Supplementary Materials for

PeatClim v1.0: A climate-driven machine-learning model for predicting potential paleo-peatland distribution and its key climate controls

Linlin Chen *et al.*

*Corresponding author. Email: linlin.chen@brsitol.ac.uk

This PDF file includes:

Table S1. WorldClim bioclimatic variables. The table lists the 19 standard bioclimatic variables (BIO1-BIO19) used as candidate predictors in this study.

Index	**Full name**
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (maximum temperature - minimum temperature))
BIO3	Isothermality (BIO2/BIO7) (×100)
BIO4	Temperature Seasonality (standard deviation ×100)
BIO5	Maximum Temperature of Warmest Month
BIO6	Minimum Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter

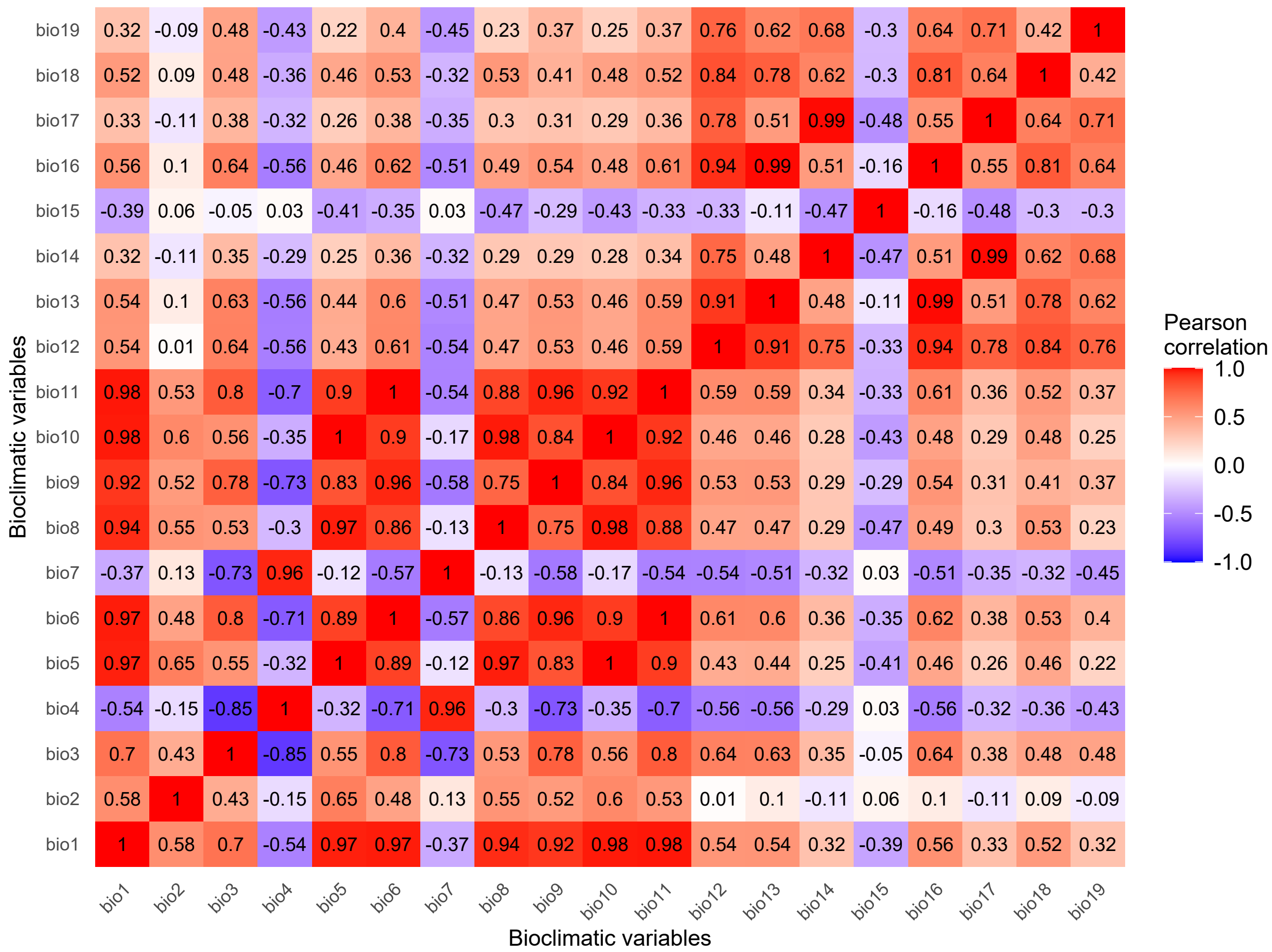


Figure S1. Pearson correlation matrix among the 19 WorldClim bioclimatic variables. Each cell shows the Pearson correlation coefficient between a pair of variables. Red and blue indicate positive and negative correlations, respectively, and colour intensity reflects the strength of the correlation.

