

A 2020 permafrost distribution map ~~over~~^{of} the Qinghai-Tibet Plateau

Yuhong Chen¹, Zhuotong Nan^{2,3}, Wenbiao Tian³, Yi Zhao⁴, Shuping ~~Zhao~~³~~Zhao~~^{2,3}, Dongkai Yang¹, Guifei Jing¹, and Fujun Niu²

¹ Hangzhou International Innovation Institute, Beihang University, Hangzhou, China

5 ² Yangtze River Delta Urban Wetland Ecosystem National Field Scientific Observation and Research Station, School of Environment and Geographic Sciences, Shanghai Normal University, Shanghai, 200234, China

³ State Key Laboratory of Climate System Prediction and Risk Management, Nanjing Normal University, Nanjing, 210023, China

10 ⁴ Key Laboratory of Water and Sediment Sciences, Ministry of Education, College of Environmental Sciences and Engineering, Peking University, Beijing, China

Correspondence to: Zhuotong Nan (nanzt@~~njnu~~^{shnu}.edu.cn)

Abstract. Permafrost on the Qinghai-Tibet Plateau (QTP) is undergoing rapid degradation, yet most existing distribution maps reflect long-term historical averages, ~~failing to capture rather than~~ the current thermal state ~~required of the ground. This temporal mismatch limits their usefulness~~ for ~~accurate~~ ecological ~~and~~ hydrological ~~modeling. To address this temporal mismatch, this study presents, and engineering applications. Here, we present~~ a 1-km ~~QTP~~^{resolution} permafrost distribution map for the ~~year~~ 2020. ~~We employed period using~~ an extended ground surface frost number model (FROSTNUM) driven by satellite-derived freezing/thawing indices. ~~To overcome the lack of~~ ~~Because no~~ concurrent field ~~surveys for parameter calibrations~~ ~~survey was available~~, we ~~implemented~~^{applied} a space-for-time substitution strategy, ~~utilizing a with~~ Random Forest regression to ~~robustly~~ estimate the empirical soil parameter (E) ~~based on~~^{from} environmental covariates. The resulting map ~~reveals~~^{shows} that ~~in 2020~~, permafrost covered approximately 1.038×10^6 km² (39.35% of the ~~plateau~~^{QTP}), while seasonally frozen ground (SFG) covered 1.466×10^6 km² (55.57%). Compared ~~to~~^{with} the 2010 baseline, the permafrost area declined by 4.8×10^4 km² (a 1.82% decrease). ~~Spatially, the degradation of~~ ~~Degradation was spatially heterogeneous: the transition from permafrost to SFG extensively occurred~~ ~~was dominant~~ in the central QTP ~~(accounting for 7.41% of the total change), and a significant marginal contraction, whereas the southern margin experienced substantial conversion~~ of SFG to non-frozen ground ~~in the southern margin (accounting for 39.62% of the total change).~~ Validations against 109 independent borehole records ~~from the 2020 period confirms the map's reliability, achieving~~ ~~yielded~~ an overall accuracy of 0.84 and a Kappa of 0.58, ~~outperforming existing 2020-period maps. This 2020 map provides an essential, up-to-date resource for quantifying the recent cryospheric shifts and supporting a temporally specific reference for engineering risk assessments~~ ~~assessment, ecological monitoring and the calibration of land surface models~~ in this ~~climate-sensitive~~^{rapidly changing} region.

30 1 Introduction

Permafrost, defined as ground that remains below 0 °C for at least two consecutive years (Lewkowicz et al., 2025), underlies approximately 24% of the exposed land surface in the Northern Hemisphere (Obu et al., 2019; Zhang et al., 1999). As a ~~critical~~^{key} component of the cryosphere, permafrost is highly sensitive to climate fluctuations and human interference

(Burke et al., 2020; Chadburn et al., 2017; Smith et al., 2022). The Qinghai-Tibet Plateau (QTP), known as the “Third Pole” and “Asian Water Tower”, hosts the largest extent of permafrost in the mid- and low-latitude regions (Zhao et al., 2004). ~~Due~~Owing to its unique high-altitude environment, QTP permafrost is characterized by relatively high temperatures and low thermal stability, making it particularly ~~responsivesensitive~~ to warming ~~trends~~ (Li et al., 2008; Yang et al., 2019; Yao et al., 2022). Under ongoing global warming, permafrost degradation on the plateau has triggered profound impacts on regional hydrological processes (Tananaev and Lotsari, 2022), carbon cycling (Mu et al., 2020), ecosystem stability (Yang et al., 2010), and the safety of critical engineering infrastructure (Hjort et al., 2022). These ~~widespreadwild-ranging~~ impacts underscore the ~~critical importance of produceingneed for~~ high-quality, up-to-date permafrost maps, which ~~serve as indispensable~~provide an ~~essential~~ foundational ~~datasets~~ for quantifying environmental changes and supporting risk assessments in cold regions (Kim et al., 2024; Obu et al., 2019).

Although numerous permafrost maps of the QTP have been compiled (Guo and Wang, 2013; Zou et al., 2017), the evolution of mapping strategies reflects ~~a continuousan ongoing~~ struggle ~~againstwith~~ data scarcity. Early maps were manually delineated on topographic bases using expert judgement and limited data sources, such as air temperature and field surveys (Cheng et al., 1996; Mi, 1990). While pioneering, these methods were ~~prone to subjectivityinevitably subjective~~. Subsequent studies adopted empirical or semi-physical models for mapping, including altitude model (Li and Cheng, 1999), mean annual ground temperature (MAGT) model (Nan et al., 2002; Ran et al., 2022), frost number model (Nan et al., 2013; Shan et al., 2022), and the temperature at the top of permafrost (TTOP) model (Obu et al., 2019; Zou et al., 2017). To overcome the sparsity of in-situ observations, ~~latestmore recent~~ semi-physical model studies ~~extensively usedhave made extensive use of~~ remote sensing and reanalysis datasets (Hachem et al., 2009; Zou et al., 2017). While these ~~extensive~~ datasets ~~mitigatehelp to fill~~ the spatial gaps ~~ofin~~ ground station networks, they ~~also~~ introduce ~~inherenttheir own~~ uncertainties. Remote sensing products, particularly land surface temperature (LST), are frequently plagued by cloud contamination, leading to significant data gaps (Yu et al., 2015). Conversely, reanalysis datasets, while spatially complete, often suffer from coarse spatial resolution and considerable uncertainty in complex terrain (Hu et al., 2019; Mao et al., 2010). More recently, physically-based land surface models (LSMs) (e.g. ~~(Zhang et al., 2019; Zhao et al., 2022))~~) and machine learning approaches (e.g. ~~(Ni et al., 2021; ZhangRan et al., 2024)), 2021)~~) have been ~~employedused~~ to capture complex process mechanisms and non-linear relationships. While LSMs can ~~theoreticallyin principle~~ simulate temporal evolution, they ~~are prone to carry~~ considerable uncertainty when upscaled to the heterogeneous QTP without extensive calibration, and ~~purepurely data-driven~~ machine learning models often suffer from spatial biases ~~dueowing~~ to the lack of representative training data across the vast plateau (Gay et al., 2023; Siewert, 2018).

Beyond these data ~~source~~limitations, a fundamental temporal deficiency persists across most existing products: ~~(Obu et al., 2019; Zou et al., 2017)~~; they largely reflect long-term historical ~~averageconditionsaverages~~ rather than the current distribution of frozen ground. Because ground observations on the QTP are sparse and unevenly distributed, ~~mostly~~ concentrated ~~mostly~~ along transportation corridors, previous mapping efforts often ~~aggregatedpooled~~ data across multiple decades to ~~maximizeincrease~~ sample size. ~~Consequently (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, ~~hindering that hinders~~ the accurate validation of land surface models and ecological assessments that require precise, ~~yearperiod-specific benchmarks~~. (Cao et al., 2023).

To address the need for a temporally specific benchmark, our previous work (Cao et al., 2023) developed a high-quality permafrost distribution map for ~~the specific year of~~ 2010. That study employed an extended ground surface frost number model (FROSTNUM), incorporating cloud-gap-filled satellite freezing/thawing indices and an empirical soil parameter (E) calibrated against extensive survey data from the 2009–2010 period. It ~~successfully~~ provided the first ~~instantaneous snapshotperiod-specific benchmark~~ of the QTP permafrost distribution and has been widely used since its publication (e.g. (Deng et al., 2024; Pan et al., 2024; Zhong et al., 2024)).

~~Nevertheless~~ However, the climate on the QTP has changed rapidly ~~duringover~~ the 2010–2020 decade (Hu et al., 2024; Zhao et al., 2025), ~~makingso~~ the 2010-period map ~~is now~~ increasingly outdated for contemporary applications. Producing a comparable map for ~~the 2020 is challengingperiod is not straightforward, however,~~ because the extensive field surveys ~~availablecarried out~~ for 2010 were not repeated in 2020. To fill this gap, ~~this study presentswe present~~ a new, high-accuracy 1-km QTP permafrost map ~~of the QTP for the 2020 period~~. We refine the FROSTNUM mapping framework (Cao et al., 2023; Hu et al., 2020) by introducing a space-for-time substitution strategy. We hypothesize that ~~whilealthough the~~ climatic drivers (freezing and thawing indices) have shifted, the ~~fundamentalunderlying statistical~~ relationship between the empirical soil parameter (E) and environmental covariates (e.g., topography, soil texture, ~~vegetation, and~~ soil moisture) ~~has remained applicableremains sufficiently stable over a decadal interval to support meaningful extrapolation~~. By ~~evaluating transfertraining machine learning strategies, comparing neural network and Random Forest (RF) approaches, we derivemodels on~~ the ~~high-quality 2010 E field and driving them with updated 2016–2020 environmental conditions, we derived the E parameter from 2010 relationships for the 2020 period~~. The resulting map provides ~~a criticalan~~ up-to-date dataset for quantifying decadal permafrost changes and supporting future high-precision modeling on the QTP.

2 Study area and data

2.1 Study area

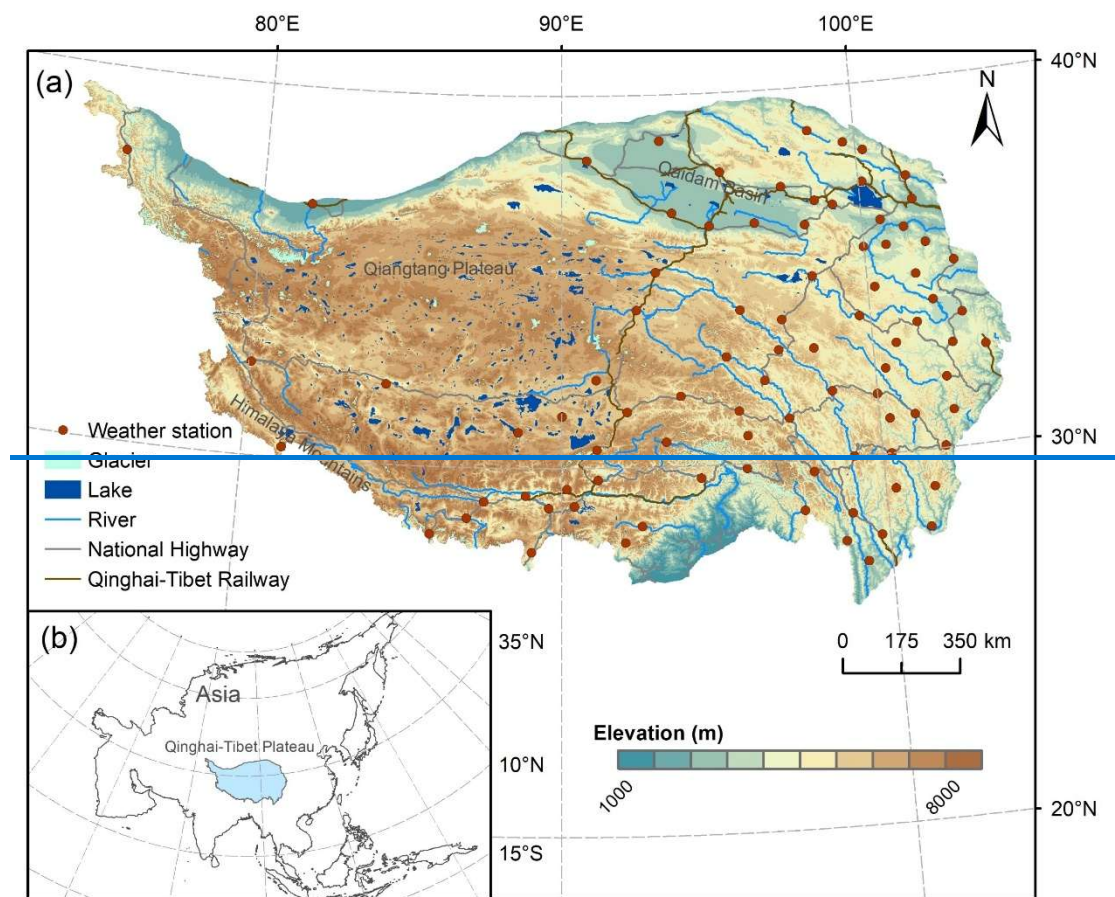
The QTP, widely ~~referred toknown~~ as the “Third Pole” and the “Asian Water Tower,” is the highest and most extensive plateau ~~in the world. Bounded withinon Earth. Located between~~ 26–40 °N and 73.5–104.5 °E, it spans an area of approximately ~~2.5 × 10⁶2.5 × 10⁶~~ km² with an average elevation exceeding 4000 m above sea level (a.s.l.) (Fig. 1). As the headwaters of major Asian river systems, including the Yangtze, Yellow River, Indus, Mekong, and Ganges, the plateau plays a ~~criticalkey~~ role in the regional hydrological cycle and ~~supportssustains~~ water resources for downstream populations.

The plateau is characterized by a distinctive high-altitude cryospheric environment, with extensive distribution of glaciers, snow cover, and permafrost (Qiu, 2008; Yao et al., 2012). The climate is governed by the interplay between the westerlies and the Asian monsoon, creating a pronounced climatic gradient. Mean annual air temperature generally ranges from -5°C and 5 °C (Zhang et al., 2019), while mean annual precipitation decreases from over 700 mm in the humid southeast to less than 50

100 mm in the arid northwest (Peng et al., 2019). Approximately 70% of annual precipitation occurs during the monsoon season from May to September (Kukulies et al., 2020; Zhu et al., 2020).

Vegetation cover transitions from alpine forests and meadows in the southeast to alpine steppes and deserts in the northwest (Zhou et al., 2025). ~~Correspondingly, permafrost~~Permafrost thickness varies ~~significantly~~accordingly, ranging from a few meters in marginal zones to approximately 350 m in the interior, with the depth of zero annual amplitude typically
105 varying between 3.5 to 17 m (Zhao et al., ~~2020~~2020). The QTP is experiencing rapid climate warming. Between 1960 and 2015, the annual mean temperature rose at a rate of 0.33 °C per decade, more than double the global warming rate (0.14 °C per decade) (Zhang et al., 2020). This accelerated warming has intensified cryospheric instability, underscoring the need for up-to-date monitoring of permafrost distribution.

~~Crucially, the QTP is experiencing rapid climate warming. Between 1960 and 2015, the annual mean temperature~~
110 ~~increased at a rate of 0.33 °C/10a, more than double the global warming rate (0.14 °C/10a) (Zhang et al., 2020). This accelerated warming has intensified cryospheric instability, necessitating up to date monitoring of permafrost distribution.~~



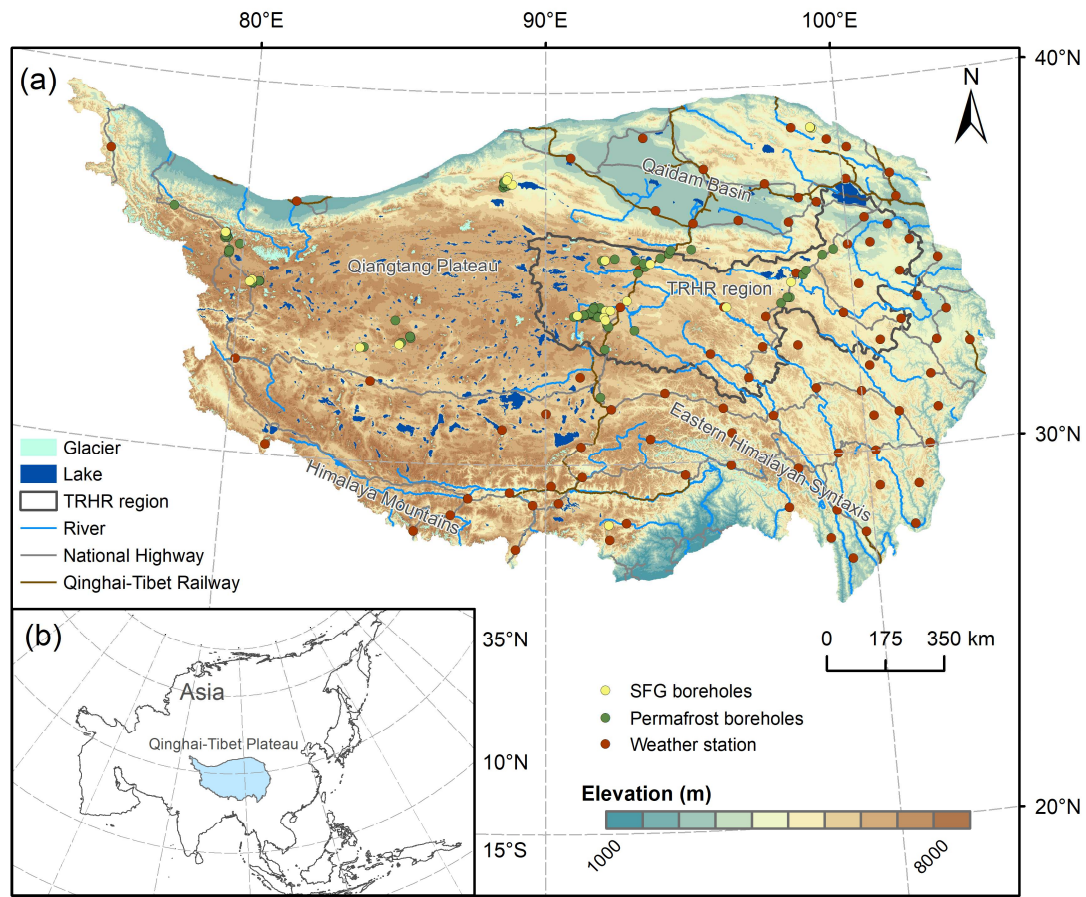


Figure 1 Topographic map and geographical context of the Qinghai-Tibet Plateau (QTP). (a) The topography of the region. The background color gradient represents elevation derived from the Shuttle Radar Topography Mission DEM. The 87 national meteorological stations (red dots) used for ground surface 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.

2.2 Remote sensing and gridded datasets

2.2.1 Thermal forcing data

The freezing and thawing indices are the primary drivers of permafrost thermal dynamics in our model are the freezing and thawing indices. To derive these indices, we utilized used the Moderate Resolution Imaging Spectroradiometer (MODIS) LST products as the basis for estimating ground surface temperature (GST). We acquired the Version 6 Terra (MOD11A1) and Aqua (MYD11A1) daily LST products, which provide four observations per day at a 1 km spatial resolution.

To determine the permafrost distribution for the ~~target year~~ 2020 ~~period~~, we processed data covering the five-year period from 2016 to 2020. This temporal selection satisfies the definition of permafrost (ground remaining frozen for at least two years) and minimizes the bias from interannual climate anomalies. ~~Since the~~Because raw satellite skin temperatures differ from the ground surface thermal regime required by the model, these data were post-processed. This included a dedicated gap-filling procedure to address cloud contamination (Chen et al., 2023) and a subsequent LST-to-GST correction to account for the thermal buffering effects of vegetation and snow. The specific algorithms for these transformations are detailed in Sect. 3.2.

2.2.2 Data for estimating local soil parameter (E)

While regional thermal dynamics are ~~driven~~governed by temperature indices, local permafrost occurrence is strongly modulated by surface and subsurface characteristics. In our model (Sect. 3.1), these local effects are aggregated into an empirical soil parameter E . To estimate E , we compiled environmental covariates representing topography, surface cover, soil properties, and moisture conditions (Karjalainen et al., 2019; Smith et al., 2022).

Topographic variables (slope, aspect, and topographic wetness index) (~~TWI~~) were ~~calculated~~derived from the Shuttle Radar Topography Mission (SRTM) 90 m Digital Elevation Model (DEM), ~~which was~~ aggregated to 1 km resolution to match the model grid. Surface cover conditions were characterized using the 1km, 16-day composite MODIS Normalized Difference Vegetation Index (NDVI) product (MOD13A2) and the daily 0.005° fractional snow cover (FSC) dataset over High Asia (Pan et al., 2024).

