

## **Author’s Response on “*Brief Communication: Evaluation and comparisons of permafrost map over Qinghai-Tibet Plateau based on inventory of in-situ evidence*”**

Bin Cao<sup>1,2</sup>, Tingjun Zhang<sup>1</sup>, Qingbai Wu<sup>3</sup>, Yu Sheng<sup>3</sup>, Lin Zhao<sup>4</sup>, and Defu Zou<sup>4</sup>

5 <sup>1</sup>Key Laboratory of Western China’s Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China

<sup>2</sup>Department of Geography & Environmental Studies, Carleton University, Ottawa K1S 5B6, Canada

<sup>3</sup>State Key Laboratory of Frozen Soil Engineering, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

10 <sup>4</sup>Cryosphere Research Station on the Qinghai-Tibet Plateau, State Key Laboratory of Cryospheric Science, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

**Correspondence:** Tingjun Zhang(tjzhang@lzu.edu.cn)

15 The authors would like to thank Peter Morse and two anonymous reviewers for the constructive feedback, and the thorough assessment of the manuscript. Below we provide a point-to-point response to each comment, reviewer comments are given in black, responses are given in blue. Additionally, we have included details of how we addressed these changes in the revised submission.

### **Response to Anonymous Referee #1**

20 Permafrost maps were released by various institutes or research teams during the past several decades. They used modeling, statistical, and other mapping techs. Basically, the maps were evaluated during processing. However, the inter-comparison, what this study was done, is required for better understanding. This study collected more than a thousand samples over the QTP. The results of this study would be useful for future permafrost studies on the QTP  
25 and broad interest to the permafrost communities.

The manuscript, however, requires a bit more work before it is acceptable for publication. For the most part, the manuscript is well written but some editing is required to improve language and increase clarity. There are a few places in the manuscript where more explanation would be helpful.

30 **Response:** The language of revised manuscript was carefully checked.

Although I have made a few comments here that I hope the authors will find useful, dealing with them may not take too much time. The authors should thoroughly proofread the revised manuscript before submission or invite a native speaker in permafrost communities to improve the language. I am willing to review the revised paper.

35

## Major:

- **Unclear description and logic (to the following results) in the Data and Methods section.**

The authors used four methods to classify permafrost or not. However, it's not enough for understanding, although this paper is a short communication.

– How deep are generally for boreholes and soil pits? 1 m, 5 m?

Response: In general, the borehole depths vary from meters to hundred meters. In this study, we used the mean annual ground temperature from boreholes, which also varies from several meters to about 20 m, to identify permafrost presence. Number of samples measured from soil pits was small (6 samples) due to the prevalent coarse soil, and their depths are between less than 1 m to about 2.5 m. We added

*“In this study, we used the mean annual ground temperature (MAGT) measured from boreholes, which varies from meters to about 20 m to identify permafrost presence or absence. Due to the prevalent coarse soil, SP was only applied in areas possible, and the depth is from less than 1 meter to about 2.5 m.”*

to clarify.

Additionally, the survey depth of all the methods is summarized in Table A1 (see below)

– It looks like this study used only  $MAGST + TO_{max} \leq 0$  as the standard. In your results, you only talked about the sites considered as permafrost. I am not sure whether these classifications (P2, L25-29) are necessary.

Response: As mentioned by Referee #2 (specific comment Page 2, line 18), the word “certainty” is changed to “confidence”. Yes, the confidence classifications were not further used in this manuscript, but only present in the inventory as supplement. Since the inventory may be useful for other researches, we would keep the classification in the inventory and move the classification description into the Appendix A (see below). We hope you agree.

The classification algorithm of confidence degree largely follows Cremonese et al. (2011) and could be summarized as

***“Appendix A: Classification algorithm of in-situ permafrost presence or absence evidence***

*For board use of the permafrost presence or absence inventory, the data confidence degree was provided (Table A1). BH and SP provide direct evidence of permafrost presence or absence based on MAGT and/or ground ice observations, and hence have high confidence (Cremonese et al., 2011). The data confidence derived from MAGST is classified based on temperature and the length of the observation period. The evaluated GPR survey result was considered as medium confidence.”*

Table A1: Classification algorithm of in-situ permafrost presence or absence evidence from various methods

Method	Indicator	Survey depth	Permafrost	Confidence degree
BH	$MAGT \leq 0$	meters to about 20 m	presence	high
SP	ground ice presence	about 1.0–2.5 m	presence	high
GST	$MAGST \leq -2 \text{ }^\circ\text{C}$ & observations $\geq 3$	0.05 or 0.1 m	presence	medium
	$MAGST \leq -2 \text{ }^\circ\text{C}$ & observations $< 3$		presence	low
	$MAGST > -2 \text{ }^\circ\text{C}$ & $MAGST + TO_{max} \leq 0 \text{ }^\circ\text{C}$		presence	low
	$MAGST < 0 \text{ }^\circ\text{C}$ & $MAGST + TO_{max} > 0 \text{ }^\circ\text{C}$		ambiguous	–
	$MAGST > 0 \text{ }^\circ\text{C}$		absence	medium
GPR	clear permafrost reflection	about 0.80–5.0 m	presence	medium

BH = borehole temperature, SP = soil pit, GST = ground surface temperature, and GPR = ground-penetrating radar.  $TO_{max}$ , the maximum thermal offset under natural conditions reported for the QTP, is  $0.79 \text{ }^\circ\text{C}$ .

– What is kind of antennas generally used in GPR survey? Also, how deep is accessed?

Response: The GPR survey was conducted using 100 and 200 MHz antennas and evaluated using direct measurements (e.g., mechanical probing, soil temperature, and soil pits) (Cao et al. , 2017). The survey depth was from about 0.8 to near 5 m depending on the active layer thickness. Authors added

“Here, GPR data from Cao et al. (2017) are measured using 100 and 200 MHz antennas depending on the active layer thickness. The GPR survey depth is from about 0.8 to near 5 m, and the data are considered as indicating the presence of permafrost only if an active layer thickness (or a clear permafrost reflection) could be established.”

to clarify.

– In section 2.3, you used DEM (3 arc second), MAAT (1 km), MASCD (~500 m), and NDVI (~250 m). I guess you extract those variables for each site in your inventory. Is it?

Response: First of all, it is moved to Section 2 as Referee #2 suggested. Yes. We extract these variables to sample sites using nearest interpolation. We added

“These climate variables were extracted to in-situ sites/plots based on nearest interpolation.”

to clarify.

You also said (P5, L2-4) “Where original field evidence of permafrost presence/absence is located within the same grid cell (30 arcsec, 1 km), they were aggregated based on their major value. For a grid with one permafrost site and one non-permafrost site, the nearer site from the grid center was used to represent the grid.” (actually, these sentences should be moved to section 2.3). Why did you have to aggregate these in-situ data to 1 km?

Response: Yes, they are located within the same grid cell of unprojected SRTM30 with a spatial resolution of 30 arcsec. The aggregation was deleted in the revised manuscript.

How did you deal with DEM, MASCD, and NDVI? Did you upscale DEM, MASCD, and NDVI to 1 km? I guess you were going to avoid conflict sites (permafrost and non-permafrost) in the same pixel. Is it?

Response: No. The 3 arcsec DEM was used to simulate the slope and aspect for the in-situ sites. The MASCD,

NDVI, and MAAT are used here to explore the representative of the inventory, and they are extracted to the sites based on nearest interpolation.

When you extracted values from different spatial resolution datasets, even if there are probably few sites in the same pixel at 1 km resolution, however, there still are three spatial datasets with higher resolution, which might bring different snow, topography, and vegetation condition to your sites. In fact, there might be different ground thermal states under the same climate and vegetation condition because of different soil wetness, soil properties, and so on. Overall, I don't think the aggregation is necessary.

Response: Yes, we agree. We omitted the evidence aggregation and conduct the evaluation with all the 1475 sites. Please note that, some statistics may be slightly different by using the original 1475 evaluation sites/plots.

Furthermore, how did you compare with the maps with different spatial scale, e.g., QTP<sub>Noah</sub> map is 10 km. Those issues were confusing and should be clarified.

Response: The evaluation was conducted at the sites we collected. Permafrost presence and absence information at evaluation sites was extracted to the evidence based on nearest from different maps. In Section 2.4 Statistics and evaluation of permafrost distribution maps, we added

*“The permafrost and absence information was extracted to in-situ sites, and...”*

to clarify.

- Misleading indicators.

PCC<sub>PF</sub>, PCC<sub>NPF</sub>, and PCC<sub>tol</sub> were used to quantify the classification accuracies of permafrost maps. To my sense, PCC<sub>PF</sub> and PCC<sub>NPF</sub> are not useful and may be misleading. When the map over-presents permafrost (i.e., much colder), PCC<sub>PF</sub> would be extremely close to 100%. Can we say this is much better? Vice versa. Thus, the description in Section 3.2 could be misleading, at least to me, and should be more cautious. I suggest removing those parts.

Response: Yes, we agree that PCC<sub>PF</sub> and PCC<sub>NPF</sub> are somehow misleading when we look at them separately without due care. On the other hand, these two indicators would be useful if they are jointly interpolated. As you mentioned, the high PCC<sub>PF</sub> together with low PCC<sub>NPF</sub> indicate the map over-presents permafrost. This information could not be indicated by either kappa coefficient nor PCC<sub>tol</sub>. For this reason, we would keep these three indicators. To reduce the misunderstanding, the PCC<sub>PF</sub> and PCC<sub>NPF</sub> are interpolated together throughout the manuscript, and the over- or less-presents permafrost was also present. Additionally, the PCC<sub>tol</sub> in Figure 2 was deleted to avoid confusion. In Section 3.2 Evaluation and comparison of existing maps, we added

*“The high PCC<sub>PF</sub> together with low PCC<sub>NPF</sub> for the IPA, QTP<sub>Noah</sub>, PZI<sub>cold</sub>, and QTP<sub>TTOP</sub> maps indicate permafrost is over-presented by them, while the PZI<sub>warm</sub> and PZI<sub>norm</sub> maps showed underestimated the permafrost over the QTP.”*

Meanwhile, do you consider the effect of the different sample volume? Because in your in-situ sites pool, number of sites with permafrost is twice as large as the sites without permafrost.

Response: Yes, the kappa coefficient, “which measures inter-rater agreement for categorical items”, was introduced here as the major indicator for map evaluation as it could largely “avoid the impact of uneven distribution of sample numbers for permafrost presence and absence”.

- More discussion?

This study found different performance in permafrost maps. It's better to discuss a little bit more about the sources of bias, such as different MAAT products. More discussion on the possible sources of the revealed differences would enhance the scientific significance. Meanwhile, it also is useful for the future permafrost map updating.