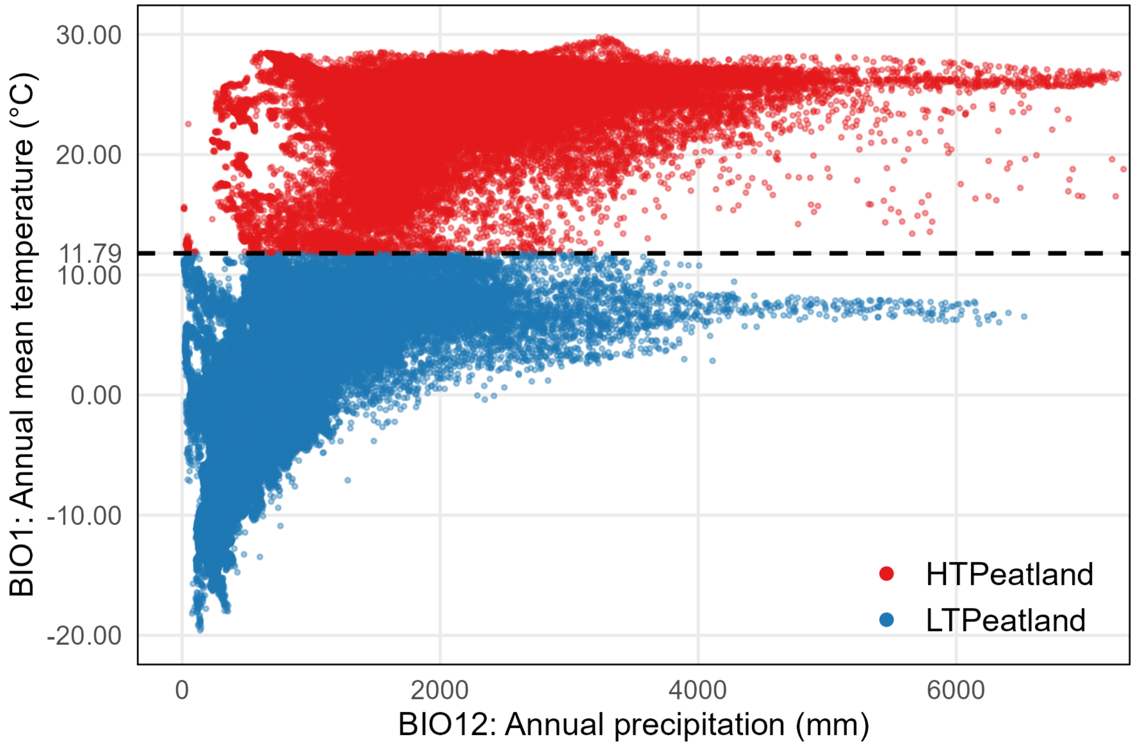
Table S2. Original data sources contributing to the Peat-DBase v.1. We began from the full Peat-DBase v.1 compilation and then applied additional quality control and 5 arc-minute regridding. The final SDM training dataset is therefore a processed derivative and does not retain all original records.

	Source name	Original amount	Retained amount
1	Bauer et al. (2024) [a]	769	308
2	Batjes et al. (2019, 2020)	94615	0
3	Beilman et al. (2009) [a]	23	0
4	Benfield et al. (2021) [a]	22	14
5	Cole et al. (2015)	3	1
6	Comas et al. (2015)	8	2
7	Crezee et al. (2022)	1558	34
8	Davies et al. (2021)	1	1
9	Davies et al. (2023a)	2	1
10	Davies et al. (2023b)	3	1
11	Gorham et al. (2012) [a]	1685	441
12	Group of Angela Gallego-Sala (unpublished data)	230	21
13	Group of Mariusz Lamentowicz (unpublished data)	249	18
14	Hribljan et al. (2023)	25	15
15	Hugelius et al. (2020) [a]	7738	3927
16	Kelly et al. (2020)	29	3
17	Keys and Henderson (1987) [a]; Thibault (1992)	20505	392
18	Lamentowicz (2005)	14	2
19	Lawson et al. (2023)	280	27
20	Manitoba Department of Natural Resources and Northern Development[a] (unpublished data); Bannatyne (1964, 1980); Treat et al. (2016)	1709	111
21	Matthew Warren (unpublished data)	276	1

22	Scottish Government (2025)	174159	558
23	Silvestri et al. (2019a, b)	63	6
24	Sun et al. (2023) [a]	146	75
25	Treat et al. (2017, 2019) [a, b]	614	291
26	Warren et al. (2012) [a]	33	2
27	Winton et al. (2025)	186	38
Summary		304945	6290

[a] Indicates sources that are confirmed to be a compilation of other datasets. [b] Five of the cores in Treat et al. (2019) were excluded due to being located in the present-day ocean. These sites were not in error, but were collected to characterize peatlands across the last glacial cycle when the sites were subaerial.

(a)



(b)

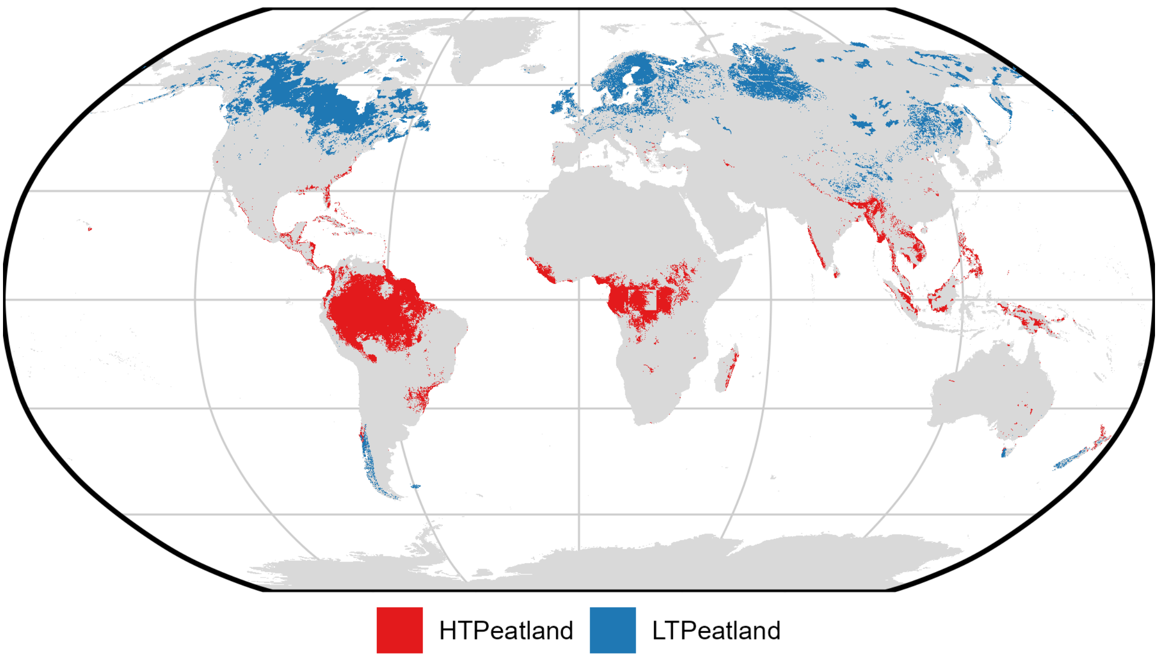


Figure S2. Climatic and spatial distributions of modern global peatlands from PEATMAP and Peat-DBase using an alternative data-conversion rule described in Section 2.3.4, in which any PEATMAP polygon touching a grid cell was classified as presence. (a) Scatterplot of peatland presences in climate space defined by annual mean temperature (BIO1; °C) and annual precipitation (BIO12; mm). (b) Global spatial distribution of peatland presences at 5 arcminute resolution. Peatlands are divided into a high-temperature peatland subset (HTPeatland; red dots and squares) and a low-temperature peatland subset (LTPeatland; blue dots and squares) using a BIO1 threshold of 11.79 °C.

