# Response to Review Comments (Referee #1)

Manuscript ESSD-2026-299

*"An Ensemble Dataset of Permafrost Thaw Conditions for Northern High Latitudes Using Open Satellite Data"*

Dear Referee #1,

Thank you for your constructive comments. Please find below our point-by-point responses and corresponding revisions. Especially, the three critical issues, "circular reasoning, common errors, and inadequate validation", are thoroughly addressed. In summary, the revised manuscript includes the following major changes:

- **Extensive search for permafrost monitoring stations:** The borehole sites increased from 26 from one network to 96 sites from four networks.
- **two new figures:** one displays the spatial patterns of ΔPP and ΔMAGT of the three published products (Obu, NIEER, ESACCI) relative to the ensemble product, and the other displays the expanded 96 borehole sites and the revised scatterplot.
- **two new subsections**: 3.1.3 "Spatial uncertainties in satellite-based permafrost products"; and 3.3.3 "Model performance and comparison with borehole records").
- **10+ new references** relevant to our revision.

Upon the ESSD review procedure, the final revised manuscript will be provided after the review window is closed.

## 1. Circular Reasoning

We appreciate your summary that this manuscript presents a "data-engineering" study, i.e., extracting a Permafrost Thaw Index (PTI) product by integrating a set of open-access, satellite-based permafrost and environmental datasets. In the big data era, the research goal of this study is to advance permafrost monitoring by taking advantage of open satellite data.

We split this valuable comment into three pieces to respond, as below.

1.1 The authors state that thermal variables, particularly Land Surface Temperature (LST) and LST trends, are the strongest predictors for the PTI. This is entirely expected. However, the open-access permafrost products used to derive the target variables (e.g., Obu et al., 2021) are themselves heavily reliant on LST and TTOP (Temperature at the Top of Permafrost) modeling.

This comment is related to how we collect the training data for the target (predictive) variable.

In reality, there is no direct, "ground-truth" of permafrost thaw measurement. Therefore, we rely on three global permafrost products (Obu, NIEER, and Permafrost_CCI) to develop our methodology. Published in peer-reviewed top journals, these three permafrost products are considered reliable representations of permafrost conditions including PP (permafrost fraction), MAGT, and ALT at a global scale. We therefore regard these products as the best available "truth" data sources, while acknowledging the uncertainties associated with their model algorithms.

We fully agree with your comment regarding the role of LST as a driving factor. As these permafrost products documented, LST is undoubtedly a primary factor regulating permafrost distribution. Building on these products, our study takes a further step to assess permafrost thaw potential. Similarly, LST is an important explanatory variable for characterizing spatial variations in thaw potential across the permafrost region.

1.2 By utilizing many features (including LST) to train an XGBoost model to predict a target variable that was fundamentally derived from LST in the first place, it is highly probable that the machine learning model is simply "reverse-engineering" the temperature-dependency assumptions of the original datasets, rather than discovering novel mechanisms. This kind of A-to-B and then B-to-A analytical approach is highly problematic.

Thanks for pointing out this logical concern. As mentioned above, temperature dependence is intrinsic to all permafrost products, including the three permafrost maps and our proposed PTI. From the following two perspectives, we address this concern to clarify that our study is not a simple A-to-B and then B-to-A loop driven by a sole LST mechanism.

First, our study mediates this logical concern by integrating the three permafrost products using a median ensemble filter. The resulting permafrost layers (particularly PP and MAGT), which are used to extract the target variable PTI, compensates for substantial prediction errors associated with any single permafrost product. By using this ensemble layer, the target variable does not rely on a single LST-based approach, thereby reducing the potential dependence on any individual permafrost product.

Secondly, LST is not the only explanatory variable in our prediction model. Our hypothesis is that, although permafrost thaw is driven by a warming climate, its spatial variability also reflects local environmental heterogeneity. By including a suite of environmental explanatory variables, the XGBoost model is expected to capture these local variations beyond LST.