Response: Yes, we agree. Our previous manuscript had partly discussed the bias from inputs for the  $QTP_{Noah}$  and IPA maps. We enhanced this part, and the inputs bias was discussed for each map as below:

*$QTP_{Noah}$  map: "Though the  $QTP_{Noah}$  map was derived using coupled land surface model (Noah), the relatively worse performance, especially for non-permafrost area ( $PCC_{NPF} = 45.9\%$ ), is likely caused by inputting coarse-scale forcing dataset ( $0.1^\circ$  resolution or  $\sim 10$  km) (Chen et al., 2011) and by the uncertainty of soil texture dataset (Yang et al., 2010)."*

*IPA map: "It is not surprising that the IPA map has fair agreement ( $k = 0.32$ ) as less observations were compiled and the method used are more suitable for high latitudes (Ran et al., 2012)."*

*$QTP_{TOP}$  map: "The  $QTP_{TOP}$  map was derived based on MODIS land surface temperature with different temporal coverage of 2003–2012 (Zou et al., 2017). Though the MODIS land surface temperature time-series gaps caused mainly by cloud were filled using the Harmonic Analysis Time-Series (HANTS) algorithm (Prince et al., 1998), the surface conditions, especially vegetation and snow cover, were ignored. In this case, land surface temperature is underestimated in high and/or dense vegetation area as it comes from the top of vegetation canopy, and is overestimated in snow covered area due to the cooling effects of snow is not considered. As a consequence, permafrost is likely overestimated in high and/or dense vegetation area and underestimated in regular snow-covered area."*

*$PZI_{global}$  map: "The MAAT used in the  $PZI_{global}$  map was statistical downscaled based on the lapse rate from the upper-air (or pressure level) temperature of NCEP, but the influences of land surface on surface air temperature, such as cold air pooling, was ignored (Cao et al., 2017). This is important as winter inversion is excepted to be common due to the prevalent mountains over the  $QTP$ . In other words, permafrost may be underestimated in valleys due to the overestimated MAAT."*

### Specific:

- P1, title: the title could be "Ground-based evaluation and inter-comparisons of permafrost maps over the Qinghai-Tibet Plateau"?

Response: We changed the title to

*"Evaluation and inter-comparisons of permafrost map over the Qinghai-Tibet Plateau based on inventory of in-situ evidence".*

As the study also provided the first inventory of permafrost presence or absence over the Qinghai-Tibet Plateau based on in-situ evidence, authors would like to reflect the inventory in the title. I hope you agree.

- P1, L3: the number, 1475, might be misleading although you collected. Because you aggregated to 1040, which excluded about 400 sites. Add a comma to 1040/1475 for consistency.

Response: The aggregation part was removed, and evaluation was conducted using all the data.

- P2, L1: “hemisphere” → hemispheric ?  
Response: Corrected.

- P2, L10: “2000” → “the 2000s”?  
Response: Corrected.

- 5
- P2, L16: insert “survey” after GPR.  
Response: Done.

- P2, L25-29: Where is so-call “high certainty” for permafrost classification? Meanwhile, it looks like this study used only  $MAGST + TO_{max} \leq 0$  as the standard. I am not sure whether these classifications are necessary. If necessary, the authors should clarify.

10 Response: As the Referee #2 mentioned, the word “certainty” is changed to “confidence”. The evidence derived from BH and SP is considered as high confidence as they provide direct information, such as mean annual ground temperature or ground ice presence. Yes. To determine permafrost presence or absence, only the function of  $MAGST + TO_{max} \leq 0$  °C was used. The confidence classifications were not used in this manuscript, but only present in the inventory as supplement. Since the inventory may be used for other related studies (e.g., permafrost simulation evaluation), and the confidence information would be useful for further selecting the data, we would keep the classification in the inventory and move the classification description into the Appendix A. Please also see our response to the major comments of “Unclear description and logic”.

15

- P3, L1: The authors should briefly clarify what kind of antennas were used and how deep is accessible.  
Response: The author added

20

“Here, GPR data from Cao et al. (2017) are measured using 100 and 200 MHz antennas depending on the active layer thickness. The GPR survey depth is from about 0.8 to near 5 m, and the data are considered as indicating the presence of permafrost only if only an active layer thickness (or a clear permafrost reflection) could be established.”

25

to clarify.

- P3, Section 2.2: It’s worth to note what climate data were used in  $QTP_{TTOP}$  and  $QTP_{Noah}$  maps. Both used the data merged MODIS temperature products and station data?

30

Response: As we mentioned in the previous submission,

“The most recent efforts were made by Zou et al. (2017) using the mean annual temperature at the top of permafrost (TTOP) model (referenced as  $QTP_{TTOP}$  map) forced by land surface temperature (or freezing and thawing indices) considering soil properties, and by Wu et al. (2018) based on Noah land surface model (referenced as  $QTP_{Noah}$  map) as well as gridded meteorological dataset (e.g., surface air temperature, radiation, and precipitation)”

35

The land surface temperature used in the  $QTP_{TTOP}$  map was calibrated based on ground observations or the station data, but only grid data was used by the  $QTP_{Noah}$  map. We changed this sentence to

40

“The most recent efforts were made by Zou et al. (2017) using the mean annual temperature at the top of permafrost (TTOP) model (referenced as  $QTP_{TTOP}$  map) forced by **calibrated (using station data)** land surface temperature (or freezing and thawing indices) considering soil properties, and by Wu et al. (2018) based on

*Noah land surface model (referenced as QTP<sub>Noah</sub> map) as well as gridded meteorological dataset, including surface air temperature, radiation, and precipitation.”*

to clarify.

- 5
- P3, L6: “(1)’ -> “(i)”  
Response: Done.
  - P3, L7: “(2)’ -> “(ii)”  
Response: Done.
- 10
- P3, L11: “... the temperature at ...” -> “...the mean annual temperature at...”  
Response: Done.
  - P4, L5: “... outline of QTP ...” -> “...outline of the QTP...”  
Response: Done.
  - P4, L24: Is the calculation of “Cohen’s kappa coefficient” too complicated? If not, please put equation(s) here and indicate what a high k means. Is there some threshold to roughly classify good, fair, or others?
- 15
- Response: “Cohen’s kappa coefficient” equations were added as

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (1)$$

where  $p_e$  and  $p_o$  are the probability of random agreement and disagreement, respectively, can be calculated as

$$p_e = \frac{(PF_T + PF_F) \times (PF_T + NPF_F)}{(PF_T + PF_F + NPF_F + NPF_T)^2} \quad (2)$$

$$p_o = \frac{(NPF_F + NPF_T) \times (PF_F + NPF_T)}{(PF_T + PF_F + NPF_F + NPF_T)^2} \quad (3)$$

Authors removed the kappa coefficient threshold description from footnote of Table 1 to this section (see below).

- 20
- ”Cohen’s kappa coefficient result is interpreted as excellent agreement for  $k \geq 0.8$ , substantial agreement for  $0.6 \leq k < 0.8$ , moderate agreement for  $0.4 \leq k < 0.6$ , fair agreement for  $0.2 \leq k < 0.4$ , and slight agreement for  $k < 0.2$ .”*
- P5, L12-13: What’s Qxx?  
Response: Did not see “Qxx”.
- 25
- P5, L15: Cao et al. (?), missing year.  
Response: It was revised to “Cao et al. (2018)”.
  - P6, L14: why is “-3 to -4 C”? Generally, -4 to -3 C?  
Response: It was revised to -4 to -3 °C.

## Response to Anonymous Referee #2

The manuscript presents a useful contribution for understanding performance of different permafrost maps at QTP. The aim of the study, methods and presented results are relatively clear, however, several parts of the text need to be clarified and part of the methods needs to be slightly extended. The manuscript has to be proofread for language and use of several terms in the manuscript can be improved. I have listed a number of specific comments below, which should improve the clarity of the text.

Response: The language of revised manuscript was carefully checked.

10 Authors should find the comments straightforward to implement.

### Specific comments:

- Page 1, line 4: change “overall accuracy of about” to “overall accuracy between”  
Response: Done.
- 15 • Page 1, line 5: omit “extremely large”. The areas are matter of scale and don’t need to be evaluated in this case. It is also not clear how this part of the sentence relates to the beginning where comparison to in-situ measurements is discussed.  
Response: Yes, they are compared in the manuscript rather than evaluated. We reformulated this part to
- 20 *“Many maps have been produced to estimate permafrost distribution over the Qinghai-Tibet Plateau, however, the estimated permafrost region ( $1.42-1.84 \times 10^6 \text{ km}^2$ ) and area ( $0.76-1.25 \times 10^6 \text{ km}^2$ ) are extremely large. The evaluation and inter-comparisons of them are poorly understood due to limited evidence.”*
- Page 1, line 6: How do you define “fragile landscapes”?  
Response: “fragile landscapes” means the areas where topography (mountains or valleys), surface conditions (e.g., vegetation cover, soil proxies, and river distribution) are spatially highly variable. The “fragile landscape” was replaced by “spatially highly variable landscape” to clarify.
- 25 • Page 2, lines 4-5: What is a large enough dataset? I assume that the evaluation datasets were large enough for the publications to be published. In the next sentence, “This would weaken their applications” sounds as the datasets were inappropriate. I would change the formulations of the both sentence to more positive. For instance: “The new larger dataset can be used to improve evaluations of the existing datasets, which would further improve their applications. . .”
- 30 Response: This part was changed to
- 35 *“Despite the increasing efforts made on permafrost mapping, existing maps over the QTP so far have not been evaluated and inter-compared with large data sets. A large amount of permafrost presence/absence evidence has been collected using a wide variety of methods (e.g., ground temperature, soil pits, and geophysics) on the QTP since the 2000s. The new larger dataset can be used to improve evaluations of the existing datasets, which would further improve their applications in permafrost and related studies, e.g., as a boundary condition for eco-hydrological model simulations.”*
- 40 • Page 2, line 16: The word evidence is used at many places in the manuscript. I’m not sure that its use is correct. It could be replaced by “information” in this case and maybe just a “validation site” elsewhere in the manuscript.  
Response: “Evidence” has been widely used for describing permafrost presence or absence “validation site”. I listed several published literatures using “evidence” below.



