Response to Referee #2

We appreciate you for your comments concerning our manuscript entitled "Permafrost temperature baseline at 15 meters depth in the Qinghai-Tibet Plateau (2010–2019)" (MS No.: essd-2024-114). Those comments are valuable in helping us improve the quality of the manuscript. We have carefully addressed all the points raised and revised the manuscript accordingly. Changes made are highlighted in blue in the revised version. Line and figure numbers refer to the updated manuscript, and a detailed point-by-point response to your comments is provided below.

General comments

1. Zou et al. present a dataset that extrapolates ground temperatures over the QTP at the depth of zero annual amplitude (here determined to be at 15 m depth). They use a support vector regression to predict ground temperatures based on nine environmental predictors. They justify this approach by claiming that this method has been shown to be superior to other supervised learning algorithms such as random forest in one study (Ran et al., 2021). While the dataset is novel in the sense that no ground temperatures at 15 m depth have been predicted with this method in the QTP, I have a few concerns about the methods used to create the dataset and the fact that a similar dataset exists on a pan-Arctic scale for the entire permafrost region through the permafrost cci ground temperature dataset. Dismissing this dataset solely on the grounds of it not reaching as deep as the dataset presented in this study is not sufficient in my opinion. Especially considering the fact that the authors claim that the DZAA ranges from 10 to 15 m in central Asia and therefore would partially be covered by the permafrost cci product. Furthermore, the R2 value of the prediction is below 0.5, meaning that less than half of the variance in ground temperature can be explained by the model. This suggests that the model could potentially be improved or a different model should be tested to see if the predictions accuracy can be increased.

Response:

We fully agree with your comment that interpreting 15 m as DZAA is evidently illogical, which was also pointed out by the other reviewer. This misunderstanding may have resulted from unclear description in our writing, as outlined below:

"*The data of MAGT at 15 m in depth are used for spatialization, considering that DZAAs generally ranges from 10 to 15 m in the QTP (Zhao et al., 2010).*"

However, the objective of our study was not to generate a map of $MAGT_{DZAA}$. To our knowledge, Ran et al. (2021, 2022) have already conducted comprehensive mapping of $MAGT_{DZAA}$ across various regions, including the QTP and the Northern Hemisphere, significant contributing to the spatial analysis of $MAGT_{DZAA}$. In contrast, our study aims to provide a fixed-depth deeper ground temperature map to support the permafrost thickness estimation, thereby avoiding the spatial variability challenges inherent to DZAA.

The selection of a depth of 15 m is based on two primary considerations. Firstly, it corresponds with the observed DZAA depth range of 10-15 m in the QTP, with the choice of the lower end of this range aimed at enhancing the stability of ground temperature readings, this particularly beneficial in areas with limited observational depth and data availability. Secondly, this depth corresponds with the extent of existing boreholes, thereby facilitating the integration of a larger dataset into the mapping process.

Therefore, the $MAGT_{15m}$ presented in this study does not overlap with the CCI products. In terms of depth, the CCI data provides permafrost temperatures at a maximum depth of 10 m, while our study presents data at a depth of 15 m. Additionally, the CCI products cover the permafrost regions of the Northern Hemisphere north of 30 °N, excluding the southernmost permafrost areas of the QTP.

Although our $MAGT_{15m}$ map exhibits similar spatial pattern to existing $MAGT_{DZAA}$ maps, it is fundamentally distinguished by its theoretical framework within permafrost research. The primary difference lies in depth: DZAA varies spatially, whereas $MAGT_{15m}$ map represents a fixed depth.

The adoption of $MAGT_{DZAA}$ as a reference and introduction for our $MAGT_{15m}$ mapping stems from the relative scarcity of research on deep permafrost temperatures. $MAGT_{DZAA}$ stands out as one of the few indexes with significant advancements in this area. In contrast, studies on deeper ground temperatures based on observed data are lacking, primarily due to the challenges of acquiring such observations. We have made substantial efforts to collect and compile the $MAGT_{15m}$ data for the period 2010-2019 to support the completion of this study.

