Global maps of soil temperature

Jonas J. Lembrechts1 | Johan van den Hoogen2 | Juha Aalto3,4 | Michael B. Ashcroft5,6 | Pieter De Frenne7 | Julia Kemppinen8
Martin Kopecky9,10 | Miska Luoto4 | Ilya M. D. Maclean11 | Thomas W. Crowther2 | Joseph J. Bailey12 | Stef Haesen13 | David H. Klinge14,15 | Pekka Niittynen4 | Brett R. Scheffers16 | Koenraad Van Meersbeek13 | Peter Aartsma17 | Otar Abdalaze18 | Mehdi Abedi19 | Rien Aerts20
Negar Ahmadian19 | Antje Ahrends21 | Juha M. Alatalo22 | Jake M. Alexander23 | Camille Nina Allonsius24 | Jan Altman9,10 | Christof Ammann25 | Christian Andres26 | Christopher Andrews27 | Jonas Ardö28 | Nicola Arriga29 | Alberto Arzac30 | Valeria Aschero31,32 | Rafael L. Assis33 | Jakob Johann Assmann34,35 | Maaike Y. Bader36 | Khadijah Bahalke19 | Peter Barančok37 | Isabel C. Barrio38 | Agustina Barros32 | Matti Barthel26 | Edmund W. Basham14 | Marijn Bauters39 | Manuele Bazzichetto40 | Luca Belelli Marchesini41 | Michael C. Bell42 | Juan C. Benavides43 | José Luis Benito Alonso44 | Bernd J. Berauer45,46 | Jarle W. Bjerke47 | Robert G. Björk48,49 | Mats P. Björkman48,49 | Katrin Björnsdóttir50 | Benjamin Blonder51 | Pascal Boeckx39 | Julia Boike52,53 | Stef Bokhorst20 | Bábara N. S. Brum54 | Josef Brúna9 | Nina Buchmann26 | Pauline Buysse55 | José Luís Camargo56 | Otávio C. Campoe57 | Onur Candan58 | Rafaella Canessa36,59 | Nicoletta Cannone60 | Michele Carbognani61 | Jofre Carnicer62,63 | Angélica Casanova-Katny64 | Simone Cesarz65,66 | Bogdan Chojnicki67 | Philippe Choler68,69 | Steven L. Chown70 | Edgar F. Cifuentes71 | Marek Čiliak72 | Tamara Contador73,74 | Peter Convey75 | Elisabeth J. Cooper76 | Edoardo Cremonese77 | Salvatore R. Curasi78 | Robin Curtis11 | Maurizio Cutini79 | C. Johan Dahlberg80,81 | Gergana N. Daskalova82 | Miguel Angel de Pablo83 | Stefano Della Chiesa84 | Jürgen Dengler65,85,86 | Bart Deronde87 | Patrice Descombes88 | Valter Di Cecco89 | Michele Di Musciamo90 | Jan Dick27 | Romina D. Dimarco91,92 | Jiri Doležal93 | Ellen Dorrepaal94

Jonas J. Lembrechts and Johan van den Hoogen should be considered joint first author.
Ivan Nijs and Jonathan Lenoir should be considered joint senior author.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. Global Change Biology published by John Wiley & Sons Ltd.
Leonardo Montagnani191 | Mikel Moriana-Armendariz76 | Umberto Morra di Cella192 | Martin Mörsdorf193 | Jonathan
R. Mosedale194 | Lena Muffler146 | Miriam Muñoz-Rojas195,196 | Jonathan A. Myers197 | Isla H. Myers-Smith82 | Laszlo Nagy198 | Marianna Nardino199 | Ilona Naujokaitis-Lewis200 | Emily Newling201
Lena Nicklas98 | Georg Niedrist202 | Armin Niessner203 | Mats B. Nilsson152 | Signe Normand34,35 | Marcelo D. Nosetto204,205 | Yann Nouvellon127,128 | Martin A. Nuñez92,206 | Romà Ogaya207,208 | Jérôme Ogée101 | Joseph Okello39,209,210 | Janusz Olejnik211 | Jørgen Eivind Olesen182 | Øystein H. Opedal212 | Simone Orsenigo190 | Andrej Palaj37
Tim Pampuch213 | Alexey V. Panov214 | Meelis Pärtel215 | Ada Pastor178 | Aníbal Pauchard111,113 | Harald Pauli165 | Marian Pavelka95 | William D. Pearse216,217 | Matthias Peichl152 | Loïc Pellissier218,219 | Rachel M. Penczykowski197 | Josep Penuelas207,208 | Matteo Petit Bon9,76,166 | Alessandro Petraglia61 | Shyam S. Phartyal220 | Gareth K. Phoenix221 | Casimiro Pio222 | Andrea Pitacco223 | Camille Pitteloud218,219 | Roman Plichta180 | Francesco Porro190 | Miguel Portillo-Estrada1 | Jérôme Poulenard222 | Rafael Poyatos63,225 | Anatoly S. Prokushkin30,214 | Radoslaw Puchalka222,226,227 | Mihai Pușcaș228,229,230 | Dajana Raduškovic1 | Krystal Randall5,231 | Amanda Ratier Backes65,131 | Sabine Remmele203 | Wolfram Remmers232 | David Renault40,233 | Anita C. Risch234 | Christian Rixen108,109 | Sharon A. Robinson5,231 | Bjorn J. M. Robroek235 | Adrian V. Rocha236 | Christian Rossi237,238 | Graziano Rossi190 | Olivier Roupsard239,240,241 | Alexey V. Rubtsov30 | Patrick Saccone165 | Clotilde Sagot242 | Jhonatan Sallo Bravo243,244 | Cinthya C. Santos245 | Judith M. Sarneel246 | Tobias Scharnweber213 | Jonas Schmeddes164 | Marius Schmidt247 | Thomas Scholten248 | Max Schuchardt46 | Naomi Schwartz249 | Tony Scott121 | Julia Seeber134,202 | Ana Cristina Segalin de Andrade245 | Tim Seipel167 | Philipp Semenchuk250 | Rebecca A. Senior251 | Josep M. Serra-Diaz252 | Piotr Sewerniak253 | Ankit Shekhar26 | Nikita V. Sidenko214 | Lukas Siebicke161 | Laura Siegwart Collier142,254 | Elizabeth Simpson216 | David P. Siqueira255 | Zuzana Sitková256 | Johan Six26 | Marko Smiljanic213 | Stuart W. Smith122,257 | Sarah Smith-Tripp258 | Ben Somers259 | Mia Vedel Sørensen122 | José João L. L. Souza260 | Bartolomeu Israel Souza261 | Arildo Souza Dias245,262 | Marko J. Spasojevic263 | James D. M. Speed264 | Fabien Spicher265 | Angela Stanisci266 | Klaus Steinbauer165 | Rainer Steinbrecher267 | Michael Steinwandter202 | Michael Stemkovski216 | Jörg G. Stephan268 |
Abstract

