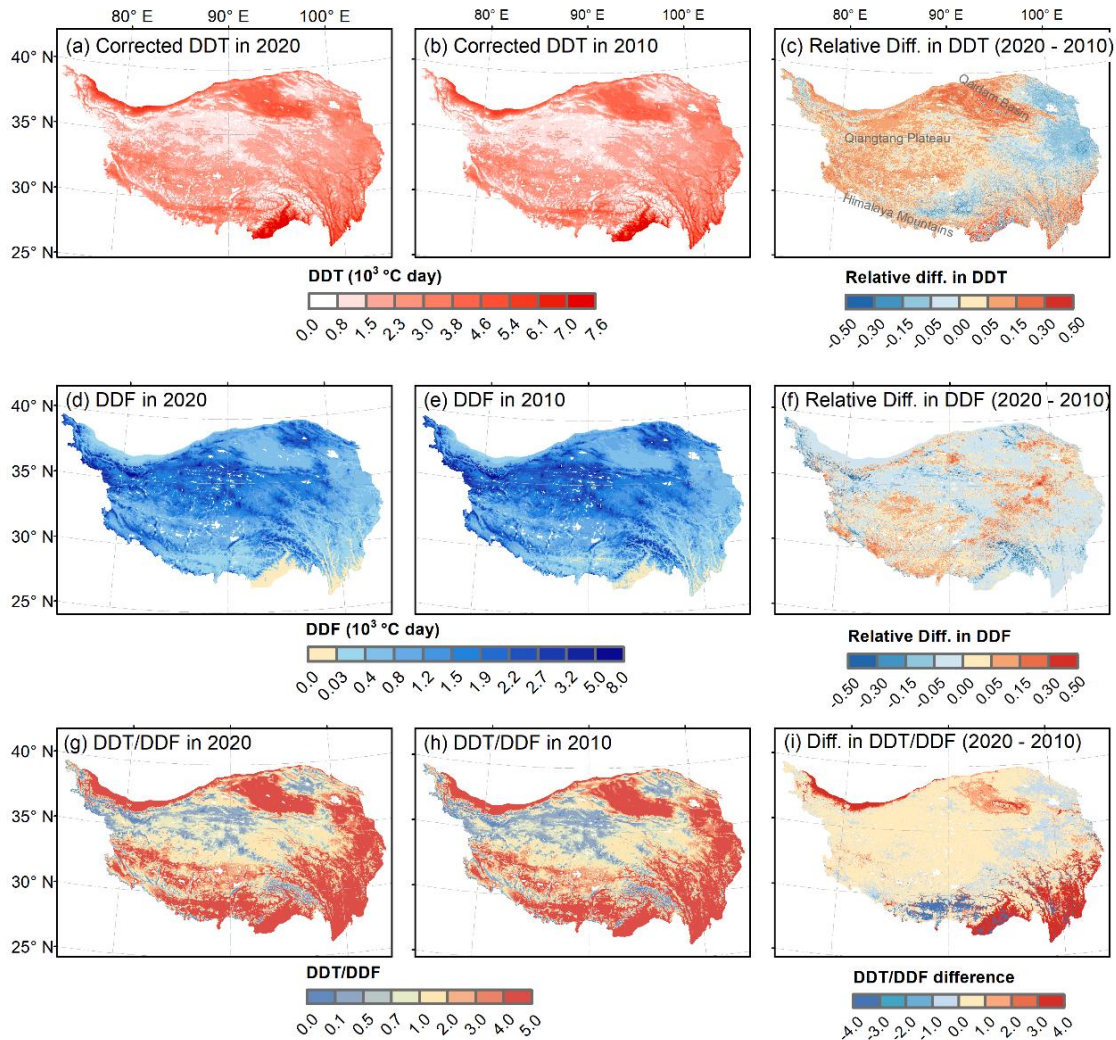


Figure R1 Spatial distribution and decadal of thawing indices (DDT) on the QTP. (a-c) display the LST-derived DDT and its relative changes, (d-f) displays the corrected DDT and its relative changes.

These additional diagnostic comparisons (raw vs corrected DDT and the station-level trends) are provided here for the reviewer’s information only and have not been included as a new figure in

the revised manuscript; the main manuscript now focuses on the corrected DDT (Fig. 3) and the supporting station evidence.

We have also revised panels (c) and (f) of Figure 3 to show relative differences for better interpretability of decadal changes.