Cremonese, E., Gruber, S., Phillips, M., Pogliotti, P., Boeckli, L., Noetzli, J., Noetzli, J., Suter, C., Bodin, X., Crepaz, A., Kellerer-Pirklbauer, A., Lang, K., Letey, S., Mair, V., Morra di Cella, U., Ravanel, L., Scapozza, C., Seppi, R Kellerer-Pirklbauer, A. (2011). Brief Communication: "An inventory of permafrost evidence for the European Alps." *The Cryosphere*, 5(3), 651–657. <https://doi.org/10.5194/tc-5-651-2011>

5

Boeckli, L., Brenning, A., Gruber, S., & Noetzli, J. (2012). A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges. *The Cryosphere*, 6(1), 125–140. <https://doi.org/10.5194/tc-6-125-2012>

10 Schmid, M.-O., Baral, P., Gruber, S., Shahi, S., Shrestha, T., Stumm, D., & Wester, P. (2015). Assessment of permafrost distribution maps in the Hindu Kush Himalayan region using rock glaciers mapped in Google Earth. *The Cryosphere*, 9(6), 2089–2099. <https://doi.org/10.5194/tc-9-2089-2015>

We would keep "evidence" in the revised manuscript, and hope you agree.

15 • Page 2, line 18: The use of word "confidence" shall be used instead of "certainty" also further in the manuscript. Response: Yes, we agree. The "certainty" was changed to "confidence".

• Page2, line 25: What are your criteria to define confidence (certainty) classes medium and low? How are these classes used further in the manuscript?

20 Response: The confidence degree was described in the manuscript and available in the inventory as supplement, however, it was not further used for map evaluation. Since the inventory may be used for other related studies (e.g., permafrost simulation evaluation), and the confidence information would be useful for further selecting the data based on research aims, we would keep the classification in the inventory and move the classification description into the Appendix A (See below).

25 The classification algorithm of confidence degree largely follows Cremonese et al. (2011) and could be summarized as

***"Appendix A: Classification algorithm of in-situ permafrost presence or absence evidence***

30 "For board use of the permafrost presence or absence inventory, the data confidence degree was provided (Table A1). BH and SP provide direct evidence of permafrost presence or absence based on MAGT and/or ground ice observations, and hence have high confidence (Cremonese et al., 2011). The data confidence derived from MAGST is classified based on temperature and the length of the observation period. The evaluated GPR survey result was considered as medium confidence.

Table A1: Classification algorithm of in-situ permafrost presence or absence evidence from various methods

Method	Indicator	Survey depth	Permafrost	Confidence degree
BH	$MAGT \leq 0$	meters to about 20 m	presence	high
SP	ground ice presence	about 1.0–2.5 m	presence	high
GST	$MAGST \leq -2 \text{ }^\circ\text{C}$ & observations $\geq 3$	0.05 or 0.1 m	presence	medium
	$MAGST \leq -2 \text{ }^\circ\text{C}$ & observations $< 3$		presence	low
	$MAGST > -2 \text{ }^\circ\text{C}$ & $MAGST + TO_{max} \leq 0 \text{ }^\circ\text{C}$		presence	low
	$MAGST < 0 \text{ }^\circ\text{C}$ & $MAGST + TO_{max} > 0 \text{ }^\circ\text{C}$		ambiguous	–
	$MAGST > 0 \text{ }^\circ\text{C}$		absence	medium
GPR	clear permafrost reflection	about 0.80–5.0 m	presence	medium

BH = borehole temperature, SP = soil pit, GST = ground surface temperature, and GPR = ground-penetrating radar.

$TO_{max}$ , the maximum thermal offset under natural conditions reported for the QTP, is  $0.79 \text{ }^\circ\text{C}$ .

- Page 3, lines 4-5: How do you define a clear permafrost reflection? The exact criteria for selection of GPR sites should be presented.

Response: Cao et al. (2017) presented detailed description of GPR data acquisition and processing, here we used the data which active layer depth was identified, and could summarized as

“Here, GPR data from Cao et al. (2017) are measured using 100 and 200 MHz antennas depending on the active layer thickness. The GPR survey depth is from about 0.8 to near 5 m, and the data are considered as indicating the presence of permafrost only if an active layer thickness (or a clear permafrost reflection) could be established.”

to clarify.

- Page 3, line 9: The IPA map shows extent of four permafrost zones and is therefore not a binary map. Present here how did you convert it in to binary map showing permafrost presence and absence.

Response: Yes, the IPA map is categorical map rather than binary. Additionally, the  $QTP_{TTOP}$  and  $QTP_{Noah}$  maps are also categorical maps. The binary map was changed to categorical map throughout the manuscript. We changed this part to

“In general, permafrost maps over the QTP could be classified as (i) categorical, using categorical classification with different permafrost types (e.g., continuous, discontinuous, sporadic, and island permafrost), seasonally frozen ground, and unfrozen ground, and (ii) continuous, using continuous probability or indices [0–1] to represent proportion of an area that is underlain by permafrost.”

to clarify.

In Section 2.4, we also added

“For map evaluation, the categorical map was aggregated to binary map by merging different permafrost types to permafrost presence [1] and by merging the others to permafrost absence [0].”

- Page 3, line 16: Please explain here how PZIcold, PZIwarm and PZInorm were derived by Gruber (2012) and what is difference between them.

Response: As we mentioned in the previous manuscript, the  $PZI_{global}$  map is derived largely based on the

heuristic-empirical relationship between PZI and mean annual air temperature (MAAT) based on generalized linear models. The model parameters are established largely based on the boundaries of continuous (PZI = 0.9 for MAAT = -8.0 °C) and isolated (PZI = 0.1 for MAAT = -1.5 °C) permafrost in the IPA map and do not use field observations. The cold and warm cases were introduced into the map to allow the propagation of uncertainty caused by input dataset and model suitability, and they differ in the parameters used. Comparing the normal case, the cold and warm variants are derived by shifting PZI and MAAT at the respective limit by  $\pm 5\%$  and  $\pm 0.5$  °C, respectively. We changed this part to

*“The model parameters are established largely based on the boundaries of continuous (PZI = 0.9 for MAAT = -8.0 °C) and isolated (PZI = 0.1 for MAAT = -1.5 °C) permafrost in the IPA map and do not use field observations. Additionally, two cases, including cold (conservative or more permafrost) and warm (anti-conservative or less permafrost), were introduced into the map to allow the propagation of uncertainty caused by input dataset and model suitability. The three cases or maps, referenced as PZI<sub>norm</sub>, PZI<sub>warm</sub>, and PZI<sub>cold</sub> maps, differ in the parameters used. Comparing the normal case, the cold and warm variants are derived by shifting PZI and MAAT at the respective limit by  $\pm 5\%$  and  $\pm 0.5$  °C, respectively.”*

- Page 3, consider moving 2.3 section before 2.2 because it is in my opinion logical continuation of the inventory of permafrost validation sites. Also consider changing the section title to “Topographical and climatological properties of the inventory (or permafrost validation) sites”

Response: The section 2.3 was moved before 2.2, and the title was changed to

*“Topographical and climatological properties of the inventory sites”.*

- Page 3, line 32: What are you referring to with “(about 500m)”?

Response: It is the spatial resolution. The sentence was changed to

*“The MASCD with a spatial resolution of about 500 m was...”*

- Page 4, line 9: Please consider extending the explanation about the difference between permafrost area and permafrost region. This concept is difficult to understand by broader permafrost community. Maybe introduce the concept of scale and ground coverage by permafrost.

Response: This part was changed to

*“Permafrost region refers to regions where permafrost exists but the entire region is not necessarily completely occupied by permafrost, while permafrost area refers to areas where are completely underlain by permafrost. For example, discontinuous permafrost regions have permafrost area ranging from 50 to 90%. In other words, in discontinuous permafrost region, there 50 to 90% of the area is underlain by permafrost, i.e., permafrost area (Zhang., 2000; Gruber , 2012).”*

- Page 4, lines 26-27: Restructure the sentences. It sounds as because of your permafrost absence/absence classification, you have 1475 sites. I assume that this is because of your site selection criteria.

Response: The sentence was changed to

*“In the inventory, there are in total 1475 permafrost presence or absence sites/plots acquired from BH, SP, GST, and GPR methods (Figure 1).”*

- Page 5, line 3: “were aggregated based on their major value”. Maybe replace with “the majority value was assigned to aggregated sites.”  
Response: Done.
- 5 • Page 5, line 15: More appropriate term for “band” would be “range”. What exactly does the word “sensitive” refer to?  
Response: “band” was replaced by “range” throughout the revised manuscript. The sentence was deleted as it does not give us too much useful information.
- Page 5, line 31: Did you mean  $QTP_{TOP}$  instead of  $PZI_{TOP}$ ?  
Response: Yes, it was corrected.
- 10 • Page 6, line 10: Again, how exactly is fragile landscape defined?  
Response: The “fragile landscape” was replaced by “*spatially highly variable landscape*”.

## References

- Cao, B., Gruber, S., & Zhang, T. (2017). REDCAPP (v1.0): Parameterizing valley inversions in air temperature data downscaled from re-analyses. *Geoscientific Model Development Discussions*, 2017, 1–35. <https://doi.org/10.5194/gmd-2017-60>
- 15 Cao, B., Gruber, S., Zhang, T., Li, L., Peng, X., Wang, K., Zheng, L., Shao, W., Guo, H. (2017). Spatial variability of active layer thickness detected by ground-penetrating radar in the Qilian Mountains, Western China. *Journal of Geophysical Research: Earth Surface*. <https://doi.org/10.1002/2016JF004018>
- Chen, Y., Yang, K., He, J., Qin, J., Shi, J., Du, J., & He, Q. (2011). Improving land surface temperature modeling for dry land of China. *Journal of Geophysical Research Atmospheres*, 116(20), 1–15. <https://doi.org/10.1029/2011JD015921>
- 20 Cremonese, E., Gruber, S., Phillips, M., Pogliotti, P., Boeckli, L., Noetzli, J., Noetzli, J., Suter, C., Bodin, X., Crepez, A., Kellerer-Pirklbauer, A., Lang, K., Letey, S., Mair, V., Morra di Cella, U., Ravel, L., Scapozza, C., Seppi, R Kellerer-Pirklbauer, A. (2011). Brief Communication: An inventory of permafrost evidence for the European Alps. *The Cryosphere*, 5, 651–657. <https://doi.org/10.5194/tc-5-651-2011>
- Gruber, S. (2012). Derivation and analysis of a high-resolution estimate of global permafrost zonation. *The Cryosphere*, 6, 221–233. <https://doi.org/10.5194/tc-6-221-2012>
- 25 Prince, S. D., Goetz, S. J., Dubayah, R. O., Czajkowski, K. P., & Thawley, M. (1998). Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using advanced very high-resolution radiometer satellite observations: Comparison with field observations. *Journal of Hydrology*, 212–213(1–4), 230–249. [https://doi.org/10.1016/S0022-1694\(98\)00210-8](https://doi.org/10.1016/S0022-1694(98)00210-8)
- Ran, Y., Li, X., Cheng, G., Zhang, T., Wu, Q., Jin, H., & Jin, R. (2012). Distribution of Permafrost in China: An Overview of Existing Permafrost Maps. *Permafrost and Periglacial Processes*, 23(4), 322–333. <https://doi.org/10.1002/ppp.1756>
- 30 Wu, X., Nan, Z., Zhao, S., Zhao, L., Cheng, G. (2018). Spatial modeling of permafrost distribution and properties on the Qinghai–Tibet Plateau. *Permafrost and Periglacial Processes*, 23(4), 86–99. <https://doi.org/10.1002/ppp.1971>
- Yang, K., He, J., Tang, W., Qin, J., & Cheng, C. C. K. (2010). On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. *Agricultural and Forest Meteorology*, 150(1), 38–46. <https://doi.org/https://doi.org/10.1016/j.agrformet.2009.08.004>
- Zou, D., Zhao, L., Sheng, Y., Chen, J., Hu, G., Wu, T., ... Cheng, G. (2017). A new map of permafrost distribution on the Tibetan Plateau. *The Cryosphere*, 11(6), 2527–2542. <https://doi.org/10.5194/tc-11-2527-2017>
- 35 Zhang, T., and Heginbottom, J., Barry, R., Brown, J. (2000). Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere. *Polar Geography*, 24(2), 126–131. <https://doi.org/10.1080/10889370009377692>