Research in global change ecology relies heavily on global climatic grids derived from estimates of air temperature in open areas at around 2 m above the ground. These climatic grids do not reflect conditions below vegetation canopies and near the ground surface, where critical ecosystem functions occur and most terrestrial species reside. Here, we provide global maps of soil temperature and bioclimatic variables at a 1-km$^2$ resolution for 0–5 and 5–15 cm soil depth. These maps were created by calculating the difference (i.e. offset) between in situ soil temperature measurements, based on time series from over 1200 1-km$^2$ pixels (summarized from 8519 unique temperature sensors) across all the world’s major terrestrial biomes, and coarse-grained air temperature estimates from ERA5-Land (an atmospheric reanalysis by the European Centre for Medium-Range Weather Forecasts). We show that mean annual soil temperature differs markedly from the corresponding gridded air temperature, by up to 10°C (mean = $3.0 \pm 2.1$°C), with substantial variation across biomes and seasons. Over the year, soils in cold and/or dry biomes are substantially warmer (+3.6 ± 2.3°C) than gridded air temperature, whereas soils in warm and humid environments are on average slightly cooler (−0.7 ± 2.3°C). The observed substantial and biome-specific offsets emphasize that the projected impacts of climate and climate change on near-surface biodiversity and ecosystem functioning are inaccurately assessed when air rather than soil temperature is used, especially in cold environments. The global soil-related
1 | INTRODUCTION

With the rapidly increasing availability of big data on species distributions, functional traits and ecosystem functioning (Bond-Lamberty & Thomson, 2018; Bruehlheide et al., 2018; Kattge et al., 2019; Kissling et al., 2018; Lenoir et al., 2020), we can now study biodiversity and ecosystem responses to global changes in unprecedented detail (Antão et al., 2020; van den Hoogen et al., 2019; Senior et al., 2019; Steidinger et al., 2019). However, despite this increasing availability of ecological data, most spatially explicit studies of ecological, biophysical and biogeochemical processes still have to rely on the same global gridded temperature data (Du et al., 2020; van den Hoogen et al., 2019; Soudzilovskaia et al., 2015). Thus far, these global gridded products are based on measurements from standard meteorological stations that record free-air temperature inside well-ventilated protective shields placed up to 2 m above-ground in open, shade-free habitats, where abiotic conditions may differ substantially from those actually experienced by most organisms (Lembrechts et al., 2020; World Meteorological Organization, 2008).

Ecological patterns and processes often relate more directly to below-canopy soil temperature rather than to well-ventilated air temperature inside a weather station. Near-surface, rather than air, temperature better predicts ecosystem functions like biogeochemical cycling (e.g. organic matter decomposition, soil respiration and other aspects of the global carbon balance) (Davis et al., 2020; Gottschall et al., 2019; Hursh et al., 2017; Jia et al., 2021; Perera-Castro et al., 2020; Pleim & Gilliam, 2009; Portillo-Estrada et al., 2016; Schimel et al., 2004). Similarly, the use of soil temperature in correlative analyses or predictive models may improve predictions of climate impacts on organismal physiology and behaviour, as well as on population and community dynamics and species distributions (Ashcroft et al., 2008; Berner et al., 2020; Kearney et al., 2009; Körner & Paulsen, 2004; Opedal et al., 2015; Scherrer et al., 2011; Schimel et al., 2004; Zellweger et al., 2020). Given the key role of soil-related processes for both above-ground and belowground parts of the ecosystem and their feedbacks to the atmosphere (Crowther et al., 2016), adequate soil temperature data are critical for a broad range of fields of study, such as ecology, biogeography, biogeochemistry, agronomy, soil science and climate system dynamics. Nevertheless, existing soil temperature products such as those from ERA5-Land (Copernicus Climate Change Service (C3S), 2019), with a resolution of 0.08 × 0.08 degrees (≈9 × 9 km at the equator) remain too coarse for most ecological applications.

The direction and magnitude of the difference or offset between in situ soil temperature and coarse-gridded air temperature products result from a combination of two factors: (i) the (vertical) micoclimatic difference between air and soil temperature and (ii) the (horizontal) mesoclimatic difference between air temperature in flat, cleared areas (i.e. where meteorological stations are located) and air temperature within different vegetation types (e.g. below a dense canopy of trees) or topographies (e.g. within a ravine or on a ridge) (De Frene et al., 2021; Lembrechts et al., 2020). In essence, the offset is thus the combination of both the vertical and horizontal differences that result from factors affecting the energy budget at the Earth’s surface, principally radiative energy: the ground absorbs radiative energy, which is transferred to the air by convective heat exchange, evaporation and spatial variation in net radiation, and lower convective conductance near the Earth’s surface results in horizontal and vertical variation in temperature (Geiger, 1950; Richardson, 1922). Both these vertical and horizontal differences in temperature vary significantly across the globe and in time as a result of environmental conditions affecting the radiation budget (e.g. as a result of topographic orientation, canopy cover or surface albedo), convective heat exchange and evaporation (e.g. foliage density, variation in the degree of wind shear caused by surface friction) and the capacity for the soil to store and conduct heat (e.g. water content and soil structure and texture) (De Frene et al., 2019; Geiger, 1950; Way & Lewkowicz, 2018; Zhang et al., 2008).

Although the physics of soil temperatures have long been well understood (Geiger, 1950; Richardson, 1922), the creation of high-resolution global gridded soil temperature products has not been feasible before, partially due to the absence of detailed global in situ soil temperature measurements (Lembrechts et al., 2020; Lembrechts & Lenoir, 2019). Recently, however, the call for micoclimatic temperature data representative of in situ conditions (i.e. microhabitat) as experienced by organisms living close to the ground surface or in the soil has become more urgent (Bramer et al., 2018). In this paper, we address this issue by generating...
global gridded maps of below-canopy and near-surface soil temperature at 1-km² resolution (in line with most existing global air temperature products). These maps are more representative of the habitat conditions as experienced by organisms living under vegetation canopies, in the topsoil or near the soil surface. They were created using the abovementioned offset between gridded air temperature data and in situ soil temperature measurements. We expect these soil temperature maps to be substantially more representative of actual microclimatic conditions than existing products as they capture relevant near-surface and belowground abiotic conditions where ecosystem functions and processes operate (Bramer et al., 2018; Daly, 2006; Körner & Hiltbrunner, 2018). Indeed, the offset between free-air (macroclimate) and soil (microclimate) temperature, and between cleared areas and other habitats, can easily reach up to ±10°C annually, even at the 1-km² spatial resolution used here (Lembrechts et al., 2019; Wild et al., 2019; Zhang et al., 2018).

