https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

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

Linlin Chen[1], Alexander Farnsworth[1,2], Paul J. Valdes[1,2]

[1]School of Geographical Sciences and Cabot Institute, University of Bristol, Bristol, BS8 1SS, UK
[2]State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

Correspondence to: Linlin Chen (linlin.chen@bristol.ac.uk)

Abstract. Peatlands and their fossilized counterpart, coal, are key indicators of past and present climate. However, tools for predicting their potential global distribution in the geological past remain limited. Here we use machine learning to build a climate-driven peatland distribution model, PeatClim v1.0, and to identify key climatic controls on peatland formation. The model is trained on bioclimatic variables in regions of modern peatland occurrence, aiming to estimate potential peatland distributions, rather than to reproduce observed maps. Results show that partitioning the global peatland dataset into low- and high-temperature subsets and training them separately improves model predictive performance and aligns better with observations. Diagnostic analysis reveals distinct dominant climatic controls for the two subsets: low-temperature peatlands (northern peatlands) are mainly controlled by annual temperature range, whereas high-temperature peatlands (tropical peatlands) are primarily controlled by annual precipitation. PeatClim v1.0 is designed for use with palaeoclimate model outputs, facilitating the prediction of potential coal deposits in Earth's history and palaeoclimate model-performance evaluation.

1 Introduction

Peatlands are major components of the global carbon and hydrological cycles and have influenced climate in both the present day and the geological past (Joosten and Clarke, 2002; Loisel et al., 2017). Although modern peatlands occupy only about 3 % of the global land surface, they store roughly one-third of global soil carbon and a substantial fraction of freshwater resources (Joosten and Clarke, 2002). Over the geological past, peat-forming ecosystems have also played an important role in the carbon cycle. Since the Last Glacial Maximum (LGM), global peatlands are estimated to have contributed a net uptake of 351 GtC over the last 21,000 years (Müller and Joos, 2020), playing a potentially significant role in glacial-interglacial cycles. In addition, peatlands and their fossilized counterpart, coal, are also widely used as indicators of past environmental and climatic conditions (Scotese et al., 2021). Regions with coal deposits are often interpreted as having been warm and humid (Scotese et al., 2021).

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

Although peat formation (or "peat initiation") depends strongly on local hydrological and topographic conditions, large-scale peatland development will also be closely linked to large-scale changes in climate and geography (Müller and Joos, 2020; Treat et al., 2019). For example, northern peatlands expanded during warm interglacial intervals and contracted during glacial cooling and permafrost development over the last 130,000 years (Treat et al., 2019). Since the LGM, tropical peatlands have been partly lost through flooding of continental shelves and partly regained through changes in temperature, precipitation, and CO_2. Large-scale latitudinal shifts in tropical peatlands were largely driven by changes in the position of the Intertropical Convergence Zone (Müller and Joos, 2020). On this basis, spatial and temporal variations in peat and coal deposits (Boucot et al., 2013; Treat et al., 2019) will generally record major shifts in peatland extent driven by climatic and geographic changes over the Phanerozoic (541 million years ago to present; Judd et al., 2024; Scotese, 2021; Scotese et al., 2021; Valdes et al., 2021). Understanding peatland distribution through deep time is therefore crucial both for reconstructing past climates, carbon-cycle dynamics and for interpreting climate-sensitive sedimentary archives. However, global peatland and coal distributions across the Phanerozoic remain incompletely resolved.

There are multiple approaches that allow the reconstruction of global peatland extent through time, each having their own limitations for reconstructing deep time spatiotemporal distributions. The measurement of modern global peatland extent primarily rely on soil and hydrological data derived from field surveys and remote sensing (Gumbricht et al., 2017; Hugelius et al., 2020a; Pontone et al., 2024; Skye et al., 2025a; Xu et al., 2018), which are costly and labour-intensive and cannot be directly extended to the geological past. Reconstructions of ancient peatland distribution rely mainly on compilations of buried peat and coal deposits (Boucot et al., 2013; Yu et al., 2010). Yet these records are temporally and spatially sparse, unevenly distributed, and strongly affected by preservation, burial, erosion, exposure, and sampling bias (Boucot et al., 2013; Yu et al., 2010). Although coal records have been used to investigate latitudinal patterns in coal distribution over the last 410 million years (Bao et al., 2023; Cao et al., 2019), such compilations alone cannot provide complete and continuous global maps of peatland distributions through deep time. In addition, the absence of coal in a given interval or region may bias paleoclimate interpretations when coal records are used as environmental indicators in isolation. A lack of coal has sometimes been overinterpreted as evidence that climatic conditions were too cold, too hot, and/or too dry to allow peatlands (and thus coals) to form.

Most global or large-scale peatland models have been developed primarily for modern, future, or late-Quaternary simulations rather than for Phanerozoic timescales (Mozafari et al., 2023; Silva et al., 2024). Peatland models typically depend on numerous predictors and boundary conditions related to topography, soils, vegetation, hydrology, radiation, and climate, many of which are unavailable or poorly constrained in deep-time settings. Moreover, these models mainly focus on hydrology, greenhouse-gas exchange, and carbon cycling, and many applications prescribe peatland extent or fraction rather than simulate peatland distribution dynamically (Silva et al., 2024). At present, the principal global peatland models available for simulating past dynamic peatland extent include the CLIMBER2-LPJ branch (e.g., Kleinen et al., 2012; Treat et al., 2019), the LPX-Bern branch and its later developments (e.g., Müller and Joos, 2020; Spahni et al., 2013; Stocker et al., 2014), and ORCHIDEE-PEAT/ORCHIDEE-MICT (e.g., Qiu et al., 2019), all of which have been applied to millennial-

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

timescale simulations. However, their extension to Phanerozoic applications is not yet established and would likely be limited by their reliance on modern-style ecohydrological parameterizations.

The strong link between large-scale peatland formation and climate suggests that long-term global peatland distribution may be predictable from past and present climate data. For example, Lottes and Ziegler (1994) identified peat-forming zones using empirical thresholds of temperature and precipitation, requiring monthly precipitation greater than 40 mm and temperature above 10 °C. While such methods can reproduce broad spatial patterns, they depend on prescribed threshold rules rather than data-driven statistical relationships, and therefore require prior assumptions about the dominant climatic controls on peatland formation, which may not hold true over geologic timescales.

To address these limitations, we present PeatClim v1.0, a climate-driven machine-learning model for global peatland distribution designed for climate-model applications, especially in deep time paleoclimate settings. PeatClim estimates potential distributions rather than attempting to reproduce observed distribution directly. Potential distribution informs where, given other valid conditions, peatland could have formed. However, they may not have formed due to other non-climate variables or may not have been preserved. This distinction is important because climatically suitable regions may not always preserve peat or coal in the geological record. By using machine learning, PeatClim v1.0 provides a fully data-driven, quantitative approach, unlike previous empirical-rule based methods. PeatClim is unique in using climate as its sole predictor, which makes it especially portable to deep-time applications. In contrast, many other peatland models require more detailed information on soil properties, terrain, geomorphology, and water table position, all of which are difficult to reconstruct robustly from the stratigraphical records alone. In this way, PeatClim provides an open source, fully data-driven and quantitative framework for exploring the potential global distribution of peatlands through the Phanerozoic.

Previous studies suggest that modern global peatlands occupy two broad climatic niches (Payne et al., 2019; Yu et al., 2009): (i) temperate to high-latitude peatlands (also referred to as northern peatlands), and (ii) tropical peatlands. However, a comprehensive assessment of the climatic controls on both niches across space and time has not yet been attempted. To test whether these niches differ in their climatic controls, we train PeatClim using either the full global peatland dataset or two temperature-based subsets, and then assess the relative importance of individual bioclimatic predictors.

Although coal formation depends on geological preservation as well as original peat accumulation, climatic conditions favourable for peat formation are expected to be a first-order control on the distribution of coal-forming environments (Lottes and Ziegler, 1994). On this basis, relationships learned by PeatClim from modern peatland occurrences and climate data may, in principle, be applied to paleoclimate model outputs to retrodict potential ancient peatland, and, more cautiously, coal-forming distributions. In this paper, however, we focus only on the methods and results for modelling modern peatland distribution. Applications to past peatland and coal distributions would be undertake in future studies.

All code and data used in this study could be found from https://zenodo.org/doi/10.5281/zenodo.20040294 (Chen et al., 2026).

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

2 Materials and Method

2.1 Climate data

Modern climate observations were sourced from the WorldClim v2.1 database, using the 19 bioclimatic variables at a resolution of 5 arc-minutes for average climatic conditions during 1970–2000 (Fick and Hijmans, 2017). These variables describe annual trends, seasonality, and extreme or limiting temperature and precipitation (Fick and Hijmans, 2017; Table S1). They were selected because peat forms from plant remains and we assume they are sensitive to biologically relevant climatic variables, such as Temperature Annual Range (BIO7) and Mean Precipitation of Warmest Quarter (BIO18). Some of the 19 bioclimatic variables are highly correlated across the global land surface (Fig. S1). For example, Annual Mean Temperature (BIO1) and Mean Temperature of Warmest Quarter (BIO10) have a pairwise Pearson correlation coefficient of 0.98, and Precipitation of Driest Month (BIO14) and Precipitation of Driest Quarter (BIO17) have a coefficient of 0.99. Worldclim has been used extensively to train ecological niche models to reconstruct a range of distributions in flora and fauna (e.g., Ribeiro et al., 2019; Tran et al., 2024; Waltari and Guralnick, 2008; Xiao et al., 2022).

2.2 Peatland data

Currently, there are several comprehensive observational global peatland datasets, including PEATMAP (Xu et al., 2017, 2018), the training dataset used in Peat-ML (Melton et al., 2022), Peat-DBase (Skye et al., 2025a, b), and The Global Peatland Database (GPD, https://greifswaldmoor.de/global-peatland-database-en.html, last access: 13 January, 2026). PEATMAP used meta-analysis to synthesize global peatland extent, including the most detailed regional extent in addition to large deposits. By contrast, the training dataset used in Peat-ML is limited to 14 extensive regions, which eliminates scattered peatlands that may be related to specific climate envelopes, such as peat formation in the midlatitudes of the Tibetan Plateau and Florida. Because of these limitations, we do not use the training dataset in Peat-ML. The latest dataset Peat-DBase focused primarily on peat depth. Its current version (v.1) has notable coverage gaps of tropical peatlands in the lower Amazon Basin, Eastern Indonesia, and Papua New Guinea, as well as northern peatlands in Eastern Russia and part of Canada, suggesting it is underestimating global peatland distribution. Peat-DBase shows less inclusive peatland cover than PEATMAP overall, especially across the boreal north, but includes additional localized coverage in some temperate regions, such as central Europe and eastern North America. The GPD focuses on peatland extent. However, it roughly divides into dominated or soil mosaic peatlands, and does not permit simple data access. In this study, we therefore used peatland occurrence data from PEATMAP (Xu et al., 2017) and Peat-DBase (Skye et al., 2025b) to train the PeatClim v1.0 model.