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., Cheng, G. (2017). A new map of permafrost distribution on the Tibetan Plateau. *The Cryosphere*, 11(6), 2527–2542. <https://doi.org/10.5194/tc-11-2527-2017>

# Brief Communication: Evaluation and **comparisons** **inter-comparisons** of permafrost map over **the** Qinghai-Tibet Plateau based on inventory of in-situ evidence

Bin Cao<sup>1,2</sup>, Tingjun Zhang<sup>1</sup>, Qingbai Wu<sup>3</sup>, Yu Sheng<sup>3</sup>, Lin Zhao<sup>4</sup>, and Defu Zou<sup>4</sup>

<sup>1</sup>Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China

<sup>2</sup>Department of Geography & Environmental Studies, Carleton University, Ottawa K1S 5B6, Canada

<sup>3</sup>State Key Laboratory of Frozen Soil Engineering, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

<sup>4</sup>Cryosphere Research Station on the Qinghai-Tibet Plateau, State Key Laboratory of Cryospheric Science, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

**Correspondence:** Tingjun Zhang (tjzhang@lzu.edu.cn)

**Abstract.** Many maps have been produced to estimate permafrost distribution over the Qinghai-Tibet Plateau (QTP), however, the **evaluation and comparisons** estimated permafrost region ( $1.42\text{--}1.84\times 10^6\text{ km}^2$ ) and area ( $0.76\text{--}1.25\times 10^6\text{ km}^2$ ) are extremely large. The evaluation and inter-comparisons of them are poorly understood due to limited evidence. Using a large number data from various sources, we present the inventory of permafrost presence/absence with 1475 sites/plots over the QTP. Based on the in-situ measurements, our evaluation results showed a wide range of map performance with the **overall accuracy of about 59–82%, and the estimated permafrost region ( $1.42\text{--}1.84\times 10^6\text{ km}^2$ ) and area ( $0.76\text{--}1.25\times 10^6\text{ km}^2$ ) are extremely large** Cohen's kappa coefficient from 0.32 to 0.58 and overall accuracy between about 55–83%. The low agreement in areas near permafrost boundary and **fragile spatially highly variable** landscapes require improved method considering more controlling factors at both medium-large and local scales.

## 10 1 Introduction

Permafrost is one of the major components of the cryosphere due to its large spatial extent. The Qinghai-Tibet Plateau (QTP), known as the Third Pole, has the largest extent of permafrost in the low-middle latitudes. Permafrost over the QTP was reported to be sensitive to global warming mainly due to high temperature ( $< -2\text{ }^\circ\text{C}$ ) (Wu and Zhang, 2008), and its distribution has strong influences on hydrological processes (e.g., Cheng and Jin (2013)), biogeochemical processes (e.g., Mu et al. (2017)), and human systems (e.g., Wu et al. (2016)).

Many maps have been produced to estimate permafrost distribution and ground ice conditions at different scales over the QTP (Ran et al., 2012). Typically, frozen ground is classified into permafrost and seasonally frozen ground, and information on the extent, i.e. the areal abundance, of permafrost is available for some of the maps (Ran et al., 2012). These maps significantly improved the understanding of permafrost distribution over the QTP, however, limited in-situ measurements and the different

classification systems ~~and as well as~~ compilation approaches used make the comparison of maps a challenge. With the availability of high-resolution spatial data sets (e.g., surface air temperature and land surface temperature), several empirical and (semi-)physical models are now applied in permafrost distribution simulations at fine scales (Nan et al., 2013; Zhao et al., 2017; Zou et al., 2017; Wu et al., 2018). Additionally, the QTP was involved in ~~hemisphere-hemispheric~~ or global maps, e.g., the  
5 Circum-Arctic Map of Permafrost and Ground-ice Conditions led by the International Permafrost Association (referenced as IPA map) (Brown, 1997), and the global permafrost zonation (~~index (PZI) map~~ (referenced as PZI<sub>global</sub> ~~map~~) derived by Gruber (2012).

Despite the increasing efforts made on permafrost mapping, existing maps over the QTP so far have not been evaluated ~~with large enough~~ and inter-compared with large data sets. ~~This would weaken~~ A large amount of permafrost presence/absence evidence has been collected using a wide variety of methods (e.g., ground temperature, soil pits, and geophysics) on the QTP since  
10 the 2000s. The new larger dataset can be used to improve evaluations of the existing datasets, which would further improve their applications in permafrost and related studies, e.g., as a boundary condition for eco-hydrological model simulations. The global warming and increasing amount of infrastructure built on permafrost add both environmental and engineering relevance to investigating permafrost distribution, and makes studies of evaluating and comparing existing permafrost maps of great im-  
15 portance.

~~A large amount of permafrost presence/absence evidence has been collected using a wide variety of methods (e.g., ground temperature, soil pits, and geophysics) on the QTP since 2000.~~ In this study, we aim

1. to provide the first inventory of permafrost presence/absence evidence for the QTP;
2. to evaluate and ~~compare~~ inter-compare existing permafrost maps on the QTP, using the new inventory data.

## 20 2 Data and methods

### 2.1 Inventory of permafrost presence/absence evidence

Four methods, including borehole temperature (BH), soil pit (SP), ground surface temperature (GST), and ground-penetrating radar (GPR) ~~survey~~, were used to acquire evidence of permafrost presence or absence (Figure 1). ~~BH and SP provide direct evidence of permafrost presence or absence based on~~, Table A1). In this study, we used the mean annual ground temperature  
25 (MAGT) ~~and/or ground ice observations, and hence have high certainty (Cremonese et al., 2011)~~ measured from borehole, which varies from meters to about 20 m to identify permafrost presence or absence. Due to the prevalent coarse soil, SP was only applied in areas possible, and the depth is from less than 1 meter to about 2.5 m. GST, referred as soil temperature at the depth of 0.05 or 0.1 m here, was used to establish permafrost presence/absence for specific sites due to the MAGT could be  
30 derived as the difference of thermal offset and mean annual ground surface temperature (MAGST) (Hasler et al., 2015). While thermal offset is spatially variable depending on soil and temperature conditions, it is relatively small on the QTP compared with northern high latitudes environments due to prevalent coarse soil and low soil moisture content. The maximum thermal offset under natural conditions reported for the QTP is 0.79 °C (referenced as maximum thermal offset, TO<sub>max</sub>) (Wu et al.,

2002, 2010; Lin et al., 2015). In this study, sites with  $MAGST + TO_{max} \leq 0$  °C are considered as permafrost sites ~~and the confidence in this is classified based on MAGST and the length of the observation period:-~~

- ~~-  $MAGST \leq -2$  °C & observations  $\geq 3$  years: medium certainty;-~~
- ~~-  $MAGST > -2$  °C &  $MAGST + TO_{max} \leq 0$  °C: low certainty;-~~
- 5 ~~-  $MAGST < 0$  °C &  $MAGST + TO_{max} > 0$  °C: ambiguous, point removed from evidence collection;-~~
- ~~-  $MAGST > 0$  °C: medium certainty of permafrost absence.-~~

. The suitability of GPR for detecting permafrost derives from the dielectric contrast between liquid water and ice (Moorman et al., 2003), and it may face the challenge of distinguishing presence of permafrost in areas with low soil moisture content (Cao et al., 2017b). Here, GPR data ~~are~~ from Cao et al. (2017b) are measured using 100 and 200 MHz antennas depending on the active layer thickness. The GPR survey depth is from about 0.8 to near 5 m, and the data are considered as indicating the presence of permafrost ~~(medium certainty) only if~~ only if an active layer thickness (or a clear permafrost reflection) could be established ~~in a published study-~~. The detailed classification algorithm of in-situ permafrost presence or absence evidence could be found from Appendix A.

## 2.2 Topographical and climatological properties of the inventory sites

15 The slope and aspect of the inventory were derived from a DEM with 3 arcsec, which is aggregated from the Global Digital Elevation Model version 2 (GDEM2) by averaging to avoid the noise in the original dataset (Cao et al., 2017a). The thermal state of permafrost and its spatial distribution result from the long-term interaction of climate and subsurface. Additionally, vegetation and snow coverage play important roles in permafrost distribution through influencing the energy exchange between the atmosphere and the ground surface (Norman et al., 1995; Zhang, 2005) . In this study, three climate variables of MAAT, 20 mean annual snow cover days (MASCDC), and the annual maximum normalized difference vegetation index (NDVI) were hence selected here to test the representative of the inventory for permafrost map evaluation. The MAAT with resolution of 1 km is from Gruber (2012) representing the referenced period 1961–1990. The MASCDC with a spatial resolution of about 500 m was derived from daily snow cover product developed by Wang et al. (2015) based on MODIS products (MOD10A1 and MYD10A1). To improve the comparison of MASCDC, it is scaled to 0–1 through dividing the total days of a given year, 25 and the mean MASCDC during 2003–2010 was produced as a predictor. The annual maximum NDVI is from MODIS/Terra Vegetation Indices 16-day product (MOD13Q1, v006) with a resolution of 250 m. The annual maximum NDVI ( $NDVI_{max}$ ) was computed for each year during 2001–2017 to approximately represent the amount of vegetation, and then aggregated as a median for the entire period to avoid sensitivity to extreme values. These climate variables were extracted to in-situ sites based on nearest interpolation. The outline of the QTP is from Zhang et al. (2002), glacier outlines is from Liu et al. (2015) 30 representing conditions in 2010, and lake data is provided by the Third Pole Environment Database.