To create the global gridded soil temperature maps introduced above, we used over 8500 time series of soil temperature measured in situ across the world's major terrestrial biomes, which are compiled and stored in the SoilTemp database (Lembrechts et al., 2020) (Figure 1a, Figure S1) and averaged into 1200 (or 1000 for the second soil layer) unique 1-km² pixels. First, to illustrate the magnitude of the studied effect, we visualized the global and biome-specific patterns in the mean annual offset between in situ soil temperature (0–5 cm and 5–15 cm depth) and coarse-scale interpolated air temperature from ERA5-Land using the average within 1 × 1 km grid cells. Hereafter, we refer to this difference between soil temperature and air temperature as the temperature offset (or offset), sensu (De Frenne et al., 2021); elsewhere called the surface offset (Smith & Riseborough, 1996, 2002). Secondly, we used a machine learning approach with 31 environmental predictor variables (including macroclimate, soil, topography, reflectance, vegetation and anthropogenic variables) to model the spatial variation in monthly temperature offsets at a 1 × 1 km resolution for all continents except Antarctica (as not covered by many of the used predictor variable layers). Using these offsets, we then calculated relevant soil-related bioclimatic variables (SBIO), mirroring the existing global bioclimatic variables for air temperature. Finally, we compared the modelled mean annual temperature (SBIO1, top-soil layer) with a similar product based on monthly ERA5L topsoil (0–7 cm) temperature with a spatial resolution of 0.08 × 0.08 degrees (∼9 × 9 km at the equator).

2 | METHODS
2.1 | Data acquisition

Analyses are based on SoilTemp, a global database of microclimate time series (Lembrechts et al., 2020). We compiled soil temperature measurements from 9362 unique sensors (mean duration 2.9 years, median duration 1.0 year, ranging from 1 month to 41 years) from 60 countries, using both published and unpublished data sources (Figure 1, Figure S1). Each sensor corresponds to one independent time series.

We used time series spanning a minimum of 1 month, with a temporal resolution of 4 h or less. Sensors of any type were included (Table S1), as long as they measured in situ. Sensors in experimentally manipulated plots, that is, plots in which microclimate has been manipulated, such as in open top chambers, were excluded. Most data (>90%) came from low-cost rugged microclimate loggers such as iButtons (Maxim Integrated, USA) or TMS4-sensors (Wild et al., 2019), with measurement errors of around 0.5–1°C (note that we are using degree Celsius over Kelvin throughout, for ease of understanding), while in a minority of cases sensors with higher meteorological specifications such as industrial or scientific-grade thermocouples and thermistors (measurement errors of less than 0.5°C) were used. Contributing data sets mostly consisted of short-term regional networks of microclimate measurements, yet also included a set (<5%) of soil temperature sensors from long-term research networks equipped with weather stations (e.g. Pastorello et al., 2017). By combining these two types of data, a much higher spatial density of sensors and broader distribution of microhabitats could be obtained than by using weather station data only.

About 68% of sensors were deployed between 2010 and 2020 and 93% between 2000 and 2020; we, thus, focus on the latter period in our analyses. Additionally, given the relatively short time frame covered by most individual sensors and thus the lack of spatially unbiased long-term time series, we were not able to test for systematic differences in the temperature offset between old and recent data sets, and thus we did not correct for this in our models. We strongly urge future studies to assess such temporal dynamics in the offset once long-term microclimate data have become sufficient and more available.

For each of the individual 9362 time series, we calculated monthly mean, minimum (5% percentile of all monthly values) and maximum (95% percentile) temperature, after checking all time series for plausibility and erroneous data. These monthly values, while perhaps not fully intercomparable between the Northern and Southern Hemispheres, are those that have traditionally been used to calculate bioclimatic variables (Fick & Hijmans, 2017). Months with more than 1 day of missing data, either at the beginning or end of the measurement period, or due to logger malfunctioning during measurement, were excluded, resulting in a final subset of 380,676 months of soil temperature time series that were used for further analyses. For each sensor with more than 12 months of data, we calculated moving averages of annual mean temperature, using each consecutive month as a starting month and calculating the mean temperature including the next 11 months. We used these moving averages to make maximal use of the full temporal extent covered by each sensor because each time series spanned a different time period, often including parts of calendar years only.

The selected data set contained sensors installed strictly belowground, measuring temperature at depths between 0 and 200 cm
below the ground surface. Sensors recording several measurements at the same site but located at different (vertical) depths were included separately (the 9362 unique sensors thus came from 7251 unique loggers).

Sensors were grouped in different soil depth categories (0–5, 5–15, 15–30, 30–60, 60–100, 100–200 cm, Table S2) to incorporate the effects of soil temperature dampening associated with vertical stratification. We limited our analyses to the topsoil (0–5 cm) and the second...
FIGURE 1  Temperature offsets between soil and air temperature differed significantly among biomes. (a) Distribution of in situ measurement locations across the globe, coloured by the mean annual temperature offset (in °C) between in situ measured soil temperature (topsoil, 0–5 cm depth) and gridded air temperature (ERA5-Land). Offsets were averaged per hexagon, each with a size of approximately 70,000 km². Mollweide projection. (b) Mean annual temperature offsets per Whittaker biome (adapted from Whittaker 1970, based on geographic location of sensors averaged at 1 km²; 0–5 cm depth), ordered by mean temperature offset and coloured by mean annual precipitation. (c–d) Distribution of sensors in 2D climate space for the topsoil (c, 0–5 cm depth, N = 4530) and the second layer (d, 5–15 cm depth, N = 3989). Colours of hexagons indicate the number of sensors at each climatic location, with a resolution of 1.2°C (x-axis) and 100 mm (y-axis). Grey dots in the background represent the global variation in climatic space (obtained by sampling 1,000,000 random locations from the CHELSA world maps). Overlay with grey lines depicts a delineation of Whittaker biomes.

soil layer (5–15 cm), as we currently lack sufficient global coverage to make accurate models at deeper soil depths (8519 time series, about 91%, came from the two upper depth layers). Due to uncertainty in the identification of these soil depths between studies (e.g. due to litter layers), no finer categorization is used.

