scientific data

OPEN

DATA DESCRIPTOR

Dataset about Warming Effects on Carbon Cycling and Greenhouse Gas Fluxes in Permafrost Ecosystems

Tao Bao[1,2], Xiyan Xu[1]✉, Gensuo Jia[1,3], Xingru Zhu[1], William J. Riley[4] & Yuanhe Yang[3,5,6]

Field observations provide direct evidence of how does carbon cycling in permafrost ecosystems respond to climate change. This study provides a comprehensive dataset on the impact of warming on carbon cycling and greenhouse gas (GHG) fluxes in permafrost ecosystems. The dataset is extracted and integrated from 132 peer-reviewed studies with 1430 paired observations across eight major permafrost ecosystems, including Arctic and subarctic tundra and wetland, and alpine meadow, steppe, tundra and wetland. This dataset includes 17 variables from experiments conducted during the growing season, covering the plant and soil carbon pools, soil nitrogen pool, and GHG (i.e., CO_2, CH_4, and N_2O) fluxes, among others. Background information on site climate conditions, vegetation and soil characteristics, and details of the warming experiments, including timing, methods, and warming magnitude, are also contained in the dataset. This dataset facilitates a comprehensive understanding of the impact of warming on carbon cycling and GHG fluxes in permafrost ecosystems, and provides supports for meta-analyses and literature reviews, remote sensing data validation, and land model development and parameterization.

Background & Summary

Permafrost underlies approximately one-fifth of the exposed land surface area in the Northern Hemisphere[1] and stores 1014–1672 Pg of organic carbon, which is roughly one-third of global soil carbon[2]. The thermal state of permafrost is particularly sensitive to changes in air temperature[3,4]. In recent decades, the Arctic and high-altitude regions have been warmed at rates two to four times faster than the global average[5]. Even if global air temperatures rise by no more than 2 °C by 2100, permafrost degradation may still occur across extensive areas[6]. Such a change would have serious consequences for local ecosystems, infrastructure, and landform stability[3,7]. Permafrost degradation leads to structural sediment collapse, exposing ancient organic matter to microbial decomposition[8]. Significant amounts of soil organic carbon previously locked in frozen organic matter would be released into the atmosphere, increasing the atmospheric concentrations of carbon dioxide (CO_2) and methane (CH_4), and activating the positive permafrost carbon-climate feedback, thereby further accelerating climate warming[9–11]. Because of this potential, some studies introduced permafrost as an important 'tipping element' of the climate system[12]. Therefore, it is urgent to quantify the potential additional carbon emissions from permafrost thaw under warming, along with the associated carbon processes[9,13]. This will facilitate the development of effective strategies to mitigate the adverse impacts of climate change.

The response of carbon processes to climate warming in permafrost regions is a complex and critical issue. Field-based warming experiments, including open-top chambers (OTCs), infrared radiators (IRs), greenhouse chambers (GCs), and soil heating cables (HCs), have provided valuable insights into these processes[11]. The difference among these methods lies in their warming targets. IRs primarily warm the soil surface, while OTCs mainly affect above-ground conditions by altering air temperature through passive convection. OTCs have become the most widely used warming method in permafrost regions across the Northern Hemisphere

[1]State Key Laboratory of Earth System Numerical Modeling and Application, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China. [2]Key Laboratory of Ecosystem Carbon Source and Sink, China Meteorological Administration (ECSS-CMA), Wuxi University, Wuxi, Jiangsu, 214063, China. [3]University of Chinese Academy of Sciences, Beijing, 100049, China. [4]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [5]State Key Laboratory of Forage Breeding-by-Design and Utilization; Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, 100093, China. [6]China National Botanical Garden, Beijing, 100093, China. ✉e-mail: xiyan.xu@tea.ac.cn

Indexes	Unit	N_Sites	N_Samples	Mean_Control	Std_Control	Mean_Warming	Std_Warming
ANPP	$g\ m^{-2}$	220	1228	211.80	225.40	228.40	262.52
BNPP	$g\ m^{-2}$	62	332	1245.08	1457.19	1425.10	2112.88
Abundance	%	132	745	50.89	35.73	51.39	37.04
Height	cm	57	297	3.80	2.28	6.20	4.07
Soil moisture	%	117	1098	34.11	24.12	31.91	25.23
Soil temperature	°C	77	685	9.94	4.11	10.76	4.41
Water Table	cm	23	84	17.67	14.67	19.67	14.32
GPP	$g\ m^{-2}\ d^{-1}$	136	776	14.55	17.44	15.35	16.97
Re	$g\ m^{-2}\ d^{-1}$	174	1055	11.45	13.28	12.95	14.36
CO_2 flux	$g\ m^{-2}d^{-1}$	129	717	3.04	10.71	2.18	11.97
CH_4 flux	$g\ m^{-2}\ yr^{-1}$	64	456	12.14	20.14	12.55	20.21
N_2O flux	$g\ m^{-2}\ yr^{-1}$	40	297	1.12	3.65	0.97	4.29
SOC	$g\ kg^{-1}$	41	167	85.97	129.59	85.29	129.04
MBC	$mg\ kg^{-1}$	47	211	761.54	1008.82	786.12	1046.22
NH_4^+-N	$mg\ kg^{-1}$	35	163	21.01	22.48	20.86	21.78
NO_3^-N	$mg\ kg^{-1}$	30	148	29.52	78.45	26.12	61.18
TN	$g\ kg^{-1}$	46	192	5.85	5.55	5.76	5.15

Table 1. Characteristics of 17 variables in the dataset before and after warming. Note: N_Sites: number of sites, N_Samples: number of observations, Mean_Control: mean value before warming, Mean_Warming: mean value after warming, Std_Control: standard deviation before warming, Std_Warming: standard deviation after warming.

due to their simplicity, low cost, and suitability for replication. GCs warm both air and soil but often limit ecological realism[14]. By applying these warming experiments, the immediate effects of increased temperatures on carbon dynamics in permafrost regions can be quantified, offering empirical data that elucidate response mechanisms[9,15].

Across permafrost regions, warming effects on ecosystem carbon balance vary with physiography (e.g., alpine vs. lowland), climate, and dominant plant functional types[16–21]. For example, some high-latitude tundra sites exhibit strengthened decomposition and net carbon release[16,19], whereas forests or high-productivity alpine systems often show partial compensation via enhanced assimilation[17,18]. Vegetation composition further modulates responses, with shrub expansion reducing net soil carbon loss but moss cover enhancing decomposition[15,20,21]. The complexity and variability in carbon responses to warming led to challenges in current permafrost-related land model development[22]. Despite continuous improvement of land model process representation, most global-scale Earth system models still poorly capture essential carbon cycling processes in permafrost regions, such as soil organic matter decomposition and methane dynamics, and many critical permafrost processes, including freeze–thaw cycles and active layer deepening, remain absent[23]. Given the exceptional warming in the Arctic and the threat to global climate mitigation goals, there is a unique urgency to synthesize diverse experimental data to understand permafrost carbon-climate feedback and improve land models.