Revised Figure 3 Spatial distribution and decadal changes of thermal forcing indices on the QTP. (a-c) *DDT* and its relative changes, (d-f) *DDF* and its relative changes. (g-i) ratio of thawing to freezing indices (*DDT/DDF*) and its decadal change. The left and middle columns present the spatial distribution for the 2020 period and 2010 period, respectively. The right column shows their decadal differences. The relative differences in (c) and (f) are calculated using baseline denominators of 2000 °C·d and 1200 °C·d, respectively, which are the regional mean values of *DDT* and *DDF* averaged across both periods. For panels (c) and (i), positive values (red) indicate

warming/instability, while negative values (blue) indicate cooling/stability. For panel (f), negative values (red) indicate winter warming (reduced freezing), while positive values (blue) indicate cooling.

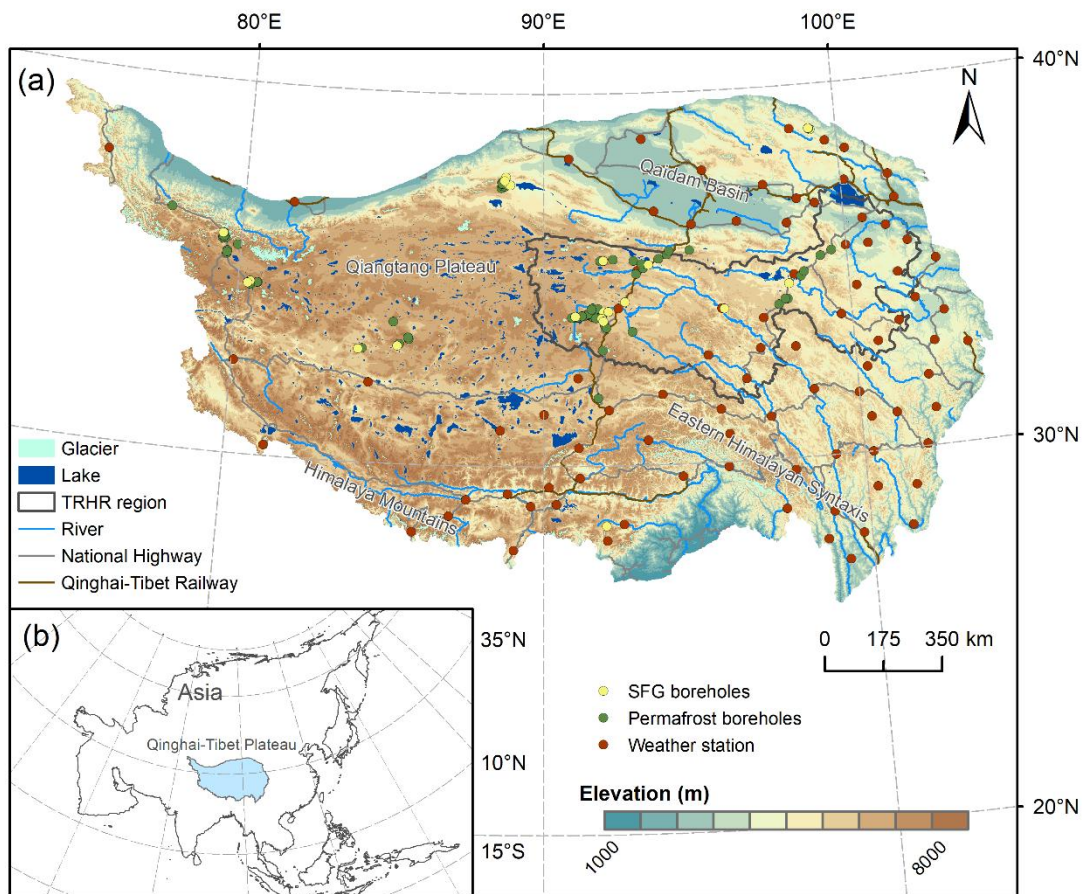
In the revised manuscript, we have updated the description of changes in *DDT* and *DDF*.