The data from PEATMAP, originally in polygon format, were converted to a 5-arcminute grid. A grid cell was classified as peatland-present when its centre point fell within a peatland polygon. Records falling outside the WorldClim land area were removed for consistency. This process yielded 140,280 peatland-presence grid cells worldwide.

Peat-DBase dataset includes coordinates for each site location. We filtered this dataset as follows: (1) retained only records with $error\_found = False$, (2) retained $sample\_duplication\_flag$ values of 1, 2, 3, 4 and 5, (3) retained only records

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

with $depth\_cm >= 30$, (4) removed records falling outside the WorldClim land area, and (5) retained only one record per 5-arcmin grid cell, selecting the record with the smallest $sample\_id$ and $location\_id$ (See Skye et al. (2025), for details of column names and values). Therefore, 6,290 peatland-presence records from Peat-DBase were retained (Table S2). Here, we used a common peat-depth threshold of 30 cm to classify peatland presence or absence (Joosten and Clarke, 2002). We included only peatland presence and excluded absence, because the PeatClim model aims to estimate the maximum potential distribution under given climatic conditions, whereas peatland absence may reflect additional non-climatic constrains, such as slope, that influence actual peatland formation. In addition, absence records in Peat-DBase show strong spatial sampling bias worldwide, with pronounced clustering in some countries, such as Czechia and Hungary, and sparse coverage in many other regions, including environmentally unsuitable areas where true absences would also be expected, which may introduce model bias. We then combined the filtered PEATMAP and Peat-DBase and retained only one record per 5-arcmin grid cell. In total, 144,259 records were retained worldwide (Fig. 1). We refer to this dataset as GlobalPeatland.

The GlobalPeatland dataset was divided into high-temperature (HTPeatland) and low-temperature (LTPeatland) subsets using k-means clustering (k=2) based on BIO1 (annual mean temperature) values, with a threshold of 11.37 °C defined as the average of the two cluster centroids (Fig. 1a). This partition assigned 12,064 grid cells to the HTPeatland subset and 132,195 grid cells to the LTPeatland subset, corresponding to 8.4 % and 91.6 % of the GlobalPeatland dataset, respectively. Most HTPeatland presences occur in the modern tropics, corresponding broadly to tropical peatlands in the literature, whereas LTPeatland presences are widespread in boreal and temperate areas, corresponding to temperate/northern peatlands (Fig. 1b). Data filtering, integration, and splitting were conducted in R (R Core Team, 2023).

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

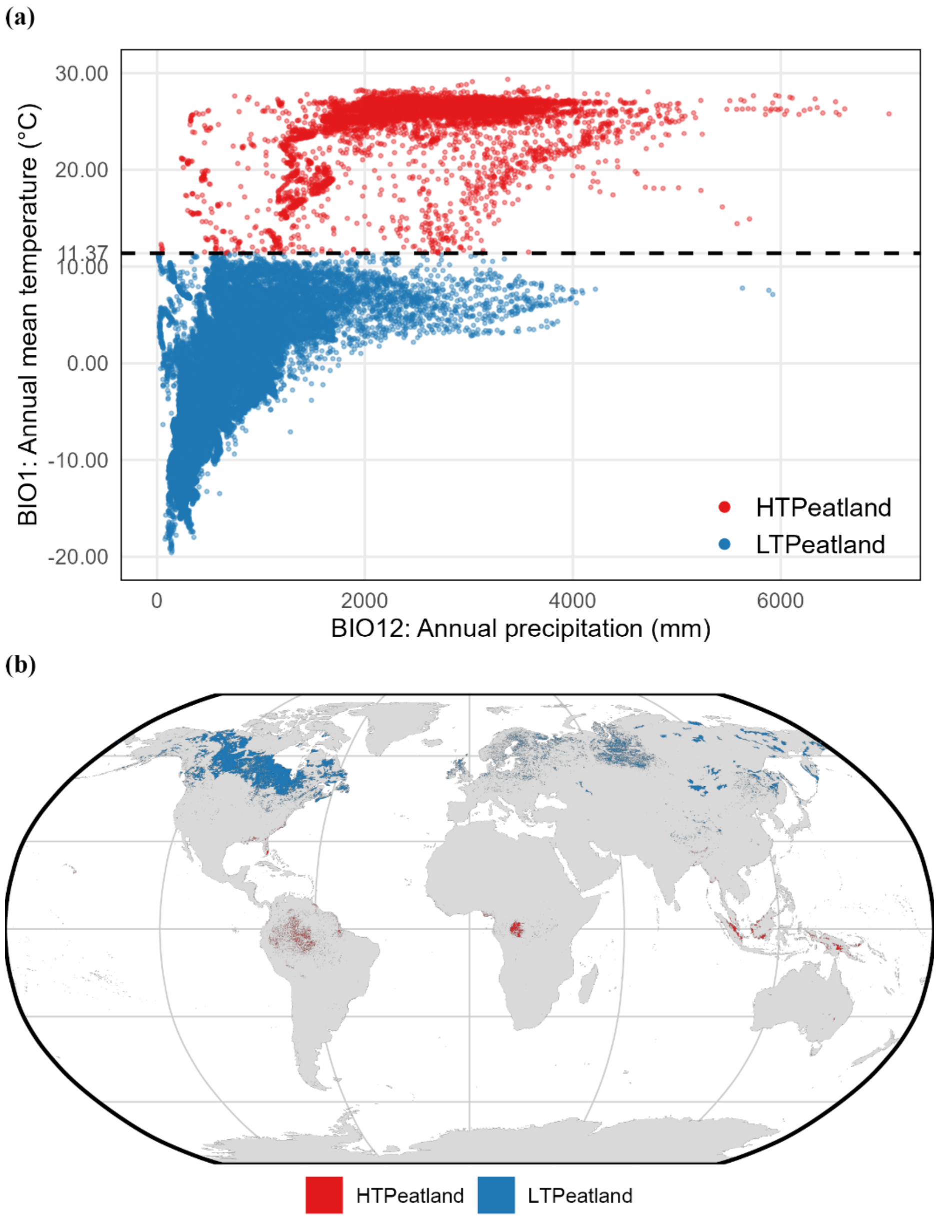


Figure 1. Climatic and spatial distributions of modern global peatlands from PEATMAP and Peat-DBase. (a) Scatterplot of peatland presencess in climate space defined by annual mean temperature (BIO1; °C) and annual precipitation (BIO12; mm). (b) Global spatial distribution of peatland presences at 5 arcminute resolution. Peatlands were divided into a high-temperature peatland subset (HTPeatland; red dots and squares) and low-temperature peatland subset (LTPeatland; blue dots and squares) using a BIO1 threshold of 11.37 °C.

2.3 Model training and evaluation

Conceptually, the PeatClim model learns relationships between peatlands and climate. It estimates the probability of peatland presence from modern peatland occurrences and climate observations. Once trained, the model can be applied to

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

predict potential peatland distributions for other time periods using the corresponding climate data. The workflow is shown in Figure 2.

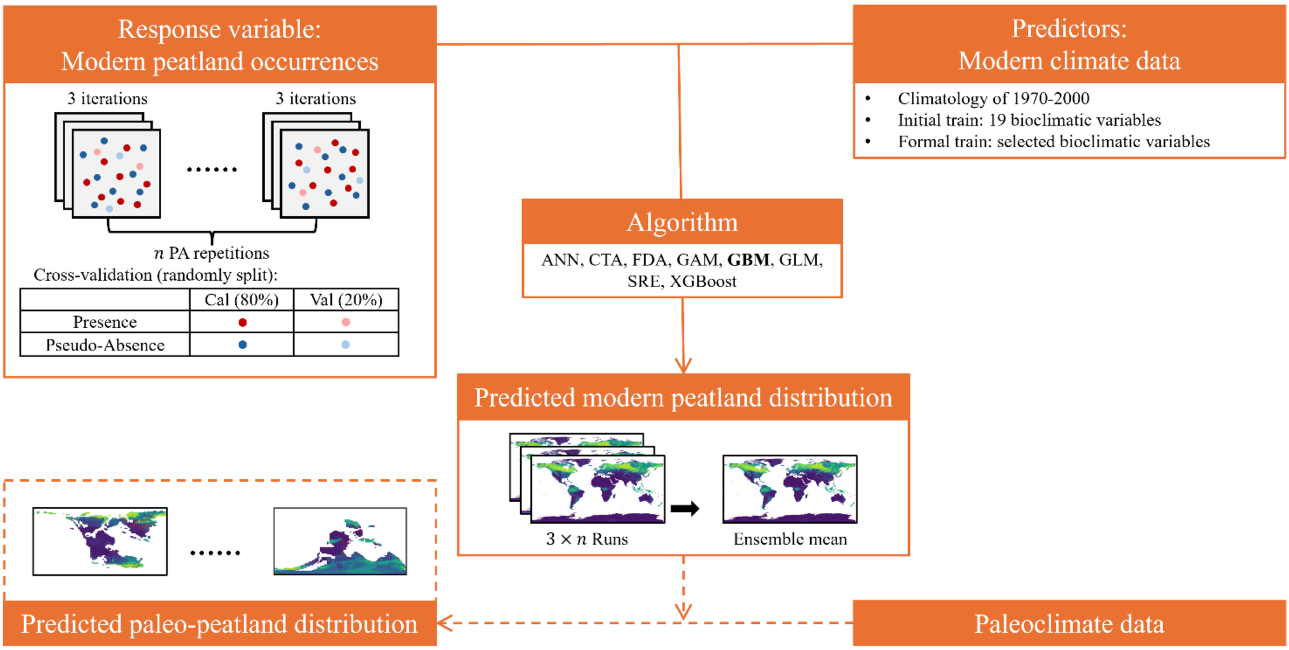


Figure 2. Schematic workflow of the climate-driven peatland distribution model (PeatClim). Modern peatland occurrences are used as the response variable and modern climate data as predictors. In each training round, the same peatland presence data are combined with $\boldsymbol{n}$ different pseudo-absence (PA) datasets. For each combination, the presence and PA data are independently and randomly split into calibration (Cal, 80 %) and validation (Val, 20 %) subsets and evaluated over three cross-validation iterations. Candidate algorithms are tested separately, after which the selected algorithm, Generalised Boosted Models (GBM), is retained. Predictions from all $\mathbf{3} \times \boldsymbol{n}$ runs are averaged to obtain the ensemble mean modern peatland distribution. The calibrated model subsequently could be applied to paleoclimate inputs to predict potential paleo-peatland distribution.

2.3.1 Model setting

Model training was conducted using the R package *biomod2* (Thuiller et al., 2025), an ensemble platform that provides various algorithms for building species distribution models (Thuiller, 2003; Thuiller et al., 2009). In the initial training to select algorithm and pseudo-absence (PA) setting, we used the filtered PEATMAP dataset and all 19 bioclimatic variables. Multiple algorithms were tested, including Artificial Neural Networks (ANN), Classification Tree Analysis (CTA), Flexible Discriminant Analysis (FDA), Generalised Additive Models (GAM), Generalised Boosted Models (GBM), Generalised Linear Models (GLM), Surface Range Envelope (SRE) and eXtreme Gradient Boosting (XGBoost).