1.3 This high autocorrelation artificially inflates the reported 91.8% accuracy, rendering it misleading.

This is a good point that we overlooked in the original manuscript. We agree that the term "accuracy" to describe the predicted PTI is misleading. Without ground-truth permafrost thaw metrics, we cannot really assess the accuracy of the predicted PTI. We revised our manuscript accordingly to avoid this misleading statement.

- In **Abstract**: The statement "*overall accuracy of 91.8%*" was removed because it did not refer to the accuracy of PTI prediction. The sentence was revised as "*…An ensemble machine learning approach, XGBoost, is employed to predict PTI across the northern high latitudes with a model performance of 91.8% accuracy...*"

The sample dataset was collected from a conceptual PTI layer built upon the three permafrost products, as outlined in Table 3 of Section 2.2.3, "Building a permafrost thaw index". We revised the manuscript to better explain the uncertainties associated with the underlying permafrost products. As below:

- In **Sub-section 2.2.3** (in **Methods**): "…*It is important to acknowledge the uncertainties associated with the conceptual PTI map when interpreting permafrost thaw conditions. In the absence of ground-truth data for quantifying PTI, this study leveraged a large training dataset and robust machine learning techniques to mitigate the effects of noises and uncertainties in the predictive variable.*"

A total of 61,325 samples were collected from the conceptual PTI layer, randomly splitting into 80% for model development (training) and 20% for model performance evaluation (validation). The confusion matrix was constructed using the validation set. We revised the manuscript to better explain the accuracy assessment:

- **Sub-section 3.3.3** (in **Results**) was renamed as "*Model performance and comparison with borehole records*". The following statement was added to address this comment:

"…*Using the validation sample set, a confusion matrix (Table 4) was constructed to evaluate the derived PTI map. The number of validation points for each PTI class is listed in the table. It should be noted that the PTI ranks at the validation points were collected from the conceptual PTI layer, which inherently contains uncertainties and noise. The derived PTI map achieved an OA of 91.79%, i.e., this high percentage of the validation samples were assigned to the same PTI classes as those in the conceptual layer. Furthermore, the derived PTI map may improve upon the conceptual layer by capturing local-scale variations from the 12 environmental variables.*"

We also revised the following paragraph in this subsection to better describe the class-specific prediction:

"*The prediction power of the XGBoost model varied with PTI classes… Given that the conceptual layer primarily represents macroscale thaw potential, these cross-rank mismatches may also reflect the model's ability to refine the PTI by incorporating environmental explanatory variables that capture local-scale variations.*"

## 2. Homogeneous Errors

2.1 The authors attempt to improve reliability by ensembling three open-access products. The fundamental prerequisite for ensembling is that the errors in the base datasets are statistically independent. In complex terrains or specific ecotype zones (e.g., coastal tundra), existing permafrost products likely fail for the exact same reasons. As rightly pointed out by Dr. Robert Way, foundational models often misrepresent permafrost distribution in regions like northeastern Canada. Ensembling datasets with the same potential biases does not mitigate uncertainty; rather, it masks the underlying issues within the final PTI product.

We highly value this comment that raises an important point that we overlooked. Because all three permafrost products rely on thermal characteristics of permafrost regions, they must share common sources of error. Our study used these products to construct the sample set for the target variable. Therefore, it cannot fully mitigate these shared errors.

We revised the manuscript to acknowledge the common sources of error inherent in these three products. In the Discussion section, a new paragraph is revised in "**3.4 Satellite-based open data for permafrost research**":