Here, we collected data from peer-reviewed studies that compared vegetation characteristics, soil properties, and greenhouse gas (GHG) emissions under control and warming conditions in the Northern Hemisphere permafrost regions. Only peer-reviewed studies meeting specific criteria were digitized and integrated into the dataset. This dataset represents the most comprehensive compilation of field-based warming experiments conducted in permafrost regions over three decades (1990–2024). It integrates 1430 paired observations (control versus warming treatment) from nearly all published studies meeting strict criteria, covering 17 key variables related to carbon cycling and GHG, i.e., CO_2, CH_4, and N_2O fluxes (Table 1). The dataset spans major Arctic, subarctic, and alpine permafrost ecosystems and covers a wide geographic range from approximately 30° to 80°N. A subset of this dataset (1,090 pairs) was used to demonstrate that warming weakens greenhouse gas sinks in alpine permafrost but enhances them in the Arctic[24]. We will continue to update and improve the dataset, and encourage others to contribute new data to enhance its coverage. Newly published data will be added to this open-source dataset. This dataset can be used for meta-analyses and literature reviews to explore research frontiers in permafrost carbon-climate feedbacks. This dataset can also serve as observational data to assist in validating research findings based on remote sensing observations. Moreover, it could improve permafrost-related parameterization schemes in land models by providing empirical evidences for model calibration and validation.

Methods

Data collection. To establish a comprehensive dataset on the responses of carbon cycles and GHG fluxes to warming in permafrost regions, we implemented strict criteria to obtain relevant data from Google Scholar (https://scholar.google.com), Web of Science (https://www.webofscience.com/), Scopus (https://www.scopus.com/) and the China National Knowledge Infrastructure (http://www.cnki.net) from January 1990 through June 2024. We also searched publications on the websites of global change laboratories and experiment networks, e.g., the Alfred-Wegener Institute (https://epic.awi.de/), the SPRUCE experiment of the Oak Ridge National Laboratory (https://mnspruce.ornl.gov), the Institute of Arctic and Alpine Research of University of Colorado Boulder

(https://www.colorado.edu/instaar/), the BioCon experiment project in University of Minnesota (https://biocon.umn.edu), and the International Arctic Research Center of University of Alaska Fairbanks (https://uaf-iarc.org) etc. These studies were performed within the permafrost extent based on information from original publications provided by the authors and Northern Hemisphere Permafrost Map (NHPM)[25] (https://doi.pangaea.de/10.1594/PANGAEA.888600).

We performed literature web scraping using the Python libraries Requests[26] with three keyword settings simultaneously, **Keyword setting 1**: "net ecosystem exchange (NEE)" or "carbon flux (C emission)" or "carbon dioxide flux (CO_2 emission)" or "methane flux (CH_4 emission)" or "nitrous oxide flux (N_2O emission)" or "greenhouse gas flux (GHG emission)" or "gross primary productivity (GPP)" or "ecosystem respiration (Re)" or "respiration" or "biomass production" or "biogeochemical processes" or "vegetation height (growth)" or "plant abundance (cover)" or "microbial biomass (MBC)" or "soil organic carbon (SOC)" or "total soil nitrogen (TN)" or "soil total nitrogen" or "aboveground (ANPP), belowground (BNPP) and net primary productivity (NPP)" or "NH_4^+-N" or "NO_3^-N"; **Keyword setting 2**: "climate change" or "warming" or "experimental warming" or "increased temperature" or "elevated temperature" or "rising temperature"; and **Keyword setting 3**: "permafrost" or "Arctic" or "High latitude" or "Alpine". The search resulted in 7713 published articles.

Then, a screening by abstract was performed. Articles that were reviews or meta-analyses were removed first. And then studies conducted in non-permafrost ecosystems or disturbed permafrost ecosystems (e.g., croplands and pasturelands), or conducted with non-manipulated warming experiments were excluded, resulting in a remaining 2315 articles.

Finally, a full-text manual screening was performed to select studies that met the following criteria: (i) warming experiments were conducted in outdoor field environments with both control and warming treatments implemented at the same temporal and spatial scales; (ii) measurements were collected during the growing season, as non–growing-season warming experiments are extremely scarce and the few available sites lack sufficient representativeness, and were therefore excluded; (iii) in studies involving multiple manipulation factors (e.g., warming combined with elevated CO_2, altered precipitation, or nutrient addition), only the warming versus control contrast was retained and all observations influenced by other treatments were discarded; (iv) when warming treatments were conducted in multiple periods in a growing season or in multiple years, each control–warming comparison corresponding to a distinct sampling period was considered as an independent paired observation; (v) multiple observations from the same study or site were also retained when they represented different response variables (e.g., CO_2, CH_4, or N_2O flux) or distinct warming–control comparisons (e.g., different warming magnitudes or warming methods) reported in the original publications, each of which was treated as an independent paired observation; and (vi) means, standard errors or standard deviations, and sample sizes for both control and warming treatments were reported. This last screening resulted in 132 selected articles (Fig. 1). If the data were only presented in figures in the articles, Engauge Digitizer software version 4.1 (https://engauge-digitizer.software.informer.com/4.1/) was used to digitalize the figures and extract data.

A total of 1430 pairs of observations from 132 published studies were collected after searching and screening. The spatial distribution of the sites where these experiments were conducted is shown in Fig. 2. The site locations and altitudes, ecosystem types, plant types, climate zones and other relevant background information were collected from relevant studies. These studies cover most of the permafrost regions in the Northern Hemisphere. Among them, 220 pairs of observations provided data with ANPP, 62 with BNPP, and 132 with plant abundance, 57 with plant height, 117 with soil moisture, 77 with soil temperature, 23 with water table, 136 with GPP, 174 with Re, 129 with NEE, 64 with CH_4 flux, 40 with N_2O flux, 41 with SOC, 47 with MBC, 35 with NH_4^+-N, 30 with NO_3^-N, 46 with TN (Table 1). In addition to these variables, the experimental design including sampling years, warming duration, magnitude of air and soil warming, and warming methods, and environmental information during the sampling period including climate variables (mean temperature (MT) and mean precipitation (MP) during the experimental period) were also extracted from the selected studies.

Data processing. The dataset covered a wide geographic range, with latitudes ranging from 30.5° to 78.9° and longitudes ranging from −157.4° to 158.5°. In cases where the original studies did not explicitly report the climate zone of the experimental sites, we derived this information from their latitude and longitude. For consistency, we adopted a simplified latitudinal classification following the conventional geographic latitude zones, defining temperate zone in 23.5–66.5 °N and cold zone in > 66.5 °N[27]. Users can reclassify climate zones according to geographic locations to meet their specific research needs, as the dataset provides detailed latitude and longitude for each site. The dataset primarily consists of growing season observation, involving eight major ecosystem types in the permafrost regions of the Northern Hemisphere: alpine meadow, alpine steppe, alpine tundra, alpine wetland, Arctic tundra, Arctic wetland, subarctic tundra, and subarctic wetland (Fig. 2a, c), and four dominant plant functional types, including herb, shrub, moss, and lichen (Fig. 2b). Ecosystem types were defined according to the classifications used in the original studies. We note that the category "alpine tundra" is rarely used in standard ecological classifications of the Qinghai–Tibet Plateau, where ecosystems are typically described as alpine meadow, steppe, or swamp meadow. However, a few studies adopted the term "alpine tundra" to emphasize the functional similarity of these ecosystems to Arctic tundra. We retained this category to remain consistent with the original studies, although it represents only a small fraction of the records and can be merged into other alpine types if need.