## 2.3 Existing maps over the QTP

Table 1 gives the summary of most widely used and recent developed permafrost maps over the QTP. In general, permafrost maps over the QTP could be classified as (i) ~~binary-categorical~~, using categorical classification with ~~permafrost-presence-for absence-0, and~~ (different permafrost types (e.g., continuous, discontinuous, sporadic, and island permafrost), seasonally frozen ground, and unfrozen ground, and (ii) continuous, using continuous probability or indices [0–1] to represent proportion of an area that is underlain by permafrost. The IPA map, which is may the most widely used ~~binary-categorical~~ map, was compiled by assembling all readily available data on the characteristics and distribution of permafrost (Ran et al., 2012). The most recent efforts were made by Zou et al. (2017) using the mean annual temperature at the top of permafrost (TTOP) model (referenced as QTP<sub>TTOP</sub> map) forced by ~~calibrated (using station data)~~ land surface temperature (or freezing and thawing indices) considering soil properties, and by Wu et al. (2018) based on Noah land surface model (referenced as QTP<sub>Noah</sub> map) as well as gridded meteorological dataset (~~e.g., including~~ surface air temperature, radiation, and precipitation). Though, these two ~~binary-categorical~~ maps are expected to be superior by using the latest measurements and advanced methods, they were evaluated using limited and narrow distributed data (~200 sites for the QTP<sub>TTOP</sub> map and 56 sites for the QTP<sub>Noah</sub> map). ~~On the other hand, the~~ The PZI<sub>global</sub> map, which gives continuous index for permafrost distribution, is derived through its heuristic-empirical relationship with mean annual air temperature (MAAT) based on generalized linear models (Gruber, 2012). The model parameters are established largely based on the boundaries of continuous ~~and isolated~~ (PZI = 0.9 for MAAT = -8.0 °C) and isolated (PZI = 0.1 for MAAT = -1.5 °C) permafrost in the IPA map and do not use field observations. Additionally, two cases, including cold (conservative or more permafrost) and warm (anti-conservative or less permafrost), were introduced into the map to allow the propagation of uncertainty caused by input dataset and model suitability. ~~A part of the QTP of the~~ PZI<sub>global</sub> ~~was evaluated~~ The three cases or maps, referenced as PZI<sub>norm</sub>, PZI<sub>warm</sub>, and PZI<sub>cold</sub> maps, differ in the parameters used. Comparing the normal case, the cold and warm variants are derived by shifting PZI and MAAT at the respective limit by ± 5% and ± 0.5 °C, respectively. The PZI<sub>global</sub> map was partly evaluated for the QTP using rock glaciers, considered as indicators of permafrost conditions, based on remote sensing imagery (Schmid et al., 2015). Rock glaciers, however are of absence in much of the QTP due to very low precipitation (Gruber et al., 2017).

## 2.4 ~~Climate variables and topography~~

~~The slope and aspect of the inventory were derived from a DEM with 3 arcsec, which is aggregated from the Global Digital Elevation Model version 2 (GDEM2) by averaging to avoid the noise in the original dataset (Cao et al., 2017a). The thermal state of permafrost and its spatial distribution result from the long-term interaction of climate and subsurface. Additionally, vegetation and snow coverage play important roles in permafrost distribution through influencing the energy exchange between the atmosphere and the ground surface (Norman et al., 1995; Zhang, 2005). In this study, three climate variables of MAAT, mean annual snow cover days (MASCD), and the annual maximum normalized difference vegetation index (NDVI) were hence selected here to test the representative of the inventory for permafrost map evaluation. The MAAT with resolution of 1 km is from Gruber (2012) representing the referenced period 1961–1990. The MASCD (about 500 m) was derived from daily~~

snow cover product developed by Wang et al. (2015) based on MODIS products (MOD10A1 and MYD10A1). To improve the comparison of MASC, it is scaled to 0–1 through dividing the total days of a given year, and the mean MASC during 2003–2010 was produced as a predictor. The annual maximum NDVI is from MODIS/Terra Vegetation Indices 16-day product (MOD13Q1, v006) with a resolution of 250 m. The annual maximum NDVI ( $NDVI_{max}$ ) was computed for each year during 2001–2017 to approximately represent the amount of vegetation, and then aggregated as a median for the entire period to avoid sensitivity to extreme values. The outline of QTP is from Zhang et al. (2002), glacier outlines is from Liu et al. (2015) representing conditions in 2010, and lake data is provided by the Third Pole Environment Database.

## 2.4 Statistics and evaluation of permafrost distribution maps

In this study, it is important to understand the difference between extent of permafrost region and permafrost area. Permafrost region is the area permafrost likely present, however, the permafrost may not be everywhere. Permafrost area is where actually refers to regions where permafrost exists but the entire region is not necessarily completely occupied by permafrost, while permafrost area refers to areas where are completely underlain by permafrost. For example, discontinuous permafrost regions have permafrost area ranging from 50 to 90%. In other words, in discontinuous permafrost region, there 50 to 90% of the area is underlain by permafrost (Zhang et al., 2000), i.e., permafrost area (Zhang et al., 2000; Gruber, 2012). To estimate permafrost region and area based on the PZI as model output, specified thresholds are required for both the extent of permafrost region and permafrost area. By following Gruber (2012), only the areas with  $PZI \geq 0.01$  were selected for further analysis, permafrost region is defined as the area with  $PZI \geq 0.1$ , and permafrost area was derived as PZI multiplied pixel area.

For map evaluation, the categorical map was aggregated to binary by merging different permafrost types as permafrost presence [1] and by merging the others as permafrost absence [0]. Evaluations of the maps with categorical types are conducted using classification accuracies (Wang et al., 2015):

$$PCC_{PF} = \frac{PF_T}{PF_T + PF_F} \times 100\% \quad (1)$$

$$PCC_{NPF} = \frac{NPF_T}{NPF_T + NPF_F} \times 100\% \quad (2)$$

$$PCC_{tol} = \frac{PF_T + NPF_T}{PF_T + PF_F + NPF_T + NPF_F} \times 100\% \quad (3)$$

where subscripts of  $T$  (True, correctly classified) and  $F$  (False, incorrectly classified) identify corrections of classification. In this case,  $PF_T$  is permafrost presence sites/plots correctly classified as permafrost, while  $PF_F$  is incorrectly classified as non-permafrost.  $NPF_T$  is permafrost absence sites correctly classified as non-permafrost, and  $NPF_F$  is incorrectly classified as permafrost.  $PCC$  is Percent percent of sites/plots Correctly Classified correctly classified, and the subscripts of  $PF$ ,  $NPF$ , and  $tol$  means permafrost, non-permafrost, and total sites/plots, respectively. For the PZI map, the PZI of 0.5 was used as the threshold of permafrost presence and absence (Boeckli et al., 2012; Azócar et al., 2017), and the above index were tested. To avoid the impact of uneven distribution of sample numbers for permafrost presence and absence, the Cohen's kappa coefficient ( $\kappa$ ), which measures inter-rater agreement for categorical items (Landis and Koch, 1977), was introduced here for map

evaluation.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (4)$$

where  $p_e$  and  $p_o$  are the probability of random agreement and disagreement, respectively, can be calculated as

$$p_e = \frac{(PF_T + PF_F) \times (PF_T + NPF_F)}{(PF_T + PF_F + NPF_F + NPF_T)^2} \quad (5)$$

$$5 \quad p_o = \frac{(NPF_F + NPF_T) \times (PF_F + NPF_T)}{(PF_T + PF_F + NPF_F + NPF_T)^2} \quad (6)$$

Cohen's kappa coefficient result is interpreted as excellent agreement for  $k \geq 0.8$ , substantial agreement for  $0.6 \leq k < 0.8$ , moderate agreement for  $0.4 \leq k < 0.6$ , fair agreement for  $0.2 \leq k < 0.4$ , and slight agreement for  $k < 0.2$ .

### 3 Results and discussion

#### 3.1 Evidence of Permafrost Presence or Absence

10 ~~Based on the classification algorithm of permafrost presence/absence, there~~ There are in total 1475 permafrost presence or absence sites/plots contained in the inventory acquired from BH, SP, GST, and GPR methods (Figure 1). Among the 1475 evidences, there are 1141 (77.4%) sites measured by BH, ~~176 (11.9184 (12.5%) sites by GST, 156 (10.6144 (9.8%) plots by GPR, and 30 (2.06 (0.4%) sites by SP~~ (Figure 1eb). There are 1012 (68.6%) permafrost presence sites/plots and 463 (31.4%) permafrost absence sites/plots. ~~Where original field evidence of permafrost presence/absence is located within the same grid cell (30 arcsec, ~1 km), they were aggregated based on their major value. For grid with one permafrost site and one non-permafrost site, the nearer site from the grid center was used to represent the grid. As a result, there are in total 1040 aggregated points/plots left for permafrost maps evaluation. These aggregated~~ These evidences extend over a large area of the QTP (latitude: 27.73–38.95.96°N, longitude: 75.06–103.57°E) (Figure 1dc). The evidence cover a wide elevation range from about 1600 m to over 5200 m, however, the majority (~~91.493.2%~~) is located between 3500 m and 5000 m. While the inventory showed an even distribution of aspects with 27.3% on the east slope, 27.9% on the south slope, 22.0% on the west slope, and 22.6% on the north slope, most of the evidence (~~96.296.1%~~) have slope angles less than 20° (Figure 1c).

Figure 1d, e and f present the coverage of ~~aggregated~~ evidence for selected climate variables, which could significantly influence permafrost distribution, ~~comparing to the entire QTP. The 1040 aggregated~~. The 1475 field sites/plots showed a relatively narrower MAAT range (-10.5–15.7 °C with Q25 lower quantile = ~~-5.9-6.0~~ °C and Q75 upper quantile = ~~-3.4-3.8~~ °C) comparing to the entire QTP with MAAT between -25.6 and 22.1 °C (Q25 lower quantile = -6.6 °C and Q75 upper quantile = -0.41 °C), and only ~~1.31.5%~~ sites/plots located in the area with MAAT < -8 °C. However, the evidence (~~84.988.2%~~) was mainly occurred in the most sensitive ~~band-range~~ (from -8 to -2 °C) of permafrost presence/absence changing with MAAT (Gruber, 2012; Cao et al., 2018). There is a slight bias for scaled MASCD coverage with little measurements (~~7.67.5%~~) in high scaled MASCD (> 0.20) area due to the harsh climate and inconvenient access. The annual maximum NDVI at evidence

sites/plots has a wide coverage for the QTP with the range of 0.05–0.88. The higher mean NDVI for evidence (0.45–0.44 at the sample sites/plots and 0.37 for the QTP) is due to the measurements are occurred in flat areas with relatively dense vegetation cover. The exploration of inventory indicated the evaluation presented in this study to be representative for most of the QTP, and may have pronounced uncertainty in steep and regular snow-covered-snow-covered regions.