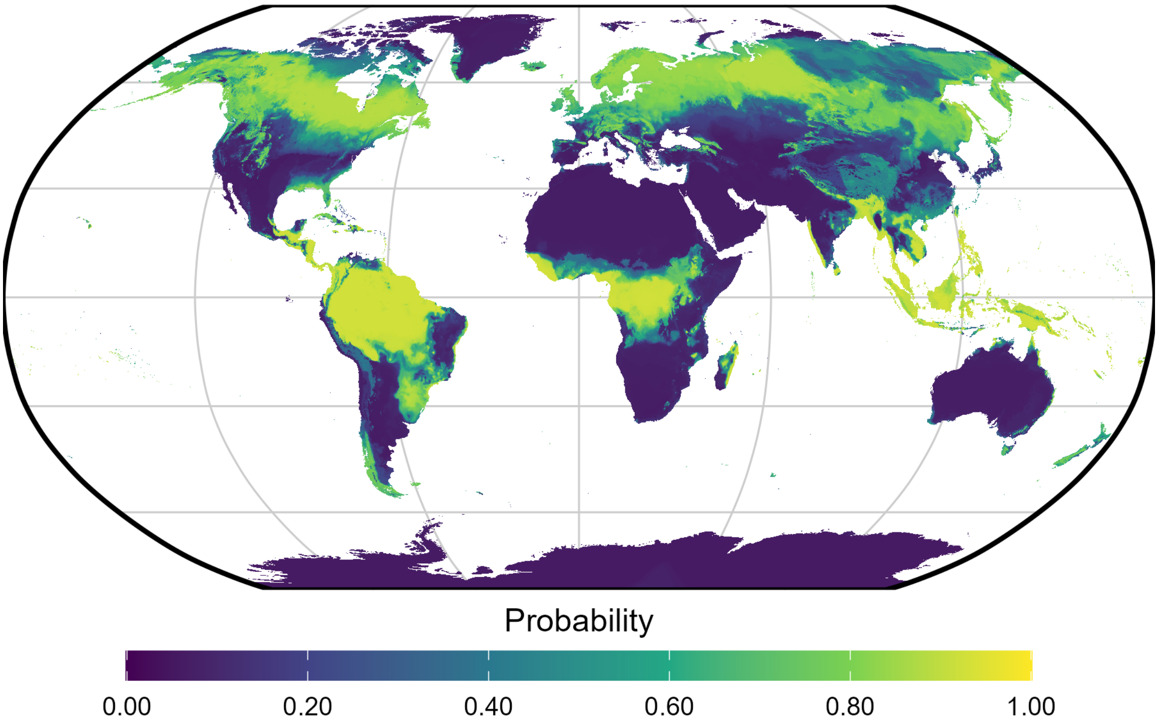


Figure S3. Predicted potential distribution of HTLTPeatland using an alternative training data-conversion rule described in Section 2.3.4 and shown in Fig. S2. The predicted probability of HTLTPeatland was calculated by combining the LTPeatland and HTPeatland model outputs using the temperature-dependent weighting scheme described by equations (1)–(3). The map shows one of the 30 runs from the underlying LTPeatland and HTPeatland models, which were trained using GBM with all 19 bioclimatic variables, nb.PA.rep=10, iteration=3, and nb.PA=236,934 for LTPeatland and nb.PA=134,829 for HTPeatland.