Among the 1430 paired observations, 876 included site-specific MT and MP values reported for the actual sampling periods of the warming experiments. For the remaining sites, we obtained experiment-period MT and MP by interpolating ERA5 climate data[28] according to each site's sampling time and geographic coordinates. We additionally provided growing-season (July–September) mean temperature (MT_GS) and precipitation (MP_GS) derived from WorldClim v2.1 (1970–2000)[29] for each site as long-term background climate information. Reported MT and MP reflect growing-season sampling periods that varied from several days to multiple

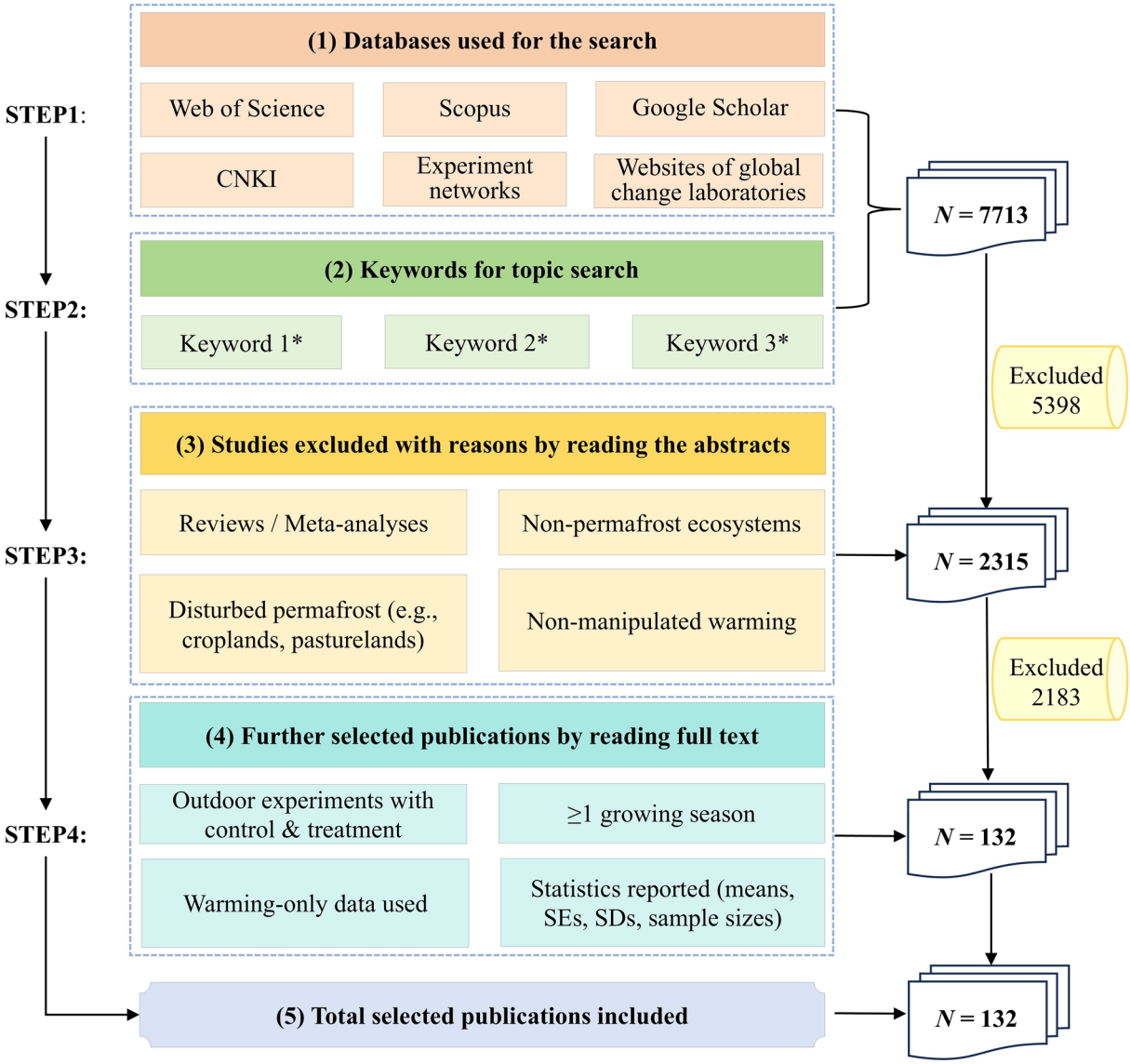

Fig. 1 Data processing approach for data collection, identification, screening, eligibility, and inclusion of the dataset.

months across studies. Across all observations, MT ranged from −15 to 17 °C, and MP from 30 to 1196 mm per sampling period (Fig. 3a,b). Alpine ecosystems generally experienced warmer conditions and higher precipitation, whereas Arctic and subarctic ecosystems predominantly occupied colder climates with lower precipitation (Fig. 3c,d). Across all observations, warming durations varied among studies, with 80% experiments lasting between 1 and 5 years but some extending over a decade (Fig. 2d). Warming magnitude (ΔT) varied among ecosystem types. Alpine wetland and alpine tundra sites generally applied stronger warming (up to 5–6 °C), whereas Arctic and subarctic wetland typically experienced lower-intensity warming (often ≤ 2 °C). These patterns are consistent with the dominance of OTC-based warming in Arctic ecosystems and the more use of IR heaters or greenhouse chambers in alpine regions. A detailed summary of ΔT distributions across ecosystem types is provided in Table 2, and the corresponding breakdown by both ecosystem type and warming method is shown in Table 3. There were also wide variations in soil temperature, soil moisture, plant characteristics (height and abundance), plant C pool (ANPP and BNPP), soil C pool (SOC and MBC), GPP, Re, GHG (CO_2, CH_4 and N_2O) fluxes, soil N pool (NH_4^+-N, NO_3^-N and TN) among different sites (Table 1).

The natural log-transformed response ratio (lnRR), referred to as the "effect size", is utilized as a robust and flexible measure to quantify responses[30]. It assesses the impact of warming treatments on variables such as GHG fluxes (CO_2, CH_4, N_2O), GPP, Re, and others. The lnRR is calculated using the formula:

$$\mathrm{lnRR} = \ln\left(\overline{X_{Warming}}/\overline{X_{Control}}\right) = \ln\overline{X_{warming}} - \ln\overline{X_{Control}} \quad (1)$$