Subsurface thermal properties were represented by soil texture data (sand and clay fractions) extracted from the China Dataset of Soil Properties for Land Surface Modeling (Shangguan et al., 2013). ~~Due to the lack of~~Because high-resolution soil moisture products are not available, we ~~utilized~~used mean annual precipitation (MAP) ~~as a proxy for regional soil moisture conditions~~, derived from the 1 km monthly precipitation dataset for China (Peng et al., 2019), ~~as a proxy for regional soil moisture conditions~~. All dynamic covariates (NDVI, FSC, and MAP) were averaged over the 2016–2020 period.

2.3 In situ observations

2.3.1 Meteorological station data

To correct the thermal offset between satellite-derived LST and the actual ground thermal regime, we utilized daily 0 cm GST observations from the Daily Meteorological Dataset of Basic Meteorological Elements of China National Surface Weather Station (1991-2020) (Zhao et al., 2024). The original dataset includes 131 stations across the QTP. We screened these stations based on the completeness of their data series, excluding sites with significant missing records during the 2016–2020 period. This ~~selection process~~ resulted in a final subset of 87 stations utilized for this study (Fig. 1). Daily GST records from 2016 to 2020 were extracted to correspond with the satellite acquisition period.

2.3.2 Borehole observations

~~Independent validation of~~We validated the mapped permafrost distribution ~~was conducted independently~~ using a compilation of 109 borehole records. ~~Due to the limited availability of~~Because deep ground temperature measurements coinciding exactly with 2020 are limited, we prioritized high-quality observations from the years immediately surrounding 2020, ~~underon~~ the assumption that deep ground thermal states remain relatively stable over such short timeframes. The primary data source was the comprehensive QTP permafrost thermal state synthesis by Zhao et al. (2021). From this dataset, we extracted records for 65 boreholes where ground temperatures were monitored at depths of 10 m and 20 m. We utilized the 2018 ground temperatures as the reference. Sites with $MAGT < 0\text{ }^{\circ}\text{C}$ at these depths were classified as permafrost.

To ~~enhance~~improve spatial coverage, we integrated three additional regional datasets: (1) 32 boreholes from the Yangtze River Source area (Li et al., 2022) with direct presence/absence observations specifically for the year 2020; (2) 6 boreholes from the Yellow River Source area (Lei et al., 2024), classified using 2017 ground temperature profiles (0–20 m); and (3) 6 sites from the Heihe River Basin observation network (Mu et al., 2022) covering the northeastern QTP, all of which are situated in seasonally frozen ground (SFG) regions (based on records from 2011–2019). In total, the validation dataset comprises 83 confirmed permafrost sites and 26 seasonally frozen ground sites, providing a spatially diverse basis for accuracy assessment.

2.4 ~~Existing permafrost datasets~~Datasets for comparison and land cover consistency

To evaluate the accuracy of our new map, we ~~utilized~~compared it against two existing permafrost datasets that provide simulations covering the ~~year~~ 2020 period. The first is the TTOP-based permafrost map (Yan et al., 2023). This map is derived from a long-term permafrost simulation dataset (1961–2020) generated at a 5-year interval using the TTOP model (Smith and Riseborough, 1996) driven by satellite LST and meteorological station observations. For this study, we extracted the permafrost distribution ~~corresponding to~~for the ~~year~~ 2020, ~~which has a period, at its native 1 km~~ spatial resolution ~~of 1 km~~.

The second dataset is a MAGT-based permafrost map derived from a ground temperature simulation product (Zou et al., 2025). This dataset provides MAGT at a depth of 15 m for the period 2010–2019, generated using a Support Vector Machine model trained on records from 231 boreholes. For ~~our~~ comparison, grid cells with a 15 m ground temperature below $0\text{ }^{\circ}\text{C}$ were classified as permafrost.

~~Additionally, we utilized~~To keep land cover definitions consistent across the compared maps, we used glacier extent data from Ye et al. (2017) and 2020 lake boundaries from Zhang et al. (2021) to exclude non-soil areas ~~from the analysis to ensure consistent land cover definitions across all maps.~~

3 Methodology

3.1 The extended FROSTNUM/COP model framework

To map the 2020-period permafrost distribution for 2020, we applied the FROSTNUM model (Cao et al., 2023; Hu et al., 2020). This semi-physical approach, previously validated for the QTP by Cao et al. (2023), determines the probability of permafrost occurrence (frost number F) based on the balance between surface freezing and thawing loads, as modulated by local ground properties. The frost number F is defined as:

$$F = \frac{\sqrt{DDF}}{\sqrt{DDF + E \cdot \sqrt{DDT}}}, \quad (1)$$

where DDF and DDT are the ground surface freezing and thawing indices ($^{\circ}\text{C} \cdot \text{d}$), respectively. 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 $F \leq 0.5$ 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).

A critical component of this framework is the determination of E . In the original methodology established (Cao et al., 2023; Hu et al., 2020), the E parameter was retrieved via an inverse optimization strategy that minimized the error between the model output and extensive concurrent field survey maps. This ensured that the 2010 map accurately reflected the specific ground conditions of that year. However, a direct application of this original optimization strategy is not feasible for 2020. This process relies strictly on the availability of, since it depends on concurrent permafrost distribution survey maps for calibration, which do not and no such maps exist for the year 2020 period. To overcome this limitation, this study introduces a novel adaptation. We hypothesize we instead adopt a space-for-time substitution strategy, on the hypothesis that while although the specific value of E at a given location may shift due to with changes in moisture or surface cover, the its fundamental statistical relationship between E and its with environmental drivers (such as topography, vegetation, soil texture, soil moisture) remains applicable over the a decadal scale. Consequently, we employ a space-for-time substitution strategy. This allows us to predict the 2020 E field based on the from relationships learned from the high-quality 2010 dataset (Fig. 2-2). 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.

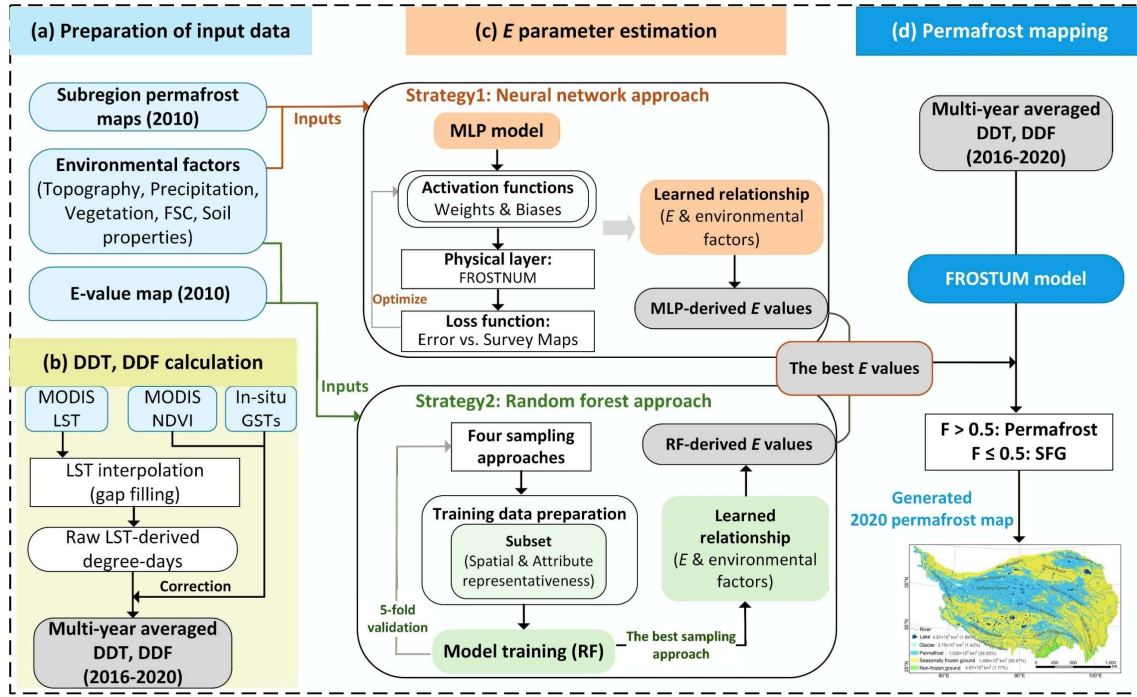


Figure 2 Schematic flowchart of the methodology to map the 2020-period permafrost distribution. The workflow comprises four distinct stages: (a) Preparation of multi-source input datasets, including satellite forcing data and environmental covariates; (b) Calculation of ground surface freezing (DDF) and thawing (DDT) indices from gap-filled MODIS land surface temperature (LST), corrected to represent ground surface conditions; (c) Estimation of the empirical soil parameter E via a space-for-time substitution strategy. This stage compares two transfer learning approaches: a neural network (MLP) approach and a Random Forest (RF) regression with optimized spatial sampling; and (d) Final permafrost distribution mapping using the extended ground surface frost number model (FROSTNUM). FSC: fractional snow cover; MODIS (Moderate Resolution Imaging Spectroradiometer); GST: ground surface temperature; NDVI: normalized difference vegetation index; SFG: seasonally frozen ground; and MLP: multilayer perceptron model.

3.2 Calculation of ground surface freezing and thawing indices

The driving variables for the model, the ground surface freezing and thawing indices (DDF and DDT), were derived from the daily MODIS LST products (MOD11A1 and MYD11A1) covering the 2016–2020 period. The processing workflow comprised gap-filling, daily aggregation, and the correction from LST to GST. Due to frequent cloud cover on the QTP, raw MODIS LST data suffer from extensive gaps. We addressed this using the solar-cloud-satellite geometry interpolation method developed by Chen et al. (2023). This approach reconstructs missing pixels by leveraging

the geometric relationship between the sun, cloud position, and satellite view angle, ~~ensuring~~ yielding high-accuracy all-weather LST estimates. Following gap-filling, the daily mean LST was calculated from the four daily instantaneous observations (two daytime and two nighttime) using a sinusoidal integration method (van Doninck et al., 2011).

~~Based on the gap-filled daily mean LST series, the~~ The raw surface freezing and thawing indices were ~~then~~ computed. ~~The freezing index (from the gap-filled daily mean LST series, DDF)~~ was calculated as the annual sum of negative degree-days. For ~~the thawing index (DDT)~~, a correction was ~~necessary~~ needed to account for the thermal offset between ~~the~~ satellite-observed LST and ~~the~~ actual GST, ~~caused~~ primarily ~~caused~~ by the insulating effect of vegetation during the growing season. Consistent with Cao et al. (2023), we corrected the LST-derived DDT using ~~a~~ multilinear regression model (Eq. (2)) that incorporates vegetation greenness and latitude. The correction is applied at 16-day intervals (corresponding to MODIS NDVI composites) to derive a final annual DDT_{GST} :

$$DDT_{i,GST} = f(DDT_{i,LST}, NDVI_i, Lat), \quad (2)$$

where $DDT_{i,GST}$ represents the corrected ground surface thawing degree-days ($^{\circ}C \cdot d$) for the i -th temporal interval (corresponding to the 16-day MODIS composites period); $DDT_{i,LST}$ is the raw thawing degree-days accumulated from positive daily mean LST values during the same i -th interval, $NDVI_i$ is corresponding NDVI value, serving as a proxy for vegetation density and its thermal insulating capacity; and Lat is the geographic latitude, which accounts for regional variations in solar radiation and incidence angle. ~~We did not apply~~ No correction ~~was applied~~ to ~~the freezing index (DDF)~~, as snow cover on the QTP is thin, ephemeral, spatially discontinuous, ~~which results~~ resulting in a negligible insulation effect at the regional scale (Wu and Zhang, 2008; Zhao et al., 2017).

3.3 Estimation of soil parameter E via space-for-time substitution

We ~~implemented~~ used a space-for-time substitution strategy to obtain E for ~~the~~ 2020 ~~without period in the absence of~~ concurrent field surveys. This ~~methodological paradigm approach~~ is ~~widely~~ well established in Earth system sciences, particularly ~~in~~ climatology, atmospheric modeling, and ecology, where spatial gradients are ~~frequently utilized~~ routinely used to infer temporal dynamics or ~~to~~ parameterize processes in data-scarce ~~scenarios~~ settings (Huang et al., 2019; Pickett, 1989). In the context of this study, ~~the approach relies on the hypothesis~~ we assume that while the specific value of E at a location may shift ~~due to~~ with changes in ~~soil~~ moisture or surface cover, the ~~fundamental~~ statistical dependencies between E and environmental covariates (soil texture, topography, and vegetation-moisture proxies) observed in 2010 remain valid for predicting the 2020 state. To ~~implement this~~ end, we developed and compared two distinct transfer learning strategies, each prioritizing a different source of training data. Both strategies utilized the same suite of predictors to estimate E , specifically: topographic variables (elevation, slope, aspect, TWI), soil texture (sand and clay fractions), vegetation conditions (NDVI), snow cover (FSC), and precipitation (MAP).

The first strategy ~~aimed to learn~~ learned the E -environment relationship directly from the high-reliability field survey maps available for five sub-regions in the 2010 study. We constructed a hybrid physics-informed neural network (Fig. 2c).

This architecture integrates a Multi-Layer Perceptron (MLP) (Fu et al., 2025; Park and Lek, 2016) with the physical FROSTNUM model in an end-to-end differentiable framework. In the forward pass, the MLP predicts the local parameter E from environmental inputs, which is immediately passed through the differentiable FROSTNUM layer to calculate the frost number and subsequent binary classification. In the backward pass, the generated classification is compared directly against the survey maps, and the classification error is backpropagated through the physical layer to update the MLP weights.

The second strategy prioritized spatial continuity and statistical robustness ~~by utilizing, using~~ the full 2010 E -field, retrieved via the original Clustering-Optimization-Prediction procedure (Cao et al., 2023), ~~as the training target~~. Unlike the first strategy, which relied on sparse survey data, this approach treated the estimation ~~of E~~ as a direct regression problem ~~using against~~ the spatially continuous 2010 baseline map. ~~While the original study aggregated the plateau into eight parameter clusters, we expanded this to sixteen classes for the training target. This refinement was necessary because restricting the target variable to only eight discrete values lose the continuous nature of the variable, thus limiting the RF's ability to learn subtle environmental gradients.~~ We trained a RF model (van Der Westhuizen et al., 2023; Wadoux et al., 2019) to learn the relationships between the environmental covariates and these "observed" E values. ~~While the original study aggregated the plateau into eight parameter clusters, we expanded this to sixteen classes for the training target.~~

~~Crucially, for~~For the second strategy, we ~~also~~ implemented a representative sampling framework to select the training data. Given the high spatial heterogeneity of the QTP, simple random sampling often fails to capture rare but physically significant parameter combinations, ~~leading to high~~which increases predictive uncertainty. We evaluated four distinct sampling techniques: k -means clustering sampling (Brus et al., 2006), Equal Range (ER) sampling (Hengl et al., 2003), Principal Component Analysis (PCA)-based sampling (Hengl et al., 2003), and Latin Hypercube Sampling (LHS) (Carré et al., 2007) across four different sample sizes (1000, 3000, 5000, and 10000 points). ~~We aimed) using 5-fold cross-validation. Our aim was~~ not only to improve computational efficiency and reduce information redundancy, but, more importantly, to minimize the uncertainty ~~associated within~~ the E -value distribution. Detailed mathematical formulations and implementation procedures for these four sampling strategies are provided in Appendix A. The ~~optimal~~-configuration, ~~determined via 5-fold cross-validation, and with~~ the ~~superior strategy~~best performance was then selected ~~from this comparison was used to drive~~generate the final FROSTNUM ~~simulation~~ E map for the 2020 ~~permafrost map~~-period. Details of the accuracy comparison among the different sampling strategies are provided in Section 4.2.

3.4 Validation ~~and uncertainty analysis~~

~~The~~We evaluated the accuracy and reliability of the ~~generated~~resulting 2020-~~period~~ permafrost map ~~were evaluated~~ through a multi-tiered validation framework ~~designed to assess~~that assessed both the intermediate modeling steps and the final spatial product. First, ~~to determine the optimal transfer learning strategy for estimating the soil parameter E ,~~we conducted a ~~rigorous~~comparative evaluation of the MLP and the RF strategies described in Sect. 3.3.3 ~~to determine the optimal transfer learning strategy for estimating E .~~ For the RF approach, we ~~utilized~~used 5-fold cross-validation to assess its ability to reproduce the 2010 baseline values, quantifying performance using the ~~Correlation Coefficient~~correlation coefficient (r), ~~Mean Absolute~~

~~Error~~mean absolute error (MAE), ~~Root-Mean-Square-Error~~root mean square error (RMSE), and percentage of variance explained. For the MLP approach, the MLP performance was validated indirectly by assessing the classification accuracy of the simulated permafrost distribution in the target subregions. Specifically, we randomly selected 70% of the pixels within these subregions for training, while the remaining 30% were reserved for validation. We also assessed the spatial consistency of the MLP-derived 2010 permafrost distribution ~~for 2010~~ by comparing it with the benchmark permafrost-map from Cao et al. (2023).

Second, we ~~performed~~carried out a ~~dual-facet~~two-part validation~~-focusing-on, assessing~~ both the accuracy of the 2020 map and the plausibility of the detected temporal changes. The spatial accuracy of the resulting 2020 permafrost distribution (generated using the selected optimal strategy), was ~~rigorously~~ assessed against the independent dataset of 109 borehole records. To perform this evaluation, we extracted the simulated permafrost status (permafrost or SFG) at the coordinates of each borehole and compared it with the in-situ ground truth. Beyond this static accuracy, we also ~~verified~~checked the ~~reasonableness~~plausibility of the simulated 2010-2020 permafrost changes~~-between 2010 and 2020~~ using 0 cm GST data from meteorological stations and available field evidence.

~~Additionally, to assess~~We also assessed the ~~map's~~map's consistency with current regional knowledge,~~-we performed~~ through a spatial inter-comparison with two existing datasets covering the same period: the TTOP-based simulation by Yan et al. (2023) and the data-driven ground temperature map by Zou et al. (2025). We ~~analyzed the~~quantified agreement in permafrost area and spatial distribution patterns, ~~specifically quantifying~~focusing on discrepancies in the zones where modeling uncertainty is typically highest. Finally, to ~~further~~diagnose the physical drivers underlying these shifts, we ~~conducted a comparative analysis of~~compared key environmental and thermal variables ~~by comparing their values in the~~between changed areas ~~versus~~and adjacent unchanged regions.

Third, we conducted a systematic uncertainty analysis for both the soil parameter E and the final permafrost map, to quantify the uncertainty arising from the choice of training sample selection. A 16-member ensemble was constructed by cross-combining four sampling strategies and four sample sizes. The stability of the predicted E values was evaluated using the ensemble mean, standard deviation, and coefficient of variance (CV). The propagation of this uncertainty into the final classification was assessed by analyzing variations in simulated permafrost areas and by mapping the permafrost agreement probability (P_{agree}), which delineates stable zones from zones of classification inconsistency. The results of this uncertainty analysis are presented in Appendix C and Section 4.4.1.