Pseudo-absence data were generated using a random strategy to indicate peatland absences for model training. For some algorithms, model performance is sensitive to both the number of PA records (nb.PA) and the number of PA repetitions (nb.PA.rep) (Barbet-Massin et al., 2012). To evaluate this, a range of nb.PA was set up across 100, 1,000, 5,000, 10,000, 140,000, and 450,000 records, representing very small, small, medium, and large PA sets, plus cases approximately equal to

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

and three times the number of presence records. The nb.PA.rep was set to 3, 10, or 20 depending on nb.PA. In each repetition, we used cross-validation with a random 80/20 split to partition the presence and PA datasets**.** During training, each setting was run for three iterations, and an unweighted scheme was adopted with unbalanced class sizes.

2.3.2 Evaluation and diagnostic

We evaluated model performance using the threshold-dependent metric True Skill Statistic (TSS; Allouche et al., 2006), and selected the classification threshold that maximized TSS. Higher TSS score indicate better model fit; and consistent scores between calibration and validation datasets indicate stable, reliable performance. We interpreted TSS using common thresholds: excellent for TSS ≥ 0.75, high accuracy for TSS ≥ 0.7, relatively satisfactory for TSS ≥ 0.5, or low accuracy for TSS < 0.5.

We further applied explainable functions in *biomod2* to interpret predictor marginal effects via response curves and to quantify predictor influence via permutation-based variable importance (Elith et al., 2005; Thuiller et al., 2009). Models showing consistent response curves across PA repetitions and iterations were considered more stable (Elith et al., 2005). Predictors with strong responses and high permutation importance were judged more influential in simulating peatland distribution. Accordingly, we included all 19 predictors for the initial training of GlobalPeatland, LTPeatland and HTPeatland, and then reduced the set to key bioclimatic variables for formal training. If any pair of predictors had $|r| > 0.90$ (Pearson), only the predictor with the higher permutation importance was retained. This reduction mitigates collinearity and overfitting, as well as reduces computational cost.

Based on TSS scores and diagnostics from the initial experiments (Fig. 4), we selected Generalised Boosted Models (GBM), 10 PA repetitions, and an nb.PA value equal to the number of presences for subsequence training. Accordingly, nb.PA was set to 144,259 for GlobalPeatland, 132,195 for LTPeatland, and 12,064 for HTPeatland training. We fixed the PA sample size because GBM performance was not sensitive to PA counts (See more in section 3.1 and Fig. 4; Barbet-Massin et al., 2012).

The bioclimatic variables selected for formal training of GlobalPeatland were Minimum Temperature of Coldest Month (BIO6), Temperature Annual Range (BIO7), Mean Temperature of Warmest Quarter (BIO10), Annual Precipitation (BIO12), Precipitation of Driest Month (BIO14) and Precipitation of Warmest Quarter (BIO18). The variables selected for LTPeatland were similar to those for GlobalPeatland but excluded BIO12. For formal training of HTPeatland, we used Temperature Annual Range (BIO7), Mean Temperature of Wettest Quarter (BIO8), Mean Temperature of Coldest Quarter (BIO11), Annual Precipitation (BIO12), and Precipitation of Driest Quarter (BIO17). The ensemble means of predicted probabilities for LTPeatland and HTPeatland were then calculated separately from 30 independent runs (10 PA repetitions × 3 iterations), using the same random seeds across models.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

2.3.3 Projection

To confirm model effectiveness, we first projected the trained model using the same 5 arc-minute WorldClim data used as predictors. To assess projection performance with other climate data, we then tested the projections using a simulated preindustrial (PI) climate data. The PI climate data were derived from the atmosphere-ocean coupled general circulation model HadCM3BL, which is designed to simulate paleoclimate throughout the Phanerozoic (Valdes et al., 2017). This paleoclimate model uses physical principles to simulate climate under different CO_2 concentrations, geographic configurations, and solar and orbital settings for different time snapshots. Its PI simulation, based on 280 ppmv CO_2, modern geography, and modern solar and orbital settings, performs moderately well relative to 21 CMIP5 models across nine atmospheric variables. Because of its application to long-timescale simulations, HadCM3BL uses a relatively coarse horizontal resolution of 3.75° longitude by 2.5° latitude (Valdes et al., 2017). To eliminate the influence of differences in spatial resolution, we further converted the WorldClim data from 5 arc-minutes to this coarser resolution and generated projections using the coarse WorldClim data.

2.3.4 Training dataset robustness

The training dataset is crucial for machine learning models (Roscher et al., 2024). In PeatClim, training data derived from PEATMAP polygons were converted into grid cells using a rule in which only polygons crossing cell centres were classified as peatland presence. We also tested an alternative conversion rule, in which any PEATMAP polygon touching a grid cell was classified as presence. This new rule produced a dataset with more records (Fig. S2) than the original training dataset, especially in the tropics where PEATMAP identifies many scattered local peatlands located along specific landscape features such as valleys and rivers. When trained on this new dataset, the model produced a similar potential peatland distribution (Fig. S3) to that generated by the original method. Although the diagnostic highlighted BIO1 for LTPeatland (Fig. S4) rather than BIO7, as in the original diagnostics, this suggests that PeatClim can robustly predict a comparable peatland distribution from either a detailed dataset or a dataset with fewer records, as long as most climate niches are covered.

2.4 Integrating HTPeatland and LTPeatland distributions

Although the HTPeatland and LTPeatland models were trained on their respective subsets, both were projected globally to produce probabilities at every grid cell. To blend these two global probability fields into a single map, we calculated weighted probabilities from the formal training results of HTPeatland and LTPeatland and denoted the integrated distribution as HTLTPeatland. The HTLTPeatland probability in each grid cell is given by

$$P_{HTLTPeatland} = w_{HTPeatland} \times P_{HTPeatland} + w_{LTPeatland} \times P_{LTPeatland} \quad (1)$$

where P_x is the probability of peatland presence predicted by the corresponding training, ranging from 0 to 1; and w_x is a logistic weight, given by

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

$$w_{HTPeatland} = \frac{1}{1 + \exp\left(-\frac{T - t_0}{\omega}\right)} \quad (2)$$

and

$$w_{LTPeatland} = 1 - w_{HTPeatland} \quad (3)$$

where T is the annual mean temperature in each grid cell, in °C; t_0 is the boundary of annual mean temperature between HTPeatland and LTPeatland subsets, which is 11.37 °C; ω defines a transition width and is set to 2 °C to provide a smooth blend across the temperature boundary (ω was chosen heuristically and not tuned). In each grid cell, the sum of $w_{HTPeatland}$ and $w_{LTPeatland}$ equals to 1.

Following these equations, the logistic weight and probability depend on how close the temperature is to the boundary t_0 (Fig. 3). In "warm" regions where $T \gg t_0$, the resulting $w_{HTPeatland} \approx 1$ and $P_{HTLTPeatland} \approx P_{HTPeatland}$; in "cool" regions where $T \ll t_0$, the resulting $w_{LTPeatland} \approx 1$ and $P_{HTLTPeatland} \approx P_{LTPeatland}$; in regions where T is near the boundary t_0, the logistic weights $w_{HTPeatland}$ and $w_{LTPeatland}$ provide soft transitions across the temperature boundary. The resulting probability of HTLTPeatland is able to maintain the individual distribution patterns of HTPeatland and LTPeatland and provides a soft blending near their boundary.

We used sensitivity analysis (Elith et al., 2005) to evaluate the predicted distribution of HTLTPeatland. Predicted HTLTPeatland grid cells were classified as peatland presence when the predicted probability was equal to or greater than a given threshold, and as absence otherwise. We evaluated four probability thresholds of 0.5, 0.6, 0.7 or 0.8. Sensitivity was defined as the proportion of observed presences that were correctly predicted, reflecting how well the map identified actual peatland presence. Correctly predicted presences, referred to as true positive (TP), were peatland occurrences in the integrated PEATMAP and Peat-DBase dataset that fell within areas predicted as peatland presence. In contrast, peatland occurrences falling within areas predicted as absence were classified as false negative (FN). Sensitivity was therefore calculated as TP/(TP+FN), with values ranging from 0 to 1. Higher sensitivity indicates better performance in identifying HTLTPeatland presence.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

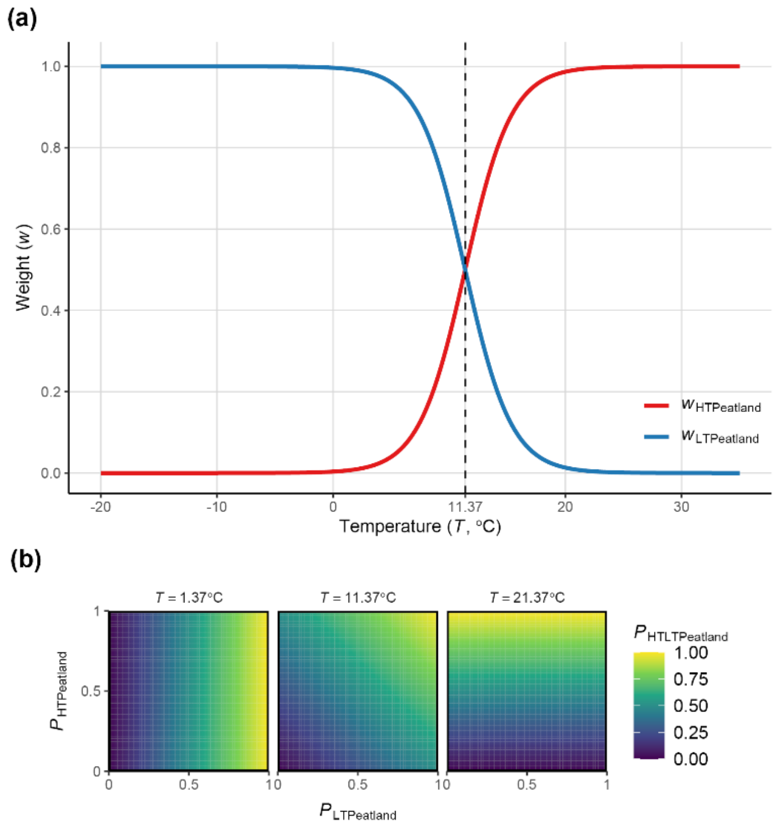


Figure 3. Weighted combination scheme for the HTLTPeatland probability. (a) Temperature-dependent weights for the HTPeatland model, $\boldsymbol{w_{\mathrm{HTPeatland}}}$ (red), and the LTPeatland model, $\boldsymbol{w_{\mathrm{LTPeatland}}}$ (blue), shown as logistic functions of temperature $\boldsymbol{T}$, with $\boldsymbol{t_0} = \mathbf{11.37}$ °C and $\boldsymbol{\omega} = \mathbf{2}$ °C; (b) Integrated weighted probability, $\boldsymbol{P_{HTLTPeatland}}$ (color shaded), for three illustrative cases: $\boldsymbol{T} \ll \boldsymbol{t_0}$, $\boldsymbol{T} = \boldsymbol{t_0}$, and $\boldsymbol{T} \gg \boldsymbol{t_0}$.