To avoid any ambiguity, we have revised the sentence as follows (Line 58-61):

"*This study aims to establish a fixed-depth deep permafrost temperature baseline using data from the QTP for a decade (2010-2019) and a machine learning approach to address the limitations associated with the use of $MAGT_{DZAA}$. Considering the availability of ground temperature records, the data of MAGT at 15 m in depth are used for spatialization.*"

The lower R^2 of the $MAGT_{15m}$ predictions in comparison to $MAGT_{DZAA}$ may be attributed to the greater depth of temperature prediction, in the context of same method and similar environmental variables. Observations reveal that DZAA in the QTP predominantly occurs at depths shallower than 15 m, especially in areas close to the permafrost boundary, where DZAA are often even shallower. For instance, in the Xidatan area, located at the northern boundary of the QTP, the DZAA is recorded to be approximately 5 to 7 meters (Liu et al., 2021). DZAA represents the maximum depth that seasonal surface temperature fluctuations can reach, and the $MAGT_{DZAA}$ values are closely related to the climatic conditions of nearby years. Utilizing contemporary or recent air temperature or ground surface temperature data (such as FDD and TDD) in predicting spatial distributions generally yields higher R^2 values.

In comparison, the $MAGT_{15m}$, due to its greater depth, is more closely linked to long-term climatic conditions, as the propagation of temperature exhibits a lag effect. In other words, the increased depth of the strata is likely the primary factor contributing to the lower R^2 of the $MAGT_{15m}$ predictions. Nevertheless, a significant relationship exists between the predicted and observed $MAGT_{15m}$ values ($p<0.001$) in this study, and both bias and RMSE, along with their standard deviations, are slightly lower than those reported in previous studies. Considering these thermal propagation characteristics, we have extended the periods for FDD and TDD to 2003-2019 and have calibrated these metrics based on observed GST data to enhance the representativeness of surface temperature variables.

I do not want to dismiss the work that the authors have put into this dataset, however I am unsure if it offers a significant contribution to the scientific community in its current state. I have a few suggestions on how to enhance the impactfulness of the paper, but I am unsure if it then still fits the scope of ESSD.

1.The SVR method has been tested by Ran et al., 2021 and found to be sufficient for their purposes. However, their R2 was 0.71 as compared to 0.48 in this study. Further, they have tested various different supervised learning algorithms to conclude that SVR is the best model to use, which is lacking in the present manuscript. Hence, I would suggest the authors also perform a test for the other models in question that can be used for this task to get a better idea of their individual performance.

Response:

Before the initial submission of this manuscript, we had already tested the methods proposed by Ran et al. (2021). The results of these methods exhibited significant differences, both in terms of statistical metrics and spatial patterns. *Table R1* presents the performance of four statistical models.

In terms of R^2, the random forest (RF) model performs the best with a 0.92 value, and the R^2 of the generalized linear model (GLM) and generalized additive model (GAM) being comparable to that of the support vector regression (SVR) model (0.47-0.48). For bias and RMSE, the RF model shows the lowest values; the RMSE of the SVR method is slightly lower than those of the GLM and GAM. From the performance (*Table R1*), the RF is undoubtedly the best model. However, examining the spatial pattern of the RF-predicted $MAGT_{15m}$ (*Fig. R1.b*), most values in the permafrost regions are concentrated between -3 °C and -0.5 °C, with a minimum of only -3.2 °C, which is not consistent with observational facts. In addition, in several seasonally frozen ground regions, such as the Qaidam Basin and southern endorheic zones, the predicted $MAGT_{15m}$ falls below 0°C, suggesting the presence of permafrost, which contradicts existing permafrost distribution maps. The results produced by the RF model may be attributed to overfitting to the observational dataset. Although parameter adjustments can enhance certain aspects of the spatial pattern of $MAGT_{15m}$, the overall outcomes still fall short of expectations.