"*Satellite imagery cannot directly detect permafrost beneath the active layer, …While different modelling approaches have been incorporated into the three permafrost products (Obu, NIEER, and Permafrost_CCI) used in this study, all three products rely heavily on thermal characteristics of satellite observations (e.g., MODIS) and ERA5 climate reanalysis. Therefore, they may share common biases in estimating permafrost conditions of PP, MAGT and ALT. Substantial discrepancies were observed among these products, and this study sought to reduce the product-specific uncertainties by generating an ensemble layer of permafrost conditions. Yet common biases among the three satellite-based products remain and could propagate into the derived PTI layer. An outstanding example is peatland permafrost. Along the coastline of Quebec-Labrador Peninsula, Canada, Wang et al. (2023b) documented the presence of permafrost complexes associated with palsa bogs and peat plateaus. Way and Lewkowicz (2016) developed a snow-modified TTOP map that revealed more heterogeneous permafrost distributions after accounting for localized variability in geomorphology and snow cover in peatlands. More specifically, Way et al. (2018) identified a uniquely large thermal offset in peatland permafrost regimes, indicating greater resilience than suggested by existing permafrost maps. In this study, we compared ground temperatures from four borehole sites at Ungava Bay with our ensemble estimates. The ensemble MAGT was consistently higher at these sites, resulting in an underrepresentation of permafrost stability along the Labrador coastline. Global peat datasets remain limited and available only at coarse resolutions. Moreover, coastal peatland permafrost is often narrow and limited in spatial extent and therefore, cannot be effectively distinguished in existing global products. For future applications, careful environmental characterization and localized permafrost modelling may help better identify, interpret and mitigate these common uncertainties.*"

Furthermore, we added a new subsection, "**3.1.3 Spatial uncertainties in satellite-based permafrost products**" to examine the spatial patterns of prediction errors among the three permafrost products relative to our ensemble product (ΔPP and ΔMAGT). This analysis also addresses a comment from Dr. Way. A new figure was added to facilitate visual comparison. To expand from your comment, we provide the revised text and the figure (Fig.5 in the revised manuscript) below.

"*The differences in PP and MAGT of the three permafrost products (source maps) relative to the ensemble product are spatially compared in Fig.5. For PP, all three source maps (Fig.5a-5c) show strong agreement with the ensemble product in the interior of the study region, corresponding to areas of continuous permafrost (PP=100%). Greater discrepancies occur farther south, in areas characterized by discontinuous, sporadic and isolated permafrost. Overall, the hot and cold spots of PP difference are localized, occurring as small patches in local areas. One notable example is the Permafrost_CCI PP map, which exhibits substantial overestimation along the southern margin of Siberia in northern Asia (red patches in Fig.5c). Moreover, all three source maps indicate a more northerly permafrost limit than that delineated by the IPA map that is shown as the background layer. This is consistent with findings reported in previous studies and may be attributed two factors: (1) satellite-based products capture finer spatial details than the IPA map that was developed largely based on expert knowledge; and (2) permafrost retreat since the publication of the IPA map that may date back to the 1950s in some regions. Our ensemble PP product provided an updated baseline representation of permafrost distribution while reducing the influence of uncertainties associated with individual source maps.*

*For MAGT, the three source maps also exhibit notable regional differences in their spatial patterns. For example, in central Siberia, the Obu map (Fig.5d) shows substantial underestimation relative to the ensemble product, whereas the NIEER map (Fig.5e) contains a concentrated cluster of strong overestimation. The Permafrost_CCI map (Fig.5f) displays several localized hot spots in northern Asia and cold spots in the High Arctic of northern Canada. By taking the median values from the three maps, our ensemble MAGT product reduced these region-specific uncertainties and biases associated with* *individual MAGT maps, thereby providing a more robust representation of permafrost thermal conditions across the northern high latitudes*."