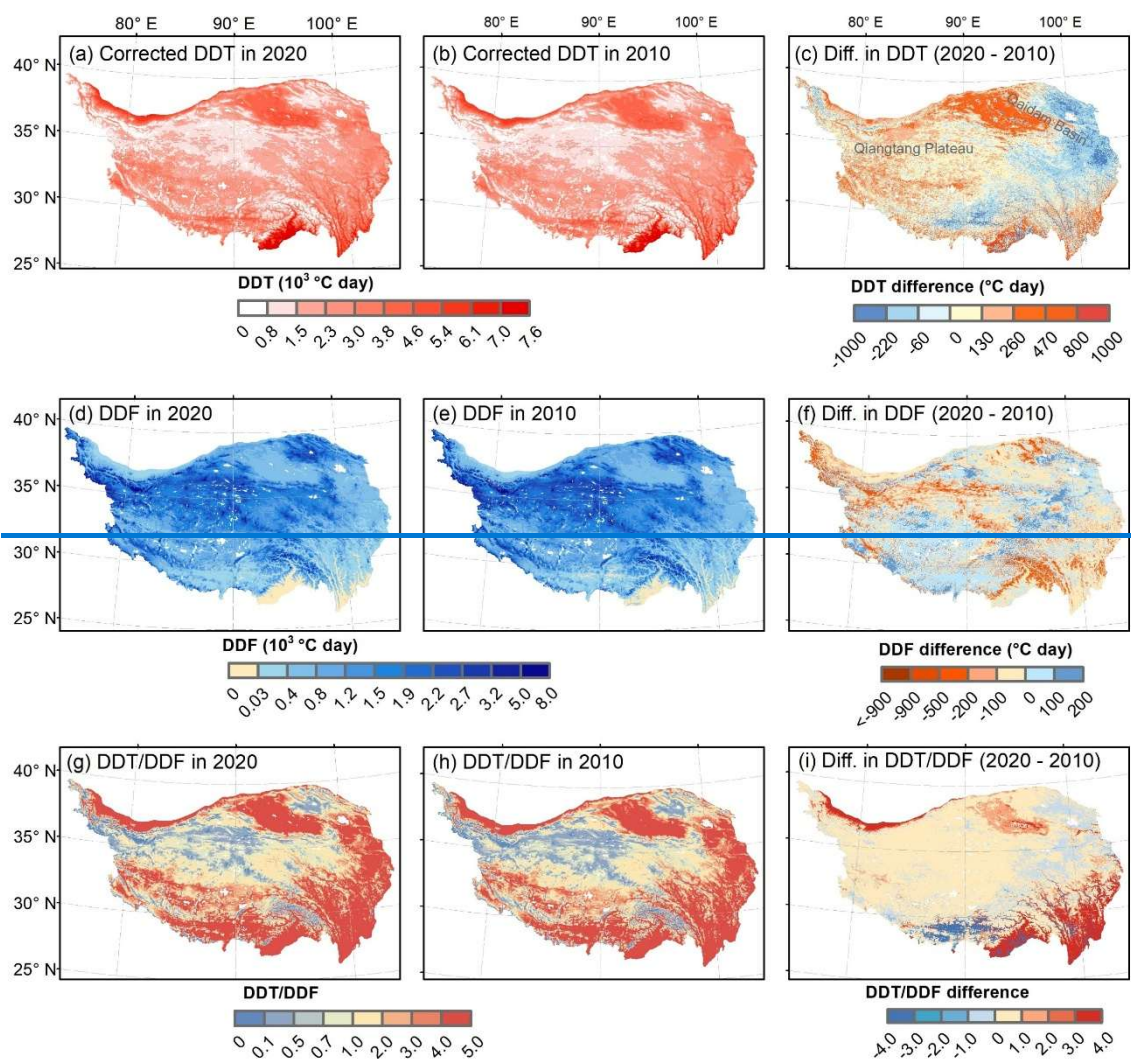
4 Results

4.1 Spatiotemporal changes in thermal forcing

The thermal forcing indices driving the permafrost model ~~exhibited~~showed distinct spatial shifts between the 2010 baseline and the 2020 target-period. The ~~ground-surface-thawing-index (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 ~~sharp~~-contrast, a ~~notable~~[slight](#) cooling trend was observed ~~in across parts of~~ the eastern QTP, ~~particularly within~~. 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). ~~Collectively, these processes appear to have dampened, or even reversed, the ground thermal response to rising air temperatures in the eastern plateau.~~ 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.

~~In parallel, the freezing index (The *DDF*)~~ reflected these regional climatic contrasts (Fig. 3 d-f). Between 2010 and 2020 ~~periods~~, pronounced winter warming (manifested as reduced *DDF* in Fig. 3f) occurred across the central QTP and was particularly intense along the southern Himalayan margin. Conversely, localized cooling (increased *DDF*) was observed in the western QTP and the eastern mountainous regions (Fig. 3f). As a result, the *DDT*-to-*DDF* ratio (*DDT/DDF*), a key climatic indicator of permafrost stability, increased markedly across the majority of the QTP from ~~the~~ 2010 to ~~the~~ 2020 ~~period~~ (Fig. 3g-i). This positive shift signals a climate trajectory unfavorable for permafrost preservation in these regions. However, the eastern region showed a decrease in the *DDT/DDF* ratio, consistent with the observed *DDT* and *DDF* cooling trend.



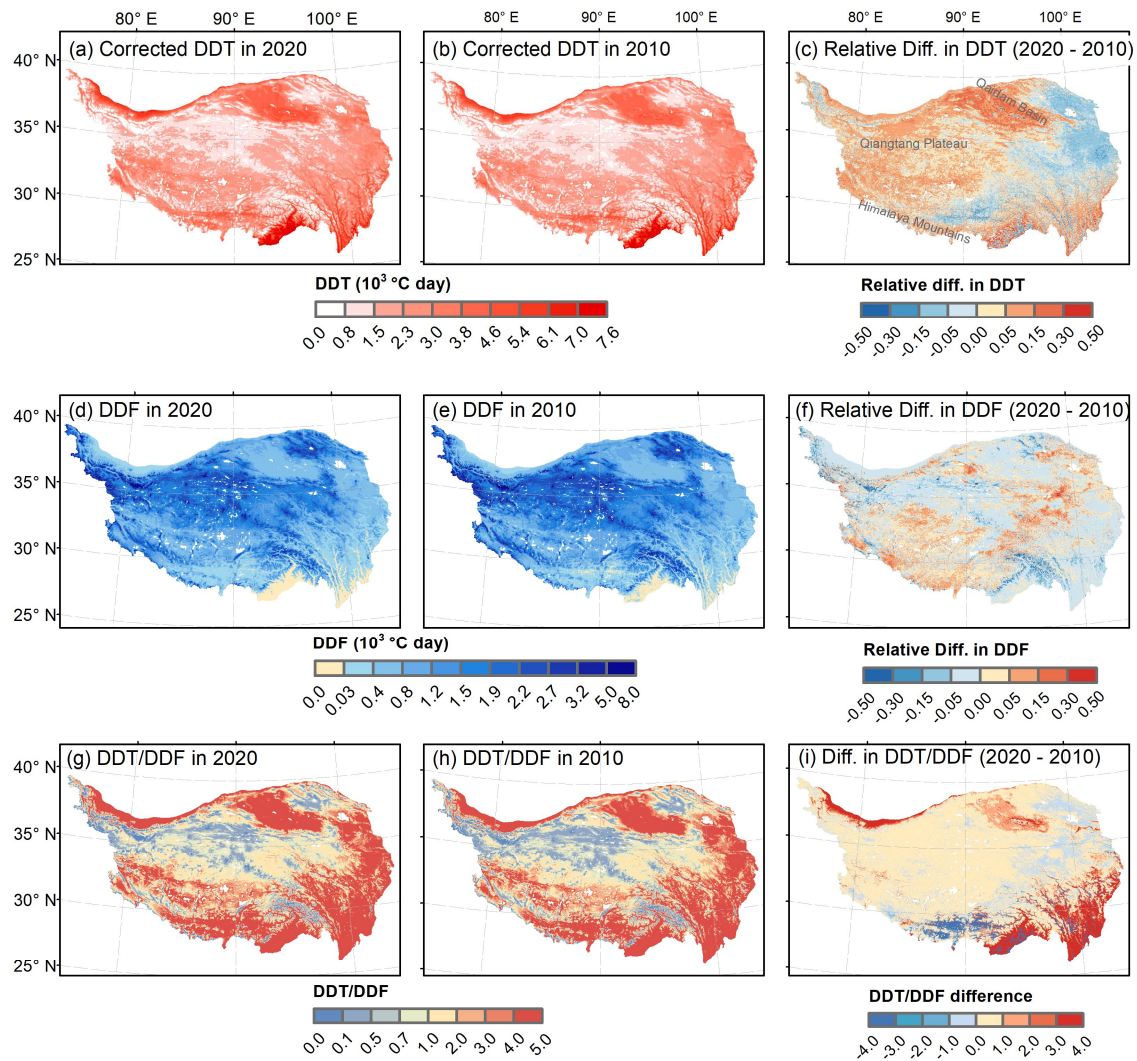


Figure 3 Spatial distribution and decadal changes of thermal forcing indices on the QTP. (a-e) displays the ground surface thawing index (and decadal change of (a-c) DDT), and its relative changes, (d-f) the freezing index, DDF and its relative changes. (g-i) the ratio of thawing to freezing indices. Left column shows the distributions in (DDT/DDF) and its decadal change. The left and middle columns present the spatial distribution for the 2020, middle column shows the distributions in period and 2010, and the period, respectively. The right column shows their decadal differences between 2020 and 2010. The relative differences in (c) and (f) are calculated using baseline denominators of $2000^{\circ}\text{C}\cdot\text{d}$ and $1200^{\circ}\text{C}\cdot\text{d}$, 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.

4.2 Estimation of the soil parameter E for the 2020 period

To ~~spatially extrapolate~~ estimate the empirical soil parameter E ~~to~~for the 2020 period, we ~~evaluated the generalization capabilities of two distinct~~ applied a space-for-time substitution strategy using two transfer learning strategies: the MLP approach and the RF approach. The comparison revealed a critical trade-off between model complexity and predictive stability. The MLP model achieved a high prediction approaches trained on the high-quality 2010 E field from Cao et al. (2023). While the MLP model achieved high accuracy (95.26%) within the training subregions. However, it overfit severely when applied to across the entire plateau-wide domain, the model exhibited severe overfitting, likely due to the equifinality inherent in training against binary classification data without sufficient physical constraints. The MLP-predicted E field was characterized by showed 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 and failed to reproduce the benchmark 2010 permafrost distribution for 2010, the resulting map deviated significantly from the rigorously validated benchmark established by of Cao et al. (2023) (see Appendix B for details). This failure likely stems from the “black-box” nature of deep learning, which, in this context, lacked effective physical constraints and consequently suffered from poor generalization capability.

In contrast, the RF-based strategy demonstrated superior robustness and suggests that unconstrained neural networks may lack sufficient physical constraints for robust spatial coherence extrapolation in heterogeneous environments such as the QTP.

The RF-based approach, in contrast, generalized much strongly. Among the four sampling strategies evaluated through 5-fold cross-validation, the k -means clustering sampling approach (utilizing method with 5000 points) produced the optimal performance performed best, achieving the lowest an RMSE (of 0.048), MAE of 0.029, and highest correlation coefficient ($r=0.862$) against explaining 74.83% of the variance in the 2010 baseline E field (Table 1). Visual inspection confirmed that the RF-derived field successfully reproduced the complex heterogeneity of soil properties across the QTP (Fig. This configuration was therefore 4a). Therefore, the RF-derived predictions were adopted to generate the E map for the final permafrost mapping. 2020 period.

Based on the RF simulations, Figure 4 presents the spatial distribution of E in 2020 exhibited significant shifts relative to 2010. We categorized areas into a “Positive group” (the 2020 period and the reference 2010. The overall spatial pattern remains broadly consistent between the two periods, although regional differences are evident. Areas with increased E , indicating higher degradation risk) and a “Negative group” (decreased E , indicating reduced risk). The analysis shows that the area of the Positive group (positive group) account for 55% of the total changed area, while areas with decreased E (negative group) account for 45%. Regions with rising E are located mainly in the central and northern plateau, whereas declining E values are more scattered, especially along the eastern and southern margins (Fig. 4c). exceeded that

To further test the robustness of the Negative group (45%), predicted E values, we performed an ensemble analysis using 16 combinations of sampling strategies and sample sizes (Appendix C). The results indicate a high level of consistency. The CV of E remains below 5% across 98.5% of the QTP, suggesting a dominance of processes conducive to permafrost

degradation (Fig. that the spatial pattern of E in the 2020 period is relatively stable with respect to the choice of training sample selection, 4e).

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

400

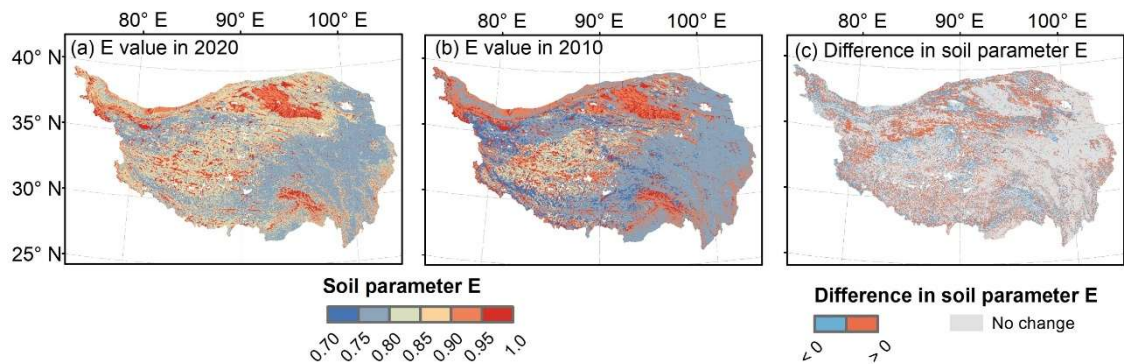


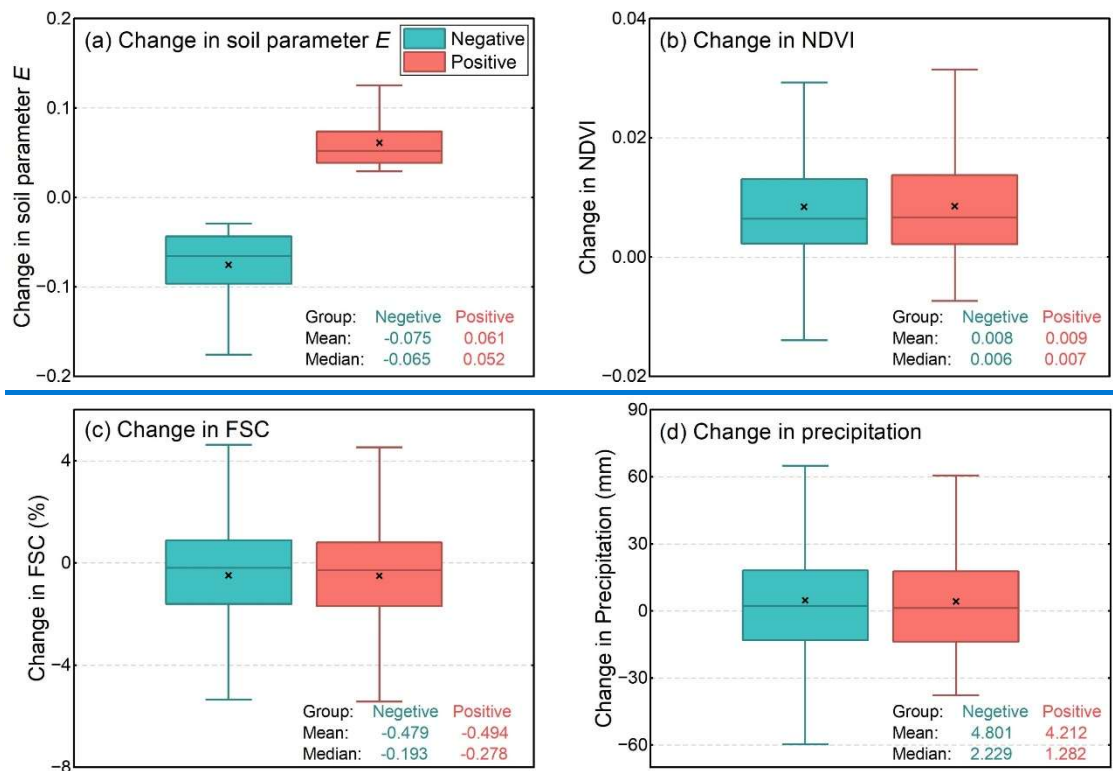
Figure 4 Spatial distribution of the soil parameter E in the 2020 period (a) and 2010 (b). (c) illustrates the difference (2020 minus 2010), where positive values (red) indicate an increase in E (conditions conducive to degradation) and negative values (blue) indicate a decrease. The 2010 E map was derived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

405

To elucidate the physical mechanisms driving these shifts in E , we analyzed the temporal variations of the environmental predictors (Fig. 5). While vegetation indices (NDVI) (Fig. 5b) and snow cover (FSC) (Fig. 5c) showed uniform increases across both groups, precipitation emerged as the distinct driver of divergence (Fig. 5d). It is important to note that due to the lack of high-resolution soil moisture products for the region, mean annual precipitation is utilized here as a proxy for regional soil moisture conditions. Although the QTP experienced a general wetting trend (Zhang et al., 2021), the regions characterized by increasing E (positive group) received significantly less precipitation increase (+1.28 mm median increase) compared to the Negative group (+2.23 mm median increase). This relative moisture deficit has profound thermal implications. Soil moisture acts as a crucial thermal buffer due to the high specific heat capacity and latent heat of fusion of water. A smaller increase in precipitation implies relatively drier soil conditions in the positive group, resulting in reduced soil thermal inertia.

410

415 Consequently, these soils are more responsive to atmospheric warming, leading to a higher E value and an accelerated thawing process.



420 **Figure 5** 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 risk (positive group) and reduced risk (negative group). In each box plot, the center 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.

4.3 The 2020-period permafrost distribution map

425 The final 2020-period permafrost distribution map for 2020, generated using the optimized RF-derived E parameter and corrected thermal forcing indices, is presented in Fig. 65. Statistical analysis shows that permafrost covers approximately $1.038 \times 10^6 \text{ km}^2$, accounting for 39.35% of the total QTP area (excluding lakes and glaciers). SFG-region is the dominant frozen ground type, occupying $1.466 \times 10^6 \text{ km}^2$ (55.57%), while non-frozen ground is restrictedconfined to the lowest marginal valleys (1.77%, $4.67 \times 10^4 \text{ km}^2$), 1.77%).

430 The spatial pattern of the 2020-period map exhibits a strong altitudinal and latitudinal zonality, consistent with the region's topoclimatic gradients. The most extensive continuous permafrost body is concentrated in the high-elevation interior,

encompassing the Qiangtang Plateau, the Hoh Xil region, and the Bayan Har Mountains. The combination of extreme elevation (>4500 m) and ~~the~~ arid, continental climate maintains a stable, negative ground thermal regime. This continuous zone ~~aets~~ ~~asforms~~ the cold core of the plateau, interrupted only by large endorheic lakes, which form localized taliks. To the north, the Qaidam Basin creates a distinct break in the permafrost continuity. Despite its high latitude, the basin's ~~significantly much~~ lower elevation and distinctive arid-desert thermal regime preclude permafrost formation, creating a vast expanse of SFG that physically isolates the permafrost of the Qilian Mountains from the main plateau body. ~~The Qilian permafrost thus exists as a large-scale "island" system, maintained strictly by the orographic uplift of the mountain range.~~

Moving southward and eastward from the Qiangtang core, ~~the~~ permafrost ~~distribution~~ becomes increasingly fragmented. In the eastern QTP, particularly in the Three-River Headwater Region, the landscape is dominated by SFG, with permafrost confined to high ridges. This ~~distribution pattern~~ is likely ~~modulated shaped~~ by the advective heat transport from the developed river networks and ~~by~~ the relatively lower latitude, rendering ~~the~~ permafrost in this region thermally fragile (Zhang et al., 2022). Similarly, along the southern and southeastern margins (the Himalayas, the Gangdise Mountains, and the Hengduan Mountains), ~~the~~ permafrost distribution is strictly topographically controlled. ~~In these regions Here~~, the combination of lower latitude and deeply dissected terrain creates sharp thermal gradients. Permafrost ~~is occurs~~ discontinuously ~~distributed~~, clinging to ~~the~~ high-altitude ridgelines and islands of extreme elevation, while the deep, incised valleys ~~maintain remain~~ non-frozen ground or SFG.

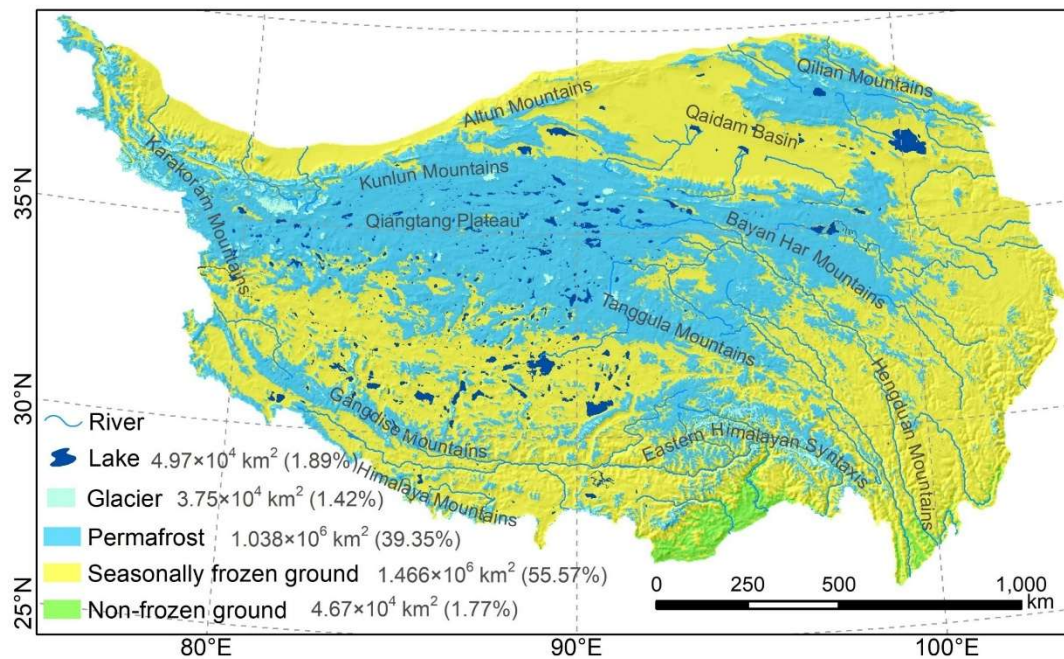


Figure 65 Map of frozen ground distribution at 1 km resolution over the QTP in the 2020 period. Areas and percentages of frozen soil types are present. ~~Glacier extent data are derived from Ye et al. (2017), and the 2020 lake boundaries are obtained from Zhang et al. (2021).~~

4.4 Accuracy assessment and inter-comparison

4.4.1 Validation against borehole observations and uncertainty analysis

Quantitative validation was performed using 109 independent borehole records from the 2020 period (Table 2). Our map achieved an overall accuracy of 84.4% and a Cohen's Kappa coefficient of 0.58, outperforming both the MAGT-based map (81.7%, $\kappa = 0.38$) and the TTOP-based map (78.9%, $\kappa = 0.28$).