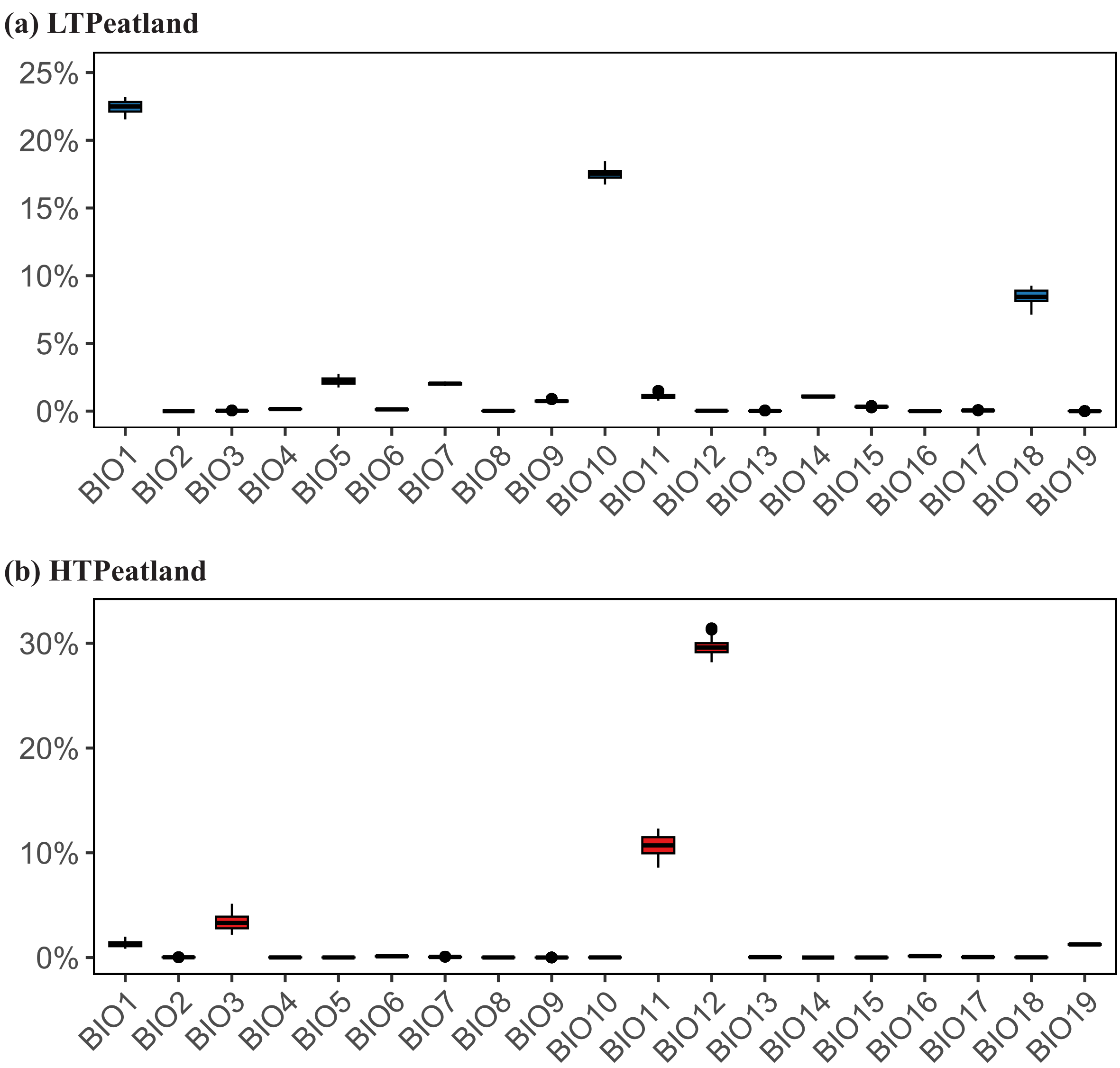


Figure S4. Permutation-based variable importance for the LTPeatland and HTPeatland models using an alternative training data-conversion rule described in Section 2.3.4 and shown in Fig. S2. Panels (a) and (b) show results for the LTPeatland and HTPeatland models, respectively. Other model settings are the same as those in Fig. S6, and the figure format is the same as that in Fig. 7.

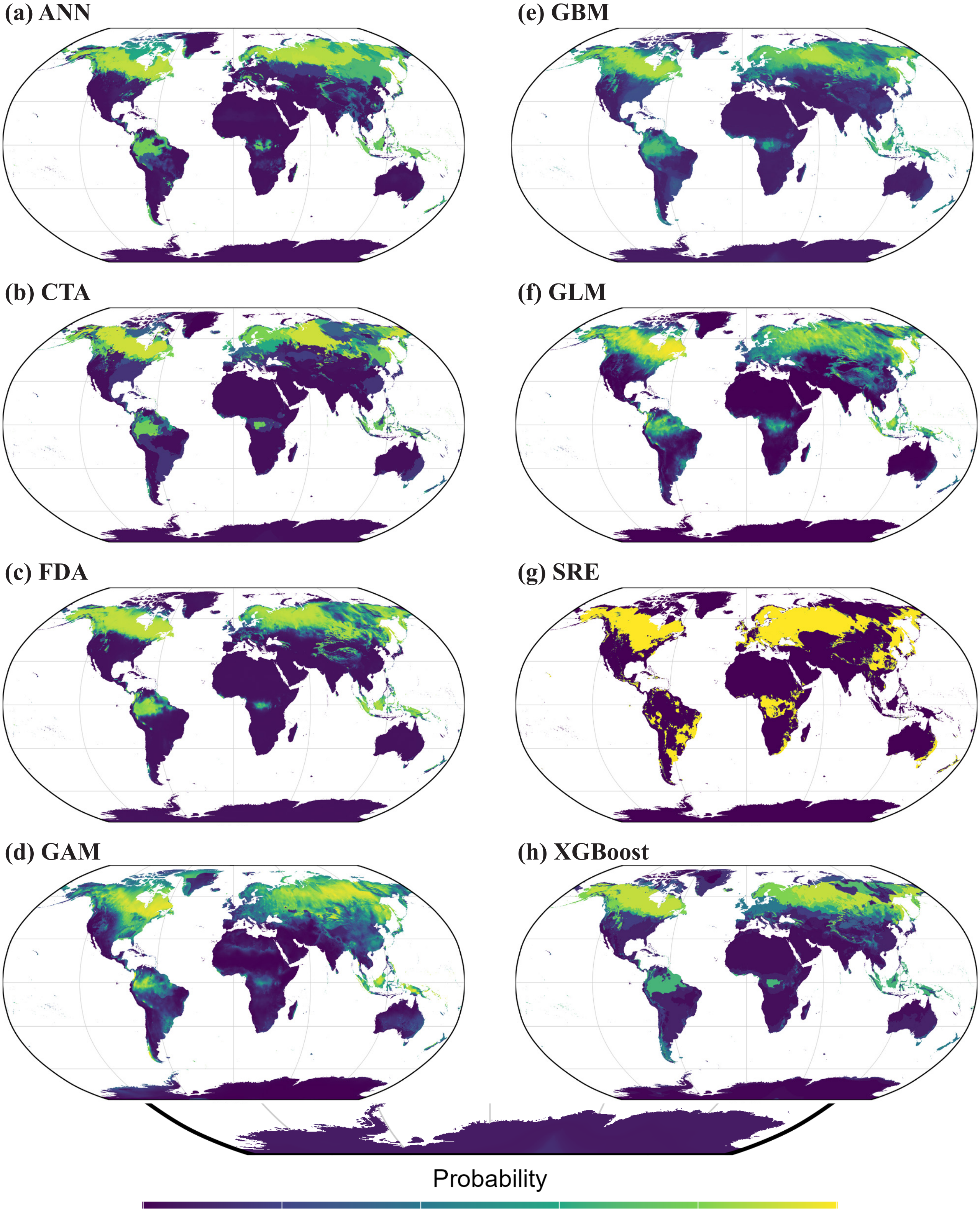


Figure S5. Predicted potential distributions from models trained using different algorithms. Panels (a–h) show one of the 30 runs for models trained with ANN, CTA, FDA, GAM, GBM, GLM, SRE, and XGBoost, respectively. All models were trained using the filtered PEATMAP dataset and 19 bioclimatic variables with nb.PA=140,000, nb.PA.rep=10, and iteration=3. Colours indicate the predicted probability of peatland presence from 0 (blue) to1 (yellow).