We tested for potential bias in temporal resolution (i.e. measurement interval) by calculating mean, minimum and maximum temperature for a selection of 2000 months for data measured every 15 min, and the same data aggregated to 30, 60, 90, 120 and 240 min. Monthly mean, minimum and maximum temperature calculated with any of the aggregated data sets differed on average less than 0.2°C from the ones with the highest temporal resolution. We were, thus, confident that pooling data with different temporal resolutions of 4 h or finer would not significantly affect our results.

2.2  Temperature offset calculation

For each monthly value at each sensor location (see Table S3 for number of data points per month), we extracted the corresponding monthly means of the 2 m air temperature from the European Centre for Medium-Range Weather (ECMWF) Forecast’s 5th reanalysis (ERA5) (from 1979 to 1981) and ERA5-Land from 1981 to 2020 (Copernicus Climate Change Service (C3S), 2019), hereafter called ERASL. The latter data set models the global climate with a spatial resolution of 0.08 × 0.08 degrees (≈9 × 9 km at the equator) with an hourly resolution, converted into monthly means using daily means for the whole month. Similarly, monthly minima and maxima were obtained from TerraClimate (Abatzoglou et al., 2018) for the period 2000 to 2020 at a 0.04 × 0.04 degrees (≈4 × 4 km at the equator) resolution. Monthly means for TerraClimate were not available, and we therefore estimated them by averaging the monthly minima and maxima. Finally, we also obtained monthly mean temperatures from CHELSA (Karger et al., 2017a, 2017b) for the period 2000 to 2013 at a 30 × 30 arc second (≈1 × 1 km at the equator) resolution. In our modelling exercises (see section 2.5 Modelling below), we opted to use the mean temperature offsets as calculated based on ERASL rather than on CHELSA. While CHELSA’s higher spatial resolution is definitely an advantage, its time period (stopping in 2013) insufficiently overlapped with the time period covered by our in situ measurements (2000–2020), soil temperature offsets based on the CHELSA data set were only used for comparative purposes.

We used TerraClimate to model offsets in monthly minimum and maximum temperature.

We calculated moving annual averages of the gridded air temperature data in the same way as for soil temperature. These were used to create annual temperature offset values following the same approach as above.

The offset between the in situ measured soil temperature in the SoilTemp database and the 2 m free-air temperature obtained from the air-temperature grids (ERASL, TerraClimate and CHELSA, hereafter called ‘gridded air temperature’) was calculated by subtracting the monthly or annual mean air temperature from the monthly or annual mean soil temperature. Positive offset values indicate a measured soil temperature higher than gridded air temperature, whereas negative offset values represent cooler soils. Similarly, monthly minimum and maximum air temperature were subtracted from minimum and maximum soil temperature, respectively. Monthly minima and maxima of the soil temperature were calculated as, respectively, the 5% lowest and highest instantaneous measurement in that month, to correct for outliers, which can be especially pronounced at the soil surface (Speak et al., 2020). As a result, patterns in minima and maxima are more conservative estimates than if we had used the absolute lowest and highest values.

Importantly, the temperature offset calculated here is a result of three key groups of drivers: (1) height effects (2 m versus 0–15 cm below the soil surface); (2) environmental or habitat effects (e.g. spatial variability in vegetation, snow or topography); and (3) spatial scale effects (resolution of gridded air temperature) (Lembrechts et al., 2020). We investigated the potential role of scale effects by comparing gridded air temperature data sources with different resolutions (ERASL, TerraClimate and CHELSA, see below; Figures S2–S3). Height effects and environmental effects are, however, not disentangled here, as the offset we propose incorporates both the difference between air and soil temperature (vertically), as well as the difference between free-air macroclimate and in situ microclimate (horizontally) in one measure (Lembrechts et al., 2020). While it can be argued that it would be better to treat both vertical and horizontal effects separately, this would require a similar database of coupled in situ air and soil temperature measurements, which is not yet available. Using in situ measured air temperature could also solve spatial mismatches (i.e. spatially averaged air temperature represents the whole 1 to 81 km² pixel, depending on pixel size, not only the exact location of the sensor). However, coupled air and soil temperature measurements are not only rare, but the air temperature measurements also have large measurement errors, especially in open habitats (Maclean et al., 2021). These errors can be up to several degrees in open habitats when using non-standardized sensors, loggers and shielding (Holden et al., 2013; Maclean et al., 2021;
Terando et al., 2017). Hence, using in situ measured air temperature without correcting for these measurement errors would be misleading.

2.3 | Global and biome-level analyses

For the purpose of visualization, annual offsets were first averaged in hexagons with a resolution of approximately 70,000 km², using the dggridR-package (version 2.0.4) in R (Barnes et al., 2017) (Figure 1). Next, we plotted mean, minimum and maximum annual soil temperature as a function of corresponding gridded air temperature from ERA5L, TerraClimate and CHELSA and used generalized additive models (GAMs, package mgcv 1.8-31; Wood, 2012) to visualize deviations from the 1:1-line (i.e. temperature offsets deviating from zero, Figures S4–S5).

All annual and monthly values within each soil depth category and falling within the same 1-km² pixel were aggregated as a mean, resulting in a total of c. 1200 unique pixels at 0–5 cm, and c. 1000 unique pixels at 5–15 cm each month, across the globe (Tables S3–S5). This averaging includes summarizing the data over space, that is, multiple sensors within the same 1-km² pixel, and time, that is, data from multi-year time series from a certain sensor, to reduce spatial and temporal autocorrelation and sampling bias. We assigned these 1-km² averages to the corresponding Whittaker biome of their georeferenced location, using the package plotbiomes (version 0.0.0.9901) in R (Figure 1c,d, Tables S4–S5 (Stefan & Levin, 2018)). We ranked biomes based on their offset and compared this with the mean annual precipitation in each biome (Figure 1b). This was done separately for each air temperature data source (ERA5L, TerraClimate and CHELSA), soil depth (0–5 cm, 5–15 cm) and time frame (ERA5L 1979–2020, 2000–2020), as well as for the offset between monthly minimum and maximum soil temperature and the minimum and maximum gridded air temperature from TerraClimate. Our analyses showed that patterns were robust to variation in spatial resolution, sensor depth, climate interpolation method and temporal scale (Figures S2–S5).