The thermal forcing indices driving the permafrost model showed distinct spatial shifts between the 2010 baseline and the 2020 period. The *DDT*, corrected to account for the thermal offset between satellite-observed skin temperature and the actual ground surface, revealed a spatially divergent pattern (Fig. 3a-c). As shown in the relative difference map (Fig. 3c), significant warming persisted in the arid northern Qaidam Basin and the western Qiantang Plateau, where sparse vegetation offers limited thermal buffering. In contrast, a slight cooling trend was observed across parts of the eastern QTP. Analysis of the raw LST-derived *DDT* indicates that most of the Three-River Headwater Region (TRHR) experienced slight warming, whereas the vegetation- and latitude-corrected *DDT* shifts the signal toward slight cooling in this region. The magnitude of this cooling is generally small, with relative differences mostly within $\pm 5\%$ across the central TRHR and reaching 5-15% in the eastern margin of the plateau. The localized cooling is likely attributable to the greening of the plateau, a phenomenon well-documented in the southeastern QTP over recent decades (Shi et al., 2023; Wang et al., 2022). Vegetation modulates the surface energy budget through three primary mechanisms: (1) reducing aerodynamic resistance and enhancing latent heat flux via evapotranspiration; (2) increasing soil water retention, which raises the thermal inertia of the soil; and (3) providing direct insulation that delays the onset of active-layer thaw (Jia et al., 2023; Ni et al., 2025; Wang et al., 2012). This interpretation is supported by local meteorological stations observations: 6 of 8 stations in the eastern margin of the TRHR recorded a decreasing trend in *DDT* during 2016-2019. This ground-based evidence indicates that the cooling signal is not merely an artifact of the satellite LST correction, but reflects a genuine change in surface thermal conditions. (Lines 325-342 in the tracked revised manuscript)

Line 294: Please indicate on this or the overview map where the “southern Himalayan margin” is.

Response: Accepted. We have updated Figure 1 to include major geographic features, including

the Himalaya Mountains, and have added boundaries for the Three-River Headwater Region and the locations of validation boreholes.



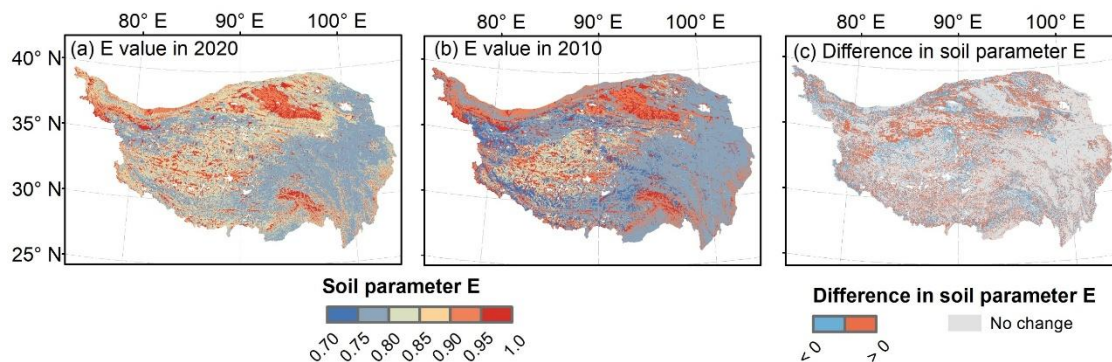
Revised Figure 1 Topographic map and geographical context of the Qinghai-Tibet Plateau (QTP). (a) The topography of the region. The background color gradient represents elevation derived from the Shuttle Radar Topography Mission DEM. The 87 national meteorological stations (red dots) used for ground surface temperature correction are shown, alongside major river systems and lakes are shown. Notably, the 83 permafrost boreholes and 26 seasonally frozen ground (SFG) boreholes for validation are largely concentrated in the Three-River Headwater Region (TRHR) and surrounding headwater areas. (b) The location of the QTP in Asia.

Line 306: Do you mean “temporally extrapolate”?

Response: Yes, this was a typographical error. We have corrected it in the revised manuscript.

Figure 4: It is not clear to me if the “ E value in 2010” refers to the values from Cao et al. (2023). Please be specific here and elsewhere in the manuscript whether the “ E values” and “Permafrost maps” from 2010 refer to the original Cao et al. publication or if they refer to your estimated E values and the thereof calculated permafrost maps.

Response: We have revised the caption of Figure 4 and the relevant text in Section 3.3 and Section 4.2 to explicitly state that the 2010 E field was derived from Cao et al. (2023), while the 2020-period E field was predicted using the space-for-time substitution strategy.



Revised Figure 4 Spatial distribution of the soil parameter E in the 2020 period (a) and 2010 (b). (c) illustrates the difference (2020 minus 2010), where positive values (red) indicate an increase in E (conditions conducive to degradation) and negative values (blue) indicate a decrease. The 2010 E map was derived from Cao et al. (2023), whereas the 2020-period E map was predicted using a space-for-time substitution strategy.