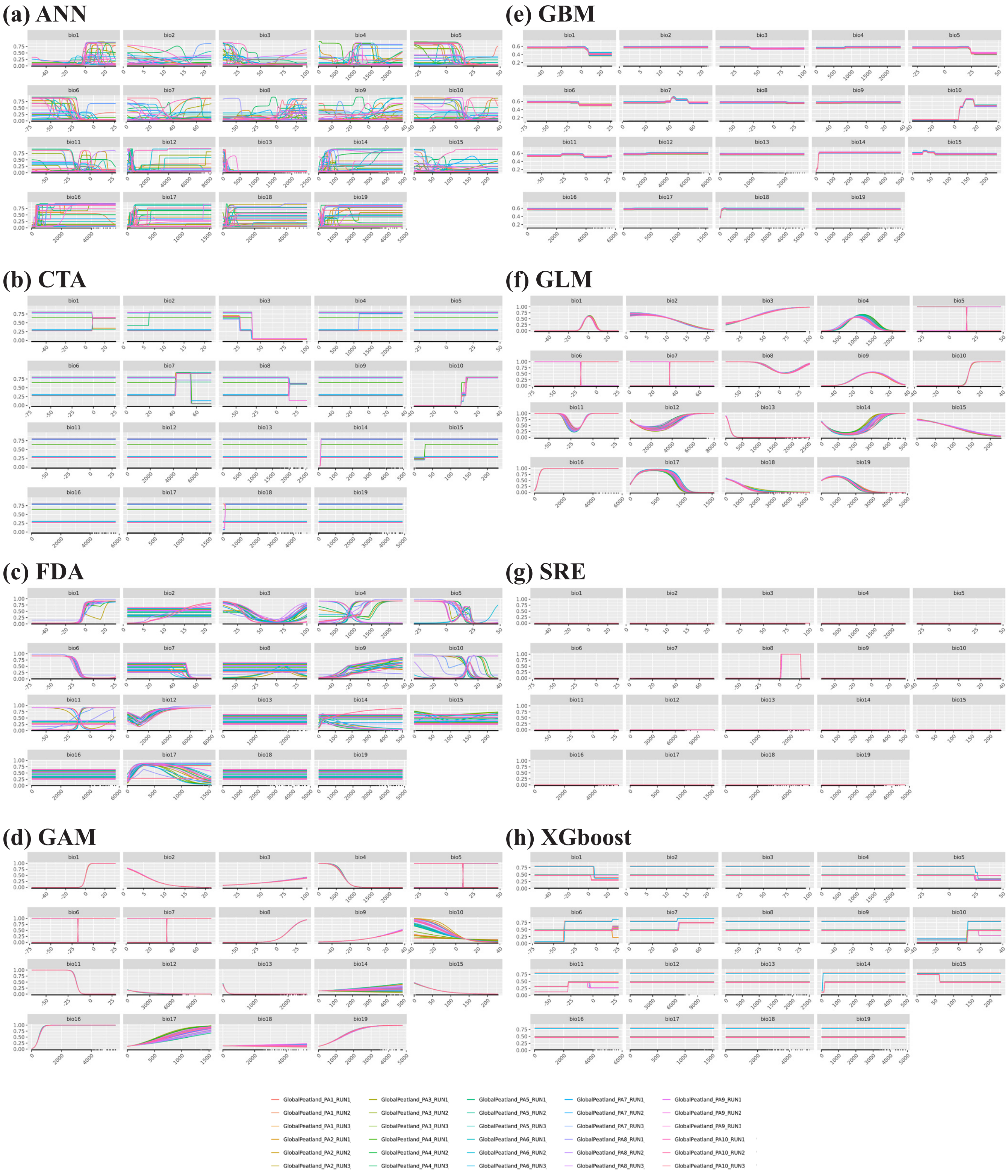


Figure S6. Response curves for models trained using different algorithms. The curves show the relationships between predicted probability of peatland presence (y axis) and predictor values (x axis). For each algorithm, 30 curves are displayed, corresponding to 30 runs. For each target predictor, response curves were generated with all other predictors fixed at their mean values. All models were trained using the filtered PEATMAP dataset and 19 bioclimatic variables with nb.PA=140,000, nb.PA.rep=10, and iteration=3.