### 5 3.2 Evaluation and comparison of existing maps

The new inventory was used to evaluate existing permafrost maps derived with different methods (Table 1). In general, these permafrost maps showed different performances, including fair agreement for the  $PZI_{warm}$  (the subscript for PZI map refers different cases and assumptions for PZI model parameters) and IPA maps, moderate agreement for the  $QTP_{Noah}$ ,  $PZI_{norm}$ , and  $PZI_{cold}$  maps, and substantial agreement for the  $PZI_{cold}$ , and  $QTP_{TTOP}$  maps, with a wide spread of  $\kappa$  from 0.24 to 0.60. 0.32 to 0.58. The high  $PCC_{PF}$  together with low  $PCC_{NPF}$  for the IPA,  $QTP_{Noah}$ ,  $PZI_{cold}$ , and  $QTP_{TTOP}$  maps indicate permafrost is over-presented by them, while the  $PZI_{warm}$  and  $PZI_{norm}$  maps showed underestimated the permafrost over the QTP. Additionally, the range of estimated permafrost region ( $1.42\text{--}1.84 \times 10^6 \text{ km}^2$ , or 30% difference) and area ( $0.76\text{--}1.25 \times 10^6 \text{ km}^2$ , or 64.4% difference) are extremely large.

The  $QTP_{TTOP}$  map achieved the best performance for permafrost distribution over the QTP with the highest  $\kappa$  (0.60, substantial-0.58, moderate agreement) and  $PCC_{tol}$  (82.0%). The 82.8%), however, patience should be taken to interpolate the map. The  $QTP_{TTOP}$  map was derived based on MODIS land surface temperature with different temporal coverage of 2003–2012 (Zou et al., 2017). Though the MODIS land surface temperature time-series gaps caused mainly by cloud were filled using the Harmonic Analysis Time-Series (HANTS) algorithm (Prince et al., 1998), the surface conditions, especially vegetation and snow cover, were ignored. In this case, land surface temperature is underestimated in high and/or dense vegetation area as it comes from the top of vegetation canopy, and is overestimated in snow-covered area due to the cooling effects of snow is not considered. As a consequence, permafrost is likely overestimated in high and/or dense vegetation area and underestimated in regular snow covered area. The  $PZI_{norm}$  and  $PZI_{cold}$  maps, judged as moderate agreement ( $\kappa = 0.57\text{--}0.56$  for the  $PZI_{norm}$  map and 0.55 for the  $PZI_{cold}$  map) with in-situ measurements, showed slightly worse performance comparing with the  $PZI_{QTP_{TTOP}}$  map. The poor performance of  $PZI_{warm}$  map and underestimation of the  $PZI_{norm}$  map indicated permafrost over the QTP is more prevalent than most of the other regions even when the climate conditions, especially the MAAT, are similar. This is very likely because the high soil thermal conductivity due to coarse soil conditions and the cooling effects of minimal snow (Zhang, 2005). Great difference of permafrost region ( $0.42 \times 10^6 \text{ km}^2$ , or 25% of the normal case) and area ( $0.49 \times 10^6 \text{ km}^2$ , or 49% of the normal case) was found for the three cases of  $PZI_{global}$  map though the upper and lower bounds of MAAT are only changed about 5% for the PZI and  $\pm 0.5 \text{ }^\circ\text{C}$  for the MAAT. The MAAT used in the  $PZI_{global}$  map was statistical downscaled based on the lapse rate from the upper-air (or pressure level) temperature of NCEP, but the influences of land surface on surface air temperature, such as cold air pooling, was ignored (Cao et al., 2017a). This is important as winter inversion is excepted to be common due to the prevalent mountains over the QTP. In other words, permafrost may be underestimated in valleys due to the overestimated MAAT. While the IPA and  $QTP_{Noah}$  maps performed slightly better (1.8–3.1% higher) for permafrost areas than the  $QTP_{TTOP}$  and  $PZI_{cold}$  maps, they suffer considerable underestimation of non-permafrost area (14.1–39.8% lower for

$PCC_{NPF}$ ). Though the QTP<sub>Noah</sub> map was derived using coupled land surface model (Noah), the relatively worse performance, especially for non-permafrost area ( $PCC_{NPF} = 49.5\%$ ), is likely caused by ~~inputting inputting~~ coarse-scale forcing dataset ( $0.1^\circ$  resolution or  $\sim 10$  km) (Chen et al., 2011) and by the uncertainty of soil texture dataset (Yang et al., 2010). ~~Great difference of permafrost region ( $0.42 \times 10^6$  km<sup>2</sup>, or 25% of the normal case) and area ( $0.49 \times 10^6$  km<sup>2</sup>, or 49% of the normal case) was found for the three cases of PZI<sub>global</sub> map though the upper and lower bounds of MAAT are only changed about 5%. It is not surprising that the IPA map has fair agreement ( $\kappa = 0.24$ ) as less observations were compiled and the method used are more suitable for high latitudes (Ran et al., 2012). The worst performance of PZI<sub>warm</sub> map (or good performance of PZI<sub>cold</sub> map) indicated permafrost is more prevalent than most of the other regions even when the climate conditions, especially the MAAT, are similar. This is very likely because the high soil thermal conductivity due to coarse soil conditions and the cooling effects of minimal snow (Zhang, 2005).~~

Spatially, the southeastern QTP of non-permafrost areas are better represented in all maps, while misclassification is prevalent in areas near the permafrost boundary and ~~fragile spatially highly variable~~ landscapes such as the sources of Yellow River (Figure 2). This is because the permafrost distribution in these areas is not only controlled by medium-large scale climate ~~condition conditions~~ (e.g., MAAT) which ~~is are~~ described by the models used, but also strongly influenced by various local factors such as peat layer, thermokarst, soil moisture, and hydrological processes. ~~Specially, about 30.3 (18.5–63.8)% for permafrost presence and about 39.2 (24.7–65.1)% for absence of the misclassification sites in the maps occurred in the MAAT band from  $-3$  to  $-4$  °C, which is expected as the threshold of permafrost presence for most areas.~~ The IPA and PZI<sub>warm</sub> maps showed a fit that is good only in some areas (e.g., southeastern for the PZI<sub>warm</sub> map and relatively colder areas for the IPA map) based on the in-situ measurements, and may not represent the permafrost distribution patterns well for the other areas beyond the measurement.

#### 4 Conclusions

We compiled an inventory of permafrost presence or absence evidence with 1475 field sites/plots obtained based on diverse methods over the QTP. With wide coverage of topography (e.g., elevation and slope aspect) and climate conditions (e.g., surface air temperature and snow cover), the inventory gives a representative baseline for site-specific permafrost occurrence. The existing permafrost maps over the QTP were better evaluated and compared with the inventory of ground-based evidence, and they showed a wide range of performance with the  $\kappa$  from 0.32 to 0.58 and overall classification accuracy ~~of about 59–8255–83%~~. The QTP<sub>TOP</sub> ~~and PZI<sub>cold</sub> maps are map~~ is recommended for representing permafrost distribution over the QTP based on our evaluations. ~~Additionally, the PZI<sub>norm</sub> and PZI<sub>cold</sub> maps with close performance are valuable alternatives for describing permafrost zonation index over the QTP.~~ The inadequate sampling is expected to result in uncertainty for map evaluation in steps and regular ~~snow covered snow covered~~ areas, and requires further investigation using systematic samples.

*Data availability.* Inventory of permafrost presence/absence is partly available as supplement, the other evidence not listed is available from the authors upon request.

## **Appendix A: Classification algorithm of in-situ permafrost presence or absence evidence**

For board use of the permafrost presence or absence inventory, the data confidence degree was provided (TableA1). BH and SP provide direct evidence of permafrost presence or absence based on MAGT and/or ground ice observations, and hence have high confidence (Cremonese et al., 2011). The data confidence derived from MAGST is classified based on temperature and the length of the observation period. The evaluated GPR survey result was considered as medium confidence.

*Author contributions.* BC carried out this study by organizing permafrost presence or absence evidence, analyzing data, performing the simulations and by structuring as well as writing the paper. TZ guided the research. QW, YS, LZ, and DZ contributed to organize the permafrost presence/absence dataset.

*Competing interests.* The authors declare that no competing interests are present.

*Acknowledgements.* The authors would like to thank Stephan Gruber and Kang Wang for their constructive suggestions. We thank Zhuotong Nan and Xiaobo Wu for providing the QTP<sub>Noah</sub> map. This study was supported by the [Strategic Priority Research Program of Chinese Academy of Sciences \(XDA20100103\)](#), the National Natural Science Foundation of China (91325202), the National Key Scientific Research Program of China (2013CBA01802), partly by the Fundamental Research Funds for the Central Universities (lzujbky\_2016\_281, 862863). We thank CMA (<http://cdc.cma.gov.cn/>) for providing the surface air and ground surface temperatures, the ASTER dataset is downloaded from United States Geological Survey (<http://gdex.cr.usgs.gov/gdex/>), glacier inventory is provided by the Environmental and Ecological Science Data Center for West China (<http://westdc.westgis.ac.cn/>), and lake inventory is from the Third Pole Environment Database (<http://www.tpedatabase.cn>).