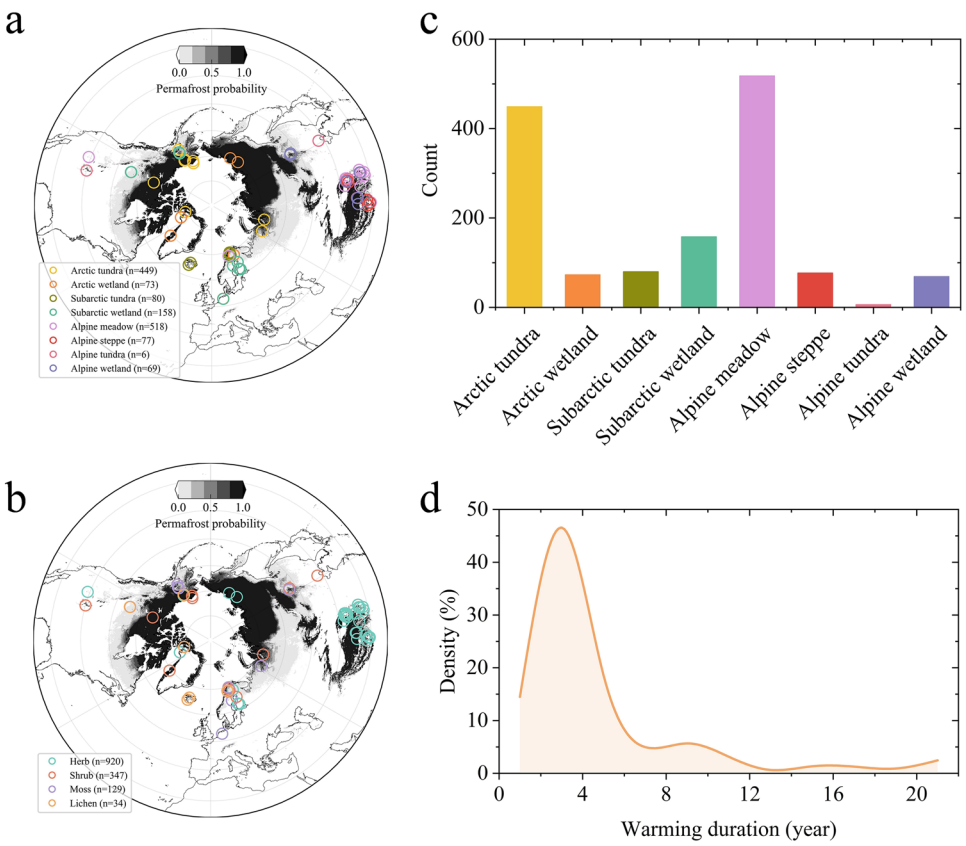

Fig. 2 Spatial distribution of 1,430 paired observations across eight major permafrost ecosystem types (**a**). Distribution of sites by dominant plant functional types (**b**). Total number of observational samples within each ecosystem type (**c**). Kernel density distribution of warming experiment duration (years) across all sites (**d**).

where $\overline{X_{Warming}}$ and $\overline{X_{Control}}$ represent the mean values of the selected variable for the warming treatment and control groups at each site, respectively. Positive and negative lnRR values indicate stimulating or inhibiting effects of warming on a measured variable X, respectively, expressed as percentage change.

The variance (Var) of lnRR is computed as:

$$\mathrm{Var} = \frac{SD_{Warming}^2}{N_{Warming} \times \overline{X_{Warming}}^2} + \frac{SD_{Control}^2}{N_{Control} \times \overline{X_{Control}}^2} \quad (2)$$

where $SD_{Warming}$ and $SD_{Control}$ are the standard deviations, and $N_{Warming}$ and $N_{Control}$ are the sample sizes of the warming treatment and control groups, respectively.

The weighting factor W for each site is derived as:

$$W = 1/Var \quad (3)$$

The "inverse variance weighting" method (Eq. 5) assigns greater weights to studies with smaller variances, enhancing the influence of more precise and stable data in study results.

Mean values, number of replications (N), and standard deviations (SD) were compiled from the corresponding publications wherever possible. In cases where standard error (SE) or coefficient of variation (CV) were reported instead of SD, SD was calculated using the following formulas:

$$SD = SE \times \sqrt{N} \quad (4)$$

$$SD = CV \times Mean \quad (5)$$

For the data without SD, SE, or CV, SD was calculated from the mean by the variance coefficient of all the datasets[31].

Data Records

The data are publicly accessible on figshare https://doi.org/10.6084/m9.figshare.29312087[32]. The following four files are included for the data:

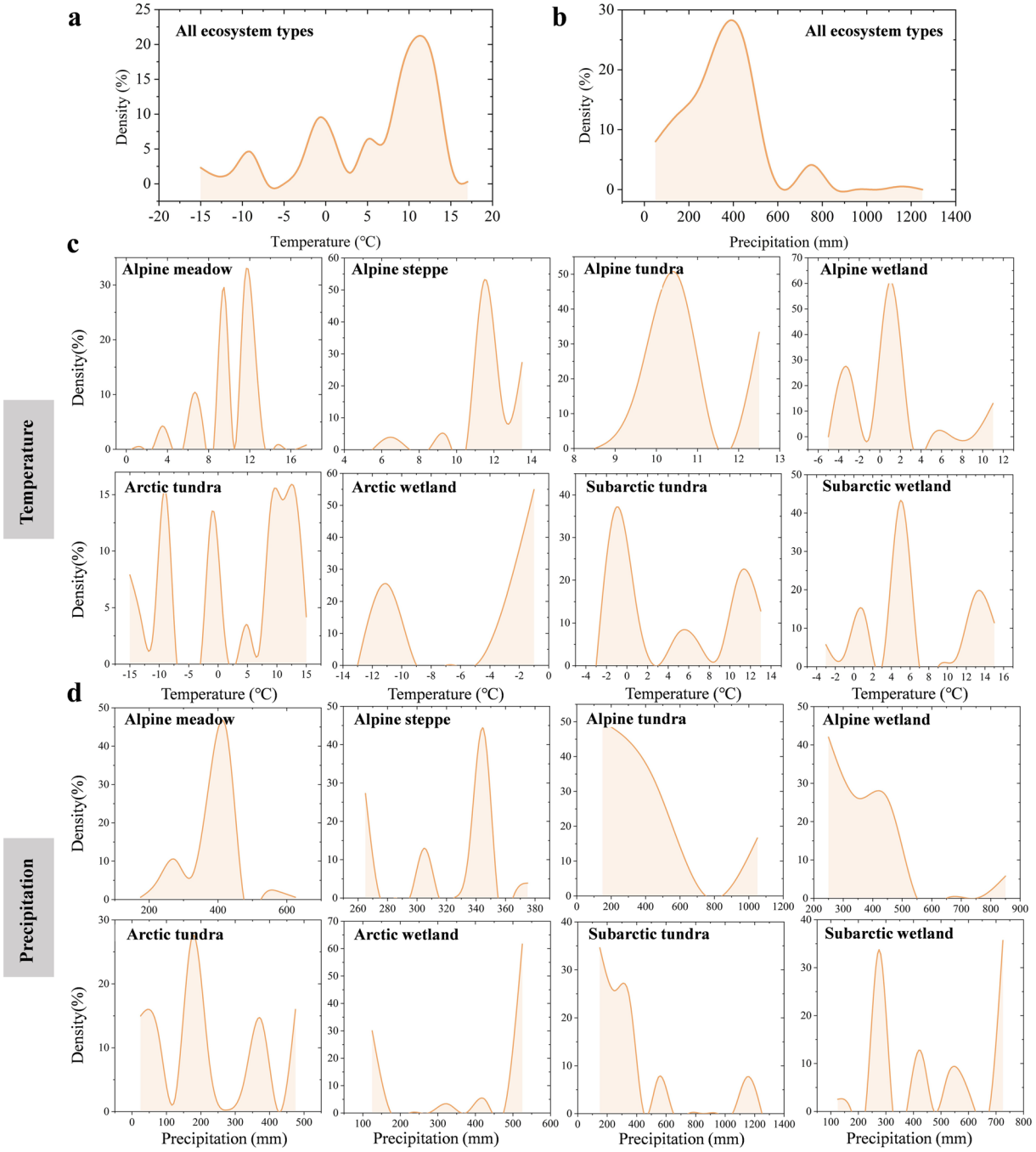

Fig. 3 Kernel density distribution of MT across all ecosystem types **(a)**. Kernel density distribution of MP across all ecosystem types **(b)**. Ecosystem-specific MT distributions for the eight permafrost ecosystem types: alpine meadow, alpine steppe, alpine tundra, alpine wetland, Arctic tundra, Arctic wetland, subarctic tundra, and subarctic wetland **(c)**. Ecosystem-specific MP distributions for the same eight ecosystem types **(d)**.