For linear-type models such as GLM and GAM, their performance is comparable to that of SVR. However, the predicted $MAGT_{15m}$ values often exhibit a seesaw effect, where lower values are predicted in high mountain areas and higher values at the permafrost margins. This seesaw effect becomes more pronounced when fewer variables are selected following collinearity analysis. After comparing the model performances and spatial patterns of different methods, we ultimately selected the SVR model for predicting $MAGT_{15m}$ in this study. Moreover, SVR is a deterministic prediction method, ensuring consistent and reproducible results with a fixed set of sample points. This choice aims to establish a methodological foundation for future analyses involving the addition of more sample points and comparisons across different input datasets.

Table R1: *Predictive performance of mean annual ground temperature at 15 m in depth ($MAGT_{15m}$) for four statistical models*.*

Performance	SVR	RF	GLM	GAM
R^2	0.48 (±0.14)	0.92 (±0.03)	0.47 (±0.13)	0.48 (±0.14)
Bias (°C)	-0.01 (±0.11)	-0.00 (±0.05)	0.01 (±0.12)	0.01 (±0.13)
RMSE (°C)	0.71 (±0.13)	0.32 (±0.05)	0.72 (±0.12)	0.72 (±0.12)

**SVR, support vector regression; RF, random forest; GLM, generalized linear model; GAM, generalized additive model. R^2, bias, and RMSE with 1 standard deviation.*

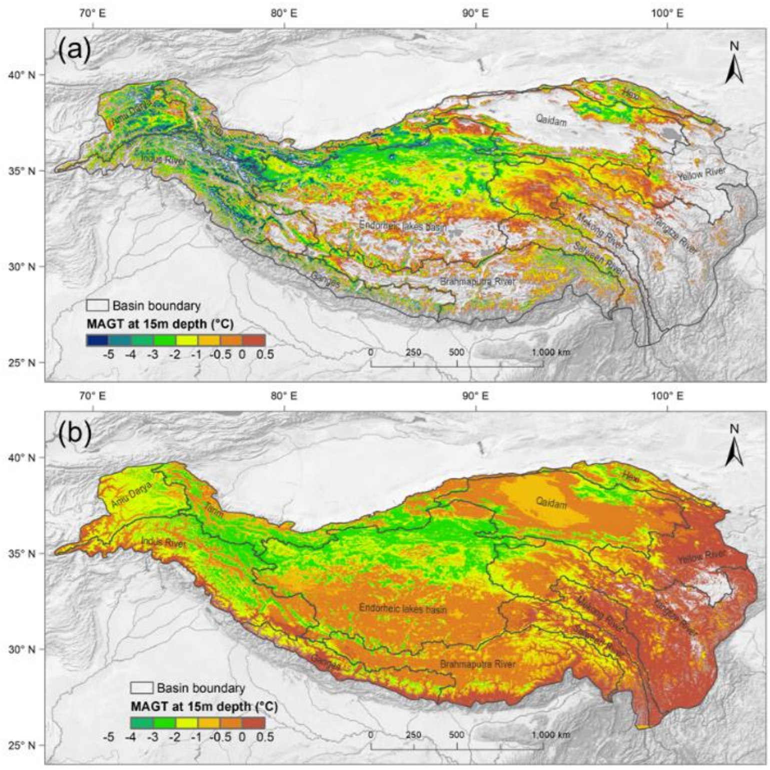

Figure R1: *Spatial distribution of predicted mean annual ground temperatures at the 15m depth ($MAGT_{15m}$) across the Qinghai-Tibet Plateau during 2010-2019, based on support vector regression (a) and random forest (b) models.*

2. Currently, the dataset is presented as a stand-alone dataset to be published in ESSD. However, the overlap with the existing permafrost cci ground temperature dataset can not be denied. My suggestion may significantly change the scope of the paper, but I wonder if it would make more sense to use the borehole data used in this study to assess how useful the ground temperatures could be to inform e.g., boundary conditions of permafrost models in the QTP. As the authors describe, the boreholes are equipped with thermistor strings, which probably means that measurements are available at several depths. This would serve as a basis to compare the borehole data directly to the ground temperature dataset at 10 m depth. A comparison to the existing dataset could then be a better motivation to conduct your own supervised learning method to improve the accuracy. However, if the R^2 is similar or higher when directly compared to the existing data (permafrost cci), there may not be a need for this since depth extrapolations of temperatures below the DZAA could be achieved with geothermal heat flux and simpler heat conduction models.