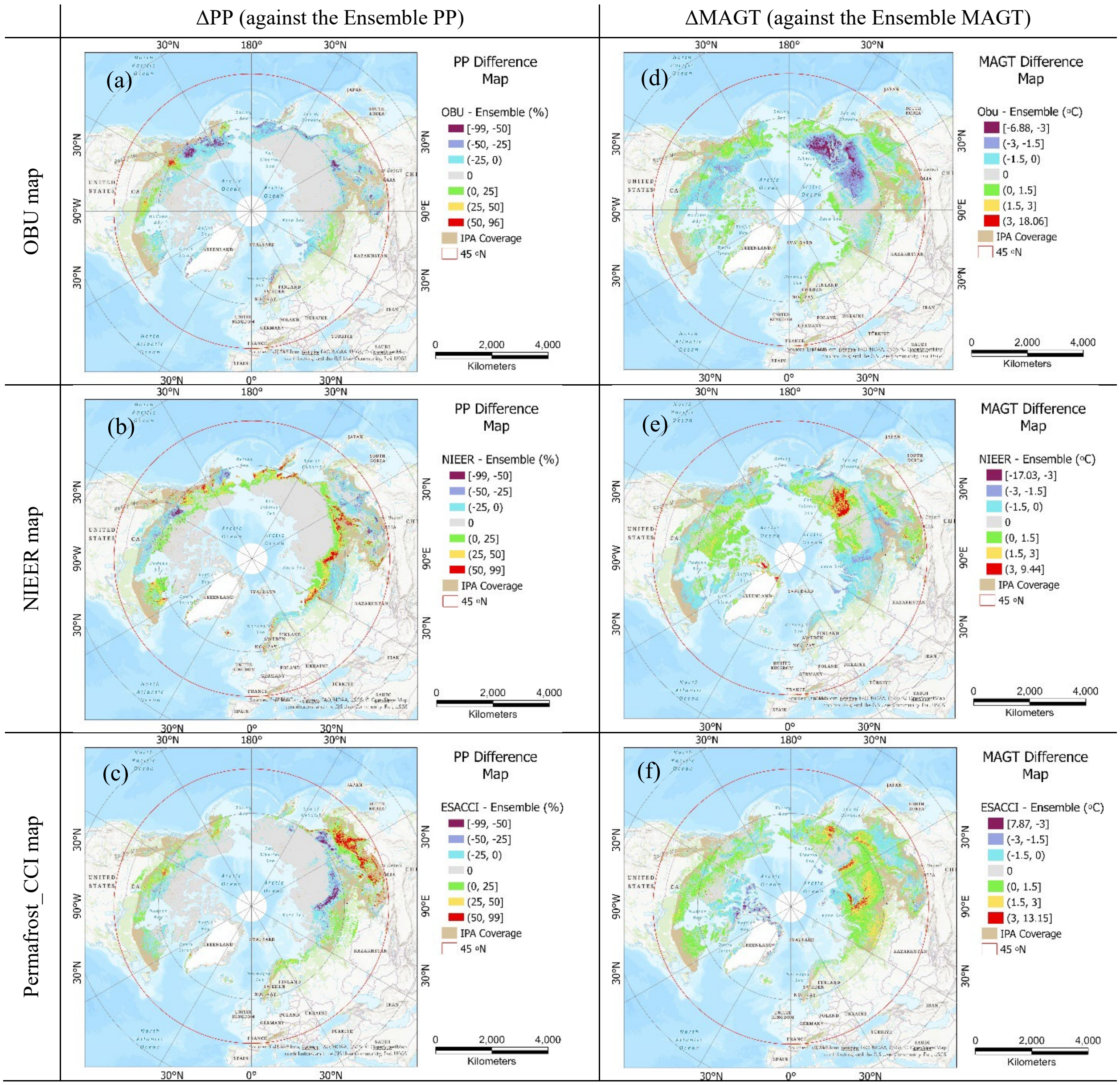



**Figure 5 Spatial differences of the three source maps (Obu, NIEER, Permafrost_CCI) relative to the ensemble product. The left panels (a–c) are the PP differences, and the right panels (d–f) are the MAGT differences. Areas in light gray (zero difference) represent locations where the ensemble value is selected. The light brown background outlines the IPA permafrost map.**

## 3. Inadequate validation

3.1 The authors project their model across the pan-Arctic but validate the predictions using only 26 sites. This is severely inadequate for evaluating a dataset of this scale. The stark contrast between the high internal model accuracy (91.8%) and the moderate rank correlation against independent borehole data (r = 0.69) suggests significant overfitting to the gridded covariates, raising critical concerns about the model's true predictive capability.

Please refer to our response to Comment #1.3 above that addresses the concern regarding the misleading model accuracy (91.8%).

Major revisions have been made to address the inadequate validation. When we initiated this study two years ago, we were only able to identify and access the GTN-P 2018 database, partly due to our limited experience. During the common discussion stage, Dr. Way referred us to the Nordicana D network, which prompted us to conduct a more extensive search for additional networks. Fortunately, several permafrost datasets were timely released in 2025, allowing us to trace more borehole records. In summary, four network sources were identified as useful for our study: the Nordicana D Network, the Swiss PERMOS Network, a European open-source dataset, and the extended GTN-P and CALM datasets released through PANGARA last year. In total, records from 96 borehole sites were compiled across the study region for validation.