To quantify uncertainty in the integrated HTLTPeatland projection, we propagated ensemble uncertainty from the HTPeatland and LTPeatland ensemble models. The coefficient of variation (CV) was used to represent relative uncertainty in predicted probability across ensemble runs. Because CV is a relative measure, uncertainty in the integrated HTLTPeatland map was propagated using standard deviations (SD) rather than by directly averaging CV values. The propagated CV of HTLTPeatland was given by

$$CV_{HTLTPeatland} = \frac{SD_{HTLTPeatland}}{P_{HTLTPeatland}} = \frac{\sqrt{w_{HTPeatland}^2 SD_{HTPeatland}^2 + w_{LTPeatland}^2 SD_{LTPeatland}^2}}{P_{HTLTPeatland}} \tag{4}$$

where CV_x and SD_x are the coefficient of variation and standard deviation of the corresponding ensemble training, respectively. The standard deviation for each component model was obtained as $SD_x = CV_x \times P_x$. $P_{HTLTPeatland}$ was calculated using Eq. (1) from the ensemble mean probabilities of HTPeatland and LTPeatland, and w_x was calculated using Eqs. (2) and (3). This uncertainty propagation assumes independence between the HTPeatland and LTPeatland ensemble uncertainties. Lower CV values indicate lower relative uncertainty and stronger agreement among ensemble predictions.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

3 Results and discussion

3.1 Evaluation of algorithm and pseudo-absence settings

We evaluated the algorithm, nb.PA, and nb.PA.rep in terms of accuracy, stability, and efficiency aspects using TSS scores, projections, and predictor response curves.

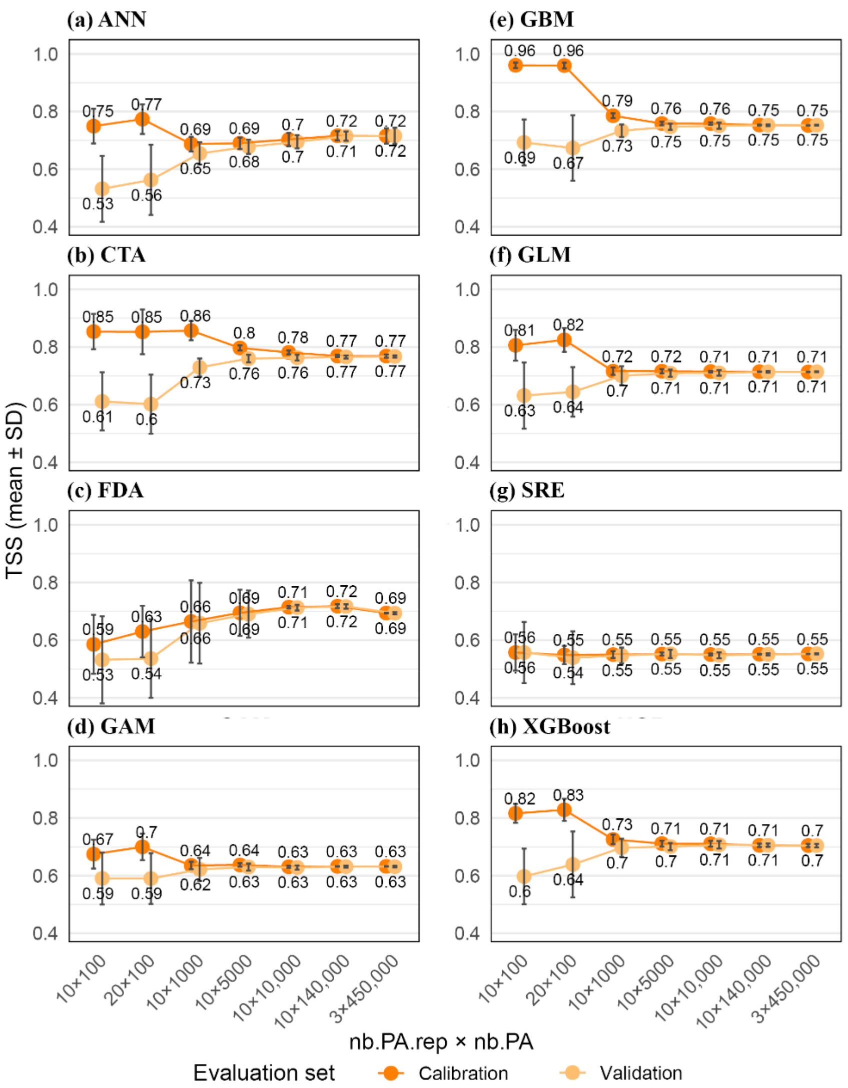


Figure 4. True Skill Statistic (TSS) for models trained under different algorithms and pseudo-absence (PA) settings. Each x-axis category denotes the number of PA repetitions and the number of PA records (**nb. PA. rep** × **nb. PA**). All models were trained with the filtered PEATMAP dataset and 19 bioclimatic variables. All models used model-specific thresholds selected to maximize TSS. Dark and light orange colors indicate the calibration and validation sets, respectively.

Figure 4 summarises TSS across algorithms and PA settings. With small PA sets (nb.PA<5,000), most algorithms (except SRE) are unstable, showing large calibration–validation gaps and large TSS standard deviations. When nb.PA≥10,000, most

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

algorithms (except SRE and GAM) attain high accuracy (TSS≥0.7) and become insensitive to further increases in nb.PA. Increasing nb.PA beyond 140,000 does not improve TSS but incurs computational cost. CTA and GBM achieve the highest TSS (≥0.75), whereas SRE consistently yields the lowest TSS (~0.55), indicating relatively satisfactory performance. On this basis, CTA and GBM, with around 10,000–140,000 PA records and 10 repetitions, emerged as leading candidates for subsequent assessment.

Another aspect of assessment is the spatial projection: how well predicted potential distributions align with observed occurrences. Although many algorithms yield comparable TSS, their spatial predictions based on 5 arc-minute WorldClim data differ (Fig. S5). Predictions from CTA, FDA, and GBM align most closely with peatland occurrences and are consistent across iterations. Take northern Eurasia as an example, ANN exhibits high variability across iterations; some runs concentrate in the northeast, whereas others are broadly spread. GAM overpredicts peatland presence in the northeastern Eurasia where peatland records are sparse. GLM and XGBoost produce broad patterns that fail to capture the more concentrated belt in north-central Eurasia. SRE overestimates peatland presences across large areas.

Predictor response curves also differ across algorithms (Fig. S6). GBM demonstrates the most consistent trends across PA repetitions and iterations, indicating stable behaviour. In contrast, ANN exhibits highly variable curves across runs. CTA shows stepwise changes due to its tree-based nature and is not very stable across runs. FDA presents smooth response curves but with substantial variability. GAM, GLM, SRE, and XGBoost present stepwise responses in some runs but little or no response in others.

In summary, based on TSS scores, spatial distributions, and response-curve assessment, GBM provides the most accurate and stable performance. Using nb.PA ranges around 10,000–140,000 and 10 PA repetitions yields the best training. These optimal settings were applied for all subsequent analyses.

3.2 Initial training with 19 bioclimatic variables

Initial HTPeatland and LTPeatland training with GBM and 19 bioclimatic variables show higher TSS than the initial GlobalPeatland training (Fig. 5). The TSS increases from about 0.75 for GlobalPeatland to about 0.80 for LTPeatland and about 0.90 for HTPeatland, highlighting the benefit of temperature-based stratification.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

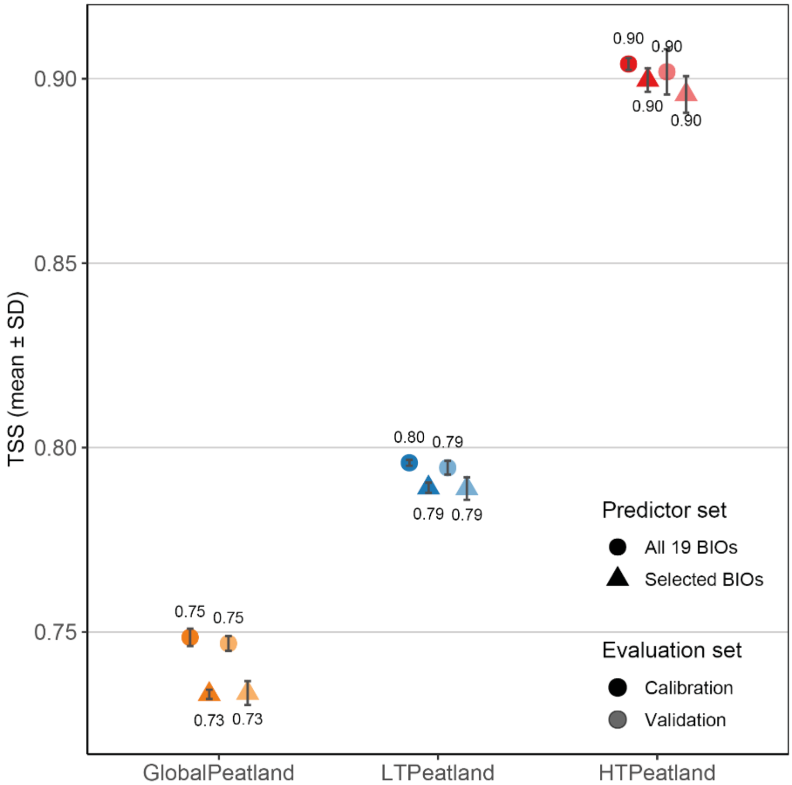
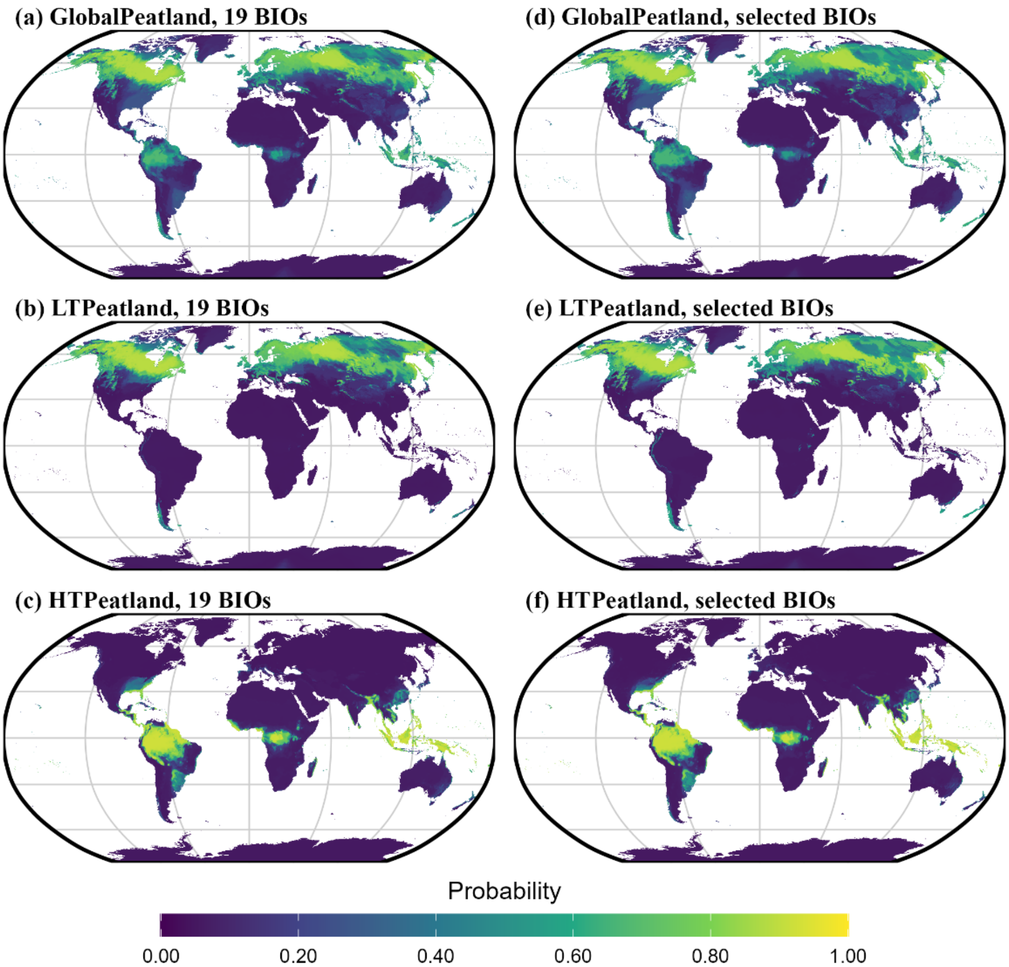