   Regardless of the scope of the final manuscript, I think a comparison to the existing datasets is crucial, considering the model in this study explains a relatively low amount of variance in the data.

Response:

As previously mentioned, the goal of this study is to map the fixed-depth deep ground temperature of permafrost on the QTP. Based on advancements in deep ground temperature research ($MAGT_{DZAA}$) and the availability of existing dataset, we selected 15 m as the mapping depth. This choice differs significantly from the CCI data, both in terms of depth and geographic coverage of the QTP. As an independent dataset, our results focus on the permafrost temperature at a depth of 15 m, which can serve as an upper boundary condition for future studies on deeper permafrost characteristics.

Due to the inability to establish a strict correspondence in depth, it is not appropriate to directly compare the results of this study with the CCI ground temperature data. While comparing borehole data with the existing 10 m depth CCI dataset is a valuable suggestion, it is somewhat outside the scope of this manuscript's objectives. However, we will consider conducting a separate evaluation in future research. Thank you very much for your insightful comments.

Modeling the regional thermal dynamics of permafrost beneath the DZAA remains to pose significant challenges for thermal conduction models. A major difficulty lies in assessing simulation uncertainty, which is one of the key motivations for adopting a fixed depth of 15 m for spatialization of ground temperature in this study. Our objective is to establish a baseline using observational data that can facilitate the comparison and evaluation of results produced by thermal conduction models.

Specific comments:

1. L56: What kind of datasets are you talking about here? Either delete the last part of

the sentence or give an overview (for example in a table) about the datasets you are talking about here.

Response:

We have added specific dataset name "ground temperature", and the revision is as follows (Line 56-57):

"*Over the past two decades, permafrost monitoring efforts on the QTP have established a substantial monitoring network and ground temperature datasets have been published (Zhao et al., 2021).*"

2. L75: Do I understand correctly that you implemented a procedure to fill temporal gaps in 78% of the data based on 22% of the observations? Please clarify.

Response:

Of the monitoring sites, 22% have maintained continuous observations over multiple years. Before establishing the relationships, we assessed an evaluation of these sites, which revealed that the $MAGT_{15m}$ range for this dataset was from -3.95 °C to 0.03 °C. This range effectively captures the essential spectrum of permafrost ground temperatures across the QTP and closely aligns with the observed thermal characteristics of permafrost in the region.

3. L77-82: From what I understand, you used 51 sites to calculate a linear trend to fill the gaps in the remaining 180 sites by assuming they all experience the same warming trend. However, your Fig. 2a clearly shows that warming trends are very different for cold vs. warm permafrost. I think applying a single warming trend that is based on 22% of the data is very problematic here. If I misunderstood this part, please clarify. Otherwise I am doubtful of the reliability of this preprocessing step.

Response:

We sincerely appreciate your detailed review; it is crucial to clarify that our methodology does not rely on a single warming trend for filling missing values. As demonstrated in *Fig. 2a*, there are significant differences in the warming rates between cold and warm permafrost; in general, cold permafrost tends to exhibit a more rapid warming rate, whereas warm permafrost warms at a comparatively slower rate. *Fig. 2b* illustrates the established relationship between $MAGT_{15m}$ values, organized from low to high temperature, and their corresponding warming rates. Based on this relationship, we subsequently calculated the warming rates for various monitoring sites using the observed $MAGT_{15m}$ values.