1. "Data_Permafrost_Carbon.xlsx" includes the data. The "Dataset" sheet included data collected from 1430 pairs of observations from 132 published studies, while the "Readme" sheet contains descriptions of the abbreviations and variables within the dataset.
2. "Data_Meta.docx" includes a summary defining each column, trait collected in the data and the unit for each variable.
3. "Data_publications.docx" contains all the 132 references of the articles for the dataset.
4. "Code.rar" includes all the Python code for data scraping, site map plotting, data validation (funnel plots), and table data generation for this study.

The file "Data_Permafrost_Carbon.xlsx" comprises systematically collected information from selected studies (refer to the Methods section). The data in the "Dataset" sheet is organized into three sections:

Ecosystem types	N_Sites	N_Samples	ΔT_mean	ΔT_median	ΔT_min	ΔT_max
Alpine meadow	376	516	1.9	2.0	0.0	5.0
Alpine steppe	62	77	2.1	2.0	1.0	3.0
Alpine tundra	5	6	3.3	3.0	1.0	5.0
Alpine wetland	46	65	3.9	3.0	1.0	6.0
Arctic tundra	301	444	2.1	2.0	0.0	5.0
Arctic wetland	15	46	1.5	2.0	1.0	2.0
Subarctic tundra	67	80	2.1	2.0	0.5	4.0
Subarctic wetland	35	136	1.4	1.0	0.0	4.0

Table 2. Warming magnitude across permafrost ecosystem types. Note: N_Sites denotes number of sites, N_Samples denotes number of observations, ΔT_mean, ΔT_median, ΔT_min, and ΔT_max represent the mean, median, minimum, and maximum warming magnitude (ΔT), respectively.

Ecosystem types	Methods	N_Samples	ΔT_mean	ΔT_median	ΔT_min	ΔT_max
Alpine meadow	GC	8	2.5	2.5	2.0	3.0
Alpine meadow	IR	81	2.0	2.0	1.0	4.0
Alpine meadow	OTC	427	1.9	2.0	0.0	5.0
Alpine steppe	IR	9	2.0	2.0	2.0	2.0
Alpine steppe	OTC	68	2.1	2.0	1.0	3.0
Alpine tundra	GC	2	5.0	5.0	5.0	5.0
Alpine tundra	OTC	4	2.5	3.0	1.0	3.0
Alpine wetland	GC	12	3.0	3.0	3.0	3.0
Alpine wetland	OTC	53	4.2	6.0	1.0	6.0
Arctic tundra	GC	125	3.5	4.0	1.0	5.0
Arctic tundra	OTC	319	1.6	2.0	0.0	4.0
Arctic wetland	OTC	46	1.5	2.0	1.0	2.0
Subarctic tundra	GC	17	3.5	3.9	3.0	4.0
Subarctic tundra	OTC	63	1.7	1.5	0.5	4.0
Subarctic wetland	GC	12	4.0	4.0	4.0	4.0
Subarctic wetland	OTC	124	1.2	1.0	0.0	2.0

Table 3. Warming magnitude by ecosystem type and experimental method. **Note:** N_Samples denotes number of observations, ΔT_mean, ΔT_median, ΔT_min, and ΔT_max represent the mean, median, minimum, and maximum warming magnitude (ΔT), respectively, within each ecosystem–method combination when the measurements were available. GCs, IRs, and OTCs refer to greenhouse chambers, infrared radiators, and open-top chambers, respectively.

The first section presents information inherent to the published article (site ID, publication year of the article, source reference of the data, and name of the first author).

The second section refers to information inherent to the warming experiment data (geographical position, ecosystem types, mean temperature, mean precipitation, dominant plant types, years of data collection, air temperature increase, duration, and warming methods).

The third section includes information of carbon cycles and greenhouse gas fluxes along with environmental variables collected during the experiments (CO_2, CH_4, and N_2O fluxes, GPP, Re, plant height and abundance, ANPP, BNPP, SOC, MBC, NH_4^+-N, NO_3^--N, TN, soil temperature, and soil moisture). For users' convenience, the dataset also includes pre-calculated lnRR and their weighing factors for each observation. These effect-size metrics can be directly used in meta-analyses or recomputed from the underlying means, standard deviations, and sample sizes.

Additionally, some variables were further classified, such as:

Climate zone was divided into temperate (23.5 to 66.5 °N), and cold (greater than 66.5 °N). Ecosystem type was divided into alpine meadow, alpine steppe, alpine tundra, alpine wetland, Arctic tundra, Arctic wetland, subarctic tundra, and subarctic wetland. The classification of ecosystem types was based on original publications provided by the authors. Dominant plant type was divided into herb, shrub, moss, and lichen. The classification of plant type was also based on original publications provided by the authors. These vegetation types are common functional groups in the permafrost regions of the Northern Hemisphere.

Methods of warming are open-top chambers (OTCs), infrared radiators (IRs) and greenhouse chambers (GCs). OTCs are open structures that harness sunlight to naturally warm the air. IRs use infrared radiation to directly heat plant surfaces, soil, or other experimental targets. GCs are enclosed structures that employ heating systems to regulate temperature. While GCs and IRs offer more controlled heating, enabling precise environmental manipulation, OTCs offer a less controlled but more ecologically relevant approach. OTCs have proven to be highly effective for field experiments in permafrost ecosystems where replicated plots are needed[14].