Figure 5. TSS for models trained using either all 19 bioclimatic variables or selected subsets. Models were trained using GBM with nb.PA.rep=10, iteration=3, and nb.PA=144,259 for GlobalPeatland, 132,195 for LTPeatland, and 12,064 for HTPeatland. The selected predictor subset comprised BIO6, BIO7, BIO10, BIO12, BIO14, and BIO18 for the GlobalPeatland model; BIO6, BIO7, BIO10, BIO14, and BIO18 for the LTPeatland model; and BIO7, BIO8, BIO11, BIO12, and BIO17 for the HTPeatland model. Dots and triangles indicate models trained with all 19 variables and the selected subsets, respectively. All models used model-specific thresholds selected to maximize TSS, with optimal thresholds of 0.53±0.02. Dark and light colors indicate the calibration and validation sets, respectively.

Potential distributions predicted for LTPeatland and HTPeatland using 5 arc-minute WorldClim data align well with observed peatland occurrences and cover distinct climate regimes (Fig. 6). LTPeatland resembles the “cool” regime in the GlobalPeatland prediction, whereas HTPeatland shows higher probabilities in the tropics and a spatial pattern similar to the corresponding tropical areas in GlobalPeatland. The high probabilities in HTPeatland are more credible than those from GlobalPeatland training, because modern observations clearly show peatland occurrences in the tropics, as well as in temperate regions. The single-group model underestimates tropical probabilities, likely reflecting data imbalance: LTPeatland accounts for 91.6 % of the global records, whereas HTPeatland comprises only 8.4 %. The subgroup with the larger percentage of the global records dominates the fit. Response curves for initial GlobalPeatland, HTPeatland and LTPeatland training show small between-run variations across PA repetitions and iterations (Fig. S7).

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026



325 **Figure 6. Predicted potential distributions from models trained using either all 19 bioclimatic variables or selected subsets.** Panels (a–c) show results from one of the 30 runs and panels (d–e) show results from ensemble mean for models trained with GlobalPeatland, LTPeatland, and HTPeatland, respectively. All panels are projected using 5 arc-minute WorldClim data. Other model settings are the same as those in Fig. 5. Colors indicate the predicted probability of peatland presence from 0 (blue) to 1 (yellow).

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

3.3 Formal training with selected bioclimatic variables

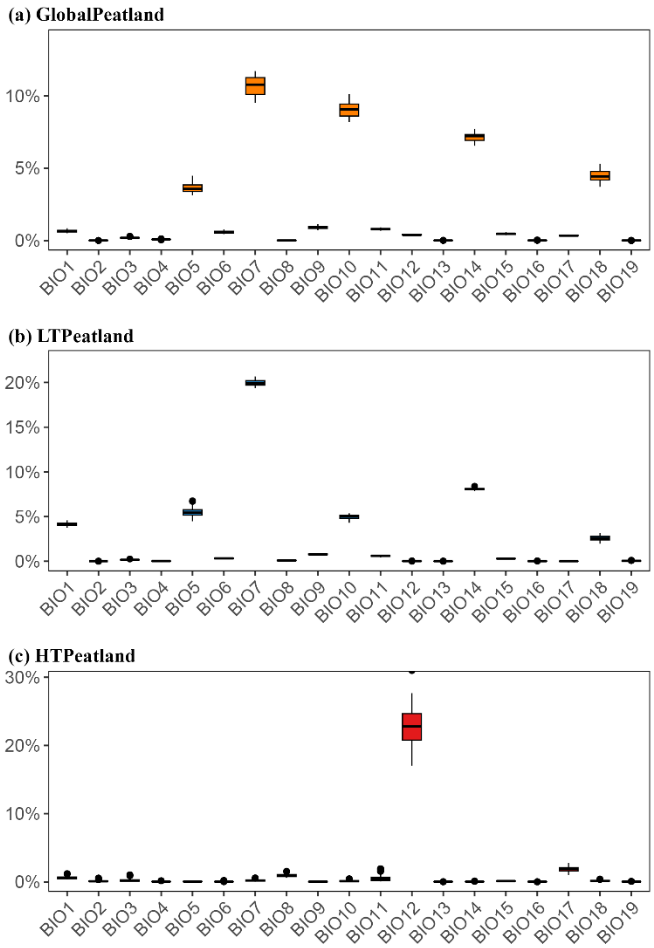


Figure 7. Permutation-based variable importance for models trained with all 19 bioclimatic variables. Panels (a)–(c) show results for the GlobalPeatland, LTPeatland, and HTPeatland models, respectively. Other model settings are the same as those for the 19-variable models in Fig. 5. Boxes indicate the interquartile range (IQR) with median lines; whiskers extend to 1.5×IQR, and points beyond the whiskers are shown as outliers. For readability, the y-axis is truncated at the 99th percentile of variable-importance values, and extreme outliers above this limit are omitted.

We selected key bioclimatic variables for formal model training using permutation-based variable importance, predictor response curves and pairwise correlations. Figure 7 shows variable importance from the initial 19-variable models across the three datasets. GlobalPeatland and LTPeatland show similar important variables, although their relative contributions differ (Fig. 7a,b). BIO7 stands out, accounting for about 10.65±0.66 % of variable importance in GlobalPeatland and about 19.97±0.31 % in LTPeatland; BIO14 and BIO10, followed by BIO5 and BIO18, also rank highly and show substantial changes in the response curves for GlobalPeatland and LTPeatland (Fig.7a,b, Fig. S7). Other predictors with relatively weak responses include BIO1 and BIO6 in these two datasets and BIO12 in GlobalPeatland (Fig. S7). Because BIO1, BIO5, and BIO10 are highly correlated (Pearson |r|>0.90; Fig. S1), we retained only BIO10, which had the highest importance. We

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

therefore selected BIO6, BIO7, BIO10, BIO12, BIO14, and BIO18 for GlobalPeatland, and BIO6, BIO7, BIO10, BIO14, and BIO18 for LTPeatland.

Predictor responses in initial HTPeatland training differs from the other two models. HTPeatland assigns the greatest importance to one dominant predictor, Annual Precipitation (BIO12; 22.87±2.94 %); while other predictors contribute little (Fig. 7). BIO12 also shows a strong response, followed by BIO8, BIO11, BIO7, and BIO17 (Fig. S7). Thus, we chose these five predictors for formal HTPeatland training. Notably, variable importance and response curves show larger between-run variation for initial HTPeatland training than for the initial GlobalPeatland and LTPeatland training, which may relate to imbalanced sample sizes.

Differences in key bioclimatic variables between LTPeatland and HTPeatland indicate that peatland formation is governed by different climatic controls in temperate-to-high-latitude regions and tropical regions. Temperate and high-latitude peatlands require distinct warm and cool seasons, as indicated by high occurrence probability when BIO7 ranges from about 34 to 56 °C (Fig. S7b). They also require the mean temperature of the warmest quarter to fall within a specific range (Fig. S7b; BIO10, about 5–18 °C), together with an adequate precipitation during the warmest quarter (Fig. S7b; BIO18, great than 259 mm) to support plant growth and organic matter accumulation. In addition, LTPeatland requires a certain level of precipitation during the driest month (Fig. S7b; BIO14, greater than 20 mm) to maintain waterlogged conditions, as well as low temperature during the coldest month (Fig. S7b; BIO6, below -22 °C), which slow decomposition. By contrast, tropical peatlands require high annual precipitation (Fig. S7c; BIO12, greater than 2625 mm). This may be because sustained, abundant precipitation facilitates organic matter accumulation via both high net primary production and suppressed decomposition rates due to high water level position. Previous studies support some of our results. For example, Hugelius et al. (2020) found that mean summer temperature, together with underlying mineral soil texture, peatland extent in the area, and photosynthetically active radiation, were among the most important variables for predicting peat depth across the Northern Hemisphere (>23° latitude). Lottes and Ziegler (1994) indicate sufficient precipitation over long periods facilitates productivity accumulation and cooler conditions favour peat preservation.