A critical advantage of our approach lies in its balanced classification capability. While both reference maps exhibited high sensitivity, identifying over 90% of confirmed permafrost sites, they suffered from significant commission errors, frequently misclassifying SFG as permafrost. In contrast, our map demonstrated a superior ability to differentiate SFG far more effectively, achieving a true negative rate of 69.2%, compared to just 34.6% for the MAGT-based map and 42.3% for the TTOP-based map. This substantial improvement indicates that our method effectively mitigates the overestimation of permafrost extent that is prevalent in existing simulations, common in existing simulations. Furthermore, the multi-member ensemble simulations demonstrates that the uncertainty in the predicted soil parameter E has a limited impact on the final permafrost classification (detailed in Appendix C). The 16 ensemble-derived permafrost maps show high spatial and areal consistency, with the estimated total permafrost area varying within a narrow range, from $1.036 \times 10^6 \text{ km}^2$ to $1.043 \times 10^6 \text{ km}^2$. This robustness is further substantiated by the permafrost agreement probability, where highly stable pixels ($P_{agree} \geq 0.8$) account for 98.85% of the entire plateau, leaving 1.15% of the region within the zone of classification inconsistency (where the 16 ensemble members disagree).

470 Table 2 Confusion matrix describing the performance of the evaluated permafrost maps (Our map, MAGT-based map (Zou et al., 2025) and TTOP-based map) (Yan et al., 2023)) against 109 borehole validation sites. MAGT: mean annual ground temperature; TTOP: temperature at the top of permafrost.

	Our map	MAGT-based map	TTOP-based map
Overall Accuracy	84.4%	81.7%	78.9%
Cohen's Kappa	0.58	0.38	0.28
True positive rate (permafrost)	74 (89.2%)	80 (96.4%)	75 (90.4%)
True negative rate (SFG)	18 (69.2%)	9 (34.6%)	11 (42.3%)
False positive rate	8 (30.8%)	17 (65.4%)	15 (57.7%)
False negative rate	9 (11.8%)	3 (3.6%)	8 (9.6%)

4.4.2 Spatial inter-comparison

475 To further ~~verify~~test the map's physical realism, we analyzed spatial inconsistencies between our results and the reference maps in ~~distinct~~several subregions (Fig. 76), specifically focusing on transition zones where modelling uncertainty is typically high. ~~Significant~~Notable discrepancies ~~were observed~~emerged in these areas, where ground observations consistently ~~supported~~avored our simulation over the reference datasets.

In the Gaize area (subregion aA in Fig. 7a6a, b) both reference maps failed to identify the seasonally frozen status of 480 borehole ZK043 (Fig. 8a7a, Fig. 9a8a), whereas our map correctly classified it as SFG. Additionally, our map correctly identified borehole ZK042 as SFG, a site also situated in this transition zone. In the northern marginal zone between the Altun and Kunlun Mountains (subregion bB in Fig. 7a6a), ground observations indicate the presence of four SFG boreholes (Fig. 8b7b). Our map correctly recognized all four, whereas the MAGT-based map identified only one (Fig. 8b7b). Although our simulation classified two confirmed permafrost boreholes in this region as SFG (Fig. 8b7b), observational records indicate 485 these sites possessed extremely warm MAGT of -0.13 °C in 2013-2016 and -0.17 °C in 2013-2018. Such values characterize are characteristic of thermally unstable permafrost that is highly susceptible to degradation, suggesting that our map likely reflectscaptures a recent transition to SFG that predates the borehole's historical baseline ~~of the borehole~~predates.

Further discrepancies were ~~identified~~found in the headwaters of the Yangtze River (subregion eC in Fig. 7a6a, b). Here, our map correctly identified four out of five confirmed SFG boreholes, whereas the MAGT-based map identified only one 490 (Fig. 8e7c) and the TTOP-based map identified three (Fig. 9b8b). While our simulation misclassified three confirmed permafrost boreholes in this region as SFG (Fig. 8e7c), it is important to note that these sites are situated within a permafrost transition zone where MAGTs range from -0.37 °C to -0.48 °C. These close to zero values suggest that the classification discrepancy may reflect real-world degradation or the inherent uncertainty in modeling such marginal thermal states. Crucially, meteorological station 56004 in this area was classified as permafrost by both reference maps but as SFG by ours (Fig. 8e7c, 495 Fig. 9b8b). Analysis of in-situ GST (0cm) data (2016-2019) at this site revealed a thawing index ($DDT \approx 2004 \text{ }^{\circ}\text{C} \cdot \text{d}$)

significantly exceeding the freezing index ($DDF \approx 1332 \text{ }^{\circ}\text{C} \cdot \text{d}$). ~~Since~~~~Because~~ a thawing index ~~exceedingthat exceeds~~ the freezing index is robust indicator of SFG, this observation provides strong evidence that our simulation offers a more reliable representation of the current thermal regime.

In the headwaters of the Yellow River (subregion ~~dD~~ in Fig. 7a6a), the SFG borehole XXH-1 was correctly classified by our map but was misclassified as permafrost by the MAGT-based map (Fig. 8d). ~~Although notable~~7d). Notable spatial inconsistencies were also observed along the southern margin of the plateau (subregion ~~eE~~ in Fig. 7b),6b), but this area was excluded from the detailed inter-comparison due to the lack of available borehole observations for validation.

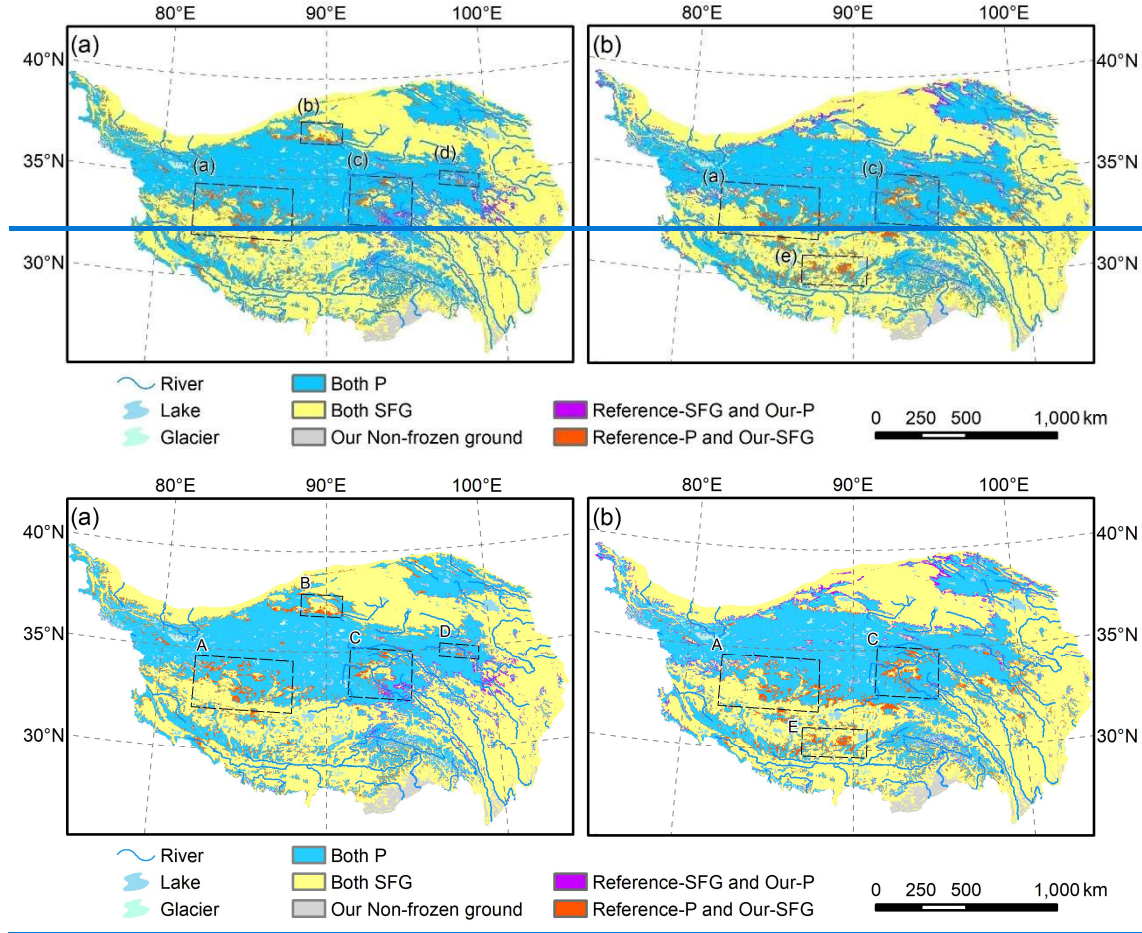


Figure 76 Spatial inconsistencies in the distribution of frozen ground between the generated 2020-period permafrost map and reference datasets: (a) the Mean Annual Ground Temperature (MAGT)-based map (Zou et al., 2025) and (b) the Temperature at the Top of Permafrost (TTOP)-based map (Yan et al., 2023). “Both P” (blue) and “Both SFG” (yellow) denote consistent classification. “Reference-P and Our-SFG” represents regions classified as permafrost by the reference map but as SFG in our map. “Reference-SFG and Our-P” indicate the opposite. The dashed boxes mark specific transition zones with obvious

510 inconsistency: (aA) Gaize and its vicinity, (bB) the Altun–Kunlun Mountains transition zone, (cC) the headwaters of the Yangtze River, (dD) the headwaters of the Yellow Rive, and (eE) the southern margin.

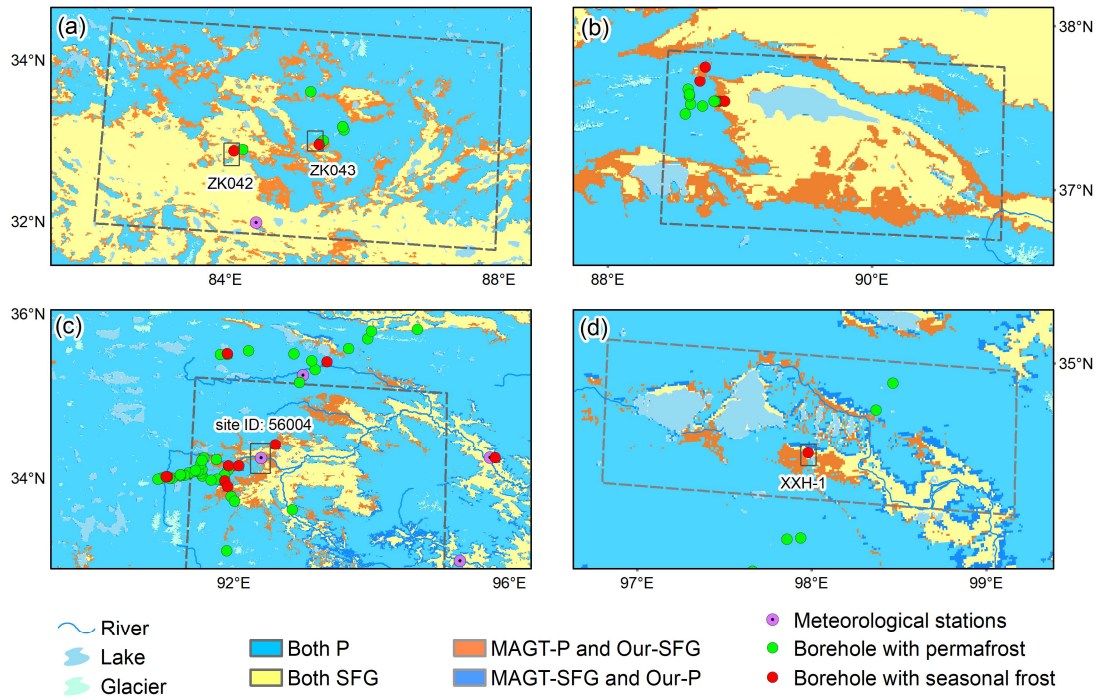
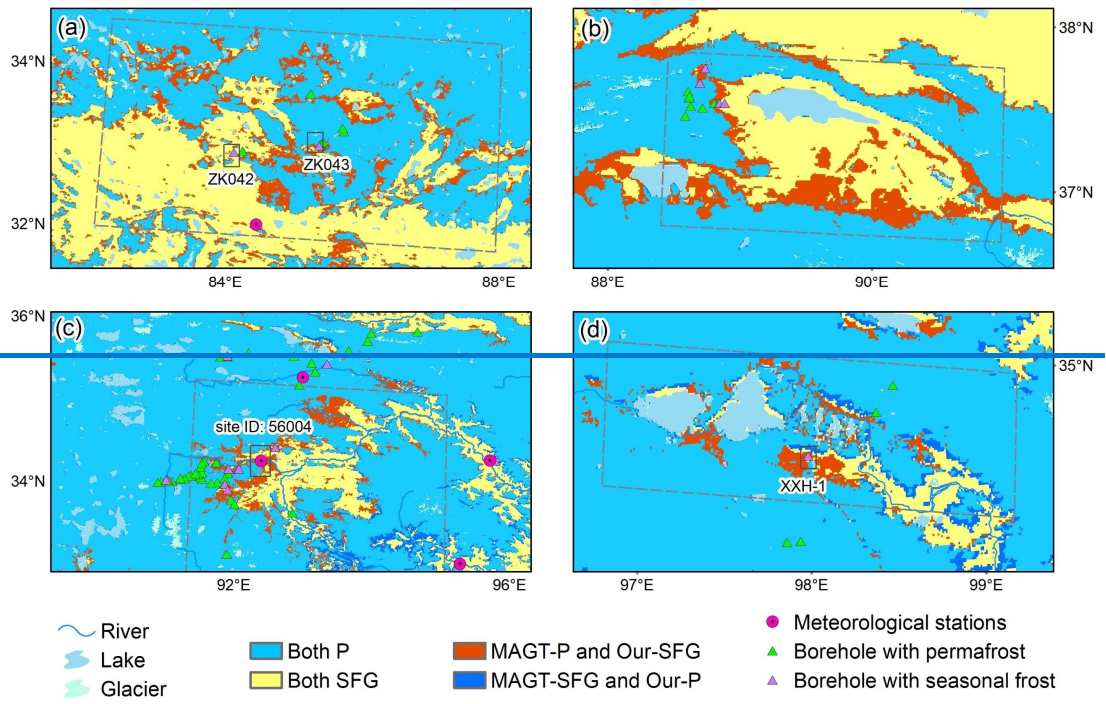


Figure 87 Detailed validation of spatial inconsistencies between the generated 2020 map and the MAGT-based map using ground-truth observations. The panels (a)-(d) correspond to the four subregions outlined in Fig. 7a. “MAGT P and Our SFG” represents regions classified as permafrost by the MAGT-based map but as SFG by our map. “MAGT SFG and Our P” indicates the opposite(A-D) outlined in Fig. 6a.

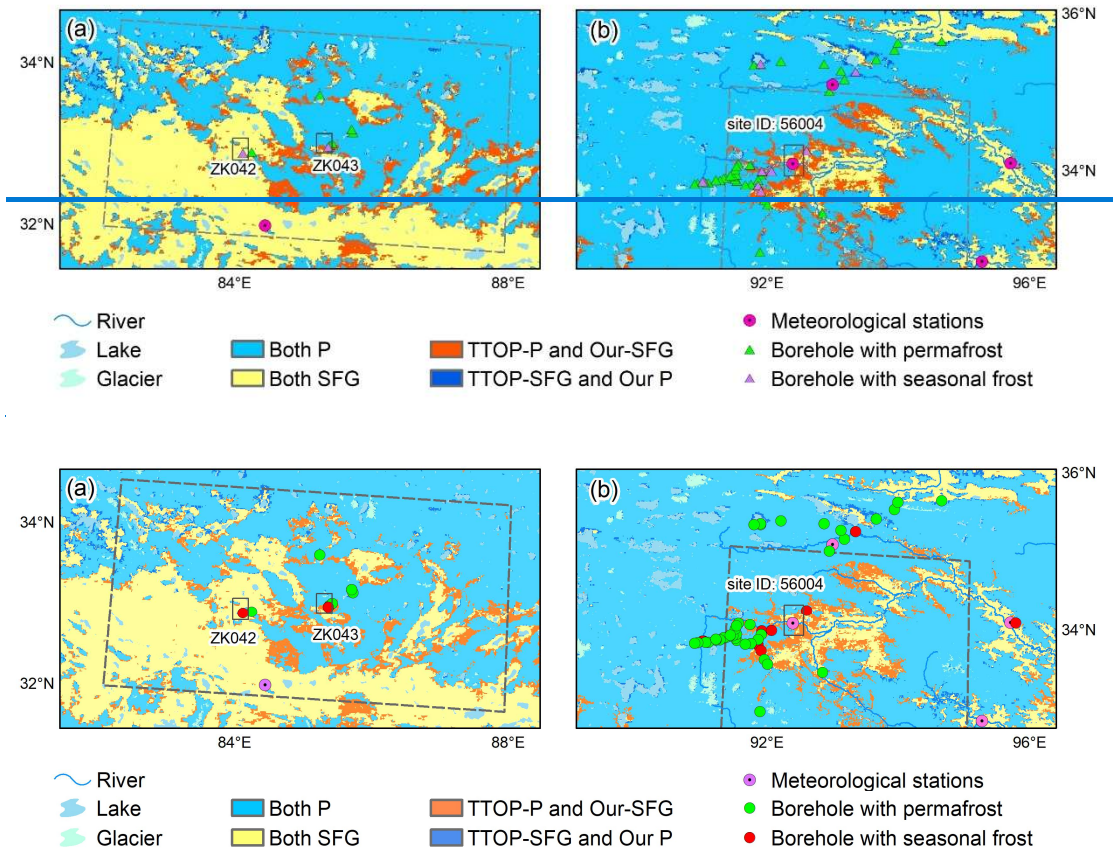


Figure 98 Detailed validation of spatial inconsistencies between the generated 2020 map and the TTOP-based map in two key transition zones. The panels (a) and (b) correspond to the subregions (aA and eC) outlined in Fig. 7b. “TTOP P and Our SFG” indicates areas identified as permafrost by the TTOP-based map but as SFG by our map. “TTOP SFG and Our P” indicates the opposite6b.

4.5 Decadal changes in frozen ground distribution (2010–2020 period)

Between the 2010 and 2020, the periods, frozen ground distribution on the QTP underwent a distinct phase of degradation (Fig. 109), characterized by a net loss of permafrost extent and a concurrent expansion of SFG and non-frozen ground (Fig. 10a9a). Spatially, these shifts were not uniform but exhibited a clear bifurcation between the plateau’s interior and its margins.

Quantitative analysis indicates that the total permafrost area decreased by approximately $4.8 \times 10^4 \text{ km}^2$ over the decade, representing a decline of 1.82% relative to the total area (Table 3). This loss was primarily driven by the conversion of unstable permafrost into SFG. Consequently, the SFG area expanded by $1.9 \times 10^4 \text{ km}^2$ (a 0.72% increase). More strikingly, ~~the area of~~ non-frozen ground ~~area~~ increased by $2.43 \times 10^4 \text{ km}^2$ (a 0.92% increase), reflecting a ~~significant rangemarked~~ contraction of the frozen ground domain at its lower elevational and latitudinal limits.

~~The spatial evolution of these~~These transitions ~~reveals~~reveal two dominant degradation patterns: ~~of degradation~~. In the central QTP region (subregion ~~aA~~ in Fig. ~~10b~~), ~~continuous or discontinuous alpine~~9b), ~~degradation of~~ permafrost ~~degraded~~ into SFG ~~occurred extensively at the permafrost-SFG transition zones~~. This process accounted for 7.41% of the total changed area and was largely concentrated in the transitional zones of the Qiangtang Plateau. ~~Conversely~~In contrast, the southern and southeastern margins (subregion ~~bB~~ in Fig. ~~10b~~9b), particularly the flanks of the Hengduan, Himalayan, and Gangdise mountains, experienced marginal retreat. The dominant transition was the conversion of SFG to non-frozen ground, which accounted for 39.62% of the total changed area (Fig. ~~10b~~9b). This shift indicates a vertical ascent of the frozen ground lower limit in response to warming.

Beyond the ground thermal state ~~itself~~, surface hydrological and glaciological features also responded to the warming climate. During this period, the glacier area decreased by $0.33 \times 10^4 \text{ km}^2$ (Ye et al., 2017), while the lake area expanded by $0.80 \times 10^4 \text{ km}^2$ (Zhang et al., 2021) (Table 3), consistent with the broader warm-wet climate trend observed across the region.