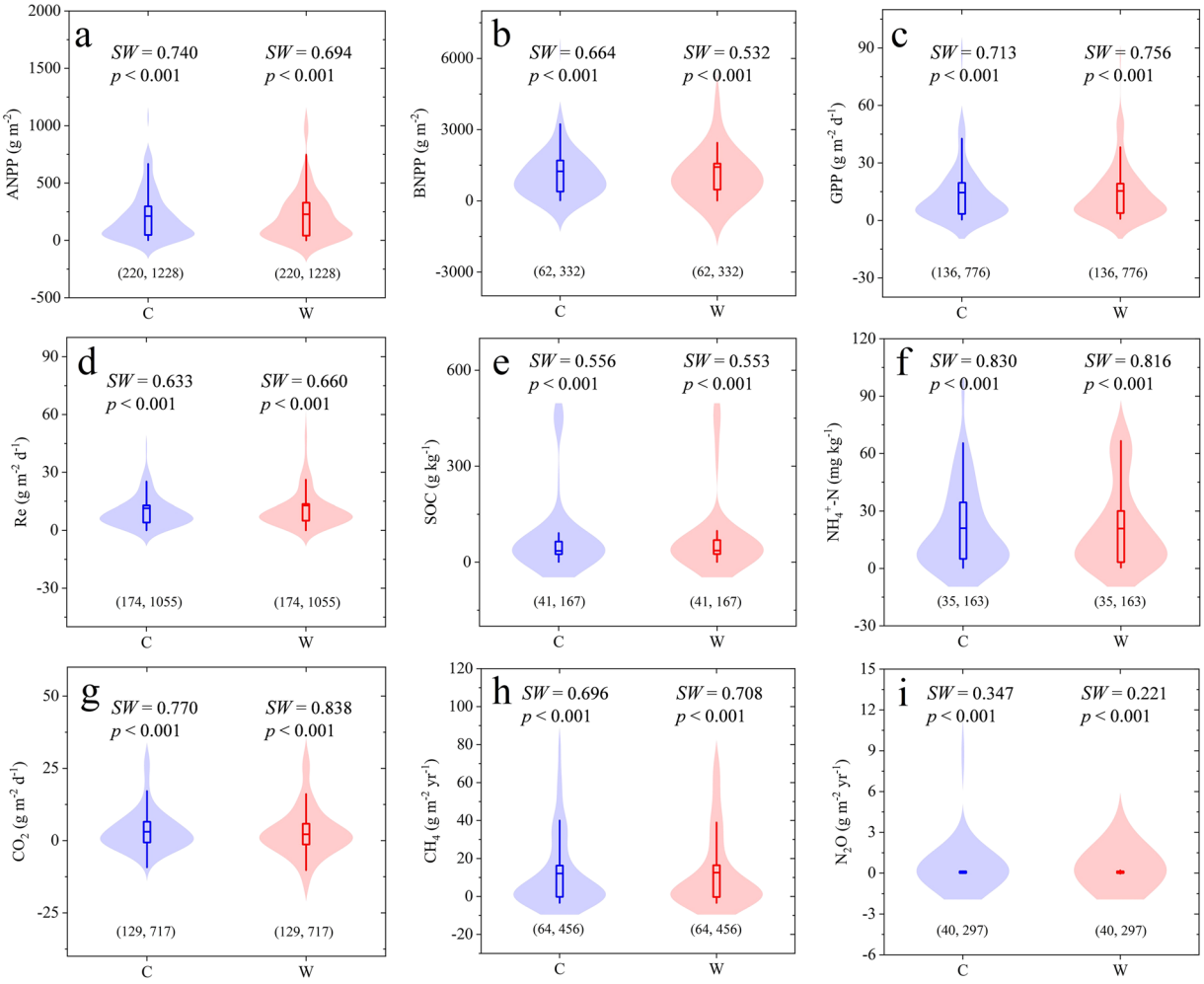

Fig. 4 Violin plots of aboveground net primary productivity (ANPP) (**a**), belowground net primary productivity (BNPP) (**b**), gross primary productivity (GPP) (**c**), ecosystem respiration (Re) (**d**), soil organic carbon (SOC) (**e**), NH_4^+-N (**f**), CO_2 (**g**), CH_4 (**h**), and N_2O (**i**) in paired control (blue) and warming (red) experiments. The width of each violin represents the kernel density estimate of the data at different levels, providing a visual summary of the distribution's shape. The box plot within each violin represents 25% and 75% quantile and the middle line represents the mean. Numbers within parentheses represent the number of sites (left) and the number of observational samples (right). Shapiro–Wilk *SW* and *p* values for each variable are shown to document the extent of departure from normality.

Technical Validation

All records in the dataset were undergone multiple stages of quality control to ensure consistency, accuracy, and reproducibility[33,34]. Each variable was checked for unit consistency and agreement with the descriptions provided in the original studies. For data extracted from figures, digitized values were cross-validated against reported summary statistics whenever available. We also performed internal consistency checks across variables (e.g., relationships among biomass, GHG fluxes, and soil properties) to identify potential outliers or transcription errors. Records failing these checks were re-examined to avoid any potential errors in the data extraction.

Data extracted from tables or figures were cross verified against original documents to mitigate transcription errors and ensure fidelity to the source. Following extraction, stringent quality control measures were applied, including assessments of data availability, logical consistency, and numerical accuracy. Special emphasis was placed on verifying the format of each data column to rectify potential mistyping or discrepancies, with standardized formats and units enforced to facilitate accurate analysis and interpretation. A total of 187 entries (13.1% of all records) required unit conversions (e.g., g/kg vs. mg/kg for soil moisture), or numeric formatting. For 63 entries where SD was not directly reported, SD values were calculated using standard statistical conversions following Eq. (4) from available SE and (5) from available CV. These conversions are algebraically equivalent to the reported statistics. All lnRR values were recalculated and validated, and the internal consistency was achieved for 100% of the effect size records.

Following data extraction, we evaluated the distributions of all variables using both visual inspection and Shapiro–Wilk tests[35]. As shown in Fig. 4, nine key variables (ANPP, BNPP, GPP, Re, SOC, NH_4^+-N, NO_3^-N,

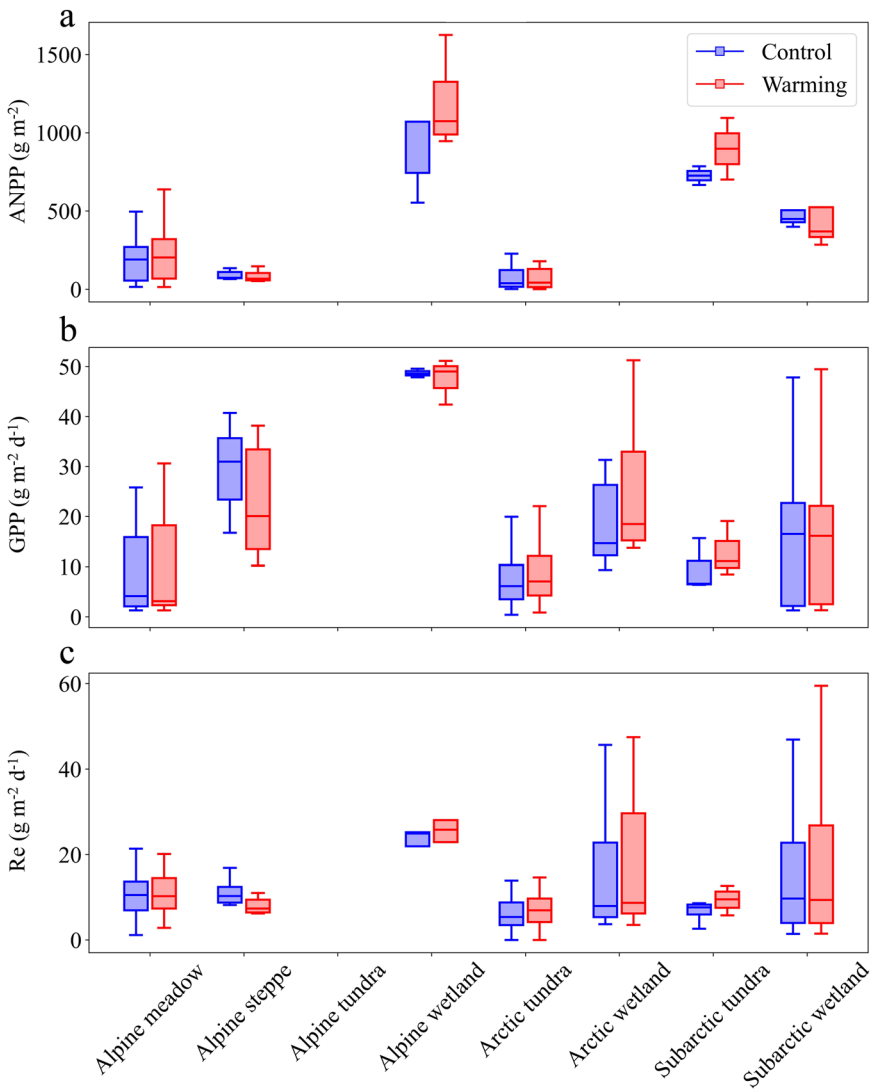

Fig. 5 Boxplots of aboveground net primary productivity (ANPP) (**a**), gross primary productivity (GPP) (**b**), and ecosystem respiration (Re) (**c**) under control (blue) and warming (red) conditions across eight permafrost ecosystem types: alpine meadow, alpine steppe, alpine tundra, alpine wetland, Arctic tundra, Arctic wetland, subarctic tundra, and subarctic wetland. Boxes represent the interquartile range with the median shown as the central line; whiskers indicate the 5th and 95th percentiles.