2.4 | Acquisition of global predictor variables

To create spatial predictive models of the offset between in situ soil temperature and gridded air temperature, we first sampled a stack of global map layers at each of the logger locations within the data set. These layers included long-term macroclimatic conditions, soil texture and physiochemical information, vegetation, radiation and topographic indices as well as anthropogenic variables. Details of all layers, including descriptions, units and source information, are described in Supplementary Data S1. In short, information about soil texture, structure and physiochemical properties was obtained from SoilGrids (version 1 [Hengl et al., 2017]), limited to the upper soil layer (top 5 cm). Long-term averages of macroclimatic conditions (i.e. monthly mean, maximum and minimum temperature, monthly precipitation) was obtained from CHELSA (version 2017 [Karger et al., 2017a]), which includes climate data averaged across 1979–2013, and from WorldClim (version 2 [Fick & Hijmans, 2017]). Monthly snow probability is based on a pixel-wise frequency of snow occurrence (snow cover >10%) in MODIS daily snow cover products (MOD10A1 & MYD10A1 [Hall et al., 2002]) in 2001–2019. Spectral vegetation indices (i.e. averaged MODIS NDVI product MYD13Q1) and surface reflectance data (i.e. MODIS MCD43A4) were obtained from the Google Earth Engine Data Catalog (developers.google.com/earth-engine/datasets) and averaged from 2015 to 2019. Landcover and topographic information were obtained from EarthEnv (Amatulli et al., 2018). Aridity index (AI) and potential evapotranspiration layers were obtained from CGIAR (Zomer et al., 2008). Anthropogenic information (population density) was obtained from the EU JRC (ghi.jrc.ec.europa.eu/ghs_pop2019.php). Aboveground biomass data were obtained from GlobBiomass (Santoro, 2018). RESOLVE ecoregion classifications were used to categorize sampling locations into biomes (Dinerstein et al., 2017). With this set of predictor variables, we included information on all different categories of drivers of soil temperature. An important variable that had to be excluded was snow depth, due to the lack of a relevant 1-km² resolution global product. The final set of predictor variables included 24 ‘static’ variables and eight monthly layers (i.e. maximum, mean, and minimum temperature, precipitation, cloud cover, solar radiation, water vapour pressure and snow cover). As cloud cover estimates were not available for high-latitude regions in the Northern Hemisphere in January and December due to a lack of daylight, we excluded cloud cover as an explanatory variable for these months (i.e. ‘EarthEnvCloudCover_MODCF_monthlymean_XX’, with XX representing the months in two-digit form Supplementary Data S1).

All variable map layers were reprojected and resampled to a unified pixel grid in EPSG:4326 (WGS84) at 30 arc-sec resolution (=1 x 1 km at the equator). Areas covered by permanent snow or ice (e.g. the Greenland ice cap or glaciated mountain ranges, identified using SoilGrids) were excluded from the analyses. Antarctic sampling points were excluded from the modelling data set owing to the limited coverage of several covariate layers in the region.

2.5 | Modelling

To generate global maps of monthly temperature offsets (Figure 2), we trained Random Forest (RF) models for each month, using the temperature offsets as the response variables and the global variable layers as predictors (Breiman, 2001; Hengl et al., 2018). We used a geospatial RF modelling pipeline as developed by van den Hoogen et al. (2021). RF models are machine learning models that combine many classification trees using randomized subsets of the data, with each tree iteratively dividing data into groups of most closely related data points (Hengl et al., 2018). They are particularly valuable here due to their capacity to uncover nonlinear relationships (e.g. due to increased decoupling of soil from air temperature in colder and thus snow-covered areas) and their ability to capture complex interactions among covariates (e.g. between snow and vegetation cover) (Olden et al., 2008). Furthermore, they may currently have advantages over mechanistic microclimate models for global modelling (Maclean &
Klinges, 2021), as the latter require highly detailed physical input parameters for calibration, and current computational barriers preclude global assessments at a $1\,\text{km}^2$ resolution and over multiple decades.

Nevertheless, we urge future endeavours to compare and potentially improve our results with estimates based on such mechanistic models.

**Figure 2** Global modelled temperature offsets between soil and air temperature show strong spatiotemporal variation across months. Modelled annual (a) and monthly (b–m) temperature offset (in °C) between in situ measured soil temperature (topsoil, 0–5 cm) and gridded air temperature. Positive (red) values indicate soils that are warmer than the air. Dark grey represents regions outside the modelling area.
We performed a grid search procedure to tune the RF models across a range of 52 hyperparameter settings (variables per split: 2–14, minimum leaf population: 2–5, in all combinations adding up to 52 models, each time with 250 trees). During this procedure, we assessed each of the 52 model’s performance using k-fold cross-validation (k = 10; folds assigned randomly, stratified per biome). The models’ mean and standard deviation values were the basis for choosing the best of all evaluated models. This procedure was repeated for each month separately for the two soil depth layers (0–5 cm, 5–15 cm), for offsets in mean, minimum and maximum temperature. The importance of predictor variables was assessed using the variable importance and ordered by mean variable importance across all models. This variable importance adds up the decreases in the impurity criterion (i.e. the measure on which the local optimal condition is chosen) at each split of a node for each individual variable over all trees in the forest (van den Hoogen et al., 2021).

### 2.6 Soil bioclimatic variables

The resulting global maps of the annual and monthly offsets between mean, minimum and maximum soil and air temperature were used to calculate relevant bioclimatic variables following the definition used in CHELSA, BIOCLIM, ANUCLIM and WorldClim (Booth et al., 2014; Fick & Hijmans, 2017; Karger et al., 2017a; Xu & Hutchinson, 2011) (Table 1, Figures 3–4). First, we calculated monthly soil mean, maximum and minimum temperature by adding monthly temperature offsets to the respective CHELSA monthly mean, maximum and minimum temperature (Karger et al., 2017a). Next, we used these soil temperature layers to compute 11 soil bioclimatic layers (SBIO, Table 1) (O’Donnell & Ignizio, 2012). Wettest and driest quarters were identified for each pixel based on CHELSA’s monthly values.