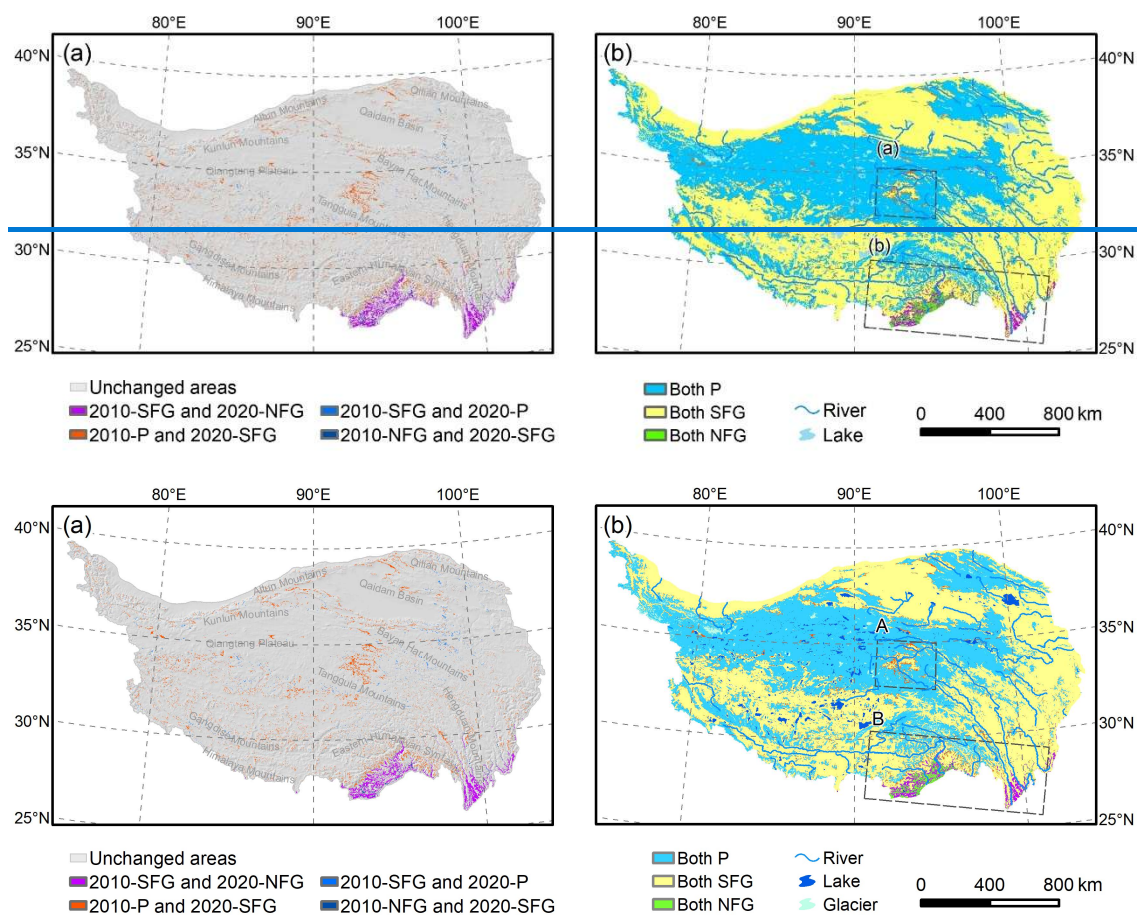


Figure 409 Spatiotemporal evolution of frozen ground types on the QTP from the 2010 to 2020 period. (a) The spatial distribution of transition types overlaid on a hillshade map. (b) The same transitions shown against the background of the 2010-period permafrost distribution. The rectangles delineate two hotspots of significant change: subregion aA (central QTP) and subregion bB (southern QTP).

Table 3 Areas and decadal changes for different frozen ground types, glaciers, and lakes on the QTP from the 2010 to 2020 period. Absolute areas are provided in km², with percentages of the total areas shown in parentheses.

	Permafrost	SFG	Non-frozen	Glacier	Lake
2010	1.086×10 ⁶ (41.17 %)	1.447×10 ⁶ (54.85%)	2.24×10 ⁴ (0.85%)	4.08×10 ⁴ (1.55%)	4.17×10 ⁴ (1.58%)
2020 period	1.038×10 ⁶ (39.35%)	1.466×10 ⁶ (55.57%)	4.67×10 ⁴ (1.77%)	3.75×10 ⁴ (1.42%)	4.97×10 ⁴ (1.89%)
Difference	-4.8×10 ⁴ (1.82%)	1.9×10 ⁴ (0.72%)	2.43×10 ⁴ (0.92%)	-0.33×10 ⁴ (0.13%)	0.80×10 ⁴ (0.31%)

5 Discussion

4.6.5.1 Mechanisms driving changes in soil parameter E variation and permafrost degradation

~~We conducted~~ The decadal changes in E offer useful insight into the factors controlling permafrost stability on the QTP. Although the plateau experienced a ~~targeted analysis~~ 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.

Soil moisture exerts strong control over the ground thermal regime because of its high specific heat capacity and latent heat of fusion. In areas where precipitation increased only modestly, soils likely remained relatively drier. Drier soils have lower thermal inertia and are therefore more responsive to atmospheric warming, which enhances ground heat uptake and drives an increase in E . In contrast, areas with greater precipitation increases likely maintained higher soil moisture levels, which increased thermal inertia and buffered the ground against warming, resulting in a decline in E . This indicates that even under a general warm-wet climate trend, spatial heterogeneity in soil moisture can drive divergent permafrost trajectories across the plateau.

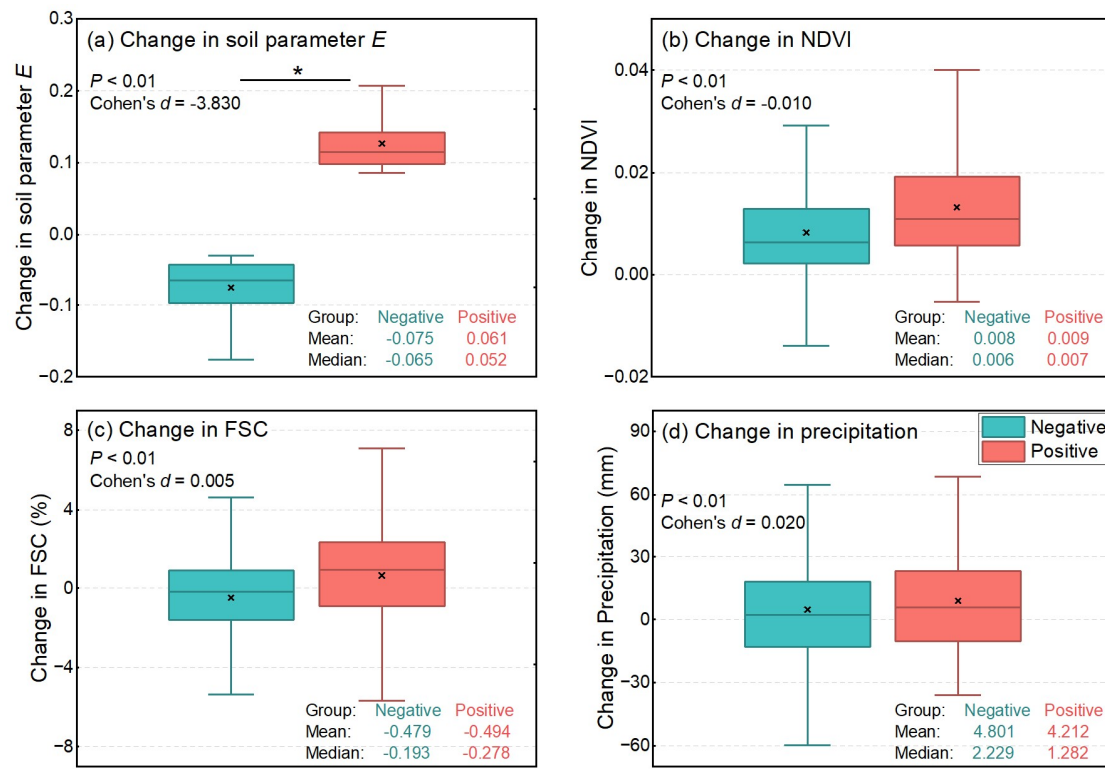


Figure 10 Distribution of decadal changes (2010-2020) in key environmental factors, categorized by the trend in soil parameter *E*. Boxplots compare changes in (a) soil parameter *E*, (b) NDVI, (c) FSC, and (d) precipitation between regions of increased degradation potential (positive group) and reduced 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 identified hotspots (subregion a, b in Fig. 10b). In-groups (Welch Two Sample t-test, $P < 0.01$ and $|\text{Cohen's } d| \geq 0.2$).

The spatial pattern of permafrost degradation lends further support to this interpretation. As shown in Fig. 11a, the dominant transition in the central QTP (Fig. 11a), the dominant process was the degradation of permafrost into SFG. This region corresponds to a critical, concentrated in the transition zone where the permafrost is inherently warm and between the two frozen ground types. This region is characterized by relatively warm permafrost, which is particularly sensitive to thermal perturbations (Zhang et al., 2022). The close spatial pattern of our correspondence between simulated degradation exhibits a strong correspondence with areas and the distribution of thermokarst lakes identified from Sentinel-2A imagery (Chen et al., 2021) and known engineering defect points along the Qinghai-Tibet Railway and Highway (Fig. 11a). As thermokarst development and engineering (Li, 2021) provides independent evidence that the observed changes reflect real

~~physical instability are direct manifestations of thawing, this co-location provides robust physical validation for the degradation areas identified by our model. Comparing the environmental variables between the~~

~~A closer comparison of degraded areas and the adjacent stable permafrost reveals regions (Fig. 12) shows that air temperature warming changes alone does not cannot explain the divergence (Fig. 12). While both regions experienced similar atmospheric warming (mean increase of 0.24 °C in degraded areas versus 0.32 °C in stable areas), the ground thermal response differed significantly (Fig. 12a). The ratio of thawing to freezing indices (DDT/DDF) increased by 0.22 in the degraded zones, more than double the increase observed in the stable zones (0.10) (Fig. 12b). This amplified thermal forcing was compounded by local surface changes, combined with the soil parameter E increased significantly significant rise in the degraded areas E (+0.079) compared to the stable areas (+ versus +0.005) (Fig. 12c). The combination of intensified net heat intake and shifting soil properties), appears to have pushed the already marginal permafrost (with MAGT, which was approaching close to 0 °C in the degraded zones (, Fig. 12d), past the freezing point, triggering the phase change to SFG.) across the thaw threshold.~~

~~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 °C·d) than in adjacent stable SFG regions (131.93 °C·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.~~

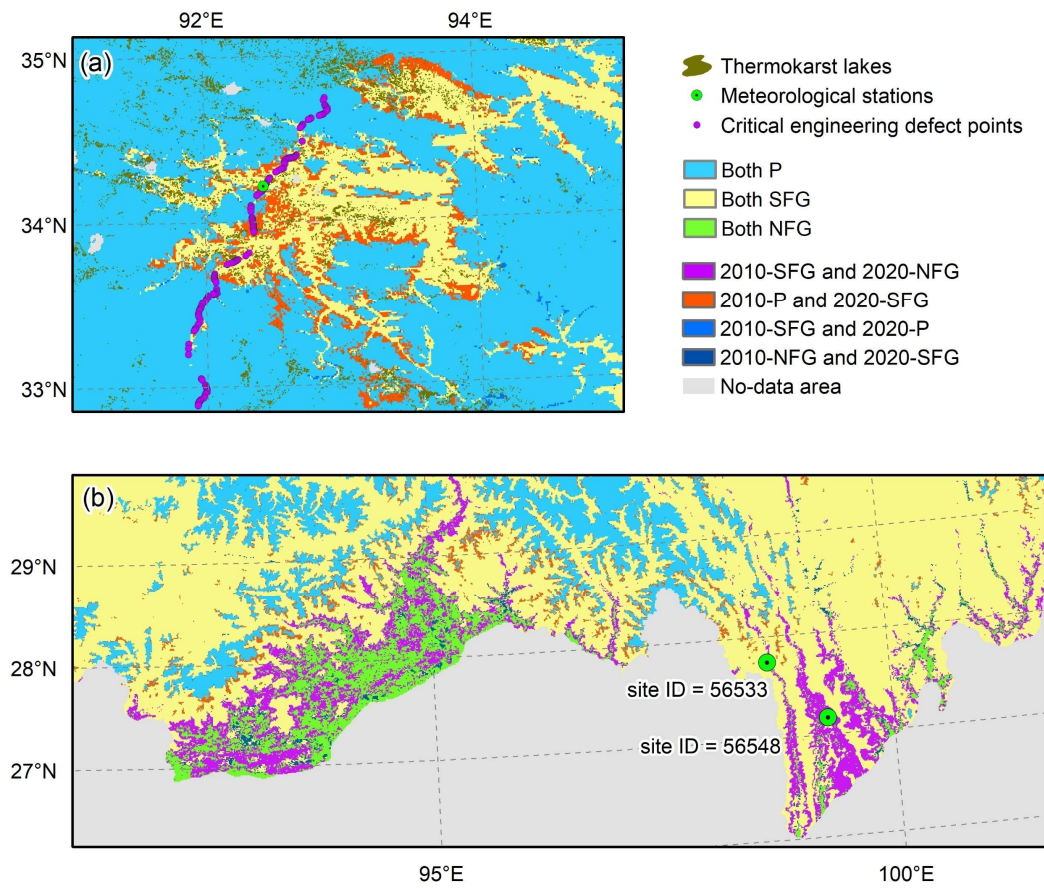


Figure 11 Distribution and changes of frozen ground types in two degradation hotspots on the QTP from 2010 to 2020 [periods](#).

Panels (a) and (b) corresponds to subregions [\(a\)A](#) and [\(b\)B](#) in Fig. [109](#), respectively. (a) Changes in the central QTP, where the dominant transition is from permafrost to SFG. [This map also shows the widespread presence of thermokarst lakes \(extracted from Sentinel-2A imagery\) \(Chen et al., 2021\) and engineering defect points, serving \(mapped via manual field surveys\) \(Li, 2021\) 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.

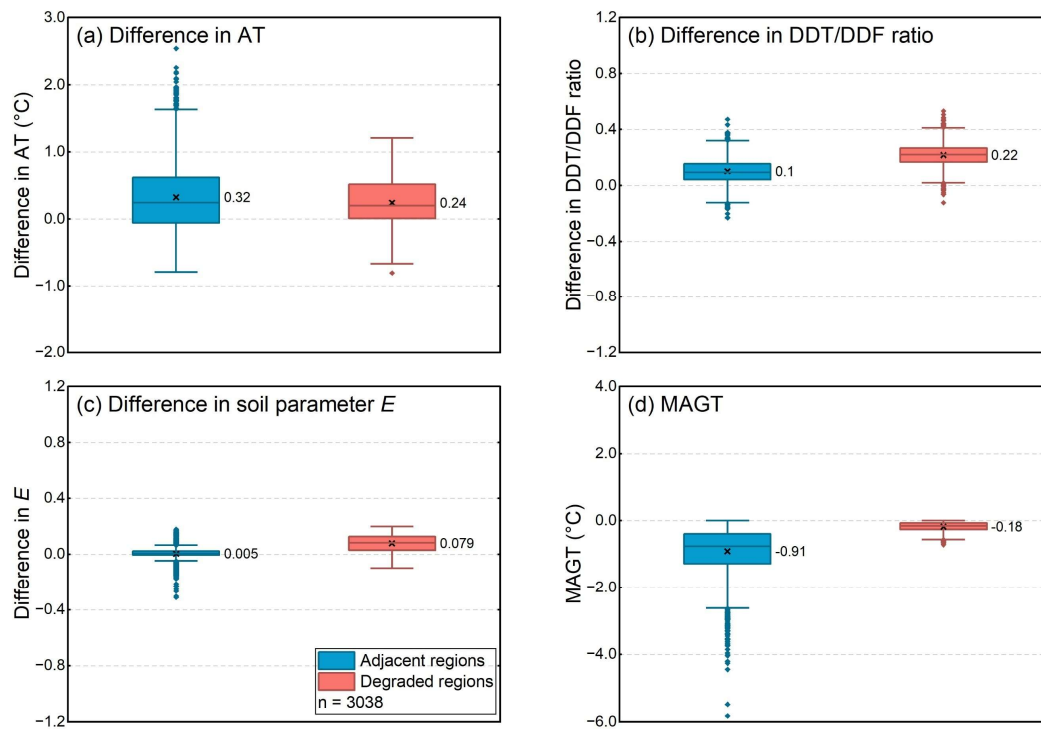


Figure 12 Comparison of key environmental variables between degraded regions (permafrost to SFG) and adjacent stable permafrost regions in subregion aA of Fig. 10b9b (Central QTP). The boxplots compare decadal changes (2010-2020) in (a) air temperature (AT), (b) the DDT/DDF ratio, and (c) the soil parameter E . For the changes shown in panels (a-c), positive values indicate an increase over the decade, while negative values indicate a decrease. Panel (d) compares the mean annual ground temperature (MAGT) in 2010-2019. In each box plot, the center line shows the median, the 'x' marks the mean, 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. Data points beyond the whiskers are plotted individually as outliers.

In contrast, the degradation pattern in the southern QTP (Fig. 11b) was characterized by the contraction of the frozen ground range, manifested as the conversion of SFG to non-frozen ground. Ground truth validation from meteorological stations 56533 and 56548 confirms this transition: observed records from 2016 to 2019 show a freezing index of zero, indicating the complete absence of ground freezing throughout the year. Mechanism analysis highlights the dominant role of summer warming in this region (Fig. 13). Unlike the central plateau, where DDF changes were influential, the southern margin saw the mean DDT increase by $310.03^{\circ}\text{C}\cdot\text{d}$ in degraded areas, compared to only $131.93^{\circ}\text{C}\cdot\text{d}$ in adjacent SFG regions (Fig. 13b). The DDF change was negligible because the DDF in these marginal zones was already minimal in 2010 (Fig. 13c). Furthermore, the increase in the local parameter E was significantly higher in the degraded region (Fig. 13d). This suggests that the loss of frozen ground here is driven by a massive increase in summer heat input combined with land surface feedbacks that effectively eliminated the shallow seasonal frost layer.

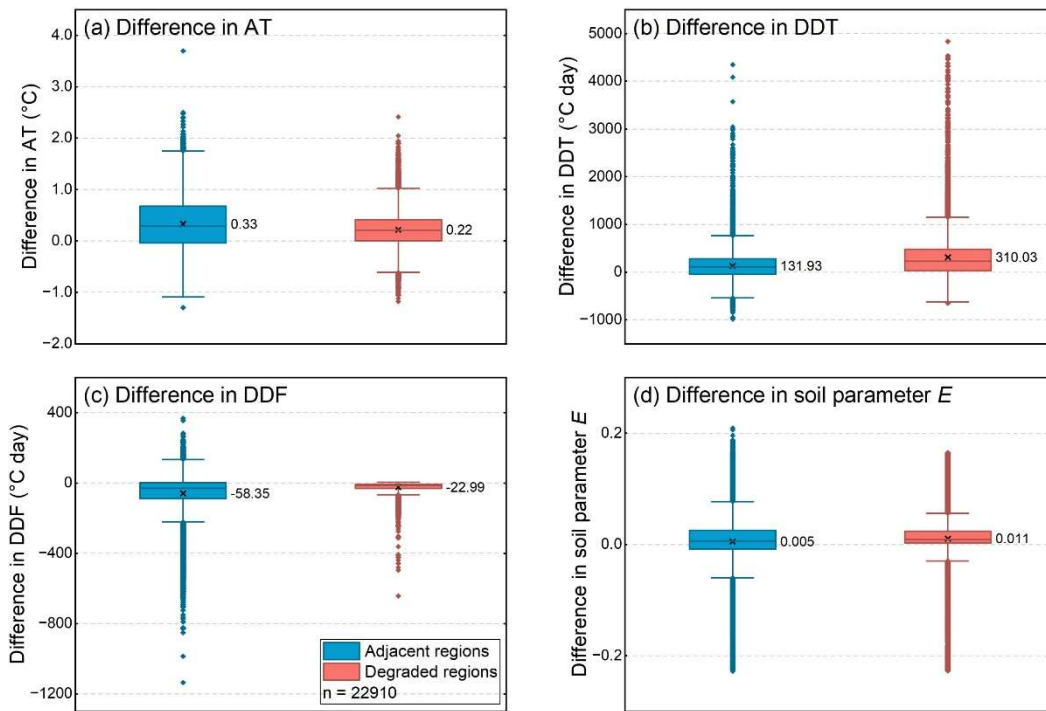


Figure 13 Comparison of key environmental variables between degraded regions (SFG to non-frozen ground) and adjacent SFG regions in subregion **bB** of Fig. 10b9b (Southern QTP). The panels compare decadal changes in (a) air temperature (AT), (b) *DDT*, (c) *DDF*, and (d) the soil parameter *E*. Positive values indicate an increase over the decade, while negative values indicate a decrease. In each box plot, the centre line shows the median, the 'x' marks the mean, the box represents the lower and upper quartiles (25th-75th percentiles), and whiskers extend to the furthest data points within 1.5 times the interquartile range. Data points beyond the whiskers are plotted individually as outliers.