## References

- Azócar, G. F., Brenning, A., and Bodin, X.: Permafrost distribution modelling in the semi-arid Chilean Andes, *The Cryosphere*, 11, 877–890, <https://doi.org/10.5194/tc-11-877-2017>, 2017.
- Boeckli, L., Brenning, A., Gruber, S., and Noetzli, J.: Permafrost distribution in the European Alps: calculation and evaluation of an index map and summary statistics, *The Cryosphere*, 6, 807–820, <https://doi.org/10.5194/tc-6-807-2012>, 2012.
- Brown, J., F. J. O. H. J. M. E.: *Circum-Arctic Map of Permafrost and Ground-ice Conditions*, 1997.
- Cao, B., Gruber, S., and Zhang, T.: REDCAPP (v1.0): parameterizing valley inversions in air temperature data downscaled from reanalyses, *Geoscientific Model Development*, 10, 2905–2923, <https://doi.org/10.5194/gmd-10-2905-2017>, 2017a.
- Cao, B., Gruber, S., Zhang, T., Li, L., Peng, X., Wang, K., Zheng, L., Shao, W., and Guo, H.: Spatial variability of active layer thickness detected by ground-penetrating radar in the Qilian Mountains, Western China, *Journal of Geophysical Research: Earth Surface*, 122, 574–591, <https://doi.org/10.1002/2016JF004018>, 2016JF004018, 2017b.
- Cao, B., Zhang, T., Peng, X., Mu, C., Wang, Q., Zheng, L., Wang, K., and Zhong, X.: Thermal Characteristics and Recent Changes of Permafrost in the Upper Reaches of the Heihe River Basin, Western China, *Journal of Geophysical Research: Atmospheres*, 0, <https://doi.org/10.1029/2018JD028442>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD028442>, 2018.
- Chen, Y., Yang, K., He, J., Qin, J., Shi, J., Du, J., and He, Q.: Improving land surface temperature modeling for dry land of China, *Journal of Geophysical Research: Atmospheres*, 116, <https://doi.org/10.1029/2011JD015921>, d20104, 2011.
- Cheng, G. and Jin, H.: Permafrost and groundwater on the Qinghai-Tibet Plateau and in northeast China, *Hydrogeology Journal*, 21, 5–23, <https://doi.org/10.1007/s10040-012-0927-2>, <https://doi.org/10.1007/s10040-012-0927-2>, 2013.
- Cremonese, E., Gruber, S., Phillips, M., Pogliotti, P., Boeckli, L., Noetzli, J., Suter, C., Bodin, X., Crepez, A., Kellerer-Pirklbauer, A., Lang, K., Letey, S., Mair, V., Morra di Cella, U., Ravel, L., Scapozza, C., Seppi, R., and Zischg, A.: Brief Communication: "An inventory of permafrost evidence for the European Alps", *The Cryosphere*, 5, 651–657, <https://doi.org/10.5194/tc-5-651-2011>, 2011.
- Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, *The Cryosphere*, 6, 221–233, <https://doi.org/10.5194/tc-6-221-2012>, 2012.
- Gruber, S., Fleiner, R., Guegan, E., Panday, P., Schmid, M.-O., Stumm, D., Wester, P., Zhang, Y., and Zhao, L.: Review article: Inferring permafrost and permafrost thaw in the mountains of the Hindu Kush Himalaya region, *The Cryosphere*, 11, 81–99, <https://doi.org/10.5194/tc-11-81-2017>, 2017.
- Hasler, A., Geertsema, M., Foord, V., Gruber, S., and Noetzli, J.: The influence of surface characteristics, topography and continentality on mountain permafrost in British Columbia, *The Cryosphere*, 9, 1025–1038, <https://doi.org/10.5194/tc-9-1025-2015>, 2015.
- Landis, J. R. and Koch, G. G.: The Measurement of Observer Agreement for Categorical Data, *Biometrics*, 33, 159–174, <http://www.jstor.org/stable/2529310>, 1977.
- Lin, Z., Burn, C. R., Niu, F., Luo, J., Liu, M., and Yin, G.: The Thermal Regime, including a Reversed Thermal Offset, of Arid Permafrost Sites with Variations in Vegetation Cover Density, Wudaoliang Basin, Qinghai-Tibet Plateau, *Permafrost and Periglacial Processes*, 26, 142–159, <https://doi.org/10.1002/ppp.1840>, 2015.
- Liu, S., Yao, X., Guo, W., Xu, J., Shangguan, D., Wei, J., Bao, W., and Wu, L.: The contemporary glaciers in China based on the Second Chinese Glacier Inventory (in Chinese with English abstract), *Acta Geographica Sinica*, 70, 3, <https://doi.org/10.11821/dlxb201501001>, 2015.

- Moorman, B. J., Robinson, S. D., and Burgess, M. M.: Imaging periglacial conditions with ground-penetrating radar, *Permafrost and Periglacial Processes*, 14, 319–329, <https://doi.org/10.1002/ppp.463>, 2003.
- Mu, C., Zhang, T., Zhao, Q., Su, H., Wang, S., Cao, B., Peng, X., Wu, Q., and Wu, X.: Permafrost affects carbon exchange and its response to experimental warming on the northern Qinghai-Tibetan Plateau, *Agricultural and Forest Meteorology*, 247, 252 – 259, <https://doi.org/https://doi.org/10.1016/j.agrformet.2017.08.009>, 2017.
- 5 Nan, Z., Huang, P., and Zhao, L.: Permafrost distribution modeling and depth estimation in the Western Qinghai-Tibet Plateau (in Chinese with English abstract), *Acta Geographica Sinica*, 68, 318, <https://doi.org/10.11821/xb201303003>, 2013.
- Norman, J., Kustas, W., and Humes, K.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, *Agricultural and Forest Meteorology*, 77, 263 – 293, *thermal Remote Sensing of the Energy and Water Balance over Vegetation*, 1995.
- 10 Prince, S., Goetz, S., Dubayah, R., Czajkowski, K., and Thawley, M.: Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using Advanced Very High-Resolution Radiometer satellite observations: comparison with field observations, *Journal of Hydrology*, 212-213, 230 – 249, [https://doi.org/https://doi.org/10.1016/S0022-1694\(98\)00210-8](https://doi.org/https://doi.org/10.1016/S0022-1694(98)00210-8), <http://www.sciencedirect.com/science/article/pii/S0022169498002108>, 1998.
- 15 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.
- Schmid, M.-O., Baral, P., Gruber, S., Shahi, S., Shrestha, T., Stumm, D., and Wester, P.: Assessment of permafrost distribution maps in the Hindu Kush Himalayan region using rock glaciers mapped in Google Earth, *The Cryosphere*, 9, 2089–2099, <https://doi.org/10.5194/tc-9-2089-2015>, <https://www.the-cryosphere.net/9/2089/2015/>, 2015.
- 20 Wang, W., Huang, X., Deng, J., Xie, H., and Liang, T.: Spatio-Temporal Change of Snow Cover and Its Response to Climate over the Tibetan Plateau Based on an Improved Daily Cloud-Free Snow Cover Product, *Remote Sensing*, 7, 169–194, <https://doi.org/10.3390/rs70100169>, 2015.
- Wu, J., Sheng, Y., Wu, Q., and Wen, Z.: Processes and modes of permafrost degradation on the Qinghai-Tibet Plateau, *Science in China Series D: Earth Sciences*, 53, 150–158, <https://doi.org/10.1007/s11430-009-0198-5>, 2010.
- 25 Wu, Q. and Zhang, T.: Recent permafrost warming on the Qinghai-Tibetan Plateau, *Journal of Geophysical Research: Atmospheres*, 113, n/a–n/a, <https://doi.org/10.1029/2007JD009539>, d13108, 2008.
- Wu, Q., Yuanlin, Z., and Yongzhi, L.: Application of the Permafrost Table Temperature and Thermal Offset Forecast Model in the Tibetan Plateau, *Journal of Glaciology and Geocryology*, pp. 24–27, 2002.
- Wu, Q., Zhang, Z., Gao, S., and Ma, W.: Thermal impacts of engineering activities and vegetation layer on permafrost in different alpine ecosystems of the Qinghai–Tibet Plateau, China, *The Cryosphere*, 10, 1695–1706, <https://doi.org/10.5194/tc-10-1695-2016>, 2016.
- 30 Wu, X., Nan, Z., Zhao, S., Zhao, L., and Cheng, G.: Spatial modeling of permafrost distribution and properties on the Qinghai–Tibet Plateau, *Permafrost and Periglacial Processes*, 29, 86–99, <https://doi.org/10.1002/ppp.1971>, 2018.
- Yang, K., He, J., Tang, W., Qin, J., and Cheng, C. C.: On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau, *Agricultural and Forest Meteorology*, 150, 38 – 46, <https://doi.org/https://doi.org/10.1016/j.agrformet.2009.08.004>, 2010.
- 35 Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, *Reviews of Geophysics*, 43, <https://doi.org/10.1029/2004RG000157>, 2005.

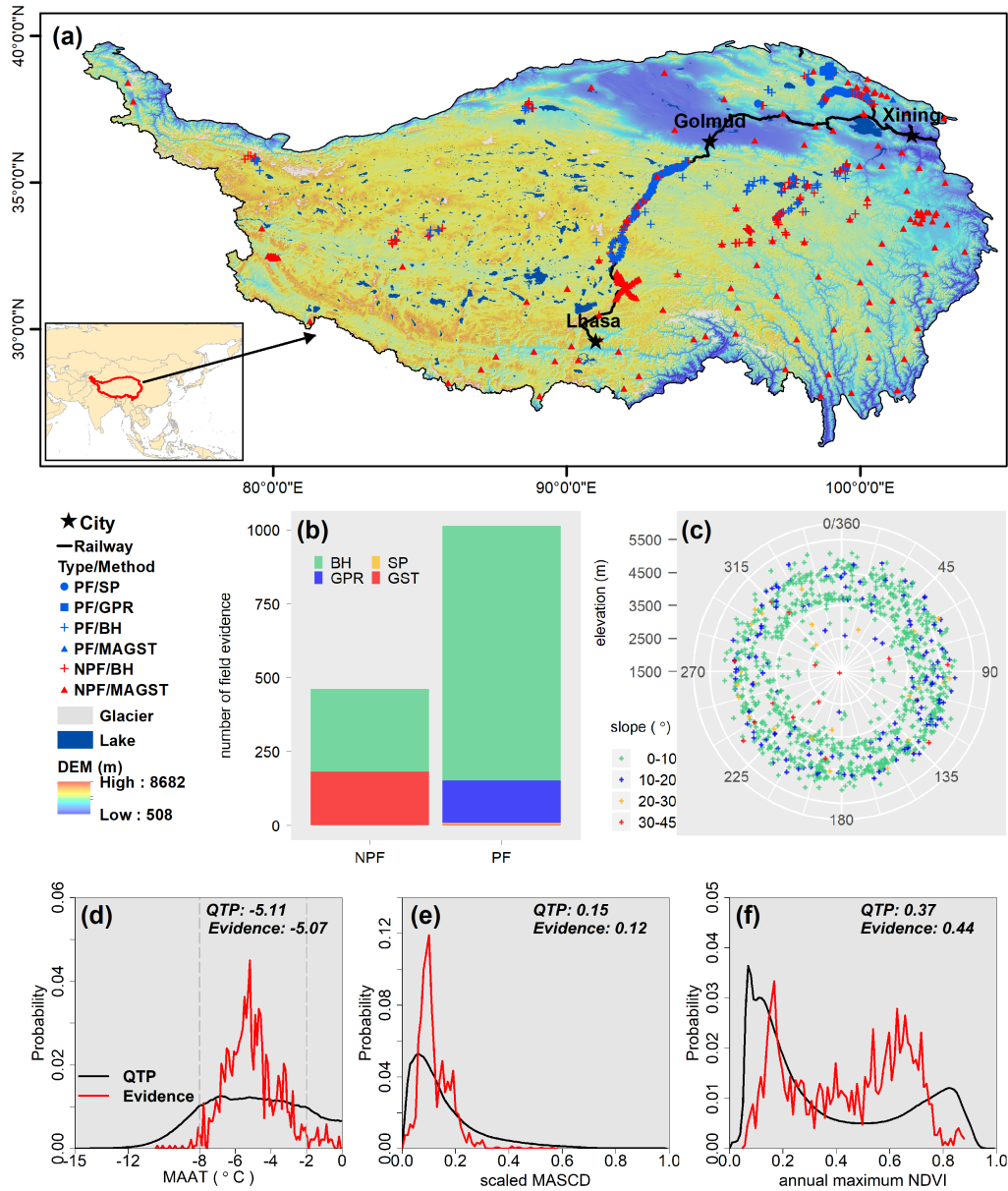


- Zhang, T., Heginbottom, J. A., Barry, R. G., and Brown, J.: Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere, *Polar Geography*, 24, 126–131, <https://doi.org/10.1080/10889370009377692>, 2000.
- Zhang, Y., Li, B., and Zheng, D.: A discussion on the boundary and area of the Tibetan Plateau in China, *Geographical Research*, 21, 1, <https://doi.org/10.11821/yj2002010001>, 2002.
- 5 Zhao, S., Nan, Z., Huang, Y., 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.
- 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.