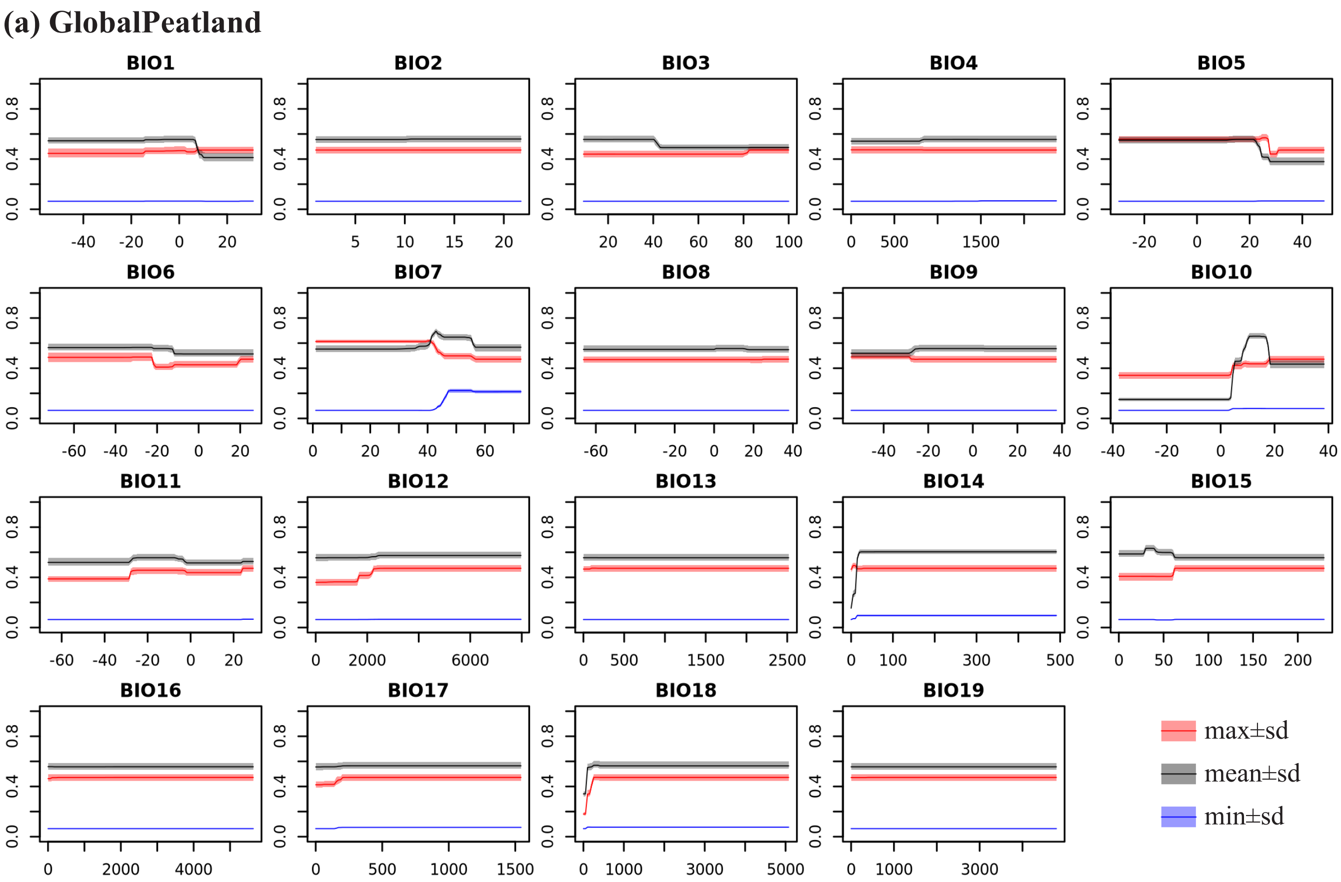
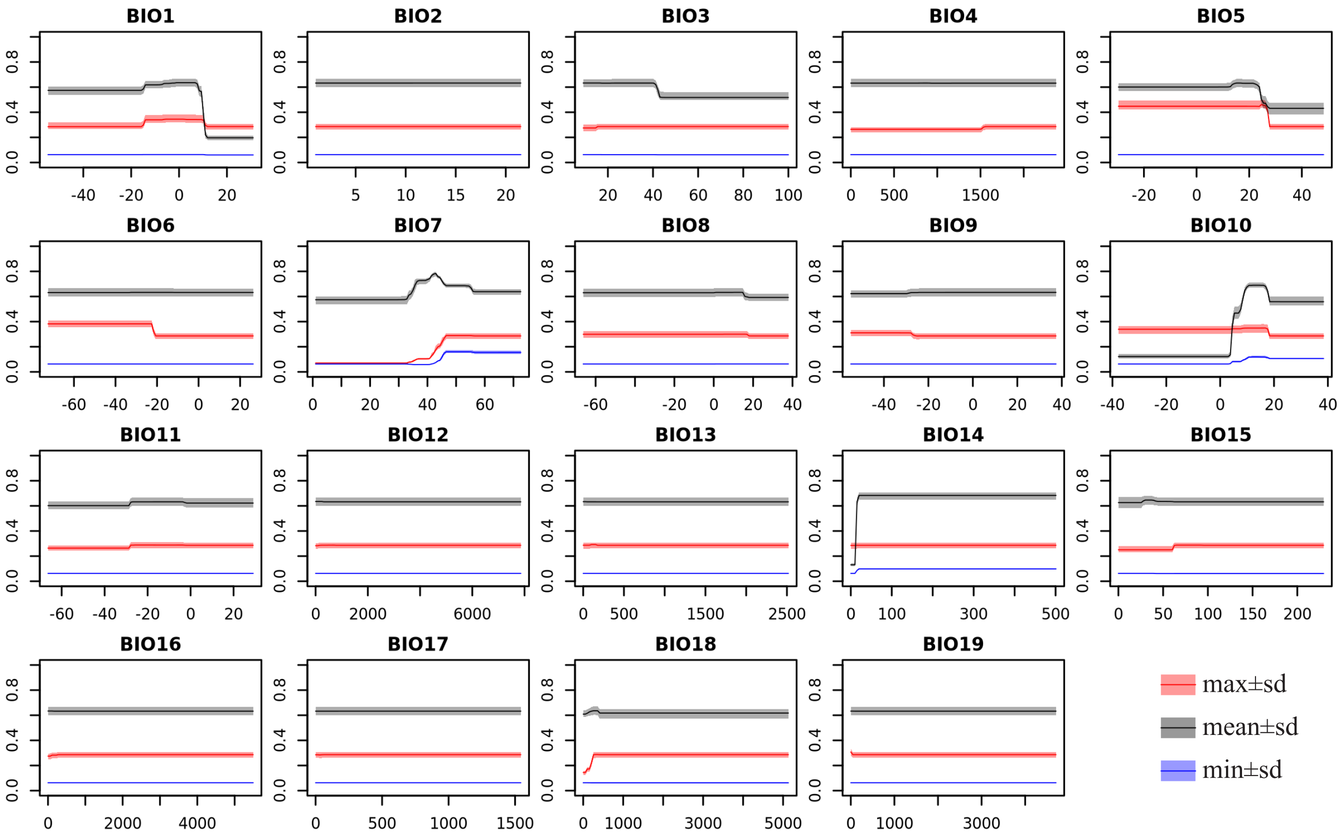


Figure S7-1. Response curves for models trained using 19 bioclimatic variables. Models trained using (a) GlobalPeatland, (b) LTPeatland, and (c) HTPeatland datasets, respectively. Lines represent the mean response across 30 runs, and the shaded areas indicate the corresponding standard deviation. For each target predictor, response curves were generated with all other predictors fixed at their maximum (red), mean (black) and minimum (blue) values. Other model settings are the same as those for the 19-variable models in Fig. 5.