51 Discussion

5.2 Practical implications of the new permafrost map

The 2020-period permafrost distribution map generated in this study demonstrates significant produced here offers real practical value for engineering risk assessment and ecological monitoring environmental management on the QTP. The spatial patterns of permafrost degradation identified show a high spatial correspondence corresponds closely with known locations of engineering defect points defects along the Qinghai-Tibet Railway and Highway. This alignment suggests that the map successfully captures the areas of subsurface thermal instability driving that contribute to infrastructure damage, providing engineers with a critical making it a useful tool for identifying high-risk zones that require may need enhanced monitoring or reinforced monitoring. The lowering of engineering measures.

Beyond its engineering applications, the permafrost table in these headwater regions fundamentally alters soil hydrology and vegetation dynamics, offering vital data. This map provides important information for assessing the resilience of the Asian Water Tower under continued warming. Changes in permafrost and SFG conditions directly affect soil hydrology, groundwater flow, and vegetation dynamics across the headwater regions of major Asian rivers. By offering a temporally specific snapshot of frozen ground conditions rather than a long-term climatological average, this dataset enables a more accurate evaluation of how recent climate change has altered hydrological processes in these critical source areas.

Beyond its practical applications, this map also fills a critical data gap in cryospheric modeling. Unlike previous mapping efforts that often aggregated research, most existing permafrost maps of the QTP are based on multi-decadal data to produce climatological averages; this study provides an instantaneous snapshot that captures the frozen ground status at a specific point in time. By offering a temporally specific and so do not reflect the current thermal state of the ground. The 2020-period map provides a temporally constrained benchmark; this dataset enables that can support more rigorous calibration of hydrological and ecological models, and validation of hydrological, ecological, and land surface models. This is especially valuable for studies that require year-specific or period-specific boundary conditions to simulate recent environmental changes or to project future permafrost evolution under different climate scenarios.

5.3 Model robustness and transferability of the mapping approach

The comparative evaluation of transfer learning strategies showed that the RF regression approach, combined with representative spatial sampling, provided robust and physically plausible estimates of E across the QTP. In contrast, the MLP approach achieved high accuracy within the training subregions but overfit severely and generated unrealistic spatial patterns when applied plateau-wide. This highlights a common challenge in applying deep learning models to large, heterogeneous regions: without strong physical constraints, such models can fit local patterns in the training data yet fail to generalize across diverse environmental conditions. The superior performance of the RF-based method here likely reflects its ability to capture complex, non-linear relationships while remaining relatively robust to noise and outliers. By treating the estimation of E as a regression problem against a spatially continuous 2010 field, rather than relying solely on sparse field survey data, the approach benefited from greater spatial coverage and reduced sampling bias. The use of k -means clustering sampling further improved model performance by ensuring that simulations reflect the current, rather than historical, thermal state of training samples represented the full range of environmental conditions across the plateau.

In this study, we overcame the lack of concurrent field surveys for 2020. We implemented a space for time substitution strategy, hypothesizing that the statistical relationships between the soil parameter and environmental covariates observed in 2010 remained valid for predicting the 2020 state. The comparative evaluation of transfer learning strategies revealed a fundamental trade-off between model complexity and predictive stability. While the hybrid neural network (MLP) theoretically offered greater flexibility, it exhibited severe overfitting and generated implausible spatial artifacts when applied to the plateau-wide domain. In contrast, the RF-based strategy demonstrated superior robustness. By treating the problem as a regression of

the 16-class E field, the RF algorithm effectively filtered noise and captured the large-scale zonal patterns of soil properties, making it the more suitable choice for this specific application.

However, 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.

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 remains, including the extended FROSTNUM framework, space-for-time substitution, and machine learning-based estimation of E , are conceptually transferable, applying them successful elsewhere depends on the availability of high-quality local calibration data and appropriate adaptation of the thermal-index correction procedures.

5.4 Limitations and future directions

Despite its overall robustness, the methodology presented in this study is subject to certain limitations. The 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.

Second, the space-for-time substitution relies on the premise strategy assumes that the interaction relationship between E and its environmental drivers and ground thermal properties is stationary remains sufficiently stable over the study period. While this assumption is appears reasonable for a over decadal timescale timescales on the QTP, it introduces may introduce greater uncertainty when the method is applied over longer time periods or under abruptly changing environmental conditions where the equilibrium between climate and soil properties might decouple. Furthermore, 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.

Third, due to the lack of updated large-scale field surveys ~~means that across~~ the 2020 QTP, the E parameter ~~for the 2020~~ period remains a modeled estimate rather than a directly calibrated observation. Although cross-validation and ensemble analysis support the robustness of these predictions, new field data would still be valuable for further constraining spatial variability under current conditions.

~~Looking forward, these limitations point to clear directions for methodological advancement. Future research should prioritize the acquisition of updated field survey data to provide direct constraints on parameter estimation. Additionally, the potential of neural network approaches should not be dismissed based on the failure of the unconstrained MLP. The overfitting issues observed here could likely be mitigated by embedding physical constraints directly into the loss function. The development of such physics-constrained learning frameworks represents a promising avenue for achieving continuous, high-resolution parameter estimation while maintaining the physical realism required for reliable cryospheric modeling.~~

61 Data availability

The dataset generated in this study, comprising the 1-km-resolution permafrost distribution of the QTP for 2020 and the associated annual ground surface freezing and thawing indices derived from MODIS data is publicly available via Figshare (Chen et al., 2026): <https://doi.org/10.6084/m9.figshare.30997375>. The sources and access details for the third-party public datasets utilized for model forcing and validation are listed in Appendix C.

These limitations point to several priorities for future research. Collecting new, spatially representative field data across the QTP would strengthen the calibration of E . In addition, developing physics-constrained machine learning models and integrating the current empirical approach with process-based simulations could help address non-stationarity and better represent abrupt permafrost changes.

76 Conclusions

This study produced a 1-km resolution permafrost distribution map ~~for of~~ the QTP for the ~~year~~ 2020 period by advancing the ~~ground surface frost number model (FROSTNUM)~~ framework. To address the challenge of data scarcity in updating the 2020 map, we ~~implemented used~~ a space-for-time substitution strategy to estimate the critical soil parameter E . Our comparative evaluation ~~demonstrated showed~~ that ~~at the~~ RF regression approach, trained on representative spatial samples ~~in from~~ the 2010 map, provided a robust and physically realistic ~~estimation of soil parameters, whereas estimates of E . The~~ unconstrained neural network (MLP) approach, ~~in contrast,~~ suffered from overfitting and ~~produced unrealistic~~ spatial artifacts.

The resulting 2020-period map ~~reveal shows~~ that permafrost covers approximately 1.038×10^6 km² (39.35% of the total QTP area), while SFG covers 1.466×10^6 km² (55.57%). Validation against 109 independent borehole records, ~~the map achieved yielded~~ an overall accuracy of 0.84 and a Cohen's Kappa of 0.58, ~~demonstrating superior performance compared to~~ outperforming existing MAGT-based and TTOP-based products.

A systematic comparison with the 2010 baseline indicates ~~a significant~~ substantial degradation of ~~the~~ frozen ground extent over the decade. The total permafrost area decreased by 4.8×10^4 km² (a 1.82% decrease), ~~primarily transitioned into~~ mainly through transition to SFG. Spatially, this degradation was concentrated in two distinct hotspots ~~driven by rising ground temperatures~~. In the central QTP, widespread degradation of unstable permafrost accounted for 7.41% of the total ~~changes~~. Conversely, ~~change~~. In contrast, the southern margins experienced a rapid contraction of the frozen ground ~~ranged~~ domain, where ~~the~~ conversion of SFG to non-frozen ground accounted for 39.62% of the total changes.

This study ~~provided a critical~~ provides an up-to-date ~~benchmark~~, high-resolution permafrost distribution map of the QTP's ~~frozen ground status~~. The ~~produced~~ map serves as ~~offers~~ a vital resource for temporally specific reference to support engineering risk assessment, ecological monitoring, and the ~~rigorous~~ calibration of land surface models in this rapidly warming QTP region.

7 Data availability

The 1-km resolution permafrost distribution of the QTP for the 2020 period and the associated annual ground surface thermal indices generated in this study are publicly available via Figshare (Chen et al., 2026): <https://doi.org/10.6084/m9.figshare.30997375>. The sources and access details for the third-party public datasets used for model forcing, environmental covariates, and validation are provided in Appendix D.

Appendix A: Sampling strategies for optimization of the Random Forest training set

We implemented a representative sampling framework ~~for training to train~~ the RF model efficiently while capturing the full heterogeneity of the QTP. The primary objective was to select a subset of training pixels from the ~~comprehensive full~~ dataset (exceeding 2 million grid cells) that minimized information redundancy while maximizing the environmental variance covered by the model. This is ~~critical~~ important for reducing predictive uncertainty, particularly for rare but physically significant permafrost conditions. We evaluated four distinct sampling algorithms: *k*-means clustering sampling (Brus et al., 2006), the Equal Range (ER) sampling (Hengl et al., 2003), Principal Component Analysis (PCA)-based sampling (Hengl et al., 2003), and Latin Hypercube Sampling (LHS) (Carré et al., 2007) across four sample sizes (1000, 3000, 5000, and 10000 points).

k-means clustering sampling: This method utilizes unsupervised learning to partition the multivariate environmental space into distinct clusters, ensuring that the training set represents distinct physical conditions rather than oversampling spatially abundant but homogeneous areas (Hartigan, 1975). We applied *k*-means algorithm to the vectors of the environmental covariates (topography, soil, vegetation, etc.). The algorithm partitions the study domain into *k* clusters by minimizing ~~the~~ within-cluster sum of squares. For each cluster, the pixel ~~located~~ closest to the geometric centroids in ~~the~~ feature space was

selected as the representative training sample. This approach inherently ~~prioritizes~~favours environmental representativeness over spatial proximity.

ER sampling: Standard random sampling often ~~results in~~introduces a majority-class bias, where the model learns the characteristics of common conditions well but fails to predict rare extremes. To mitigate this, we employed ER sampling, which stratifies the sampling domain based on the distribution of key environmental gradients (Hengl et al., 2003). The core principle of ER is to divide the entire range of a predictor variable into several equal-width strata (or histogram slices). Samples are then allocated within these strata proportionally to the data density. For example, considering a variable with a normal distribution divided into 5 clusters across its standard statistical range (-3σ to $+3\sigma$), the stratification limits correspond to cumulative probabilities of 3.6%, 27.4%, 72.6%, 96.4%, and 100%. ~~Consequently, the~~The resulting sampling weights ~~are~~ assigned to these strata are ~~therefore~~ 0.036, 0.238, 0.452, 0.238, and 0.036, respectively. This method remains robust even for skewed distributions, ~~an advantage it retains over simple equal-frequency binning~~. In our implementation, we generated sample allocations for each environmental factor independently, dividing the total target sample size equally ~~among the~~across factors. ~~We then aggregated these~~These individual sample sets ~~were then aggregated~~ into a combined training pool. Finally, ~~we applied~~ a deduplication step ~~was applied~~ to remove any overlapping points. Testing confirmed that this deduplication ~~resulted in a negligible reduction in~~reduced the final count only negligibly, thereby preserving the intended sample size.

PCA-based sampling: The environmental covariates ~~driving that~~drive permafrost distribution often exhibit high multicollinearity. The PCA-based method (Hengl et al., 2003) address this by transforming the original, correlated predictors into a set of principle components. Sampling was then conducted in this reduced PC space to maximize the coverage of environmental variability. In our implementation, we first performed a PCA analysis on all environmental factors and retained the components that cumulatively explained 95% of the total variance. The total sample size was then allocated to these selected PCs in proportion to the variance ~~they each~~ explained. For example, given a target of 10,000 samples, if the first PC accounted for 64% of the variance, 6,400 sample points were allocated to it. ~~The selection of specific~~Specific points ~~was achieved~~were then selected by stratifying each retained PC using the ER method and randomly drawing the allocated number of samples from within those strata.

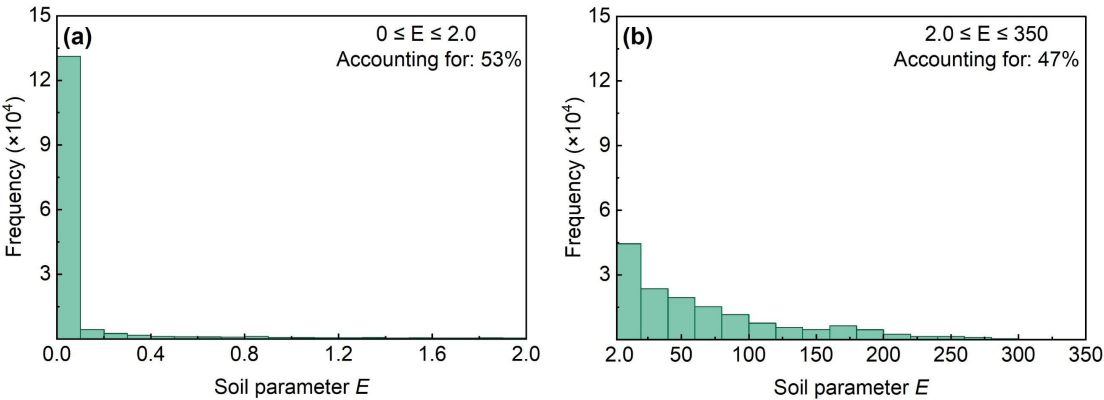
LHS sampling: LHS is a constrained Monte Carlo scheme widely recognized for its efficiency in sampling when no prior soil samples exist, ~~relying solely on and only~~ ancillary data are available (Carré et al., 2007). It is a stratified-random procedure designed to ensure ~~that~~ the full range of every variable is sampled with uniform probability. The method ~~operates~~works by dividing the cumulative distribution of each of the k environmental variables into n equiprobable intervals (where n is the target sample size). A single value is then randomly selected from each interval for every variable. ~~Finally, these~~These n values for each variable are finally randomly paired to create n unique, multivariate sample points. This process ensures that the marginal distribution of each variable is maximally stratified, ~~guaranteeing~~so that the training set covers the entire multivariate feature space without the clustering or gaps common in simple random sampling.

We ~~conducted an analysis to evaluate~~evaluated the validity of the ~~soil parameter~~ E predicted by the ~~Multi-Layer Perceptron~~(MLP) model. ~~The~~This evaluation focused on both the statistical distribution of the predicted values and the spatial realism of the resulting permafrost simulation.

As shown in Fig. B1, the MLP-predicted E values exhibit a highly irregular statistical distribution that deviates from
805 physical expectations. Approximately 53% of the total pixels fall within the range of 0-2.0 (Fig. B1a), with the highest frequency heavily skewed towards the 0-0.1 interval. Conversely, the remaining 47% of the values are scattered across an unrealistically wide range of 2.0-350 (Fig. B1b). This extreme bifurcation suggests the model failed to learn the continuous physical properties of the soil, instead overfitting to the binary classification targets.

Spatially, the MLP-derived E field (Fig. B2a) displays significant inconsistencies when compared to the reference E
810 distribution derived from the Particle Swarm Optimization algorithm with boundary constraints (Fig. B2b), which served as the validated baseline in Cao et al. (2023). Due to the absence of physical constraints, the MLP-derived E values exhibit extreme spatial heterogeneity and magnitude deviations of several orders.

The practical consequences of these artifacts ~~are evident~~become clear when the MLP-derived E field is used to reconstruct the ~~2010~~ permafrost distribution ~~for 2010~~. The resulting map (Fig. B2c) exhibits a substantial visual and spatial mismatch with
815 the rigorously validated 2010 baseline map (Fig. B2d). Given these fundamental discrepancies in both statistical distribution and spatial ~~patterns~~pattern, we conclude that the unconstrained MLP approach does not accurately reflect actual soil conditions. ~~Consequently, this~~This method was ~~therefore~~ excluded from the final mapping workflow. Future applications of deep learning in this context would likely ~~require the integration of~~need physical constraints or penalty terms to ensure physically realistic parameter ~~estimation~~estimates.



820 Figure B1. Frequency histograms of the soil parameter E predicted by the MLP model. (a) Distribution of values within the range of $0 \leq E \leq 2.0$, accounting for 53% of total pixels. (b) Distribution of values within the range of $2.0 \leq E \leq 350$, accounting for the remaining 47%.

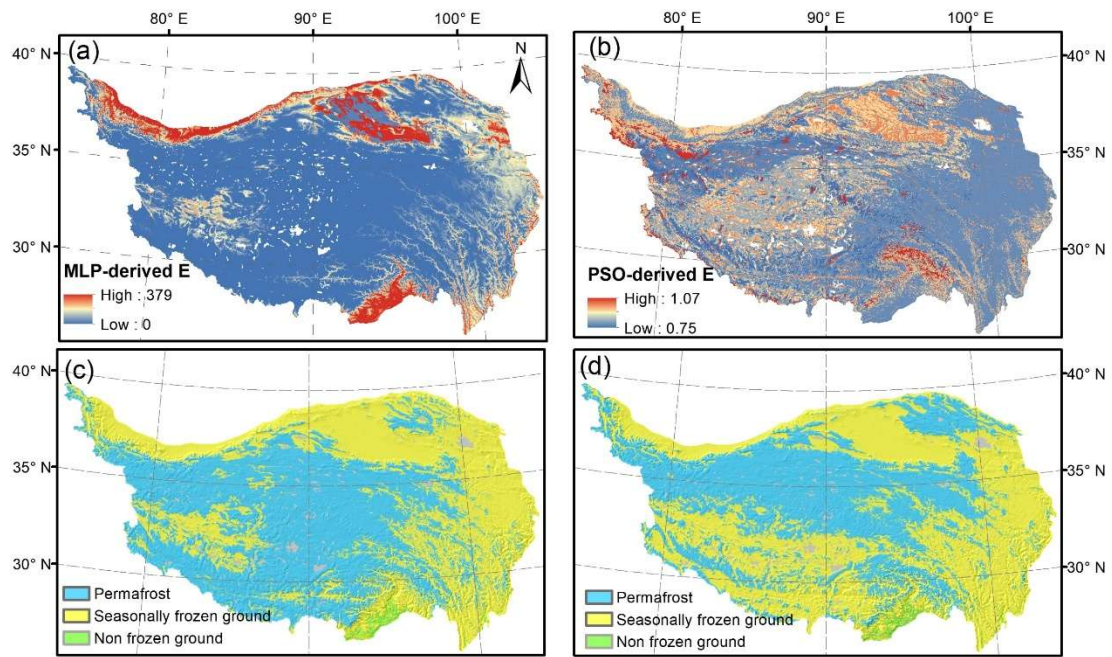


Figure B2: Comparison of the spatial distribution of soil parameter E and the resulting frozen ground types. (a) Spatial distribution of E derived from the MLP model. (b) Spatial distribution of E derived from the Particle Swarm Optimization (PSO) algorithm (Cao et al., 2023). (c) Frozen ground distribution for 2010 simulated using the MLP-derived E . (d) Validated baseline frozen ground map for the year 2010: (Cao et al., 2023).

Appendix C: Sources of open datasets used This study utilized several publicly available datasets. The daily MODIS LST Uncertainty assessment of soil parameter E and emissivity products (MOD11A1 and MYD11A1 version 6) (<https://doi.org/10.5067/MODIS/MOD11A1.006>, <https://doi.org/10.5067/MODIS/MYD11A1.006>) and the NDVI product (MOD13A1) (<https://doi.org/10.5067/MODIS/MOD13A1.006>) are provided by NASA. Topographic data was sourced from the SRTM 90m DEM (version 4; (Reuter et al., 2007)) (<https://egidaresi.community/data/srtm-90m-digital-elevation-database-v4-1/>, accessed on January 6, 2026). Climatic data and surface data: the 1 km monthly precipitation dataset for China (Peng et al., 2019) (<https://doi.org/10.5281/zenodo.3185722>); the 0.005° daily Fractional Snow Cover Dataset Over High Asia (Pan et al., 2024) (<https://doi.org/10.11888/Cryos.tpdc.272503>); and soil properties from the China Data Set of Soil Properties for Land Surface modeling (Shangguan et al., 2013) (<http://globalechange.bnu.edu.cn/research/soil2>, accessed on January 6, 2026). Ground truth and comparison datasets included: the Daily Meteorological Dataset of Basic Meteorological Elements of China National Surface Weather Station (1991–2020) from China National Meteorological Information Centre (Zhao et al., 2024; available upon registration and request at <https://data.cma.cn/> due to the data provider's licensing policies); borehole ground temperature data from Zhao et al. (2021) (<https://doi.org/10.11888/Geocry.tpdc.271107>); a permafrost distribution map

To assess the uncertainties associated with the choice of training sample selection, an uncertainty analysis was conducted for both the soil parameter E and the final permafrost map across the QTP. The analysis was based on 16 model configurations generated from combining four sampling strategies with four sample sizes.