Accordingly, we substantially expanded **Subsection 3.3.3** and renamed it "*3.3.3 Model performance and comparison with borehole records*". Figure 10 was also revised to include two panels: a distribution map of all borehole sitess (Fig.10a) and the revised scatterplot (Fig.10b). With the inclusion of a much larger in-situ dataset, the Spearman' r increased from 0.69 to 0.84, indicating stronger agreement between the borehole observations and the predicted PTI.

The complete revision is copied here for your reference:

> *"…This study searched all publicly available permafrost station networks to access those valuable observational records. To ensure a consistent representation of mean MAGT and MAGT|ALT trends, only stations with both MAGT and ALT records spanning at least three years were included in the analysis. Stations with only one of the two permafrost variables available were excluded.*
>
> *Globally, two major worldwide networks are available for monitoring permafrost conditions and changes: the Global Terrestrial Network for Permafrost (GTN-P, 2018; GTN-P, 2021) and the Circumpolar Active Layer Monitoring (CALM) program (Nelson et al., 2021). GTN-P was established by IPA in 1980 and currently involves 33 countries, with 1,487 permafrost-temperature sites and 256 active-layer-thickness sites using boreholes at various depths from 0 – 20 m. CALM was established in 1991 in a voluntary basis and currently involves 15 countries, monitoring more than 125 sites, approximately 60 of which measure ALT on grids ranging from 1 ha to 1 km$^2$. In 2025, GTN-P MAGT and CALM ALT datasets for the Northern Hemisphere were released through PANGARA (Wieczorek et al. 2025; Streletskiy et al. 2025), although the periods of record vary among sites, and GTN-P and CALM monitoring sites are often not geographically co-located. All stations located in the northern high latitudes (> 45$^o$N) were examined in this study. The two measurements were paired when the GTN-P and CALM sites were geographically close (while the collection years often vary). A total of 57 paired GTN-P/CALM stations were identified from the 2025 datasets. The distances between the paired GTN-P and CALM sites range from 0 to 47.02 km. For these stations, the annual MAGT and ALT values were obtained, from which the mean MAGT and both trends of MAGT and ALT were calculated.*
>
> *The GTN-P and CALM stations are operated by national and regional monitoring programs, through which additional ground-temperature measurements at more stations have been collected. Although MAGT and ALT records are not always directly available as for the GTN-P/CALM stations, many of*

*these borehole sites are equipped with thermistor cables that record ground temperatures at multiple depths in a daily or finer interval. Using these vertical borehole temperature profiles, we derived annual MAGT and ALT values as described below. MAGT was defined as the mean annual ground temperature at the depth of zero annual temperature amplitude. For each site, we examined its daily temperature profiles to identify this depth, which was approximated as the depth at which the difference between the maximum and minimum daily temperatures over a year was less than 0.15 °C. When no measurement depth satisfied this criterion, the deepest available measurement depth was used as a proxy for the zero-amplitude depth. Some stations only released monthly temperature profiles, which may result in larger uncertainties in the calculation. ALT was determined as the annual maximum depth of the 0°C isotherm, calculated by linear interpolation between measurement depths that bracketed the 0 °C isotherm within each daily thermal profile. Only days on which the shallowest sensor recorded a temperature above 0 °C and a positive-to-negative temperature crossing with depth was present were considered valid. The maximum interpolated depth of the 0 °C isotherm among all valid days within a year was taken as the annual ALT. Once the annual MAGT and ALT values were extracted for each site, the mean MAGT and both trends of MAGT and ALT were calculated.*

*Three regional monitoring networks were identified during the course of this study, providing additional 39 stations with paired MAGT and ALT:*