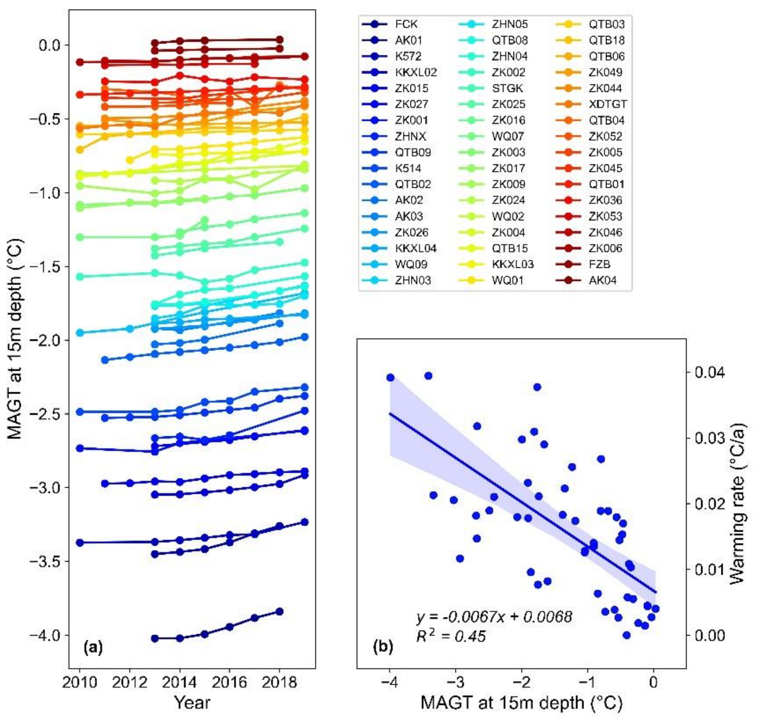

Figure 2: *Warming rates of $MAGT_{15m}$ during 2010-2019 (a) and the relationship between warming rates and the average $MAGT_{15m}$ (b).*

4. Eq 1. An R^2 of 0.45 does not create a lot of trust into your interpolation method (see comment above).

Response:

Although the R^2 value of 0.45 is relatively modest, statistical analysis reveals that the relationship between predicted and observed $MAGT_{15m}$ values is highly significant ($p<0.001$). At present understanding, the magnitude of $MAGT_{15m}$ is the dominant factor controlling the warming rate of $MAGT_{15m}$. However, in addition to $MAGT_{15m}$, the warming rate may also be closely related to permafrost characteristics (e.g., soil texture and ground ice content) and active layer properties (e.g., soil moisture and active layer thickness), as well as the magnitude of climate change. At this stage of the research, given the lack of more detailed or accurate site-specific observations of permafrost and its environmental characteristics, we primarily attribute the variations in the warming rate to differences in $MAGT_{15m}$.

5. L107: I am not very familiar with SVR, but is a 90/10 a typical split for this method? I was expecting a 80/20 or even a 70/30 split since you do not have a very large dataset. Can you provide the model performances with different splits? And how high is the risk for overfitting with the 90/10 split?

Response:

In the SVR method, a 90/10 split ratio is commonly used, as referenced in Ran et al.

(2021), and determined based on the sample size of this study. Considering your suggestion, we further evaluated model performance using 80/20 and 70/30 split ratios, as shown in *Table R2*. The R^2 and RMSE across the three split ratios (90/10, 80/20, and 70/30) ranged from 0.45 to 0.48 and 0.71 to 0.73, respectively, with a bias of -0.01 in all cases. These results indicate that there are no significant variances between the three split ratios when using the SVR method, thereby supporting the validity of the 90/10 split.

Table R2: *Predictive performance of the support vector regression (SVR) model across various split ratios.*

Split ratio (%)	R^2	Bias (°C)	RMSE (°C)
90/10	0.48 (±0.14)	-0.01 (±0.11)	0.71 (±0.13)
80/20	0.46 (±0.09)	-0.01 (±0.07)	0.72 (±0.08)
70/30	0.45 (±0.07)	-0.01 (±0.06)	0.73 (±0.07)

6. L132: "high accuracy" is inappropriate here. How do you determine it is "high"? The indicators you are describing are not creating a lot of confidence.

Response:

The term "high accuracy" was inappropriate, as you suggested, we have removed the relevant description, as follows (Line 137-139):

"The cross-validation of 1000 runs demonstrated that the mean values of the three statistical indicators, i.e., bias, root-mean-square error (RMSE), and coefficient of determination (R^2) were -0.01 °C (±0.11 °C), 0.71 °C (±0.13 °C), and 0.48 (±0.14), respectively."