CO_2, CH_4, and N_2O) exhibited skewed and long-tailed distributions under both control and warming conditions, indicating significant variation among experimental sites. The Shapiro–Wilk statistics ($SW = 0.22–0.84$; $p < 0.001$ in all cases) indicated significant departures from normality. To ensure data accuracy, we placed particular attention on values falling within the upper and lower ~3% of each distribution, where digitization errors, unit inconsistencies, or transcription mistakes are most likely to occur. These values were re-checked against the original tables, figures, or supplementary datasets to ensure their consistency with the published measurements. Ecosystem-level comparisons further revealed differences among ecosystem types, with alpine wetland showing the highest ANPP and GPP and alpine steppe and Arctic tundra exhibiting much lower values under both control and warming conditions (Fig. 5).

Finally, we conducted publication bias tests on the dataset. Conducting publication bias tests is crucial to ensure the robustness and validity of the dataset. Publication bias, characterized by the tendency of studies with statistically significant or positive results to be more readily published, can distort the overall effect size estimation and compromise the reliability of study conclusions[36]. We evaluated publication bias using funnel plots (that is, the scatter plots of the effect size (lnRR) against their sample size and Egger's test[37]. The funnel plots exhibited symmetric distribution of all experimental studies around a mean effect size, forming a "funnel" shape. Additionally, Egger's regression test yielded a p-value greater than 0.05, indicating the absence of statistically significant publication bias (Fig. 6). These findings indicate that the lnRR-based effect sizes derived from the dataset are robust and not significantly affected by publication bias.

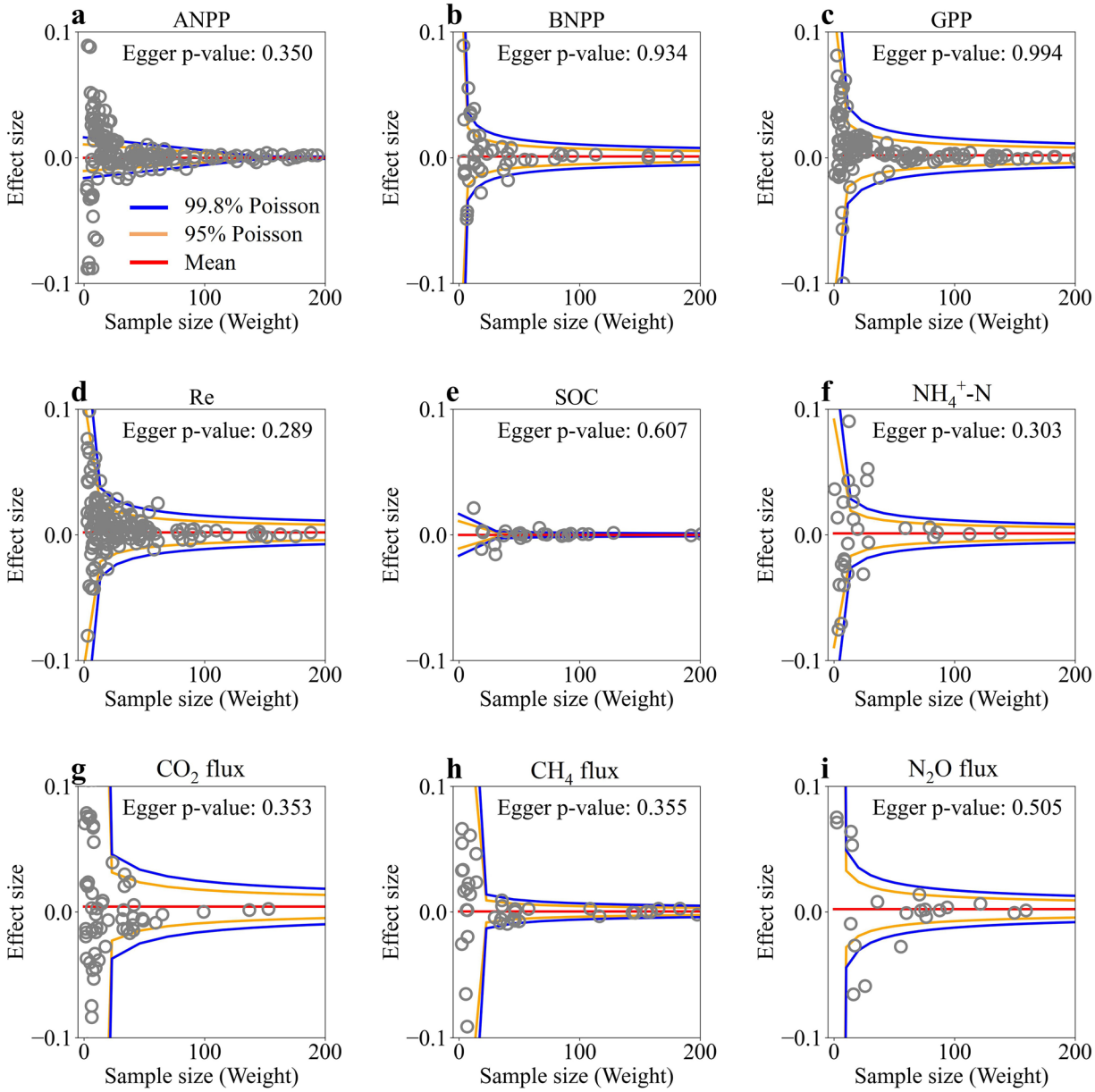

Fig. 6 Funnel plots of aboveground net primary productivity (ANPP) (**a**), belowground net primary productivity (BNPP) (**b**), gross primary productivity (GPP) (**c**), ecosystem respiration (Re) (**d**), soil organic carbon (SOC) (**e**), NH_4^+-N (**f**), CO_2 (**g**), CH_4 (**h**), and N_2O (**i**) under warming. Results of publication bias tests using Egger's test and all the above results were displayed with $p > 0.05$, indicating the absence of publication bias.

Usage Notes

This current dataset is valuable for conducting studies of mechanisms at large spatial scales, bibliometric research, literature reviews, and meta-analyses to assess the impact of warming on carbon cycles and greenhouse gas fluxes in permafrost ecosystems. It provides harmonized *in situ* measurements derived from 1430 paired observations collected across Arctic and alpine permafrost regions, broadly spanning approximately 30°to 80°N, and capturing warming experiments over three decades during the growing season. By converting all warming experiment observations into a unified effect-size metric, the dataset ensures cross-study comparability and provides a consistent basis for reliable synthesis across sites. While the dataset offers extensive coverage of warming effects on permafrost carbon cycling, the measurements predominantly represent growing-season conditions.

In addition, its standardized structure and clearly defined variables support reproducible analyses and enable the dataset to be integrated with other environmental datasets. This dataset can also provide ground validation data for studies based on remote sensing observations, support the development, improvement, and parameterization of land process models for permafrost ecosystems, and serve as a reference for decisions related to sustainable development and mitigating climate change.

The dataset presented in this study will be dynamic and subject to updates. We hope that our study will also raise awareness of the urgent need for open and shared datasets, especially in permafrost regions, which are extremely vulnerable to climate change and the data are particularly scarce.