(b) LTPeatland



(c) HTPeatland

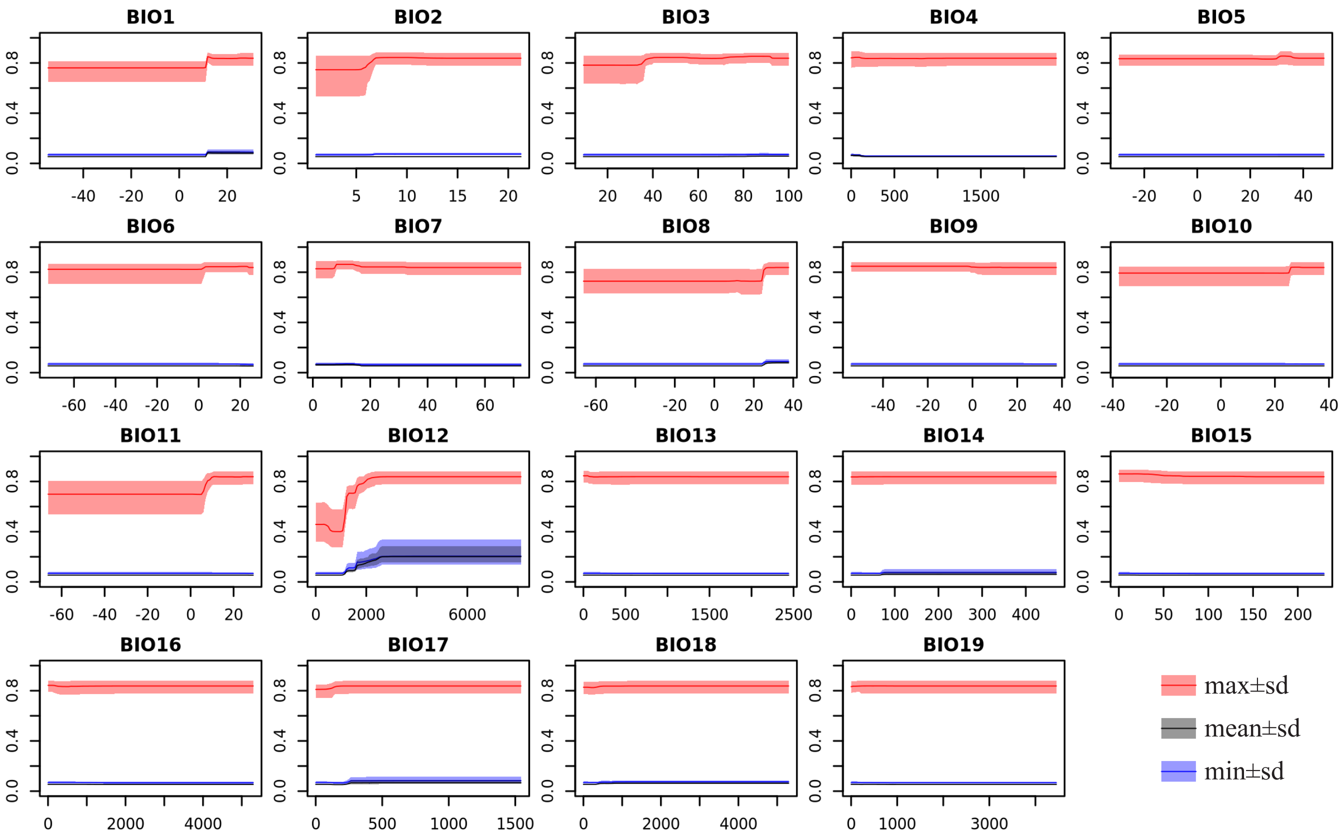


Figure S7-2. Response curves for models trained using 19 bioclimatic variables. (continue)

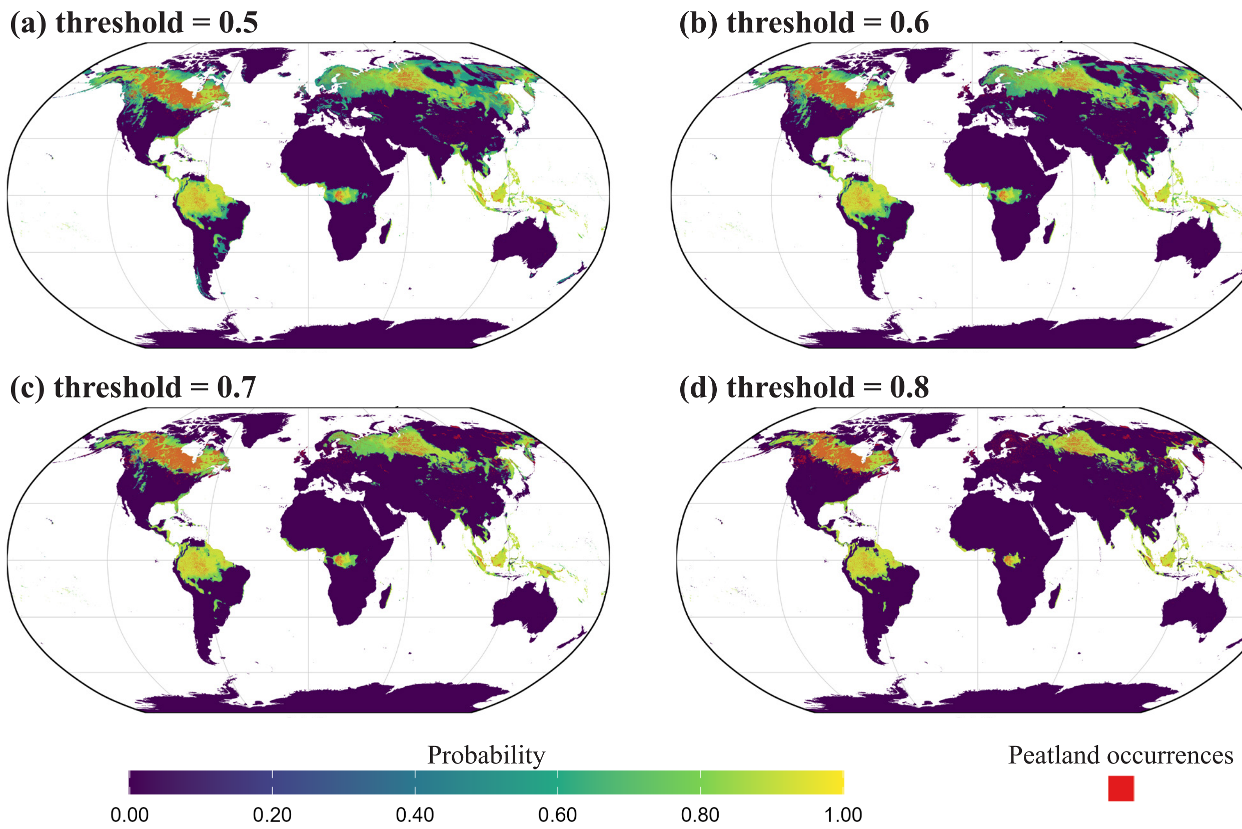


Figure S8. Predicted potential distribution of HTLTPeatland under different probability thresholds. Panels show predicted peatland presence defined by probabilities equal to or greater than thresholds of (a) 0.5, (b) 0.6, (c) 0.7, and (d) 0.8. Red squares represent peatland occurrences from the integrated PEATMAP and Peat-DBase dataset. Other model settings are the same as in Fig. 8.