7. Fig. 3: Please add a label for the red line either in the figure or in the caption.

Response:

We have added the label for the red line in *Fig.3* as per your suggestion.

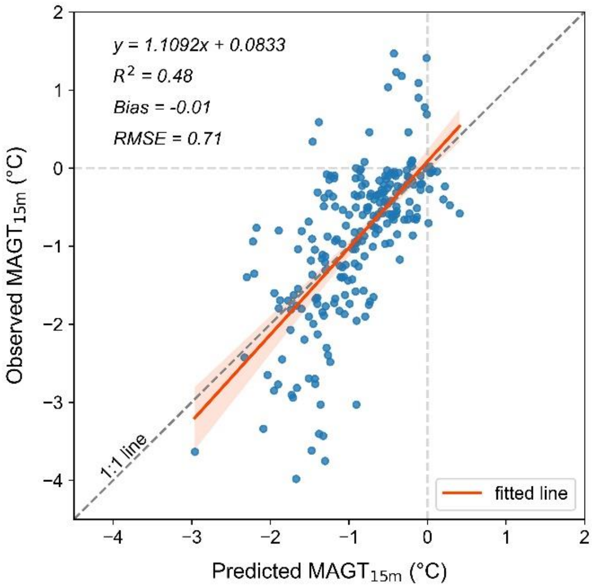

Figure 3: *Relationship between predicted and observed mean annual ground temperatures at 15 m depth ($MAGT_{15m}$) in permafrost regions on the Qinghai-Tibet Plateau during 2010-2019.*

8. Fig. 4: Maybe I have missed it in the text with all the numbers, but did you say that you are masking all values > 0°C? It looks like the final dataset only shows values < 0°C. Is that because you do not have confidence in non-frozen conditions? Are you assuming that there is no permafrost in regions with T > 0°C? Please clarify this throughout your results section.

Response:

In *Fig. 4* of the original manuscript, we have presented only values > 0°C to highlight the $MAGT_{15m}$ in the permafrost region of the QTP, which is the most significant result of this study. To show the prediction results for areas with positive temperatures, we included the complete set of predictions for the entire QTP in *Fig. R2*.

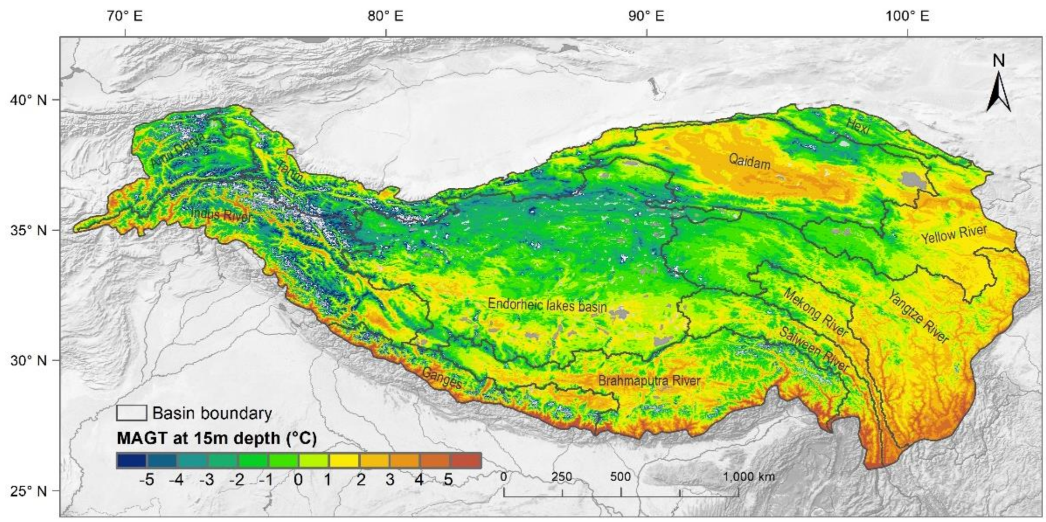

Figure R2: Spatial distribution of predicted mean annual ground temperatures at the 15m depth across the Qinghai-Tibet Plateau.