Data availability

All data supporting this Data Descriptor are openly available on figshare https://doi.org/10.6084/m9.figshare.29312087[32].

Code availability

The dataset and related code are available on the figshare repository[32]. For any inquiries regarding code understanding or data usage, users can contact the corresponding author.

Received: 30 September 2025; Accepted: 8 January 2026;
Published online: 14 January 2026

References

1. Obu, J. How much of the earth's surface is underlain by permafrost? *J. Geophys. Res. Earth Surf.* **126**, e2021JF006123 (2021).
2. Mishra, U. *et al*. Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks. *Sci. Adv.* **7**, eaaz5236 (2021).
3. Biskaborn, B. K. *et al*. Permafrost is warming at a global scale. *Nat. Commun.* **10**, 264 (2019).
4. Zhu, X., Jia, G. & Xu, X. Accelerated rise in wildfire carbon emissions from Arctic continuous permafrost. *Sci. Bull.* (2024).
5. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.* **5**, 424–430 (2015).
6. Chadburn, S. *et al*. An observation-based constraint on permafrost loss as a function of global warming. *Nat. Clim. Chang.* **7**, 340–344 (2017).
7. Schuur, E. A. *et al*. Permafrost and climate change: carbon cycle feedbacks from the warming Arctic. *Annu. Rev. Environ. Resour.* **47**, 343–371 (2022).
8. Turetsky, M. R. *et al*. Carbon release through abrupt permafrost thaw. *Nat. Geosci.* **13**, 138–143 (2020).
9. Schuur, E. A. *et al*. Climate change and the permafrost carbon feedback. *Nature* **520**, 171–179 (2015).
10. Schaefer, K., Lantuit, H., Romanovsky, V. E., Schuur, E. A. & Witt, R. The impact of the permafrost carbon feedback on global climate. *Environ. Res. Lett.* **9**, 085003 (2014).
11. Bao, T., Jia, G. & Xu, X. Weakening greenhouse gas sink of pristine wetlands under warming. *Nat. Clim. Chang.* **13**, 462–469 (2023).
12. Lenton, T. M. *et al*. Climate tipping points—too risky to bet against. *Nature* **575**, 592–595 (2019).
13. Jones, M. C. *et al*. Past permafrost dynamics can inform future permafrost carbon–climate feedbacks. *Commun. Earth Environ.* **4**, 272 (2023).
14. Welshofer, K. B., Zarnetske, P. L., Lany, N. K. & Thompson, L. A. Open-top chambers for temperature manipulation in taller-stature plant communities. *Methods Ecol. Evol.* **9**, 254–259 (2018).
15. Bao, T., Jia, G. & Xu, X. Warming enhances dominance of vascular plants over cryptogams across northern wetlands. *Glob. Change Biol.* **28**, 4097–4109 (2022).
16. Oberbauer, S. F. *et al*. Tundra CO_2 fluxes in response to experimental warming across latitudinal and moisture gradients. *Ecol. Monogr.* **77**, 221–238 (2007).
17. Natali, S. M., Schuur, E. A., Webb, E. E., Pries, C. E. H. & Crummer, K. G. Permafrost degradation stimulates carbon loss from experimentally warmed tundra. *Ecology* **95**, 602–608 (2014).
18. Na, L., Genxu, W., Yan, Y., Yongheng, G. & Guangsheng, L. Plant production, and carbon and nitrogen source pools, are strongly intensified by experimental warming in alpine ecosystems in the Qinghai–Tibet Plateau. *Soil Biol. Biochem.* **43**, 942–953 (2011).
19. Biasi, C. *et al*. Initial effects of experimental warming on carbon exchange rates, plant growth and microbial dynamics of a lichen-rich dwarf shrub tundra in Siberia. *Plant Soil* **307**, 191–205 (2008).
20. Wilcox, E. J. *et al*. Tundra shrub expansion may amplify permafrost thaw by advancing snowmelt timing. *Arct. Sci.* **5**, 202–217 (2019).
21. Ladrón de Guevara, M. *et al*. Warming reduces the cover, richness and evenness of lichen-dominated biocrusts but promotes moss growth: insights from an 8 yr experiment. *New Phytol.* **220**, 811–823 (2018).
22. Grosse, G., Goetz, S., McGuire, A. D., Romanovsky, V. E. & Schuur, E. A. Changing permafrost in a warming world and feedbacks to the Earth system. *Environ. Res. Lett.* **11**, 040201 (2016).
23. Schädel, C. *et al*. Earth system models must include permafrost carbon processes. *Nat. Clim. Chang.* **14**, 114–116 (2024).
24. Bao, T. *et al*. Climate-carbon feedback tradeoff between Arctic and alpine permafrost under warming. *Sci. Adv.* **11**, eadt8366 (2025).
25. Obu, J. *et al*. Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km^2 scale. *Earth-Sci. Rev.* **193**, 299–316 (2019).
26. Chandra, R. V. & Varanasi, B. S. *Python Requests Essentials*. (Packt Publishing, 2015).
27. Nie, S., Fu, S., Cao, W. & Jia, X. Comparison of monthly air and land surface temperature extremes simulated using CMIP5 and CMIP6 versions of the Beijing Climate Center climate model. *Theor. Appl. Climatol.* **140**, 487–502 (2020).
28. Hersbach, H. *et al*. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
29. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
30. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).
31. Hou, E. *et al*. Global meta-analysis shows pervasive phosphorus limitation of aboveground plant production in natural terrestrial ecosystems. *Nat. Commun.* **11**, 637 (2020).
32. Bao, T. *et al*. Dataset about Warming Effects on Carbon Cycling and Greenhouse Gas Fluxes in Permafrost Ecosystems. Figshare. https://doi.org/10.6084/m9.figshare.29312087 (2025).
33. Gomez, F. *et al*. A dataset for soil organic carbon in agricultural systems for the Southeast Asia region. *Sci. Data* **11**, 374 (2024).
34. Li, X. *et al*. A global dataset of biochar application effects on crop yield, soil properties, and greenhouse gas emissions. *Sci. Data* **11**, 57 (2024).
35. Royston, P. Approximating the Shapiro–Wilk W-test for non-normality. *Stat. Comput.* **2**, 117–119 (1992).
36. Thornton, A. & Lee, P. Publication bias in meta-analysis: its causes and consequences. *J. Clin. Epidemiol.* **53**, 207–216 (2000).
37. Koricheva, J., Gurevitch, J. & Mengersen, K. *Handbook of Meta-analysis in Ecology and Evolution*. (Princeton Univ. Press, 2013).

Acknowledgements
We thank all the authors whose work was included in our dataset. This study is funded by National Key R&D Program of China (2022YFF0801904 to X.X. and 2024YFF1307603 to T.B.), and the Open Research Fund of the Key Laboratory of Ecosystem Carbon Source and Sink, China Meteorological Administration (ECSS-CMA), NUIST (ECSS-CMA202402 to T.B.), and Natural Science Foundation of China (#42206254 to T.B.). W.J.R. was supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area, Office of Biological and Environmental Research of the U.S. Department of Energy Office of Science. Lawrence Berkeley National Laboratory (LBNL) is managed by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

Author contributions
T.B., X.X. and G.J. conceived this paper. T.B. and X.Z. extracted and integrated the data from studies into Data_Permafrost_Carbon.xlsx. T.B. and X.X. summarized the dataset and drafted the manuscript. All authors revised and approved the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to X.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.