### 2.7 Model uncertainty

To assess the uncertainty in the monthly models, we performed a stratified bootstrapping procedure, with total size of the bootstrap samples equal to the original training data (van den Hoogen et al., 2021). Using biomes as a stratification category, we ensured the samples included in each of the bootstrap training collections were proportionally representative of each biome’s total area. Next, we trained RF models (with the same hyperparameters as selected during the grid-search procedure) using each of 100 bootstrap iterations. Each of these trained RF models was then used to classify the predictor layer stack, to generate per-pixel 95% confidence intervals and standard deviation for the modelled monthly offsets (Figure 5a, Figure S6a). The mean R² value of the RF models for the monthly mean temperature offset was 0.70 (from 0.64 to 0.78) at 0–5 cm and 0.76 (0.63–0.85) at 5 to 15 cm across all 12 monthly models. Mean RMSE of the models was 2.20°C (1.94–2.51°C) at 0–5 cm, and 2.06°C (1.67–2.35°C) at 5–15 cm.

Importantly, model uncertainty as reported in Figure 5a and Figure S6a comes on top of existing uncertainties in (1) in situ soil temperature measurements and (2) the ERA5L macroclimate models as used in our models. However, both of those are usually under 1°C (Copernicus Climate Change Service (C3S), 2019; Wild et al., 2019).

To assess the spatial extent of extrapolation, which is necessary due to the incomplete global coverage of the training data, we first performed a principal component analysis (PCA) on the full environmental space covered by the monthly training data, including all explanatory variables as used in the models, and then transformed the composite image into the same principal components’ (PC) spaces as of the sampled data (van den Hoogen et al., 2019). Next, we created convex hulls for each of the bivariate combinations from the first 10 to 12 PCs, covering at least 90% of the sample space variation, with the number of PCs depending on the month. Using the coordinates of these convex hulls, we assessed whether each pixel fell within or outside each of these convex hulls and calculated the percentage of bivariate combinations for which this was the case (Figure 5b, Figure S6b). This process was repeated for each month and for each of the two soil depths separately.

These uncertainty maps are important because one should be careful with extrapolation beyond the range of conditions covered by the environmental variables included in the original calibration data set, especially in the case of non-linear patterns such as modelled here. The maps are provided as spatial masks to

| Bioclimatic variable | Meaning |
|----------------------|---------|
| SBIO1                | annual mean temperature |
| SBIO2                | mean diurnal range (mean of monthly (max temp - min temp)) |
| SBIO3                | isothermality (SBIO2/SBIO7) (x100) |
| SBIO4                | temperature seasonality (standard deviation x100) |
| SBIO5                | max temperature of warmest month |
| SBIO6                | min temperature of coldest month |
| SBIO7                | temperature annual range (SBIO5-SBIO6) |
| SBIO8                | mean temperature of wettest quarter |
| SBIO9                | mean temperature of driest quarter |
| SBIO10               | mean temperature of warmest quarter |
| SBIO11               | mean temperature of coldest quarter |
remove or reduce the weighting of the pixels for which predictions are beyond the range of values covered by the models during calibration. To assess this further, we used a spatial leave-one-out cross-validation analysis to test for spatial autocorrelation in the data set (Figure S7) (van den Hoogen et al., 2021). This approach trains a model for each sample in the data set on all remaining samples, excluding data points that fall within an increasingly large buffer around that focal sample. Results show lowest confidence for May to September at 5–15 cm, likely driven by uneven global coverage of data points.

Finally, we compared the modelled mean annual temperature (SBIO1, topsoil layer) with a similar product based on monthly ERA5L topsoil (0–7 cm) temperature with a spatial resolution of 0.08 × 0.08 degrees (=9 × 9 km at the equator, Copernicus Climate Change Service (C3S), 2019). The corresponding SBIO1 based on ERA5L was calculated using the means of the monthly averages for each month over the period 1981 to 2016, and averaging these 12 monthly values into one annual product. We then visualized spatial differences between SBIO1 and ERA5, as well as differences across the macroclimatic gradient, to identify mismatches between both data sets.
All geospatial modelling was performed using the Python API in Google Earth Engine (Gorelick et al., 2017). The R statistical software, version 4.0.2 (R Core Team, 2020), was used for data visualizations. All maps were plotted using the Mollweide projection, which preserves relative areas, to avoid large distortions at high latitudes.

2.8 | Sources of uncertainty

The temporal mismatch between the period covered by CHELSA (1979–2013) and our in situ measurements (2000–2020) prevented us from directly using CHELSA climate to calculate the temperature offsets used in our models. This temporal mismatch might affect the offsets calculated here because the relationship between temperature offset and macroclimate will change through time as the climate warms. Similarly, inter-annual differences in offsets due to specific weather conditions cannot be implemented in the used approach. However, we are confident that, at the relatively coarse spatial (1 km²) and temporal (monthly averages) resolution we are working at, our results are sufficiently robust to withstand these temporal issues, given that we found high consistency in offset patterns between the different time frames and air temperature data sets examined (Figures S2–S5). Nevertheless, we strongly urge future research to disentangle...
these potential temporal dynamics, especially given the increasing rate at which the climate is warming (GISTEMP Team, 2021; Xu et al., 2018).

Similarly, a potential bias could result from the mismatch in method and resolution between ERA5L—used to calculate the temperature offsets—and CHELSA, which was used to create the bioclimatic variables. However, even though temperature offsets have slightly larger variation when based on the coarser-grained ERA5L-data than on the finer-grained CHELSA-data, Figures S2–S5 show that relationships between soil and air temperature are largely consistent in all biomes and across the whole global temperature gradient. Therefore, the larger offsets created additional random scatter, yet no consistent bias.

Finally, we acknowledge that the 1-km² resolution gridded products might not be representative of conditions at the in situ measurement locations within each pixel. This issue could be particularly significant for different vegetation types (here proxied at the pixel level using total aboveground biomass (unit: tons/ha i.e., Mg/ha, for the year 2010; Santoro, 2018) and NDVI (MODIS NDVI product MYD13Q1, averaged over 2015–2019). To verify this, we compared a pixel’s estimated aboveground biomass with the dominant in situ habitat (forest versus open) surrounding the sensors in that pixel (Table S6). Importantly, all sensors installed in forests fell indeed in pixels with more than 1 ton/ha aboveground biomass. Similarly, 75% or more of sensors in open terrain fell in pixels with biomass estimates of less than 1 ton/ha. Only in the temperate woodland biome was the match between in situ habitat estimates and pixel-level aboveground biomass lower, with less than 95% of sensors in forested locations correctly placed in pixels with more than 1 ton/ha biomass, and less than 50% of open terrain sensors in pixels with less than 1 ton/ha biomass. While our predictions will thus not be accurate for locations within a pixel that largely deviate from average conditions (e.g. open terrain in pixels identified as largely forested, or vice versa), they should be largely representative for those pixel-level averages.