Subsequent training with the selected bioclimatic variables yielded TSS values comparable to the 19-variable training (Fig. 5). Calibration TSS with 19 versus 5 variables were 0.75 vs 0.74 for GlobalPeatland, 0.80 vs 0.79 for LTPeatland, and 0.91 vs 0.91 for HTPeatland. Moreover, the reduced-set predicted distributions closely matching the full-set predictions (Fig. 6). Taken together, TSS and spatial agreement support the robustness of the tailored models.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

3.4 HTLTPeatland potential distribution

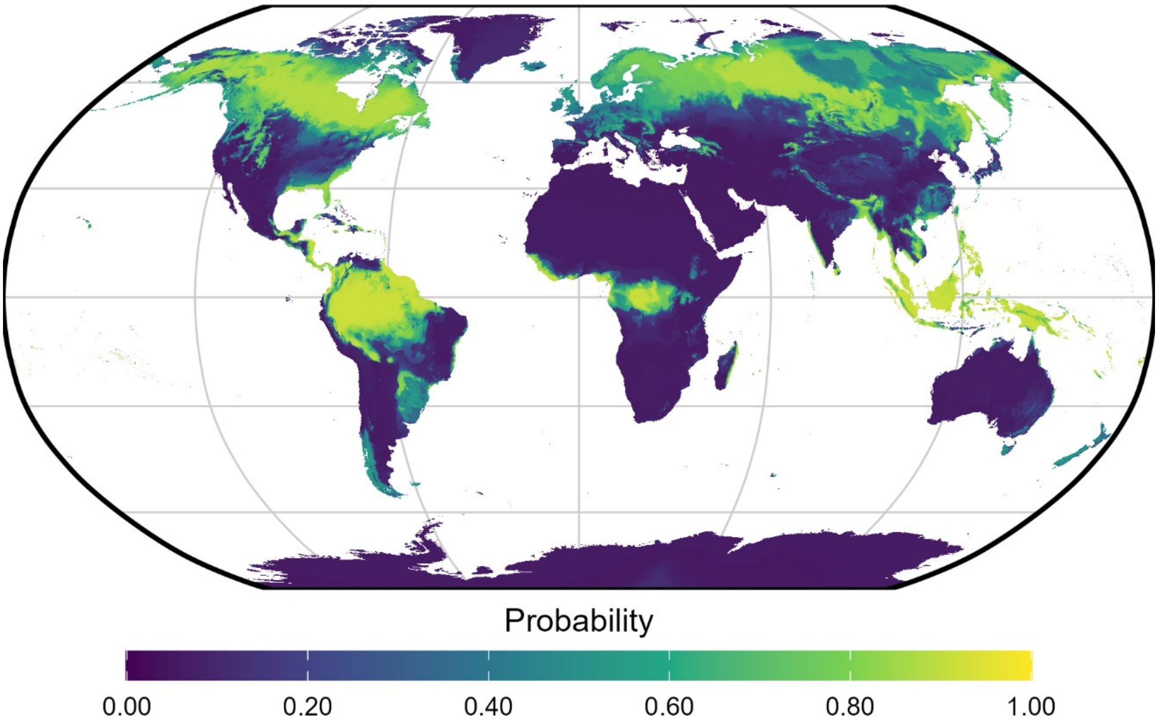


Figure 8. Predicted potential distribution of HTLTPeatland. The predicted probability of HTLTPeatland was calculated by integrating the LTPeatland and HTPeatland ensemble model outputs using the temperature-dependent weighting scheme described by equations (1)–(3). The underlying LTPeatland and HTPeatland models were trained with selected bioclimatic variables and projected using 5 arc-minute WorldClim data, as shown in Fig. 6(e) and (f).

Table 1. Evaluation of HTLTPeatland potential distribution under different classification thresholds. The table reports the number of true positives (TP), false negatives (FN), and sensitivity values based on comparison with the integrated PEATMAP and Peat-DBase dataset.

Threshold	True Positive (TP)	False Negative (FN)	Sensitivity
0.5	137,462	6,797	0.953
0.6	130,427	13,832	0.904
0.7	115,274	28,985	0.799
0.8	94,544	49,715	0.655

Figure 8 shows the predicted potential distribution of HTLTPeatland, calculated by applying a temperature-dependent weighting scheme to the LTPeatland and HTPeatland predicted probabilities derived from 5 arc-minute WorldClim data. The predicted HTLTPeatland distribution shows good agreement with the integrated PEATMAP and Peat-DBase dataset (Table 1, Fig. S8, and Fig.1b). Using a probability threshold of 0.5 to classify predicted peatland as presence (probability≥0.5) or

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

absence (probability<0.5), the model correctly identifies 137,462 peatland occurrences, corresponding to a sensitivity of 0.953, while 6,797 occurrences fell within grid cells predicted as absence. Sensitivity decreased with increasing threshold but remained 0.655 even at a threshold of 0.8.

The HTLTPeatland potential distribution shows a spatial pattern comparable with observed peatlands in the integrated PEATMAP and Peat-DBase dataset. Extensive high-probability areas concentrate in the boreal and subarctic regions, including boreal North America (notably the Hudson Bay Lowlands and adjacent Canadian lowlands), Fennoscandia, West Siberian Lowland and parts of the Russian Far East, as well as tropics in the Amazon lowlands, the Congo Basin, and the Maritime Southeast Asia. Smaller, regional clusters are also apparent in the U.S. Atlantic and Gulf Coastal Plains, Central American lowlands, the Gulf of Guinea coastal belt, the eastern Himalayan foothills, and part of the southeastern South America.

The HTLTPeatland high-probability peatlands are, however, broader than observed peatland presences. This discrepancy is expected because the model represents climate-constrained potential, indicating regions where modern climate conditions are suitable for peatland initiation with a given probability. Yet local controls (e.g., geomorphology, topography, and substrate) and human impacts (e.g., drainage and rewetting) can limit or eliminate peatland occurrence. For instance, steep terrain is unlikely to maintain peat even under favourable climate, and heavily urbanized landscapes may no longer support peatlands despite climatic suitability.

Ensemble uncertainty was generally low across the integrated HTLTPeatland projection, with elevated CV values mainly occurring along distribution margins and fragmented transition zones (Fig. S9). The median CV was 0.0018 across the global mapped area and 0.0007 within predicted peatland presence areas (probability ≥ 0.6), indicating strong ensemble agreement in the final peatland prediction.

The HTLTPeatland projection based on the coarse-resolution WorldClim data (Fig. S10a) largely resembles that generated using the 5 arc-minute WorldClim data (Fig.8). The HTLTPeatland projection using the simulated PI climate data (Fig. S10b) shows a broadly similar pattern to the WorldClim-based projections, but differs in several regions. Specifically, the PI-based prediction predicts reduced probabilities in the West Siberian Lowland and more restricted high-probability areas in northern North America and northern South America, but locally higher probabilities in southwestern Central Asia. These differences may arise from differences among climate datasets, or they may reflect the complex climatic controls on peat initiation. This highlights the need for caution when interpreting projections based on different climate datasets.

4 Conclusions

In this study, we present PeatClim v1.0, a climate-driven machine-learning model that estimates the probability of peatland presence from modern peatland occurrences and a comprehensive set of bioclimatic variables. The global peatland dataset is either treated as a single group or partitioned into low- and high-temperature subgroups based on annual mean temperature. The two subgroups are trained separately, and their predictions are integrated into a single global probability map.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

We find that: (1) among all algorithms tested, GBM demonstrates the highest accuracy and stability in simulating global peatland distributions; (2) separate training of the two subgroups improves model performance relative to single-group training; and (3) for LTPeatland, Temperature Annual Range (BIO7) is the most influential predictor, followed by Precipitation of Driest Month (BIO14) and Mean Temperature of Warmest Quarter (BIO10), whereas for HTPeatland, Annual Precipitation (BIO12) plays the dominant role. These results highlight the value of both algorithm selection and ecological stratification in enhancing the predictive modelling of peatland distributions under modern scenarios.

The strength of our model is that it only requires climate inputs and does not require detailed soil or geomorphological data. It is predicting potential peatlands. However, there are some limitations. The climate data used to train PeatClim are climatological averages from 1970 to 2000, whereas present-day peatlands are formed in different time periods, some may even date back to 50 thousand years ago (Treat et al., 2019). This temporal mismatch could introduce predictive bias. But we consider this bias to be negligible because the extent of most large peatlands is stable on a decadal timescale. PeatClim marks input peatland occurrences as presence or absence, omitting information on fractional cover. A purely random cross-validation strategy may inflate performance in spatially autocorrelated data. Random generation of pseudo-absences may also introduce sampling bias.

Although PeatClim will inevitably overestimate observed peatland extent because it predicts climate-constrained potential rather than actual distribution, it is useful for predicting global potential peatland distributions under present-day conditions and for quantitatively investigating peatland evolution in the geological past using palaeoclimate model outputs. The workflow also has broader applications, in that climate-driven distribution models can be extended to other lithologies (e.g., evaporites and bauxites), enabling the construction of comprehensive lithological distribution maps for evaluating palaeoclimate model performance.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

Code and data availability.

All code and data used in this study is archived at PeatClim (v1.0.1; https://doi.org/10.5281/zenodo.20040294; Chen et al., 2026). External input datasets were obtained from their original repositories, are not redistributed here, and are cited in the reference list. The repository also includes documentation describing the required external inputs, workflow, configuration settings, software environment, and package versions.

Author contributions.

Conceptualization: P.J.V., A.F., and L.C. Data curation: L.C. Formal analysis: L.C. Funding acquisition: L.C., P.J.V., and A.F. Investigation: L.C. Methodology: L.C., P.J.V., and A.F. Software: L.C. Supervision: P.J.V. and A.F. Visualization: L.C. Writing (original draft preparation): L.C. Writing (review and editing): L.C., A.F., and P.J.V.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations**.**

Acknowledgements. We thank Hélène Blancheteau (France) for helpful advice on the use of the biomod2 R package. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol (http://www.bristol.ac.uk/acrc/).

Funding. This work was supported by the China Scholarship Council grant 202204910010 to L.C. and the NSFGEO-NERC grant NE/X015505/1 to P.J.V. and A.F.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

References

Allouche, O., Tsoar, A., and Kadmon, R.: Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS), Journal of Applied Ecology, 43, 1223–1232, https://doi.org/10.1111/j.1365-2664.2006.01214.x, 2006.

Bannatyne, B. B.: Preliminary survey of bogs for peat moss in southeastern Manitoba, Manitoba Department of Mines and Natural Resources, Winnipeg, 1964.

Bannatyne, B. B.: Sphagnum Bogs in Southern Manitoba and their Identification by Remote Sensing, Manitoba Department of Energy and Mines, 1980.

Bao, X., Hu, Y., Scotese, C. R., Li, X., Guo, J., Lan, J., Lin, Q., Yuan, S., Wei, M., Li, Z., Man, K., Yin, Z., Han, J., Zhang, J., Wei, Q., Liu, Y., Yang, J., and Nie, J.: Quantifying climate conditions for the formation of coals and evaporites, National Science Review, 10, nwad051, https://doi.org/10.1093/nsr/nwad051, 2023.

Barbet-Massin, M., Jiguet, F., Albert, C. H., and Thuiller, W.: Selecting pseudo-absences for species distribution models: how, where and how many?, Methods in Ecology and Evolution, 3, 327–338, https://doi.org/10.1111/j.2041-210X.2011.00172.x, 2012.

Batjes, N. H., Ribeiro, E., and van Oostrum, A.: Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019), https://doi.org/10.17027/ISRIC-WDCSOILS.20190901, 2019.

Batjes, N. H., Ribeiro, E., and Van Oostrum, A.: Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019), Earth Syst. Sci. Data, 12, 299–320, https://doi.org/10.5194/essd-12-299-2020, 2020.