1) *Nordicana D Network, 2010-2022*

   *The Nordicana D8 SILA (Borehole and Near-Surface Ground Temperature) Network comprises 46 boreholes across the Nunavik–Nunavut region of Canada (Allard et al., 2024). Each thermistor-equipped borehole records ground temperature at multiple depths from 3 to 20 m, with data available at four temporal resolutions (hourly, daily, monthly, yearly). We utilized the daily records to extract the MAGT and ALT values. Among all boreholes in this network, only 19 provided temperature time series that met our data requirements. The remaining sites were excluded because of insufficient borehole depth, significant data loss, or anomalous records in daily temperature profiles.*

2) *Swiss Permafrost Monitoring Network (PERMOS), 1997-2024*

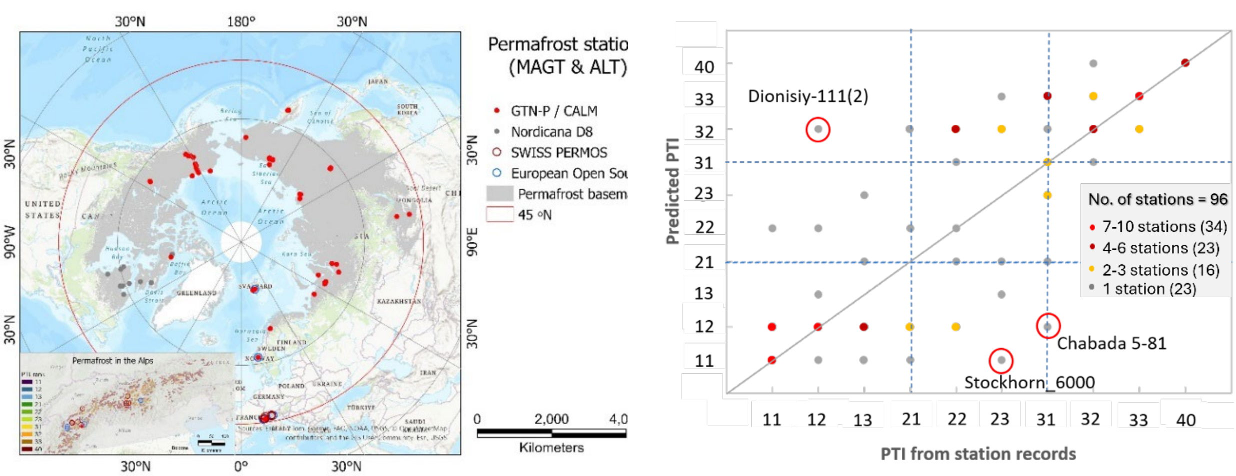
   *Published by the Swiss Permafrost Monitoring Network, this database contains a series of permafrost variables at established borehole sites in the Swiss Alps mountains (PERMOS 2025). Our study utilized its daily ground-temperature time series recorded by borehole thermistors. The database contains 23 borehole sites in Alps permafrost regions. The PERMOS sites are spatially clustered, with multiple boreholes often established within 1 km of one another. Some boreholes have included in the GTN-P network. In this case, to avoid spatial redundancy and duplication, we selected the borehole with the longest record to best represent permafrost conditions. After data screening, 10 borehole sites were retained. We extracted their annual MAGT and ALT values following the procedure described above.*

3) *Open-source European mountain permafrost ground temperature dataset (Noetzli et al. 2024a), 1996-2022*

   *Noetzli et al. (2024b) compiled a ground-temperature dataset from 64 boreholes located within or near European mountain permafrost regions, with observations collected from multiple national permafrost observation networks and institutions. The study reported a warming rate at a depth of 10 m exceeding 1 °C per decade in the region, suggesting accelerated ground warming and permafrost thaw in European mountains. Only monthly ground-temperature records are provided in this European open-source dataset. Among the 38 boreholes located within our study region, some sites are from established networks such as PERMOS and GTN-P, while some boreholes are located too close to each other (< 1 km). Because only monthly temperature profiles are available, accurately calculating the depth of zero annual amplitude is challenging. In the European Alps as observed in the PERMOS database, the borehole depths are mostly 10 or 20 m and the depth of zero annual amplitudes is commonly near the deepest borehole measurement point. Following this observation, we approximated*