In section 3.3:

The second strategy prioritized spatial continuity and statistical robustness, using the full 2010 E -field retrieved via the original Clustering-Optimization-Prediction procedure (Cao et al., 2023) as the training target. (Lines 268-269 in the tracked revised manuscript)

In section 4.2:

To estimate the empirical soil parameter E for the 2020 period, we applied a space-for-time substitution strategy using two transfer learning approaches trained on the high-quality 2010

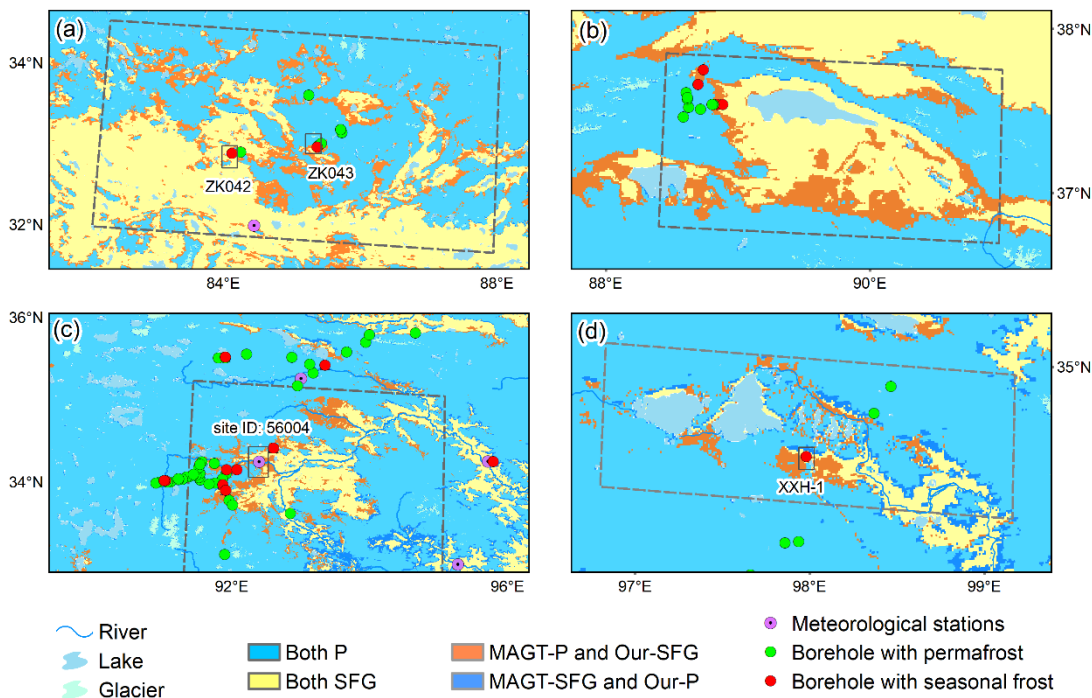
E field from Cao et al. (2023). (Lines 362-364 in the tracked revised manuscript)

Line 366: It's unclear to me where the “Three-River Headwater Region” is. Please include the placenames used in the text in the relevant figure of in the overview map in Figure 1.

Response: We have updated Figure 1 to clearly delineate the Three-River Headwater Region.

Line 407: You state that you identify “four out of five confirmed SFG boreholes”, but there appears to be eight purple “borehole with seasonal frost” markers in Figure 8c.

Response: We thank the reviewer for catching this inconsistency. We have redrawn the figure (Figure 7 in the revised manuscript) with improved symbology and have ensured that the grey dashed boxes and borehole counts are consistent with the text. The revised figures now clearly show five SFG boreholes within the focused subregion.

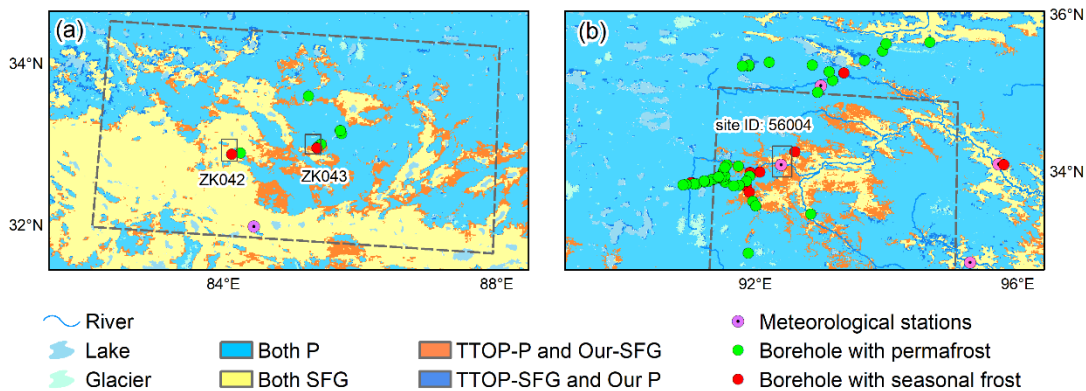


Revised Figure 7 Detailed validation of spatial inconsistencies between the generated 2020 map and the MAGT-based map using ground-truth observations. The panels (a)-(d) correspond to the four subregions outlined in Fig. 7a. “MAGT-P and Our-SFG” represents regions classified as permafrost by the MAGT-based map but as SFG by our map. “MAGT-SFG and Our-P” indicates

the opposite.