850 The predicted soil parameter E shows a high degree of consistency among the 16 ensemble members (Fig. C1). The ensemble mean of E (Fig. C1a) exhibits a spatial pattern that is highly consistent with the E map derived from the optimal sampling scheme used in the main analysis (Fig. 4a). ~~—based on the Top of Permafrost (TTOP) from Yan et al. (2023) (<https://doi.org/10.11888/Cryos.tpdc.300955>); and ground temperature data at 15 m depth for 2010–2019 from Zou et al. (2025) (<https://doi.org/10.11888/Cryos.tpdc.301165>). Our 2020 permafrost distribution map and associated data are available at Figshare (Chen et al., 2026): <https://doi.org/10.6084/m9.figshare.30997375>.~~ The standard deviation (SD) remains low across most regions of the plateau, generally below 0.04 (Fig. C1b). The coefficient of variance (CV) further demonstrates 855 the stability of the predicted E values (Fig. C1c). Overall, 98.5% of pixels have CV values below 5%, with most pixels concentrated between 1.5% and 2.5% (Fig. C1d). Relatively higher CV values (5%–7.5%) are mainly found in the Qiangtang Plateau, where environmental conditions are more heterogeneous.

860 The uncertainty in the predicted soil parameter E has a limited influence on the final permafrost classification, resulting in a highly consistent permafrost map among the 16 ensemble simulation (Fig. C2). Spatially, the ensemble-derived permafrost maps show a high degree of consistency (Fig. C2a). Most areas are consistently classified as either permafrost or SFG across all 16 simulations. Areas with inconsistent classifications are limited and mainly occur along the boundaries between stable permafrost and SFG, where the classification is more sensitive to variations in the predicted E parameter.

865 The area estimates are also highly consistent among the 16 ensemble simulations (Fig. C2b). The total permafrost area ranges from $1.036 \times 10^6 \text{ km}^2$ to $1.043 \times 10^6 \text{ km}^2$, while the SFG area ranges from $1.461 \times 10^6 \text{ km}^2$ to $1.468 \times 10^6 \text{ km}^2$. The small spread among the ensemble members indicates that uncertainties in the predicted E parameter have a limited effect on regional-scale area estimates. This consistency is further confirmed by the distribution of permafrost agreement probability (P_{agree} , Fig. C2c). Stable SFG ($P_{agree} \leq 0.2$) and stable permafrost ($P_{agree} \geq 0.8$) account for 56.69% and 42.16% of the total pixels, respectively. In contrast, only 1.15% of pixels fall within the transition zone ($0.2 < P_{agree} < 0.8$). The concentration of pixels near agreement probabilities of 0 and 1 indicates a high level of consistency among the ensemble members and demonstrates 870 the robustness of the mapped permafrost distribution.

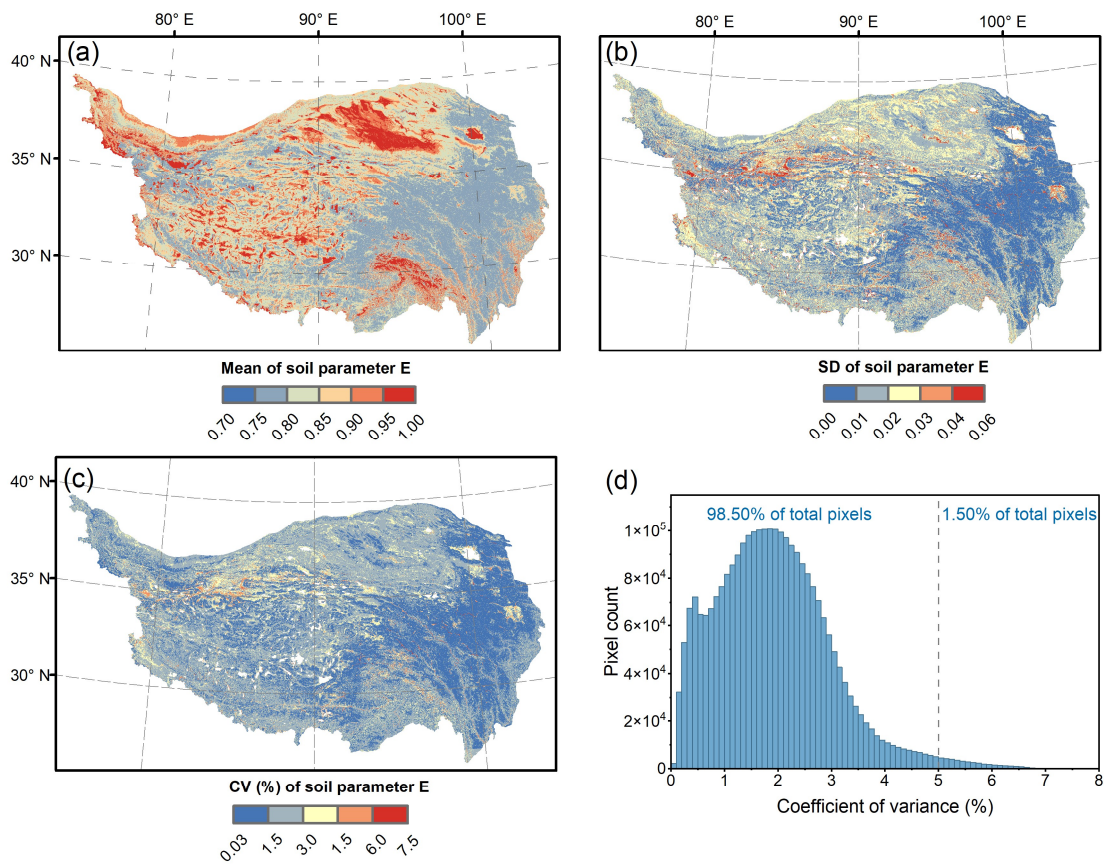


Figure C1. Uncertainty assessment of the soil parameter E derived from 16 model configurations combining different sampling strategies and sample sizes. (a) Ensemble mean of E ; (b) Standard deviation (SD) of E ; (c) Coefficient of variance (CV, %) of E ; and (d) frequency distribution of pixel-level 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.

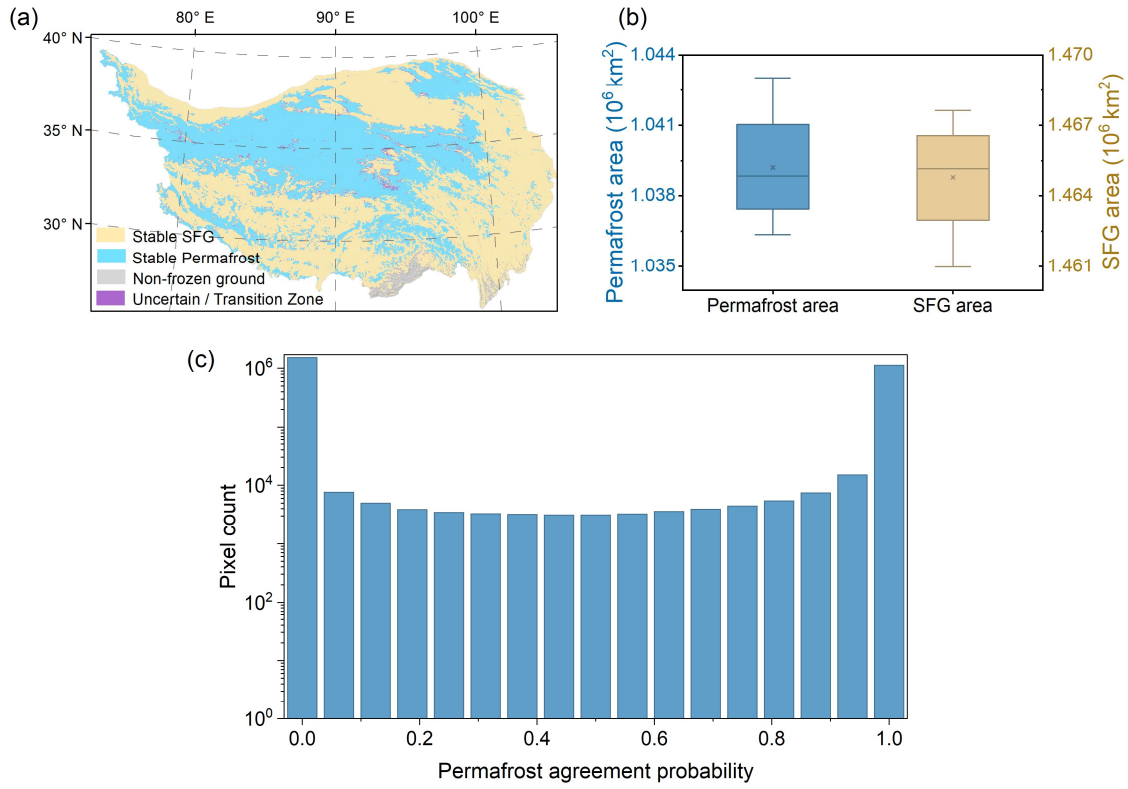


Figure C2. Uncertainty assessment of simulated permafrost distributions derived from 16 parameter E maps. (a) Spatial classification consistency of the 16 ensemble members. Blue pixels indicate areas consistently classified as permafrost by all ensemble members (16/16), yellow pixels indicate areas consistently 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 ($\log_{10}$). 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.

Appendix D: Sources of datasets used

Table D1. Sources of open datasets used

<u>Category</u>	<u>Dataset</u>	<u>Reference / Source</u>	<u>Access Information</u>
<u>Remote Sensing</u>	<u>MODIS LST and emissivity (version 6)</u>	<u>MOD11A1 & MYD11A1</u>	<u>https://doi.org/10.5067/MODIS/MOD11A1.006</u> <u>https://doi.org/10.5067/MODIS/MYD11A1.006</u>
<u>Remote Sensing</u>	<u>MODIS NDVI</u>	<u>MOD13A1</u>	<u>https://doi.org/10.5067/MODIS/MOD13A1.006</u>
<u>Topography</u>	<u>SRTM 90m DEM (version 4)</u>	<u>Reuter et al. (2007)</u>	<u>https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/</u> , accessed on January 6, 2026
<u>Climate</u>	<u>1 km Monthly Precipitation Dataset for China</u>	<u>Peng et al. (2019)</u>	<u>https://doi.org/10.5281/zenodo.3185722</u>
<u>Cryosphere</u>	<u>0.005° daily Fractional Snow Cover Dataset Over High Asia</u>	<u>Pan et al. (2024)</u>	<u>https://doi.org/10.11888/Cryos.tpd.272503</u>
<u>Soil</u>	<u>China Data Set of Soil Properties for Land Surface Modeling</u>	<u>Shangguan et al. (2013)</u>	<u>http://globalchange.bnu.edu.cn/research/soil2/</u> , accessed on January 6, 2026
<u>Meteorology</u>	<u>Daily Meteorological Dataset of Basic Meteorological Elements of China National Surface Weather Station (1991-2020)</u>	<u>Zhao et al. (2024)</u>	<u>Available upon registration at https://data.cma.cn/</u>
<u>Validation</u>	<u>Borehole ground temperature data</u>	<u>Zhao et al. (2021)</u>	<u>https://doi.org/10.11888/Geocry.tpd.271107</u>
<u>Validation</u>	<u>TTOP-based permafrost distribution map</u>	<u>Yan et al. (2023)</u>	<u>https://doi.org/10.11888/Cryos.tpd.300955</u>
<u>Validation</u>	<u>Ground temperature at 15 m depth (2010-2019)</u>	<u>Zou et al. (2025)</u>	<u>https://doi.org/10.11888/Cryos.tpd.301165</u>
<u>This study</u>	<u>1 km permafrost distribution map and thermal indices for the 2020 period</u>	<u>Chen et al. (2026)</u>	<u>https://doi.org/10.6084/m9.figshare.30997375</u>

895 Conceptualization, Z.N.; Methodology, Z.N. and Y.C.; Formal analysis, Y.C., W.T. and Y.Z.; Validation, Y.C. and Y.Z.;
Writing—original draft, Y.C. and Z.N.; Writing—review and editing, Z.N., Y.C., S.Z., D.Y., G.J. and F.N.; Supervision, Z.N.;
Funding acquisition, Z.N. and Y.C.. All authors have read and agree to the published version of the manuscript.

The contact author has declared that none of the authors has any competing interests.

The authors gratefully acknowledge the National Tibetan Plateau Data Center and NASA for providing the datasets used in this study. Sincere appreciation is also extended to the scientists who participated in the Second Tibetan Plateau Scientific Expedition and Research campaign for their dedicated efforts in collecting valuable in situ data across the Qinghai-Tibet Plateau.

This research was supported by the grant of National Natural Science Foundation of China (No. 42571149). Yuhong Chen was also supported by the Postdoctoral Research Funding of Hangzhou International Innovation Institute of Beihang University (Grant No. 2025BKZ075).

915

References

- Brus, D. J., de Gruijter, J. J., and van Groenigen, J. W.: Chapter 14 Designing spatial coverage samples using the k-means clustering algorithm, in: *Developments in Soil Science*, edited by: P. Lagacherie, A.B McBratney, and M. Voltz, Elsevier, Amsterdam, Netherlands, 183-192, [https://doi.org/10.1016/S0166-2481\(06\)31014-8](https://doi.org/10.1016/S0166-2481(06)31014-8), 2006.
- Burke, E. J., Zhang, Y., and Krinner, G.: Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, *The Cryosphere*, 14, 3155-3174, <https://doi.org/10.5194/tc-14-3155-2020>, 2020.
- Cao, Z., Nan, Z., Hu, J., Chen, Y., and Zhang, Y.: A new 2010 permafrost distribution map over the Qinghai–Tibet Plateau based on subregion survey maps: a benchmark for regional permafrost modeling, *Earth Syst. Sci. Data*, 15, 3905-3930, <https://doi.org/10.5194/essd-15-3905-2023>, 2023.
- Carré, F., Mcbratney, A. B., and Minasny, B.: Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping, *Geoderma*, 141, 1-14, <https://doi.org/10.1016/j.geoderma.2007.01.018>, 2007.
- Chadburn, S. E., Burke, E. J., Cox, P. M., Friedlingstein, P., Hugelius, G., and Westermann, S.: An observation-based constraint on permafrost loss as a function of global warming, *Nat. Clim. Chang.*, 7, 340-344, <https://doi.org/10.1038/nclimate3262>, 2017.
- Chen, X., Mu, C., Jia, L., Li, Z., Fan, C., Miu, M., Peng, X., and Wu, X.: *Thermokarst lakes on the Qinghai-Tibet Plateau* (2018), <https://doi.org/10.11888/Geocry.tpd.271205>, 2021.
- Chen, Y., Nan, Z., Cao, Z., Ou, M., and Feng, K.: A stepwise framework for interpolating land surface temperature under cloudy conditions based on the solar-cloud-satellite geometry, *ISPRS-J. Photogramm. Remote Sens.*, 197, 292-308, <https://doi.org/10.1016/j.isprsjprs.2023.02.004>, 2023.
- Chen, Y., Nan, Z., Tian, W., Zhao, Y., Zhao, S., Yang, D., Jing, G., and Niu, F.: A 2020 permafrost distribution map over the Qinghai-Tibet Plateau, Figshare [data set], <https://doi.org/10.6084/m9.figshare.30997375>, 2026.
- Cheng, G., Li, S., and Nan, Z.: Map of permafrost on the Qinghai-Tibet Plateau (1:3,000,000) (1983-1996), National Tibetan Plateau Data Center [data set], <https://doi.org/10.11888/Geocry.tpd.270014>, 1996.
- Deng, H., Zhang, Z., and Wu, Y.: Accelerated permafrost degradation in thermokarst landforms in Qilian Mountains from 2007 to 2020 observed by SBAS-InSAR, *Ecol. Indic.*, 159, 111724, <https://doi.org/10.1016/j.ecolind.2024.111724>, 2024.
- Fu, Y., Horton, R., and Heitman, J.: Neural network estimation of thermal conductivity across full saturation for various soil types, *Comput. Electron. Agric.*, 235, 110321, <https://doi.org/10.1016/j.compag.2025.110321>, 2025.
- Gay, B. A., Pastick, N. J., Züfle, A. E., Armstrong, A. H., Miner, K. R., and Qu, J. J.: Investigating permafrost carbon dynamics in Alaska with artificial intelligence, *Environ. Res. Lett.*, 18, 125001, <https://doi.org/10.1088/1748-9326/ad0607>, 2023.
- Guo, D. and Wang, H.: Simulation of permafrost and seasonally frozen ground conditions on the Tibetan Plateau, 1981–2010, *J. Geophys. Res.-Atmos.*, 118, 5216-5230, <https://doi.org/10.1002/jgrd.50457>, 2013.
- Hachem, S., Allard, M., and Duguay, C.: Using the MODIS land surface temperature product for mapping permafrost: an

- application to northern Québec and Labrador, Canada, *Permafrost and Periglacial Processes*, 20, 407-416, <https://doi.org/10.1002/ppp.672>, 2009.
- Hartigan, J.: *Clustering algorithms*, Wiley, New York, USA, 364 pp., ISBN 0-471-35645-X, 1975.
- 955 Hengl, T., Rossiter, D. G., and Stein, A. G.: Soil sampling strategies for spatial prediction by correlation with auxiliary maps, *Australian Journal of Soil Research*, 41, 1403-1422, <https://doi.org/10.1071/SR03005>, 2003.
- Hjort, J., Streletskiy, D., Doré, G., Wu, Q., Bjella, K., and Luoto, M.: Impacts of permafrost degradation on infrastructure, *Nat. Rev. Earth Environ.*, 3, 24-38, <https://doi.org/10.1038/s43017-021-00247-8>, 2022.
- Hu, G., Zhao, L., Sun, Z., Zou, D., Xiao, Y., Liu, G., Du, E., Wang, C., Wang, Y., Wu, X., Wang, L., and Zhao, Y.:
960 Spatiotemporal characteristics and variability in the thermal state of permafrost on the Qinghai-Tibet Plateau, *Permafrost and Periglacial Processes*, 35, 143-156, <https://doi.org/10.1002/ppp.2219>, 2024.
- Hu, G., Zhao, L., Wu, X., Li, R., Wu, T., Su, Y., and Hao, J.: Evaluation of reanalysis air temperature products in permafrost regions on the Qinghai-Tibetan Plateau, *Theor. Appl. Climatol.*, 138, 1457-1470, <https://doi.org/10.1007/s00704-019-02888-8>, 2019.
- 965 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.
- Huang, X., Tang, G., Zhu, T., Ding, H., and Na, J.: Space-for-time substitution in geomorphology: A critical review and conceptual framework, *J. Geogr. Sci.*, 29, 1670-1680, <https://doi.org/10.1007/s11442-019-1684-0>, 2019.
- 970 Jia, Y., Chen, S., Wu, M., Gu, Y., Wei, P., Wu, T., Shang, Z., Wang, S., and Yu, H.: Improved permafrost stability by revegetation in extremely degraded grassland of the Qinghai-Tibetan Plateau, *Geoderma*, 430, 116350, <https://doi.org/10.1016/j.geoderma.2023.116350>, 2023.
- Karjalainen, O., Luoto, M., Aalto, J., and Hjort, J.: New insights into the environmental factors controlling the ground thermal regime across the Northern Hemisphere: a comparison between permafrost and non-permafrost areas, *The Cryosphere*, 13,
975 693-707, <https://doi.org/10.5194/tc-13-693-2019>, 2019.
- Kim, K. Y., Haagensohn, R., Kansara, P., Rajaram, H., and Lakshmi, V.: Augmenting daily MODIS LST with AIRS surface temperature retrievals to estimate ground temperature and permafrost extent in High Mountain Asia, *Remote Sens. Environ.*, 305, 114075, <https://doi.org/10.1016/j.rse.2024.114075>, 2024.
- Kukulies, J., Chen, D., and Wang, M.: Temporal and spatial variations of convection, clouds and precipitation over the Tibetan
980 Plateau from recent satellite observations. Part II: Precipitation climatology derived from global precipitation measurement mission, *Int. J. Climatol.*, 40, 4858-4875, <https://doi.org/10.1002/joc.6493>, 2020.
- Lei, W., Luo, D., Chen, F., Liu, J., Peng, Y., Li, S., and Shen, Q.: Changes in the thermal regime of permafrost in the Headwater Area of the Yellow River in 1979-2018 based on the HYDRUS model, *Advances in Earth Science*, 39, 1-13, <https://doi.org/10.11867/j.-issn.1001-8166.2024.084>, 2024 (in Chinese).
- 985 Lewkowicz, A. G., O'Neill, H. B., Wolfe, S. A., Roy-Léveillé, P., Roujanski, V. E., Hoefer, E., Gruber, S., Brooks, H., Rudy,