*MAGT for each site in the European open-source dataset as the 12-month average of ground temperatures at the deepest available depth. ALT values were obtained from nearby CALM stations. The distances between the ground-temperature site and CALM site ranged from 0 to 41.77 km. After removing duplicate or closely spaced sites, 10 boreholes with data of sufficient quality were retained for our analysis.*

*From these global and regional networks, 96 stations with valid borehole records were retained for quality assessment of the predicted PTI map (Fig.10). Fig.10a displays their spatial distributions across the permafrost basemap, with a minimum distance of 1 km between any two stations. The PTI ranks at these stations were extracted using the same ranking system as presented in Table 3. Their PTI ranks ranged from 11 to 33 because borehole records did not consider Rank 40. For stations with Rank 33 but located within the isolated patches delineated by the IPA map, the PTI rank was re-assigned to 40. In the scatterplot of the station-extracted PTI and the predicted PTI (Fig.10b), the PTI at 40 out of 96 points matched exactly (lying on the 1:1 line) and 25 points differed by only one rank (close to the 1:1 line). The comparisons in Fig.10b yield a ranked correlation coefficient (Spearman's r) of 0.84, which is statistically significant based on a student's t test ($p < 0.001$). The strong goodness of fit indicates the validity of the predicted PTI map for assessing permafrost thaw conditions of permafrost in the study region. Three stations, as marked in Fig.10b, exhibited the largest discrepancies between the station-measured and predicted PTI values.*




(a) (b)

**Figure 10 Borehole sites used for quality assessment of the predicted PTI (a) and scatterplot comparing the borehole-derived and predicted PTI values (b). The inset in (a) shows the borehole sites in the Alps overlaid on the PTI map. Coloured dots in (b) indicate multiple stations.**

*Two stations revealed substantial underestimation of permafrost thaw by the predicted PTI. Chabada 5-81, registered as GTN-P #1115, had a station-derived PTI rank of 31, compared with a predicted PTI rank of 12. The station lies in a low-elevation, flat alluvial plain of the Lena River near Yakutsk, Russia. In the 2025 GTN-P database, this borehole has a long 36-year monthly record (1982–2018) with an average MAGT of -0.597 °C and no apparent trends on either MAGT or ALT. It is noteworthy that a nearby station, Chabada 8-82 (GTN-P #0950), is located only 4 km to the northeast. This station had a similar observation period but a substantially lower MAGT of -2.523 °C, also without apparent trends on MAGT or ALT, resulting in a station-measured PTI rank of 21. This rank is more consistent with the predicted PTI rank of 12. Based on the environmental maps, both stations are situated within the extensive Lena River Floodplain, characterized by relatively homogeneous, deciduous coniferous forest within the East Siberian taiga. The predicted PTI map (Fig.9a) similarly shows an extensive homogeneous area classified as Rank 12. Therefore, although the model underestimated thaw* *conditions at Chabada 5-81, the predicted PTI appears to reasonably represent the broader regional pattern of permafrost thermal conditions.*

*In contrast to the Chabada stations, Stockhorn_6000 from the European open-source dataset represents alpine permafrost at an elevation of 3,410 m in the Swiss Alps. Based on its 20-year monthly records (2001-2022), we calculated an average MAGT of -2.428 °C with trends observed on both variables (PTI = 23). These observations are generally consistent with other borehole records in this region. The predicted PTI, however, identified a small patch of PTI Rank 11 at the location of this station. As shown in the inset of Fig.10a, permafrost in the Alps is distributed at high elevations and is characterized by greater thaw potential (predicted PTI ranks 32–40). The surrounding ecoregion is European-Mediterranean montane forest, with dominant land covers ranging from grasslands to barren tundra and glaciers. This station is located within a transition zone of barren tundra and glacier, which may attribute to the discrepancy between the borehole-derived and satellite predictions. Further investigation using localized, fine-scale environmental data is needed to determine where such extremely isolated patches of highly stable permafrost are present, or whether the discrepancy reflects limitations in the spatial resolution of satellite open data.*