Figure 8 & 9: I find it hard to distinguish the borehole markers, please consider revising the symbols for readability.

Response: Both figures have been redrawn with larger, more distinct symbols for better readability.



Revised Figure 8 Detailed validation of spatial inconsistencies between the generated 2020 map and the TTOP-based map in two key transition zones. The panels (a) and (b) correspond to the subregions (A and C) outlined in Fig. 6b.

Line 446: it is unclear where the classifications “continuous and discontinuous permafrost” come from, and if they are relevant for this context.

Response: We have removed the terms “continuous” and “discontinuous permafrost” from the manuscript and now refer to the “transition zone between permafrost and SFG in the plateau hinterland.”

The corrected sentence now reads:

In the central QTP region (subregion A in Fig. 9b), degradation of permafrost into SFG occurred extensively at the permafrost-SFG transition zones. (Lines 533-535 in the tracked revised manuscript)

Line 466: Where do the datasets of thermokarst lakes and engineering instabilities come from? Remote sensing, field mapping?

Response: We have added the data sources in both the main text and the caption of Figure 11: thermokarst lakes were extracted from Sentinel-2A imagery by Chen et al. (2021), and engineering defect points were obtained from field surveys compiled by Li (2021).

The close spatial correspondence between simulated degradation areas and the distribution of thermokarst lakes identified from Sentinel-2A imagery (Chen et al., 2021) and known engineering defect points (Li, 2021) provides independent evidence that the observed changes reflect real physical instability. (Lines 583-586 in the tracked revised manuscript)

The revised caption now reads:

Figure 11 Distribution and changes of frozen ground types in two degradation hotspots on the QTP from 2010 to 2020 periods. Panels (a) and (b) corresponds to subregions A and B in Fig. 9, respectively. (a) Changes in the central QTP, where the dominant transition is from permafrost to SFG. Thermokarst lakes (extracted from Sentinel-2A imagery) (Chen et al., 2021) and engineering defect points (mapped via manual field surveys) (Li, 2021) are superimposed on the map to serve as proxies for permafrost instability. (b) Changes in the southern QTP, characterized by the transition from SFG to non-frozen ground. The observed 0cm GST values from two meteorological stations (ID: 56533, 56548) provide validation for the model results.

References:

- [1] Chen, X., Mu, C., Jia, L., Li, Z., Fan, C., Mu, M., Peng, X., Wu, X. (2021). Thermokarst lakes on the Qinghai-Tibet Plateau (2018). Third Pole Environment Data Center. <https://doi.org/10.11888/Geocry.tpsc.271205>.
- [2] Li, G. (2021). Survey data of major frozen soil engineering diseases in South Asia channel and Himalayas (2020-2021). Third Pole Environment Data Center. <https://doi.org/10.11888/Cryos.tpsc.271880>.

Line 492-502 and elsewhere: I'm surprised that so much attention is given to the transition of seasonally frozen ground to non-frozen ground. While this is an interesting transition, I think this can be deemphasized as the manuscript is about permafrost mapping.

Response: We agree with the reviewer's assessment. We have substantially condensed and

refocused this discussion to maintain emphasis on permafrost dynamics, while retaining only the essential validation from the two meteorological stations.

In the southern QTP, the primary process was the conversion of SFG to non-frozen ground (Fig. 11b). This transition was accompanied by a much larger increase in the thawing index in degraded areas ($310.03\text{ }^{\circ}\text{C}\cdot\text{d}$) than in adjacent stable SFG regions ($131.93\text{ }^{\circ}\text{C}\cdot\text{d}$), while changes in the freezing index were minimal (Fig. 13). The marked increase in E (Fig. 13d) in degraded southern regions suggests that land-surface feedbacks further accelerated the loss of seasonal frost. These contrasting patterns indicate that the dominant mechanisms differ between the central and southern parts of the plateau. (Lines 598-602 in the tracked revised manuscript)