- A. C., Koenig, C. E. M., Brown, N., and Bonnaventure, P. P.: (Eds.): Glossary of Permafrost Science and Engineering, Canadian Permafrost Association, Toronto, Canada, 188 pp., <http://doi.org/10.3138/cpa-gpse>, 2025.
- [Li, G.: 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.tpdc.271880, 2021.](https://doi.org/10.11888/Cryos.tpdc.271880)
- 990 Li, X. and Cheng, G.: A GIS-aided response model of high-altitude permafrost to global change, *Science in China Series D: Earth Sciences*, 42, 72-79, <https://doi.org/10.1007/BF02878500>, 1999.
- Li, X., Cheng, G., Jin, H., Kang, E., Che, T., Jin, R., Wu, L., Nan, Z., Wang, J., and Shen, Y.: Cryospheric change in China, *Glob. Planet. Change*, 62, 210-218, <https://doi.org/10.1016/j.gloplacha.2008.02.001>, 2008.
- Li, Z., Zhao, L., Liu, G., Zou, D., Wang, L., Yang, B., Du, E., Hu, G., Zhou, H., and Wang, C.: Analysis of soil moisture content in the active layer in the permafrost region of the Tuotuo River source in frozen season, *Journal of Glaciology and Geocryology*, 44, 56-68, <https://doi.org/10.7522/j.issn.1000-0240.2022.0001>, 2022.
- 995 Mao, J., Shi, X., Ma, L., Kaiser, D. P., Li, Q., and Thornton, P. E.: Assessment of reanalysis daily extreme temperatures with China's homogenized historical dataset during 1979–2001 using probability density functions, *J. Clim.*, 23, 6605-6623, <https://doi.org/10.1175/2010JCLI3581.1>, 2010.
- 1000 Mi, D.: Map of snow, ice and frozen ground in China, *Journal of Glaciology and Geocryology*, 12, 175-181, <https://doi.org/10.7522/j.issn.1000-0240.1990.0054>, 1990.
- Mu, C., Abbott, B. W., Norris, A. J., Mu, M., Fan, C., Chen, X., Jia, L., Yang, R., Zhang, T., Wang, K., Peng, X., Wu, Q., Guggenberger, G., and Wu, X.: The status and stability of permafrost carbon on the Tibetan Plateau, *Earth-Sci. Rev.*, 211, 103433, <https://doi.org/10.1016/j.earscirev.2020.103433>, 2020.
- 1005 Mu, C., Peng, X., Du, R., Liu, H., Jin, H., Liang, B., Mu, M., Sun, W., Fan, C., Wu, X., Frauenfeld, O. W., and Zhang, T.: A comprehensive dataset for earth system models in a permafrost region: meteorological, permafrost, and carbon observations (2011–2020) in Northeastern Qinghai-Tibet Plateau, *Earth System Science Data Discussions*, 2022, 1-32, <https://doi.org/10.5194/essd-2022-347>, 2022.
- Nan, Z., Huang, P., and Zhao, L.: Permafrost distribution modeling and depth estimation in the Western Qinghai-Tibet Plateau, *Acta Geographica Sinica*, 68, 318-327, <https://doi.org/10.11821/xb201303003>, 2013 (in Chinese).
- 1010 Nan, Z., Li, S., and Liu, Y.: Mean annual ground temperature distribution on the Tibetan Plateau: permafrost distribution mapping and further application, *Journal of Glaciology and Geocryology*, 24, 142-148, <https://doi.org/10.7522/j.issn.1000-0240.2002.0024>, 2002.
- Ni, J., Du, Y., Wu, T., Mo, Q., Zhang, X., and Su, Y.: Hydrothermal variations of active layer in permafrost regions and the influence of vegetation in the Qinghai-Tibetan Plateau, *Theor. Appl. Climatol.*, 156, 12, <https://doi.org/10.1007/s00704-024-05286-x>, 2025.
- 1015 Ni, J., Wu, T., Zhu, X., Hu, G., Zou, D., Wu, X., Li, R., Xie, C., Qiao, Y., Pang, Q., Hao, J., and Yang, C.: Simulation of the present and future projection of permafrost on the Qinghai-Tibet Plateau with statistical and machine learning models, *J. Geophys. Res.-Atmos.*, 126, e2020JD033402, <https://doi.org/10.1029/2020JD033402>, 2021.

- 1020 Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M. O., Lewkowicz, A. G., Panda, S. K., Romanovsky, V., Way, R. G., Westergaard-Nielsen, A., Wu, T., Yamkhin, J., and Zou, D.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km² scale, *Earth-Sci. Rev.*, 193, 299-316, <https://doi.org/10.1016/j.earscirev.2019.04.023>, 2019.
- 1025 Pan, F., Jiang, L., Wang, G., Pan, J., Huang, J., Zhang, C., Cui, H., Yang, J., Zheng, Z., Wu, S., and Shi, J.: MODIS daily cloud-gap-filled fractional snow cover dataset of the Asian Water Tower region (2000–2022), *Earth Syst. Sci. Data*, 16, 2501-2523, <https://doi.org/10.5194/essd-16-2501-2024>, 2024.
- Pan, Y., Li, X., and Li, S.: Effects of different soil thermal conductivity schemes on the simulation of permafrost on the Tibetan Plateau, *Geoderma*, 442, 116789, <https://doi.org/10.1016/j.geoderma.2024.116789>, 2024.
- 1030 Park, Y. [S.](#) and Lek, S.: Chapter 7 - Artificial neural networks: Multilayer perceptron for ecological modeling, in: *Developments in Environmental Modelling*, edited by: Sven Erik Jørgensen, Elsevier, Amsterdam, Netherlands, 123-140, <https://doi.org/10.1016/B978-0-444-63623-2.00007-4>, 2016.
- Peng, S., Ding, Y., Liu, W., and Li, Z.: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, *Earth Syst. Sci. Data*, 11, 1931-1946, <https://doi.org/10.5194/essd-11-1931-2019>, 2019.
- 1035 Pickett, S. T.: Space-for-time substitution as an alternative to long-term studies, in: *Long-Term Studies in Ecology*, edited by: Gene E. Likens, Springer, New York, USA, 110-135, https://doi.org/10.1007/978-1-4615-7358-6_5, 1989.
- Qiu, J.: China: The third pole, *Nature*, 454, 393-396, <https://doi.org/10.1038/454393a>, 2008.
- Ran, Y., Li, X., Cheng, G., Che, J., Aalto, J., Karjalainen, O., Hjort, J., Luoto, M., Jin, H., Obu, J., Hori, M., Yu, Q., and Chang, X.: New high-resolution estimates of the permafrost thermal state and hydrothermal conditions over the Northern Hemisphere, *Earth Syst. Sci. Data*, 14, 865-884, <https://doi.org/10.5194/essd-14-865-2022>, 2022.
- 1040 [Ran, Y., Li, X., Cheng, G., Nan, Z., Che, J., Sheng, Y., Wu, Q., Jin, H., Luo, D., and Tang, Z.: Mapping the permafrost stability on the Tibetan Plateau for 2005-2015. *Sci. China Earth Sci.*, 64, 62-79, <https://doi.org/10.1007/s11430-020-9685-3>, 2021.](#)
- [Ran, Y., Li, X., Cheng, G., Zhang, T., Wu, Q., Jin, H., and Jin, R.: Distribution of permafrost in China: an overview of existing permafrost maps, *Permafrost and Periglacial Processes*, 23, 322-333, <https://doi.org/10.1002/ppp.1756>, 2012.](#)
- 1045 Reuter, H. I., Nelson, A., and Jarvis, A.: An evaluation of void-filling interpolation methods for SRTM data, *Geogr. Inf. Syst.*, 21, 983-1008, <https://doi.org/10.1080/13658810601169899>, 2007.
- Shan, W., Zhang, C., Guo, Y., and Qiu, L.: Mapping the thermal state of permafrost in Northeast China based on the surface frost number model, *Remote Sens.*, 14, 3185, <https://doi.org/10.3390/rs14133185>, 2022.
- Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., Ji, D., Ye, A., Yuan, H., Zhang, Q., Chen, D., Chen, M., Chu, J.,
- 1050 Dou, Y., Guo, J., Li, H., Li, J., Liang, L., Liang, X., Liu, H., Liu, S., Miao, C., and Zhang, Y.: A China data set of soil properties for land surface modeling, *J. Adv. Model. Earth Syst.*, 5, 212-224, <https://doi.org/10.1002/jame.20026>, 2013.
- Shi, S., Wang, P., Zhan, X., Han, J., Guo, M., and Wang, F.: Warming and increasing precipitation induced greening on the northern Qinghai-Tibet Plateau, *Catena*, 233, 107483, <https://doi.org/10.1016/j.catena.2023.107483>, 2023.

- Siewert, M. B.: High-resolution digital mapping of soil organic carbon in permafrost terrain using machine learning: a case study in a sub-Arctic peatland environment, *Biogeosciences*, 15, 1663-1682, <https://doi.org/10.5194/bg-15-1663-2018>, 2018.
- Smith, M. W. and Riseborough, D. W.: Permafrost monitoring and detection of climate change, *Permafrost and Periglacial Processes*, 7, 301-309, [https://doi.org/10.1002/\(SICI\)1099-1530\(199610\)7:4<301::AID-PPP231>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-1530(199610)7:4<301::AID-PPP231>3.0.CO;2-R), 1996.
- Smith, S. L., O'Neill, H. B., Isaksen, K., Noetzli, J., Romanovsky, V. E., and Anonymous: The changing thermal state of permafrost, *Nat. Rev. Earth Environ.*, 3, 10-23, <https://doi.org/10.1038/s43017-021-00240-1>, 2022.
- Tananaev, N. and Lotsari, E.: Defrosting northern catchments: Fluvial effects of permafrost degradation, *Earth-Sci. Rev.*, 228, 103996, <https://doi.org/10.1016/j.earscirev.2022.103996>, 2022.
- van Der Westhuizen, S., Heuvelink, G. B., and Hofmeyr, D. P.: Multivariate random forest for digital soil mapping, *Geoderma*, 431, 116365, <https://doi.org/10.1016/j.geoderma.2023.116365>, 2023.
- van Doninck, J., Peters, J., De Baets, B., De Clercq, E. M., Ducheyne, E., and Verhoest, N. E. C.: The potential of multitemporal Aqua and Terra MODIS apparent thermal inertia as a soil moisture indicator, *Int. J. Appl. Earth Obs. Geoinf.*, 13, 934-941, <https://doi.org/10.1016/j.jag.2011.07.003>, 2011.
- Wadoux, A. M., Brus, D. J., and Heuvelink, G. B.: Sampling design optimization for soil mapping with random forest, *Geoderma*, 355, 113913, <https://doi.org/10.1016/j.geoderma.2019.113913>, 2019.
- Wang, G., Liu, G., Li, C., and Yang, Y.: The variability of soil thermal and hydrological dynamics with vegetation cover in a permafrost region, *Agric. For. Meteorol.*, 162-163, 44-57, <https://doi.org/10.1016/j.agrformet.2012.04.006>, 2012.
- Wang, Y., Lv, W., Xue, K., Wang, S., Zhang, L., Hu, R., Zeng, H., Xu, X., Li, Y., and Jiang, L.: Grassland changes and adaptive management on the Qinghai-Tibetan Plateau, *Nat. Rev. Earth Environ.*, 3, 668-683, <https://doi.org/10.1038/s43017-022-00330-8>, 2022.
- Wu, Q. and Zhang, T.: Recent permafrost warming on the Qinghai-Tibetan Plateau, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/10.1029/2007JD009539>, 2008.
- Yan, D., Zheng, X., Feng, M., Liang, S., Hu, Z., Kuang, X., and Feng, Y.: Frozen ground change data set in the Tibetan Plateau (1961-2020), National Tibetan Plateau/Third Pole Environment Data Center [data set], <https://doi.org/10.11888/Cryos.tpd.300955>, 2023.
- Yang, M., Nelson, F. E., Shiklomanov, N. I., Guo, D., and Wan, G.: Permafrost degradation and its environmental effects on the Tibetan Plateau: A review of recent research, *Earth-Sci. Rev.*, 103, 31-44, <https://doi.org/10.1016/j.earscirev.2010.07.002>, 2010.
- Yang, M., Wang, X., Pang, G., Wan, G., and Liu, Z.: The Tibetan Plateau cryosphere: Observations and model simulations for current status and recent changes, *Earth-Sci. Rev.*, 190, 353-369, <https://doi.org/10.1016/j.earscirev.2018.12.018>, 2019.
- Yao, T., Bolch, T., Chen, D., Gao, J., Immerzeel, W., Piao, S., Su, F., Thompson, L., Wada, Y., and Wang, L.: The imbalance of the Asian water tower, *Nat. Rev. Earth Environ.*, 3, 618-632, <https://doi.org/10.1038/s43017-022-00299-4>, 2022.
- Yao, T., Thompson, L., Yang, W., Yu, W., Gao, Y., Guo, X., Yang, X., Duan, K., Zhao, H., Xu, B., Pu, J., Lu, A., Xiang, Y.,

- Kattel, D. B., and Joswiak, D.: Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings, *Nat. Clim. Chang.*, 2, 663-667, <https://doi.org/10.1038/nclimate1580>, 2012.
- 1090 Ye, Q., Zong, J., Tian, L., Cogley, J. G., Song, C., and Guo, W.: Glacier changes on the Tibetan Plateau derived from Landsat imagery: mid-1970s – 2000–13, *J. Glaciol.*, 63, 273-287, <https://doi.org/10.1017/jog.2016.137>, 2017.
- Yu, W., Nan, Z., Wang, Z., Chen, H., Wu, T., and Zhao, L.: An effective interpolation method for MODIS land surface temperature on the Qinghai-Tibet Plateau, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 8, 4539-4550, <https://doi.org/10.1109/JSTARS.2015.2464094>, 2015.
- 1095 Zhang, G., Nan, Z., Hu, N., Yin, Z., Zhao, L., Cheng, G., and Mu, C.: Qinghai-Tibet Plateau permafrost at risk in the late 21st century, *Earth Future*, 10, <https://doi.org/10.1029/2022EF002652>, 2022.
- Zhang, G., Nan, Z., Wu, X., Ji, H., and Zhao, S.: The Role of Winter Warming in Permafrost Change Over the Qinghai-Tibet Plateau, *Geophys. Res. Lett.*, 46, 11261-11269, <https://doi.org/10.1029/2019GL084292>, 2019.
- Zhang, G., Nan, Z., Zhao, L., Liang, Y., and Cheng, G.: Qinghai-Tibet Plateau wetting reduces permafrost thermal responses to climate warming, *Earth Planet. Sci. Lett.*, 562, 116858, <https://doi.org/10.1016/j.epsl.2021.116858>, 2021.
- 1100 Zhang, G., Ran, Y., Wan, W., Luo, W., Chen, W., Xu, F., and Li, X.: 100 years of lake evolution over the Qinghai-Tibet Plateau, *Earth Syst. Sci. Data*, 13, 3951-3966, <https://doi.org/10.5194/essd-13-3951-2021>, 2021.
- Zhang, G., Yao, T., Xie, H., Yang, K., Zhu, L., Shum, C. K., Bolch, T., Yi, S., Allen, S., Jiang, L., Chen, W., and Ke, C.: Response of Tibetan Plateau lakes to climate change: Trends, patterns, and mechanisms, *Earth-Sci. Rev.*, 208, 103269, <https://doi.org/10.1016/j.earscirev.2020.103269>, 2020.
- 1105 Zhang, ~~M., Li, R., Pei, W., Zhou, Y., Li, G., and Yang, S.: Permafrost degradation risk evaluation in the Qinghai Tibet Plateau under climate change based on machine learning models, *J. Geophys. Res. Atmos.*, 129, e2023JD039611, <https://doi.org/10.1029/2023JD039611>, 2024.~~
- ~~Zhang,~~ T., Barry, R. G., Knowles, K., Heginbottom, J. A., and Brown, J.: Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere, *Polar Geogr.*, 23, 132-154, <https://doi.org/10.1080/10889379909377670>, 1999.
- Zhao, J., Zhao, L., Sun, Z., Hu, G., Zou, D., Xiao, M., Liu, G., Pang, Q., Du, E., Li, Z., Wu, X., Xiao, Y., Wang, L., and Zhang, W.: The thermal state of permafrost under climate change on the Qinghai-Tibet Plateau (1980-2022): a case study of the West Kunlun, *The Cryosphere*, 19, 4211-4236, <https://doi.org/10.5194/tc-19-4211-2025>, 2025.
- 1115 Zhao, J., Zhao, L., Sun, Z., Niu, F., Hu, G., Zou, D., Liu, G., Du, E., Wang, C., Wang, L., Qiao, Y., Shi, J., Zhang, Y., Gao, J., Wang, Y., Li, Y., Yu, W., Zhou, H., Xing, Z., Xiao, M., Yin, L., and Wang, S.: Simulating the current and future northern limit of permafrost on the Qinghai-Tibet Plateau, *The Cryosphere*, 16, 4823-4846, <https://doi.org/10.5194/tc-16-4823-2022>, 2022.
- Zhao, L., Ping, C., Yang, D., Cheng, G., Ding, Y., and Liu, S.: Changes of climate and seasonally frozen ground over the past 30 years in Qinghai-Xizang (Tibetan) Plateau, China, *Glob. Planet. Change*, 43, 19-31, <https://doi.org/10.1016/j.gloplacha.2004.02.003>, 2004.
- 1120

- Zhao, L., Zou, D., Hu, G., Du, E., Pang, Q., Xiao, Y., Li, R., Sheng, Y., Wu, X., and Sun, Z.: Changing climate and the permafrost environment on the Qinghai–Tibet (Xizang) plateau, *Permafrost and Periglacial Processes*, 31, 396–405, <https://doi.org/10.1002/ppp.2056>, 2020.
- 1125 Zhao, L., Zou, D., Hu, G., Wu, T., Du, E., Liu, G., Xiao, Y., Li, R., Pang, Q., Qiao, Y., Wu, X., Sun, Z., Xing, Z., Sheng, Y., Zhao, Y., Shi, J., Xie, C., Wang, L., Wang, C., and Cheng, G.: A synthesis dataset of permafrost thermal state for the Qinghai–Tibet (Xizang) Plateau, China, *Earth Syst. Sci. Data*, 13, 4207–4218, <https://doi.org/10.5194/essd-13-4207-2021>, 2021.
- 1130 Zhao, S. P., Nan, Z. T., Huang, Y. B., and Zhao, L.: The application and evaluation of simple permafrost distribution models on the Qinghai–Tibet Plateau, *Permafrost and Periglacial Processes*, 28, 391–404, <https://doi.org/10.1002/ppp.1939>, 2017.
- Zhao, Y., Liao, J., Zhang, Q., Chen, J., Gong, X., Shi, Y., Shi, M., Yang, D., Fan, S., Zhou, X., Cao, L., and Hu, K.: Development of China Ground Climate Normal Value Dataset from 1991 to 2020, *Chinese Journal of Atmospheric Sciences*, 48, 555–571, <https://doi.org/10.3878/j.issn.1006-9895.2204.22010>, 2024 (in Chinese).
- 1135 Zhong, L., Lei, H., and Yang, J.: Development of a distributed physics-informed deep learning hydrological model for data-scarce regions, *Water Resour. Res.*, 60, e2023WR036333, <https://doi.org/10.1029/2023WR036333>, 2024.
- Zhou, G., Ren, H., Zhang, L., Lv, X., and Zhou, M.: Annual vegetation maps in the Qinghai–Tibet Plateau (QTP) from 2000 to 2022 based on MODIS series satellite imagery, *Earth Syst. Sci. Data*, 17, 773–797, <https://doi.org/10.5194/essd-17-773-2025>, 2025.
- 1140 Zhu, Y., Sang, Y., Chen, D., Sivakumar, B., and Li, D.: Effects of the South Asian summer monsoon anomaly on interannual variations in precipitation over the South-Central Tibetan Plateau, *Environ. Res. Lett.*, 15, 124067, <https://doi.org/10.1088/1748-9326/abc71b>, 2020.
- Zou, D., Zhao, L., Hu, G., Du, E., Liu, G., Wang, C., and Li, W.: Permafrost temperature baseline at 15 m depth on the Qinghai–Tibetan Plateau (2010–2019), *Earth Syst. Sci. Data*, 17, 1731–1742, <https://doi.org/10.5194/essd-17-1731-2025>, 2025.
- 1145 Zou, D., Zhao, L., Sheng, Y., Chen, J., Hu, G., Wu, T., Wu, J., Xie, C., Wu, X., Pang, Q., Wang, W., Du, E., Li, W., Liu, G., Li, J., Qin, Y., Qiao, Y., Wang, Z., Shi, J., and Cheng, G.: A new map of permafrost distribution on the Tibetan Plateau, *The Cryosphere*, 11, 2527–2542, <https://doi.org/10.5194/tc-11-2527-2017>, 2017.