*One station revealed a substantial overestimation of thaw. Dionisiy-111(2), recognized as GTN-P #1710, is located on the Chkotka Peninsula in northeastern Russia. The Arctic peninsula is predominantly covered by herbaceous and low shrub tundra within the Bering tundra ecoregion. The borehole site is situated on an isolated massif of Mount Dionisiy rising above the surrounding Anadyr lowland. Based on its 26-year annual records (1987-2012), the site had an average MAGT of -4.0 °C with increasing MAGT trend but no ALT trend. Its station-derived PTI was 12, whereas the predicted PTI rank was 32 that was relatively uniform across the peninsula. This discrepancy may attribute to the site's localized topographic conditions, which may not be adequately represented by the spatial resolution of satellite open data in this study.*

*Overall, these three stations demonstrated that local environmental conditions can substantially influence the PTI rankings. Additionally, where trends were detected, the satellite-based MAGT trends consistently indicated warming across the study region, in agreement with the widespread permafrost warming observed globally (Biskaborn et al., 2019). However, limited studies have suggested that high-elevation permafrost may exhibit short-term cooling trends driven by local factors such as snow and topography (Luetschg et al., 2008). Among the 96 stations examined in this study, a few sites also exhibited weak decreasing MAGT trends. Future studies incorporating finer-scale local environmental variables may help explain these site-specific MAGT trends and improve the representation of local permafrost conditions.*

3.2 Furthermore, as a data-driven algorithm, XGBoost lacks explicit physical constraints (such as heat conduction equations or soil-water characteristic curves). It is therefore prone to failure in discontinuous zones, alpine permafrost, or other complex environments.

We fully agree. XGBoost is a machine learning algorithm instead of a physics-based model. Its extrapolation capability depends on the environmental conditions represented bu the explanatory variables in the training data. Therefore, its predictive performance may decrease in heterogeneous environments that are poorly represented in the training dataset.

We sought to mitigate this limitation by constructing a big training dataset comprising 61,325 samples. A stratified random sampling design was adopted to ensure a balanced representation of sample points, such that the PTI in complex environments including sporadic permafrost, isolated patches and alpine permafrost were adequately represented in the training samples.

In our model development, all explanatory variables were standardized to ensure consistency in their numerical scales and remove the outliers. This procedure is described in the manuscript as:

"*…Each variable was standardized to a range of [0,100] using a min-max linear stretch. Exceptionally, the ensemble land cover map remained categorical, retaining its original class codes 10–220 (21 classes). The peat map was re-coded as 0 (no peat), 50 (peat in soil) and 100 (peat dominated)…*"

In **Subsection 3.4** "*Satellite-based open data for permafrost research*", a piece of discussion was added:

"*However, XGBoost is a data-driven algorithm. Rather than explicitly solving equations governing heat transfer, hydrological processes, or soil freezing/thawing, its extrapolation capability depends on the environmental conditions represented by the explanatory variables in training samples. Its predictive performance may decrease when the explanatory variables are not sufficiently representative, or when the training dataset inadequately captures the complex and heterogeneous environments such as sporadic permafrost, and isolated patches, and alpine permafrost zones.*"

## 4. Specific issues (minor comments)

Thank you for pointing out these grammatical errors. We have corrected them accordingly. In addition, we carefully proofread the manuscript multiple times and did our best to identify and correct other grammatical errors. The formatting styles were also standardized. For example, the "Northern Hemisphere" was capitalized throughout the manuscript.

We also thoroughly revised the verb tenses to ensure consistency. The past tense was used to describe our data analysis conducted and results obtained in this study, whereas the present tense was used to describe the three published permafrost products, as they are established and publicly available datasets.

For the comment about Figure 2 – "Please use a consistent scale to better highlight the differences and changes between the different diagrams". We may have misunderstood this comment, but Figure 2 (as well as all figures in this manuscript) already uses a constant scale across all four diagrams (panels). A constant legend was also applied to facilitate direct comparison. The only difference is the minimum and maximum values within each map. These values differ because they reflect the actual ranges of map values. We would appreciate further clarification if you intended a different aspect of the figure.