Positive $MAGT_{15m}$ does not necessarily means absence of permafrost because of extensive and increasing presence of supra-permafrost subaerial talik, especially to the east of the QTEC from Golmud-Lhasa and along the engineering lines. Thus, a criterion of subzero $MAGT_{15m}$ for judging the occurrence of permafrost may underestimate the permafrost extent. We have considered keeping some positive $MAGT_{15m}$ values to ensure coverage of most permafrost exist regions. However, due to the high variability in geothermal gradients of the permafrost base, determining an appropriate positive MAGT threshold proved challenging. After carefully reviewing both your comments and those of the other reviewer, we have followed the conventions of previous studies and retained regions with $MAGT_{15m}$ < 0.5 °C in the revised manuscript, to encompass areas where talik is more likely to be widespread. To ensure the reliability of permafrost temperature analysis, we did not reanalyze data with $MAGT_{15m}$ > 0 °C in the *Result section* of the revised manuscript. As an alternative, we have included a discussion of regions with $MAGT_{15m}$ > 0 °C, as outlined below (Line 255-257):

"*Additionally, permafrost may still persist in areas where $MAGT_{15m}$ exceeds 0 °C. Statistical analysis reveals that the areas with $MAGT_{15m}$ within the ranges of 0-0.1 °C and 0-0.2 °C cover approximately 0.05×10^6 km^2 and 0.10×10^6 km^2, respectively.*"

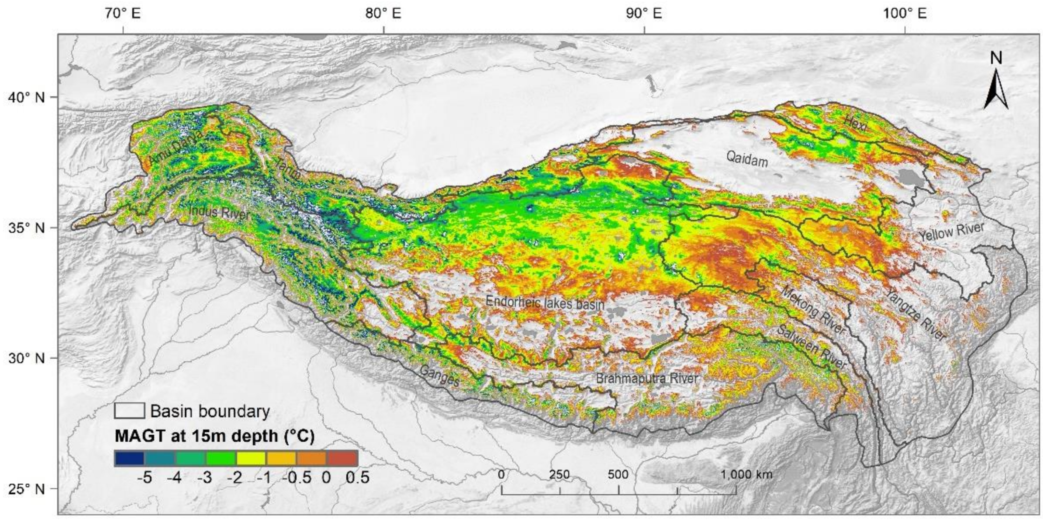

***Figure 4:** Spatial distribution of predicted mean annual ground temperatures at the 15m depth ($MAGT_{15m}$) across the Qinghai-Tibet Plateau during 2010-2019.*

9. Section 3.2.2: This section is very difficult to read. Would it be possible to put all those numbers into a table, refer to it in the text and focus on the conceptual characteristics only?