Bauer, I. E., Davies, M. A., Bona, K. A., Hararuk, O., Shaw, C. H., Thompson, D. K., Kurz, W. A., Webster, K. L., Garneau, M., McLaughlin, J. W., Packalen, M. S., Prystupa, E., Sanderson, N. K., and Tarnocai, C.: Peat profile database from peatlands in Canada, Ecology, 105, e4398, https://doi.org/10.1002/ecy.4398, 2024.

Beilman, D. W., MacDonald, G. M., Smith, L. C., and Reimer, P. J.: Carbon accumulation in peatlands of West Siberia over the last 2000 years, Global Biogeochemical Cycles, 23, https://doi.org/10.1029/2007GB003112, 2009.

Benfield, A. J., Yu, Z., and Benavides, J. C.: Environmental controls over Holocene carbon accumulation in *Distichia muscoides*-dominated peatlands in the eastern Andes of Colombia, Quaternary Science Reviews, 251, 106687, https://doi.org/10.1016/j.quascirev.2020.106687, 2021.

Boucot, A. J., Xu, C., Scotese, C. R., and Morley, R. J.: Phanerozoic Paleoclimate: An Atlas of Lithologic Indicators of Climate, SEPM (Society for Sedimentary Geology), Tulsa, Oklahoma, U.S.A., https://doi.org/10.2110/sepmcsp.11, 2013.

Cao, W., Williams, S., Flament, N., Zahirovic, S., Scotese, C., and Müller, R. D.: Palaeolatitudinal distribution of lithologic indicators of climate in a palaeogeographic framework, Geological Magazine, 156, 331–354, https://doi.org/10.1017/S0016756818000110, 2019.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

Chen, L., Valdes, P., and Farnsworth, A.: PeatClim: A climate-driven machine-learning model for predicting potential paleo-peatland distribution and its key climate controls, , https://doi.org/10.5281/zenodo.20040294, 2026.

Cole, L. E. S., Bhagwat, S. A., and Willis, K. J.: Long-term disturbance dynamics and resilience of tropical peat swamp forests, Journal of Ecology, 103, 16–30, https://doi.org/10.1111/1365-2745.12329, 2015.

Comas, X., Terry, N., Slater, L., Warren, M., Kolka, R., Kristiyono, A., Sudiana, N., Nurjaman, D., and Darusman, T.: Imaging tropical peatlands in Indonesia using ground-penetrating radar (GPR) and electrical resistivity imaging (ERI): implications for carbon stock estimates and peat soil characterization, Biogeosciences, 12, 2995–3007, https://doi.org/10.5194/bg-12-2995-2015, 2015.

Crezee, B., Dargie, G. C., Ewango, C. E. N., Mitchard, E. T. A., Emba B., O., Kanyama T., J., Bola, P., Ndjango, J.-B. N., Girkin, N. T., Bocko, Y. E., Ifo, S. A., Hubau, W., Seidensticker, D., Batumike, R., Imani, G., Cuní-Sanchez, A., Kiahtipes, C. A., Lebamba, J., Wotzka, H.-P., Bean, H., Baker, T. R., Baird, A. J., Boom, A., Morris, P. J., Page, S. E., Lawson, I. T., and Lewis, S. L.: Mapping peat thickness and carbon stocks of the central Congo Basin using field data, Nat. Geosci., 15, 639–644, https://doi.org/10.1038/s41561-022-00966-7, 2022.

Davies, M. A., Blewett, J., Naafs, B. D. A., and Finkelstein, S. A.: Ecohydrological controls on apparent rates of peat carbon accumulation in a boreal bog record from the Hudson Bay Lowlands, northern Ontario, Canada, Quaternary Research, 104, 14–27, https://doi.org/10.1017/qua.2021.22, 2021.

Davies, M. A., Mclaughlin, J. W., Packalen, M. S., and Finkelstein, S. A.: Holocene carbon storage and testate amoeba community structure in treed peatlands of the western Hudson Bay Lowlands margin, Canada, Journal of Quaternary Science, 38, 92–106, https://doi.org/10.1002/jqs.3465, 2023a.

Davies, M. A., McLaughlin, J. W., Packalen, M. S., and Finkelstein, S. A.: Using Holocene paleo-fire records to estimate carbon stock vulnerabilities in Hudson Bay Lowlands peatlands, FACETS, 8, 1–26, https://doi.org/10.1139/facets-2022-0162, 2023b.

Elith, J., Ferrier, S., Huettmann, F., and Leathwick, J.: The evaluation strip: A new and robust method for plotting predicted responses from species distribution models, Ecological Modelling, 186, 280–289, https://doi.org/10.1016/j.ecolmodel.2004.12.007, 2005.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International Journal of Climatology, 37, 4302–4315, https://doi.org/10.1002/joc.5086, 2017.

Gorham, E., Lehman, C., Dyke, A., Clymo, D., and Janssens, J.: Long-term carbon sequestration in North American peatlands, Quaternary Science Reviews, 58, 77–82, https://doi.org/10.1016/j.quascirev.2012.09.018, 2012.

Gumbricht, T., Roman-Cuesta, R. M., Verchot, L., Herold, M., Wittmann, F., Householder, E., Herold, N., and Murdiyarso, D.: An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor, Global Change Biology, 23, 3581–3599, https://doi.org/10.1111/gcb.13689, 2017.

Hribljan, J. A., Hough, M., Lilleskov, E. A., Suarez, E., Heckman, K., Planas-Clarke, A. M., and Chimner, R. A.: Elevation and temperature are strong predictors of long-term carbon accumulation across tropical Andean mountain peatlands, Mitig Adapt Strateg Glob Change, 29, 1, https://doi.org/10.1007/s11027-023-10089-y, 2023.

Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., Marushchak, M., Olefeldt, D., Packalen, M., Siewert, M. B., Treat, C., Turetsky, M., Voigt, C., and Yu, Z.: Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw, Proc. Natl. Acad. Sci. U.S.A., 117, 20438–20446, https://doi.org/10.1073/pnas.1916387117, 2020a.

Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., Marushchak, M., Olefeldt, D., Packalen, M., Siewert, M. B., Treat, C., Turetsky, M., Voigt, C., and Yu, Z.: Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw, Proceedings of the National Academy of Sciences, 117, 20438–20446, https://doi.org/10.1073/pnas.1916387117, 2020b.

Joosten, H. and Clarke, D.: Wise use of Wise use of peatlands peatlands, International Mire Conservation Group and International Peat Society, 2002.

Judd, E. J., Tierney, J. E., Lunt, D. J., Montañez, I. P., Huber, B. T., Wing, S. L., and Valdes, P. J.: A 485-million-year history of Earth's surface temperature, Science, 385, eadk3705, https://doi.org/10.1126/science.adk3705, 2024.

Kelly, T. J., Lawson, I. T., Roucoux, K. H., Baker, T. R., and Honorio Coronado, E. N.: Patterns and drivers of development in a west Amazonian peatland during the late Holocene, Quaternary Science Reviews, 230, 106168, https://doi.org/10.1016/j.quascirev.2020.106168, 2020.

Keys, D. and Henderson, R. E.: An investigation of the peat resources of New Brunswick, New Brunswick Department of Natural Resources and Energy, 1987.

Kleinen, T., Brovkin, V., and Schuldt, R. J.: A dynamic model of wetland extent and peat accumulation: results for the Holocene, Biogeosciences, 9, 235–248, https://doi.org/10.5194/bg-9-235-2012, 2012.

Lamentowicz, M.: Geneza torfowisk naturalnych i seminaturalnych w Nadleśnictwie Tuchola ((The origin and development of natural and seminatural peatlands in Tuchola Forest Inspectorate), Bogucki Wydawnictwo Naukowe, Poznań, 2005.

Lawson, I. T., Åkesson, C., Baker, T. R., Cordova Oroche, C. J., Dargie, G. C., del Aguila-Pasquel, J., Grandez Ríos, J., Hastie, A., Honorio Coronado, E. N., Mitchard, E. T., Roucoux, K. H., Reyna Huaymacari, J., and Williams, M.: Peat depths from the pastaza-marañón basin, amazonian peru, 2019-2020, https://doi.org/10.5285/ab13a06f-392f-4bc6-b1bf-06dd8b020307, 2023.

Loisel, J., van Bellen, S., Pelletier, L., Talbot, J., Hugelius, G., Karran, D., Yu, Z., Nichols, J., and Holmquist, J.: Insights and issues with estimating northern peatland carbon stocks and fluxes since the Last Glacial Maximum, Earth-Science Reviews, 165, 59–80, https://doi.org/10.1016/j.earscirev.2016.12.001, 2017.

Lottes, A. L. and Ziegler, A. M.: World peat occurrence and the seasonality of climate and vegetation, Palaeogeography, Palaeoclimatology, Palaeoecology, 106, 23–37, https://doi.org/10.1016/0031-0182(94)90003-5, 1994.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

Melton, J. R., Chan, E., Millard, K., Fortier, M., Winton, R. S., Martín-López, J. M., Cadillo-Quiroz, H., Kidd, D., and Verchot, L. V.: A map of global peatland extent created using machine learning (Peat-ML), Geosci. Model Dev., 15, 4709–4738, https://doi.org/10.5194/gmd-15-4709-2022, 2022.

Mozafari, B., Bruen, M., Donohue, S., Renou-Wilson, F., and O'Loughlin, F.: Peatland dynamics: A review of process-based models and approaches, Science of The Total Environment, 877, 162890, https://doi.org/10.1016/j.scitotenv.2023.162890, 2023.

Müller, J. and Joos, F.: Global peatland area and carbon dynamics from the Last Glacial Maximum to the present – a process-based model investigation, Biogeosciences, 17, 5285–5308, https://doi.org/10.5194/bg-17-5285-2020, 2020.

Payne, R. J., Ring-Hrubesh, F., Rush, G., Sloan, T. J., Evans, C. D., and Mauquoy, D.: Peatland initiation and carbon accumulation in the Falkland Islands, Quaternary Science Reviews, 212, 213–218, https://doi.org/10.1016/j.quascirev.2019.03.022, 2019.

Pontone, N., Millard, K., Thompson, D. K., Guindon, L., and Beaudoin, A.: A hierarchical, multi-sensor framework for peatland sub-class and vegetation mapping throughout the Canadian boreal forest, Remote Sensing in Ecology and Conservation, 10, 500–516, https://doi.org/10.1002/rse2.384, 2024.

Qiu, C., Zhu, D., Ciais, P., Guenet, B., Peng, S., Krinner, G., Tootchi, A., Ducharne, A., and Hastie, A.: Modelling northern peatland area and carbon dynamics since the Holocene with the ORCHIDEE-PEAT land surface model (SVN r5488), Geoscientific Model Development, 12, 2961–2982, https://doi.org/10.5194/gmd-12-2961-2019, 2019.

R Core Team: R: A language and environment for statistical computing, 2023.

Ribeiro, M. M., Roque, N., Ribeiro, S., Gavinhos, C., Castanheira, I., Quinta-Nova, L., Albuquerque, T., and Gerassis, S.: Bioclimatic modeling in the Last Glacial Maximum, Mid-Holocene and facing future climatic changes in the strawberry tree (Arbutus unedo L.), PLoS ONE, 14, e0210062, https://doi.org/10.1371/journal.pone.0210062, 2019.