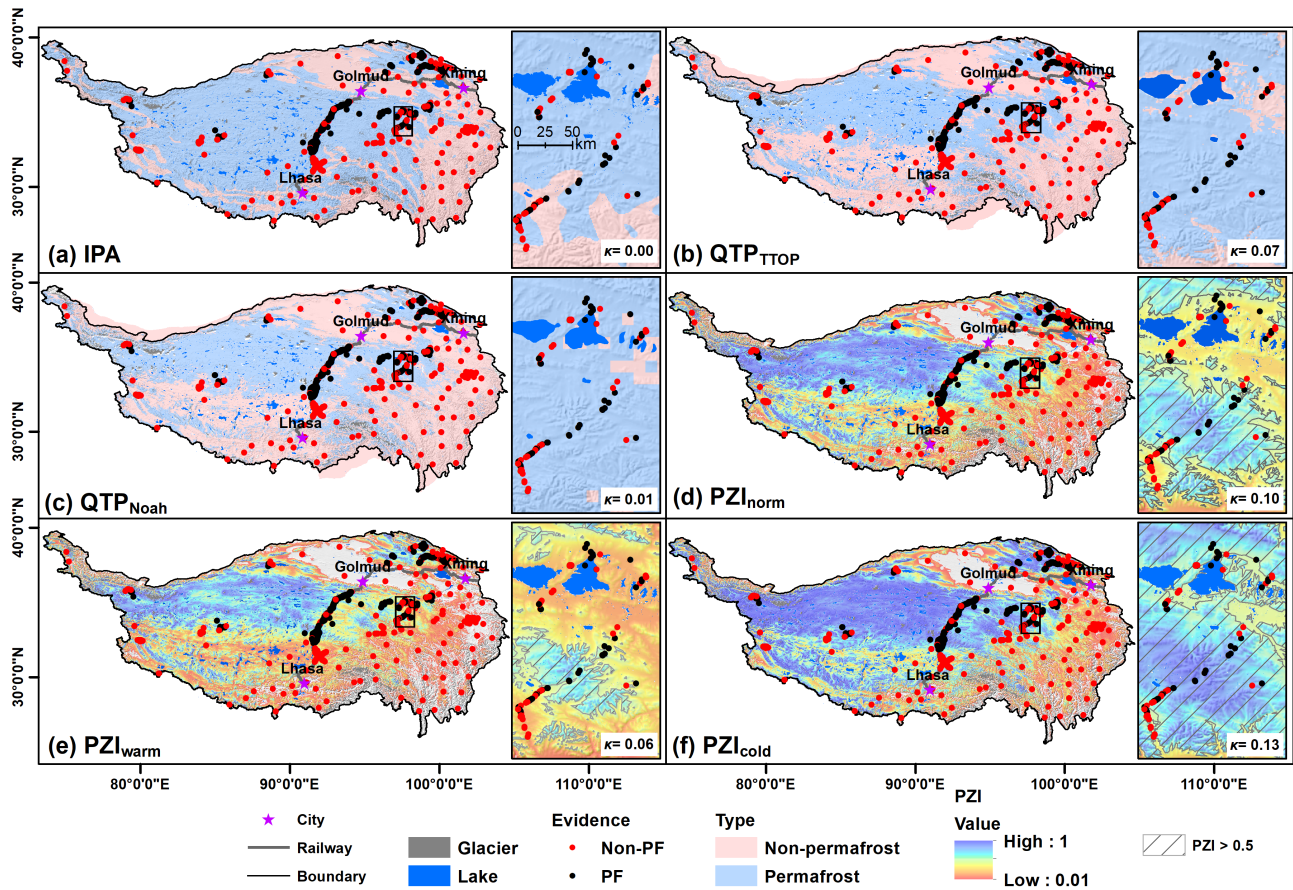
**Table 1.** Summary and evaluation of existing permafrost maps over the Qinghai-Tibet Plateau

Name	IPA	QTP <sub>TROP</sub>	QTP <sub>Noah</sub>	PZI <sub>norm</sub>	PZI <sub>warm</sub>	PZI <sub>cold</sub>
Year	1997	2017	2018	2012	2012	2012
Method	–	semi-physical model	physical model	heuristic GLM	heuristic GLM	heuristic GLM
Classification Criteria	categorical	categorical	categorical	continuous	continuous	continuous
Scale	1:10,000,000	~1 km	0.1° (~10 km)	~1 km	~1 km	~1 km
$PCC_{PF}$ [%]	<del>96.7</del> 97.3	<del>93.6</del> 93.9	<del>95.4</del> 96.4	<del>74.6</del> 76.6	<del>34.1</del> 35.3	<del>94.0</del> 94.3
$PCC_{NPF}$ [%]	<del>23.8</del> 21.9	<del>63.6</del> 58.6	<del>49.5</del> 45.9	<del>85.4</del> 82.6	<del>99.0</del> 98.5	<del>59.3</del> 54.0
$PCC_{tot}$ [%]	<del>68.5</del> 73.8	<del>82.0</del> 82.8	<del>77.8</del> 80.7	<del>78.8</del> 78.5	<del>59.2</del> 55.1	<del>80.6</del> 81.7
$\kappa$	<del>0.24</del> 0.32	<del>0.60</del> 0.58	<del>0.49</del> 0.52	<del>0.57</del> 0.56	<del>0.28</del> 0.36	<del>0.57</del> 0.55
PF Region [ $10^6$ km <sup>2</sup> ]	1.63	–	–	1.68	1.42	1.84
PF Area [ $10^6$ km <sup>2</sup> ]	–	1.06 ± 0.09	1.13	1.00	0.76	1.25
Reference	Brown (1997)	Zou et al. (2017)	Wu et al. (2018)	Gruber (2012)	Gruber (2012)	Gruber (2012)

Evaluations are conducted using 1475 in-situ measurements of permafrost presence or absence. GLM = generalized linear model. PF means permafrost. Norm (normal), warm, and cold means different cases and assumptions of parameters for PZI simulations in PZI<sub>global</sub> map, details could be found from Table 1 of Gruber (2012). Criteria of continuous means permafrost distribution is compiled as PZI range of [0.01–1]. Some bias is expected for permafrost areas of QTP<sub>TROP</sub> and QTP<sub>Noah</sub> as different QTP boundaries, lake and glacier data are used (Figure 2).



**Figure 1.** (a) The location of the QTP, and in-situ permafrost presence/absence evidence distribution over the QTP, superimposed on the background of digital elevation model (DEM) with 3-30 arcsec. (b) Number of original field evidence located in permafrost absence (NPF) and presence (PF) regions. BH means field evidence measured by borehole drilling, GPR means ground-penetrating radar, SP means soil pit, and MAGST means mean annual ground surface temperature. (c) Distribution of aggregated field evidence in terms of elevation (radius), slope (colored), and aspect (0/360° represents North). Spread of evidence (red line) for the climate variable of (d) MAAT, (e) scaled MASCD, and (f) annual maximum NDVI comparing to the entire QTP (black line). Numbers in (d), (e), and (f) are mean values. Only the sites/plots with MAAT < 0 °C, which is precondition for permafrost presence, were present in (d).



**Figure 2.** The permafrost classification results at in-situ evidence sites/plots in the (a) IPA, (b) QTP<sub>TTOP</sub>, (c) QTP<sub>Noah</sub>, (d) PZI<sub>norm</sub>, (e) PZI<sub>warm</sub>, and (f) PZI<sub>cold</sub> maps.  $\kappa$  and  $PCC$  are the evaluation results for the selected fragile-spatially highly variable landscapes (marked by black box). All the maps are re-sampled to the unprojected grid of SRTM30 DEM with a resolution of 30 arcsec ( $\sim 1$  km) to avoid maps bias of with different resolutions, geographic projection, and format. The boundary of QTP used in this study is marked by black line. Binary-Categorical classification is used for the QTP<sub>TTOP</sub>, QTP<sub>Noah</sub>, and IPA maps, while continuous PZI was present for the PZI<sub>norm</sub>, PZI<sub>warm</sub>, PZI<sub>cold</sub> maps. The blank part in PZI maps is area with PZI < 0.01. The  $\kappa$  and  $PCC_{tol}$  present in right small figures were evaluated in the selected areas with 84 106 evidence.

**Table A1.** Classification algorithm of in-situ permafrost presence or absence evidence from various methods

Method	Indicator	Survey depth	Permafrost	Confidence degree
BH	$MAGT \leq 0 \text{ } ^\circ\text{C}$	meters to about 20 m	presence	high
SP	ground ice presence	about 1.0–2.5 m	presence	high
GST	$MAGST \leq -2 \text{ } ^\circ\text{C}$ & observations $\geq 3$	0.05 or 0.1 m	presence	medium
	$MAGST \leq -2 \text{ } ^\circ\text{C}$ & observations $< 3$		presence	low
	$MAGST > -2 \text{ } ^\circ\text{C}$ & $MAGST + TO_{\max} \leq 0 \text{ } ^\circ\text{C}$		presence	low
	$MAGST < 0 \text{ } ^\circ\text{C}$ & $MAGST + TO_{\max} > 0 \text{ } ^\circ\text{C}$		ambiguous	–
	$MAGST > 0 \text{ } ^\circ\text{C}$		absence	medium
GPR	clear permafrost reflection	about 0.80–5.0 m	presence	medium

BH = borehole temperature, SP = soil pit, GST = ground surface temperature, and GPR = ground-penetrating radar.  $TO_{\max}$ , the maximum thermal offset under natural conditions reported for the QTP, is 0.79 °C.