3 | RESULTS

3.1 | Biome-wide patterns in the temperature offset

We found positive and negative temperature offsets of up to 10°C between in situ measured mean annual topsoil temperature and gridded air temperature (mean = 3.0 ± 2.1°C standard deviation, Figure 1, 0–5 cm depth; 5–15 cm is available in Figures S2, S5). The magnitude and direction of these temperature offsets varied considerably within and across biomes. Mean annual topsoil temperature was on average 3.6 ± 2.3°C higher than gridded air temperature in cold and/or
dry biomes, namely tundra, boreal forests, temperate grasslands and subtropical deserts. In contrast, offsets were slightly negative in warm and wet biomes (tropical savannas, temperate forests and tropical rainforests) where soils were, on average, 0.7 ± 2.7°C cooler than gridded air temperature (Figure 1b, Figures S2 and S5; note, however, the lower spatial coverage in these biomes in Figure 1a,c,d, Table S4). Temperature offsets in annual minimum and maximum temperature amounted to c. 10°C maximum. While annual soil temperature minima were on average higher than corresponding gridded air temperature minima in all biomes, temperature offsets of annual maxima followed largely the same biome-related trends as seen for the annual means, albeit with the higher variability expected for temperature extremes (Figures S2g, S2h, S4g, S4h). Using different air temperature data sources did not alter the annual temperature offset and biome-related patterns (see Methods and Figures S2-S5).

Soils in the temperate seasonal forest biome were on average 0.8°C (±2.2°C) cooler than air temperature within 1-km² grid cells of forested habitats, and 1.0°C (±4.0°C) warmer than the air within 1-km² grid cells of non-forested habitats, resulting in a biome-wide average of 0.5°C (Table S7). Similar patterns were observed in other biomes.

3.2 Temporal and spatial variation in temperature offsets

Our RF outputs highlighted a strong seasonality in monthly temperature offsets, especially towards higher latitudes (Figure 2). High-latitude soils were found to be several degrees warmer than the air (monthly offsets of up to 25°C) during their respective winter months, and cooler (up to 10°C) in summer months, both at 0–5 cm (Figure 2) and 5–15 cm (Figure S8) soil depths. In the tropics and subtropics, soils in dry biomes (e.g. in the Sahara Desert or southern Africa) were predicted to be warmer than air throughout most of the year, while soils in mesic biomes (e.g. tropical biomes in South America, central Africa and Southeast Asia) were modelled to be consistently cooler, at both soil depths. These global gridded products were then used to create temperature-based global bioclimatic variables for soils (SBIO, Figure 3, Figure S9).

3.3 Global variation in soil temperature

We observed 17% less spatial variation in mean annual soil temperature globally (expressed by the standard deviation) than in air temperature, largely driven by the positive offset between soil and air temperature in cold environments (Figure 4). Importantly, our machine learning models slightly (up to 1°C, or around 10% of variation) underestimated temperature offsets at both extremes of the temperature gradient at the 1-km² resolution (Figure S10) and likely even more in comparison with finer-resolution products. Estimates of the reduction in variation across space are thus conservative, especially in the coldest biomes. The reduction in spatial temperature variation was observed in all cold and cool biomes, with tundra and boreal forests having both a significant positive mean temperature offset and a reduction of 20% and 22% in variation, respectively (Figure 4c). In the warmest biomes (e.g. tropical savanna and subtropical desert), however, we found an increase in variation of, on average, 10%.

Our bootstrap approach to validate modelled monthly offsets indicated high consistency among the outcomes of 100 bootstrapped models (Figure 5, Figure S6a), with standard deviations in most months and across most parts of the globe around or below ±1°C. One exception to this was the temperature offset at high latitudes of the Northern Hemisphere during winter months (standard deviation up to ±5°C in the 0–5 cm layer). Predictive performance was comparable across biomes, although with large variation in data availability (Figure S11).

The importance of predictor variables in the RF models was largely consistent across months. Macroclimatic variables such as incoming solar radiation as well as long-term averages in air temperature and precipitation were by far the most influential explanatory variables in the spatial models of the monthly temperature offset (Figures S12 and S13).

We highlight that the current availability of in situ soil temperature measurements is significantly lower in the tropics (Table S5), where our model had to extrapolate temperatures beyond the range used to calibrate the model (Figure 5b, Figure S6b).

Finally, our comparison with a mean annual soil temperature product derived from the coarse-resolution ERA5L topsoil temperature showed that spatial variability, for example, driven by topographic heterogeneity, is much better captured here than in the coarser resolution of the ERA5L-based product (Figure 6c-e). Nevertheless, our predictions at the coarse scale showed to be condensed within a 5°C range of values from the ERA5L-predictions, for more than 95% of pixels globally. Noteworthy, our predictions resulted in consistently cooler soil temperature predictions than topsoil conditions provided by ERA5L across large areas, such as the boreal and tropical forest biomes (Figure 6a,b). Additionally, our models predicted lower values for SBIO1 than ERA5L in all regions with mean annual soil temperature below 0°C, except for a few locations around Greenland and Svalbard (Figure 6a,b).

4 DISCUSSION

4.1 Global patterns in soil temperature

We observed large spatiotemporal heterogeneity in the global offset between soil and air temperature, often in the order of several degrees annually and up to more than 20°C during winter months at high latitudes. These values are in line with empirical data from regional studies (Lembrechts et al., 2019; Obu et al., 2019; Zhang et al., 2018). Both annual and monthly offsets showed clear discrepancies between cold and dry versus warm and wet biomes. The modelled monthly offsets covaried strongly negatively with both long-term averages in free-air temperature and solar radiation, linking to the well-known decoupling of soil from air temperature due to snow (for cold extremes in cold and...
However, the secondary importance of variables related to precipitation and soil structure hints to the additional distinction between wet and dry biomes at the warm end of the temperature gradient. There, buffering due to shading, evapotranspiration and the specific heat of water (mostly against warm extremes in warm and wet biomes) results in cooler soil temperature (De Frenne et al., 2013; Geiger, 1950; Grünberg et al., 2020; Grundstein et al., 2005; Hennon et al., 2010; Wang & Dickinson, 2012), while such buffering is not as strong in warm and dry biomes due to the lower water availability (Greiser et al., 2018; Wang & Dickinson, 2012; Zhou et al., 2021). As such, these results highlight strong macroclimatic impacts on the soil microclimate across the globe (see also De Frenne et al., 2019), yet with soil temperature importantly non-linearly related to air temperature at the global scale. This confirms that the latter is not sufficient as a proxy for temperature conditions near or in the soil. With our soil-specific global bioclimatic products, we have provided the means to correct for these important region-specific, non-linear differences between soil and air temperature at an unprecedented spatial resolution.