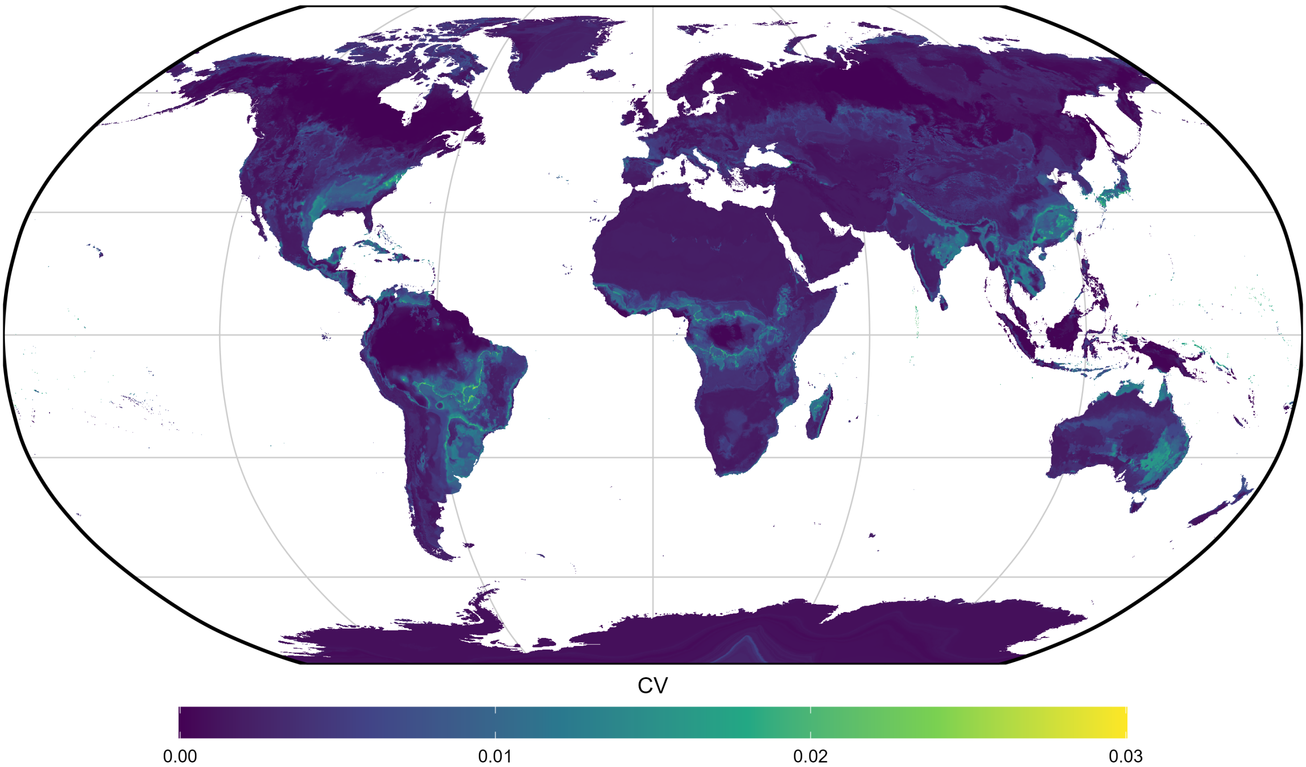


Figure S9. Coefficient of variation (CV) of the integrated HTLTPeatland prediction. The map shows spatial variation in relative uncertainty after propagating uncertainty from the HTPeatland and LTPeatland ensemble models through the temperature-weighted integration scheme. Low CV values indicate strong agreement among ensemble prediction. Across the global mapped area, the median CV was 0.0018 and the 95th percentile was 0.0082; within predicted peatland presence areas (probability≥0.6), the median CV was 0.0007 and the 95th percentile was 0.0046.

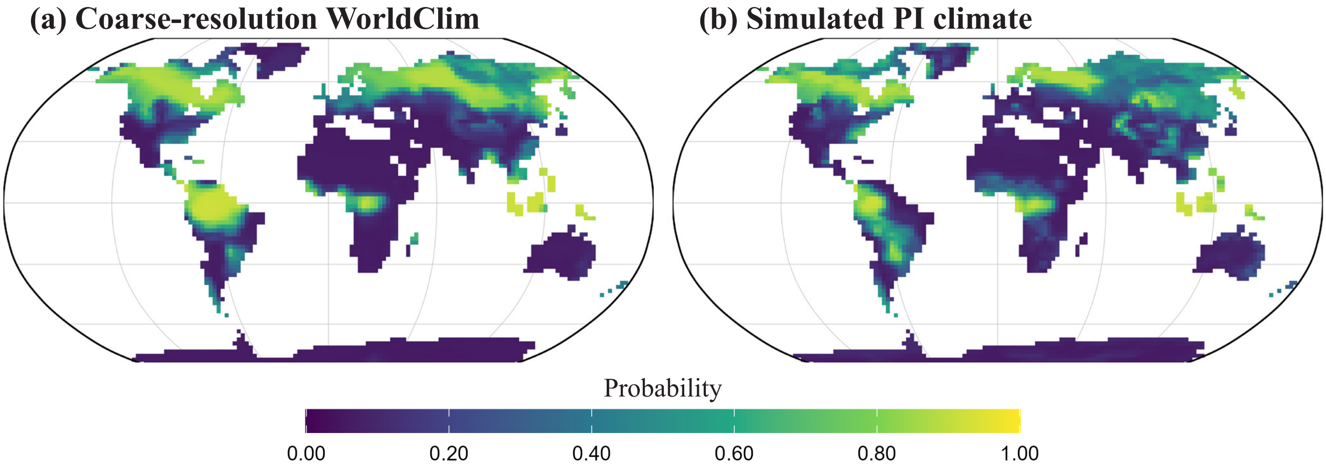


Figure S10. Potential distribution of HTLTPeatland projected using different climate datasets. (a) Projected using WorldClim data resampled to a horizontal resolution of 3.75° longitude by 2.5° latitude; (b) Projected using simulated preindustrial (PI) climate data derived from the HadCM3BL model. The underlying LTPeatland and HTPeatland models were the same as those used in Figure 8.