Response:

In the classification system based on $MAGT_{DZAA}$, permafrost can be divided into three types: cold (≤ -3.0 °C), cool (-3 to -1.5 °C), and warm (> -1.5 °C) permafrost (Ran et al., 2022). However, there are significant differences in both depths and values between $MAGT_{15m}$ and $MAGT_{DZAA}$, and using this classification system in the *Results* section may lead to confusion. Although we have not placed all the relevant numbers into a single table then directly referenced their conceptual characteristics, we have made efforts to simplify the numerical descriptions to enhance the readability of the text of the *Section 3.2.2*.

10. Fig 7.: What are the units in the figure legend? I assume °C?

Response:

We have added the legend name and unit (°C) in the revised manuscript.

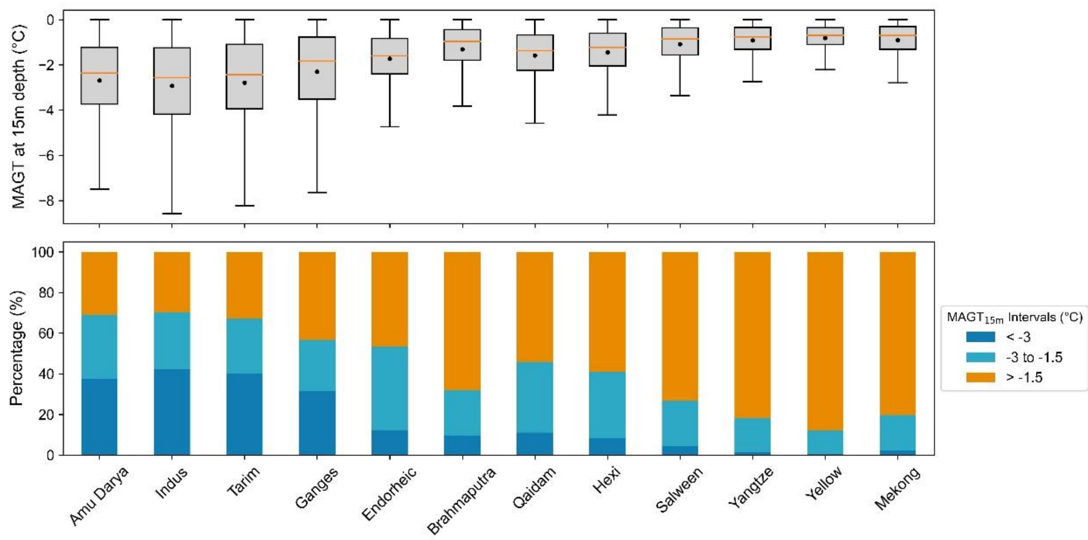

Figure 7: *Distribution (a) and percentage of area in three intervals (b) of MAGT at 15 m depth ($MAGT_{15m}$) in 12 basins of the Qinghai-Tibet Plateau during 2010-2019.*

11. L259-261: This sentence is very confusing and I am not able to follow it. Please see Biskaborn et al., 2019, which you are already citing, for an example on how to describe the difference between warming of "cold" and "warm" permafrost and how it relates to latent heat consumption.Might be worth citing Gruber 2012 (cited later in paper) in your introduction where you discuss TP permafrost maps.

Response:

Considering the incomplete conclusions, regional misalignment, and the style of the ESSD journal (also recommended by the other reviewer), we have decided to remove this section of text. This revision aims to maintain a clear focus on the QTP region and the data structure and functionality presented in the manuscript.

References:

Liu, G., Xie, C., Zhao, L., Xiao, Y., Wu, T., Wang, W., and Liu, W.: Permafrost warming near the northern limit of permafrost on the Qinghai-Tibetan Plateau during the period from 2005 to 2017: A case study in the Xidatan area, Permafr. Periglac., 32, 323–334, doi:10.1002/ppp.2089, 2021.

Ran, Y., Li, X., Cheng, G., Nan, Z., Che, J., Sheng, Y., Wu, Q., Jin, H., Luo, D., Tang, Z., and Wu, X.: Mapping the permafrost stability on the Tibetan Plateau for 2005–2015, Sci. China Earth Sci., 64, 62–79, doi:10.1007/s11430-020-9685-3, 2021.

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, doi:10.5194/essd-14-865-2022, 2022.