Roscher, R., Russwurm, M., Gevaert, C., Kampffmeyer, M., Dos Santos, J. A., Vakalopoulou, M., Hänsch, R., Hansen, S., Nogueira, K., Prexl, J., and Tuia, D.: Better, Not Just More: Data-centric machine learning for Earth observation, IEEE Geoscience and Remote Sensing Magazine, 12, 335–355, https://doi.org/10.1109/MGRS.2024.3470986, 2024.

Scotese, C. R.: An Atlas of Phanerozoic Paleogeographic Maps: The Seas Come In and the Seas Go Out, Annu. Rev. Earth Planet. Sci., 49, 679–728, https://doi.org/10.1146/annurev-earth-081320-064052, 2021.

Scotese, C. R., Song, H., Mills, B. J. W., and Van Der Meer, D. G.: Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years, Earth-Sci. Rev., 215, 103503, https://doi.org/10.1016/j.earscirev.2021.103503, 2021.

Scottish Government: Peatland Action – peat depth point data (Dataset updated 26 May 2025), 2025.

Silva, M. P., Healy, M. G., and Gill, L.: Reviews and syntheses: A scoping review evaluating the potential application of ecohydrological models for northern peatland restoration, Biogeosciences, 21, 3143–3163, https://doi.org/10.5194/bg-21-3143-2024, 2024.

Silvestri, S., Rosemary, K., Andrea, V., Curtis, R., Gusti, A., Noah, D., Neal, F., and Xavier, C.: Field-lab data and analyses results West Kalimantan, Indonesia, https://doi.org/10.5281/zenodo.3572061, 2019a.

Silvestri, S., Knight, R., Viezzoli, A., Richardson, C. J., Anshari, G. Z., Dewar, N., Flanagan, N., and Comas, X.: Quantification of Peat Thickness and Stored Carbon at the Landscape Scale in Tropical Peatlands: A Comparison of Airborne Geophysics and an Empirical Topographic Method, Journal of Geophysical Research: Earth Surface, 124, 3107–3123, https://doi.org/10.1029/2019JF005273, 2019b.

Skye, J., Melton, J. R., Goldblatt, C., Saumier, L., Gallego-Sala, A., Garneau, M., Winton, R. S., Bahati, E. B., Benavides, J. C., Fedorchuk, L., Imani, G., Kagaba Kairumba, C., Kansiime, F., Lamentowicz, M., Mbasi, M., Wochal, D., Czerwiński, S., Landowski, J., Landowska, J., Maire, V., Väliranta, M. M., Warren, M., Cole, L. E. S., Davies, M. A., Lilleskov, E. A., Sun, J., and Wang, Y.: Peat-DBase v.1: a compiled database of global peat depth measurements, Earth System Science Data, 17, 7313–7330, https://doi.org/10.5194/essd-17-7313-2025, 2025a.

Skye, J., Melton, J., Goldblatt, C., Saumier, L., Gallego Sala, A., Garneau, M., Winton, S. R., Bahati, E. B., Benavides, J. C., Fedorchuk, L., Imani, G., Kagaba Kairumba, C., Kansiime, F., Lamentowicz, M., Mbasi, M., Wochal, D., Czerwiński, S., Landowski, J., Landowska, J., Maire, V., Väliranta, M. M., Warren, M., Cole, L. E. S., Davies, M. A., Lilleskov, E. A., Sun, J., and Wang, Y.: Peat-DBase v.1: A Compiled Database of Global Peat Depth Measurements, https://doi.org/10.5281/zenodo.17991992, 2025b.

Spahni, R., Joos, F., Stocker, B. D., Steinacher, M., and Yu, Z. C.: Transient simulations of the carbon and nitrogen dynamics in northern peatlands: from the Last Glacial Maximum to the 21st century, Climate of the Past, 9, 1287–1308, https://doi.org/10.5194/cp-9-1287-2013, 2013.

Stocker, B. D., Spahni, R., and Joos, F.: DYPTOP: a cost-efficient TOPMODEL implementation to simulate sub-grid spatio-temporal dynamics of global wetlands and peatlands, Geoscientific Model Development, 7, 3089–3110, https://doi.org/10.5194/gmd-7-3089-2014, 2014.

Sun, J., Gallego-Sala, A., and Yu, Z.: Topographic and climatic controls of peatland distribution on the Tibetan Plateau, Sci Rep, 13, 14811, https://doi.org/10.1038/s41598-023-39699-x, 2023.

Thibault, J.: The New Brunswick Peatland Database, New Brunswick Department of Natural Resources and Energy Development, 1992.

Thuiller, W.: BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change, Global Change Biology, 9, 1353–1362, https://doi.org/10.1046/j.1365-2486.2003.00666.x, 2003.

Thuiller, W., Lafourcade, B., Engler, R., and Araújo, M. B.: BIOMOD – a platform for ensemble forecasting of species distributions, Ecography, 32, 369–373, https://doi.org/10.1111/j.1600-0587.2008.05742.x, 2009.

https://doi.org/10.5194/egusphere-2026-2755
Preprint. Discussion started: 21 July 2026

Thuiller, W., Georges, D., Gueguen, M., Engler, R., Breiner, F., Lafourcade, B., Patin, R., and Blancheteau, H.: biomod2: Ensemble Platform for Species Distribution Modeling, , R package version 4.2-6-2, https://doi.org/10.32614/CRAN.package.biomod2, 2025.

Tran, D. V., Tominaga, A., Pham, L. T., and Nishikawa, K.: Ecological niche modeling shed light on new insights of the speciation processes and historical distribution of Japanese fire-bellied newt *Cynops pyrrhogaster* (Amphibia: Urodela), Ecological Informatics, 79, 102443, https://doi.org/10.1016/j.ecoinf.2023.102443, 2024.

Treat, C. C., Jones, M. C., Camill, P., Gallego-Sala, A. V., Garneau, M., Harden, J. W., Hugelius, G., Klein, E. S., Kokfelt, U., Kuhry, P., Loisel, J., Mathijssen, P. J. H., O'Donnell, J. A., Oksanen, P. O., Ronkainen, T. M., Sannel, A. B. K., Talbot, J., Tarnocai, C., and Väliranta, M.: (Table S2) Peat properties for cores including sampling depths, bulk density, carbon, nitrogen, LOI, dominant vegetation type, plant macrofossils, peat environmental class, https://doi.org/10.1594/PANGAEA.863695, 2016.

Treat, C. C., Broothaerts, N., Dalton, A. S., Dommain, R., Douglas, T., Drexler, J., Finkelstein, S. A., Grosse, G., Hope, G., Hutchings, J. A., Jones, M. C., Kleinen, T., Kuhry, P., Lacourse, T., Lähteenoja, O., Loisel, J., Notebaert, B., Payne, R. J., Peteet, D. M., Sannel, A. B. K., Stelling, J., Strauss, J., Swindles, G. T., Talbot, J., Tarnocai, C., Verstraeten, G., Williams, C. J., Xia, Z., Yu, Z., and Brovkin, V.: (Table S2) Global dataset of peatland basal ages, https://doi.org/10.1594/PANGAEA.873065, 2017.

Treat, C. C., Kleinen, T., Broothaerts, N., Dalton, A. S., Dommain, R., Douglas, T. A., Drexler, J. Z., Finkelstein, S. A., Grosse, G., Hope, G., Hutchings, J., Jones, M. C., Kuhry, P., Lacourse, T., Lähteenoja, O., Loisel, J., Notebaert, B., Payne, R. J., Peteet, D. M., Sannel, A. B. K., Stelling, J. M., Strauss, J., Swindles, G. T., Talbot, J., Tarnocai, C., Verstraeten, G., Williams, C. J., Xia, Z., Yu, Z., Väliranta, M., Hättestrand, M., Alexanderson, H., and Brovkin, V.: Widespread global peatland establishment and persistence over the last 130,000 y, Proceedings of the National Academy of Sciences, 116, 4822–4827, https://doi.org/10.1073/pnas.1813305116, 2019.

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, Geosci. Model Dev., 10, 3715–3743, https://doi.org/10.5194/gmd-10-3715-2017, 2017.

Valdes, P. J., Scotese, C. R., and Lunt, D. J.: Deep ocean temperatures through time, Clim. Past, 17, 1483–1506, https://doi.org/10.5194/cp-17-1483-2021, 2021.

Waltari, E. and Guralnick, R.: Ecological niche modelling of montane mammals in the Great Basin, North America: Examining past and present connectivity of species across basins and ranges, Journal of Biogeography, 36, 148–161, https://doi.org/10.1111/j.1365-2699.2008.01959.x, 2008.

Warren, M. W., Kauffman, J. B., Murdiyarso, D., Anshari, G., Hergoualc'h, K., Kurnianto, S., Purbopuspito, J., Gusmayanti, E., Afifudin, M., Rahajoe, J., Alhamd, L., Limin, S., and Iswandi, A.: A cost-efficient method to assess carbon stocks in tropical peat soil, Biogeosciences, 9, 4477–4485, https://doi.org/10.5194/bg-9-4477-2012, 2012.

Winton, R. S., Benavides, J. C., Mendoza, E., Uhde, A., Hastie, A., Honorio Coronado, E. N., Hernandez Ortega, A. G., Paukku, S., Mullins, B., del Aguila Pasquel, J., Aymard-Corredor, G. A., Baker, T. R., Draper, F. C., Flores Llampazo, G., Herrera, R., Phillips, O. L., Reyna Huaymacari, J. M., ter Steege, H., Stropp, J., Lawson, I. T., Gallego-Sala, A. V., Boom, A., Wehrli, B., and Hoyt, A. M.: Widespread carbon-dense peatlands in the Colombian lowlands, Environ. Res. Lett., 20, 054025, https://doi.org/10.1088/1748-9326/adbc03, 2025.

Xiao, S., Li, S., Wang, X., Chen, L., and Su, T.: *Cedrus* distribution change: past, present, and future, Ecological Indicators, 142, 109159, https://doi.org/10.1016/j.ecolind.2022.109159, 2022.

Xu, J., Morris, P. J., Liu, J., and Holden, J.: PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis., https://doi.org/10.5518/252, 2017.

Xu, J., Morris, P. J., Liu, J., and Holden, J.: PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis, CATENA, 160, 134–140, https://doi.org/10.1016/j.catena.2017.09.010, 2018.

Yu, Z., Beilman, D., and Jones, M.: Sensitivity of Northern Peatland Carbon Dynamics to Holocene Climate Change, Washington DC American Geophysical Union Geophysical Monograph Series, 184, 55–69, https://doi.org/10.1029/2008GM000822, 2009.

Yu, Z., Loisel, J., Brosseau, D. P., Beilman, D. W., and Hunt, S. J.: Global peatland dynamics since the Last Glacial Maximum, Geophysical Research Letters, 37, https://doi.org/10.1029/2010GL043584, 2010.