### 4.2 Drivers of the temperature offset

Our empirical modelling approach enabled us to accurately map global patterns in soil temperature. In doing so we did not aim to disentangle the mechanisms governing the temperature offset: such an endeavour would require modelling the biophysics of energy exchange at the soil surface across biomes (Kearney et al., 2013; Geiger, 1950; Grünberg et al., 2020; Grundstein et al., 2005; Hennon et al., 2010; Wang & Dickinson, 2012), while such buffering is not as strong in warm and dry biomes due to the lower water availability (Greiser et al., 2018; Wang & Dickinson, 2012; Zhou et al., 2021). As such, these results highlight strong macroclimatic impacts on the soil microclimate across the globe (see also De Frenne et al., 2019), yet with soil temperature importantly non-linearly related to air temperature at the global scale. This confirms that the latter is not sufficient as a proxy for temperature conditions near or in the soil. With our soil-specific global bioclimatic products, we have provided the means to correct for these important region-specific, non-linear differences between soil and air temperature at an unprecedented spatial resolution.

**Figure 6** The mean annual soil temperature (SBIO1, 1 x 1 km resolution) modelled here is consistently cooler than ERA5L (9 x 9 km) soil temperature in forested areas. (a) Spatial representation of the difference between SBIO1 based on our model and based on ERA5L soil temperature data. Negative values (blue colours) indicate areas where our model predicts cooler soil temperature. Dark grey areas (Greenland and Antarctica) are excluded from our models. Asterisk in Scandinavia indicates the highlighted area in panels d to f (see below). (b) Distribution of the difference between SBIO1 and ERA5L along the macroclimatic gradient (represented by SBIO1 itself) based on a random subsample of 50,000 points from the map in a). Red line from a Generalized Additive Model (GAM) with $k = 4$. (c-e) High-resolution zoomed panels of an area of high elevational contrast in Norway (from 66.0–66.4°N, 15.0–16.0°E) visualizing SBIO1 (c), ERA5L (d) and their difference (e), to highlight the higher spatial resolution as obtained with SBIO1.

---

**FIGURE 6** The mean annual soil temperature (SBIO1, 1 x 1 km resolution) modelled here is consistently cooler than ERA5L (9 x 9 km) soil temperature in forested areas. (a) Spatial representation of the difference between SBIO1 based on our model and based on ERA5L soil temperature data. Negative values (blue colours) indicate areas where our model predicts cooler soil temperature. Dark grey areas (Greenland and Antarctica) are excluded from our models. Asterisk in Scandinavia indicates the highlighted area in panels d to f (see below). (b) Distribution of the difference between SBIO1 and ERA5L along the macroclimatic gradient (represented by SBIO1 itself) based on a random subsample of 50,000 points from the map in a). Red line from a Generalized Additive Model (GAM) with $k = 4$. (c-e) High-resolution zoomed panels of an area of high elevational contrast in Norway (from 66.0–66.4°N, 15.0–16.0°E) visualizing SBIO1 (c), ERA5L (d) and their difference (e), to highlight the higher spatial resolution as obtained with SBIO1.

---

**FIGURE 6** The mean annual soil temperature (SBIO1, 1 x 1 km resolution) modelled here is consistently cooler than ERA5L (9 x 9 km) soil temperature in forested areas. (a) Spatial representation of the difference between SBIO1 based on our model and based on ERA5L soil temperature data. Negative values (blue colours) indicate areas where our model predicts cooler soil temperature. Dark grey areas (Greenland and Antarctica) are excluded from our models. Asterisk in Scandinavia indicates the highlighted area in panels d to f (see below). (b) Distribution of the difference between SBIO1 and ERA5L along the macroclimatic gradient (represented by SBIO1 itself) based on a random subsample of 50,000 points from the map in a). Red line from a Generalized Additive Model (GAM) with $k = 4$. (c-e) High-resolution zoomed panels of an area of high elevational contrast in Norway (from 66.0–66.4°N, 15.0–16.0°E) visualizing SBIO1 (c), ERA5L (d) and their difference (e), to highlight the higher spatial resolution as obtained with SBIO1.
Importantly, many of the predictor variables used in our study (e.g. long-term averages in macroclimatic conditions or solar radiation) are unlikely to represent direct causal relationships underlying the temperature offset, but may rather indirectly relate to many ensuing factors that affect the functioning of ecosystems at fine spatial scales which, in turn, feedback on local temperature offsets, such as energy and water balances, snow cover, wind intensity and vegetation cover (De Frenne et al., 2021). For example, while increased solar radiation itself would theoretically result in soils warming more than the air, high solar radiation at the global scale often coincides with high vegetation cover blocking radiation input to the soil, thus correlating with relatively cooler soils (De Frenne et al., 2021). Our results highlight, however, that the complex relationship between microclimatic soil temperature and macroclimatic air temperature is predictable across large spatial extents thanks to broad scale patterns, even if this is governed by a multitude of local-scale factors involving fine spatiotemporal resolutions. Nevertheless, the predictive quality of our models was lower in high latitude regions, where high variation in the in situ measured offsets—likely driven by the interactions between snow, local topography and vegetation—reduced predictive power of the models at the 1-km² resolution (Greiser et al., 2018; Grünberg et al., 2020; Myers-Smith et al., 2020; Niittynen et al., 2020; Way & Lewkowicz, 2018).

4.3 | Implications for microclimate warming

Our results highlight clear biome-specific differences in mean annual temperature between air and soil temperatures, as well as a significant reduction in the spatial variation in temperature in the soil or near the soil surface, especially in cold and cool biomes (Figure 4). These patterns remain even despite the presence of often strongly opposing monthly offset trends (Figure 2). The observed correlation between long-term averages in macroclimatic conditions and the annual temperature offset illustrates that soil temperature is unlikely to warm at the same rate as air temperature when macroclimate warms. Indeed, one degree of air temperature warming could result in either a bigger or smaller soil temperature change, depending on where along the macroclimatic gradient this is happening. These effects might be seen in cold biome soils most strongly, as they not only experience the largest (positive) temperature offsets and reductions in climate range compared to air temperature (Figure 4b,c), but they are also expected to experience the strongest magnitude of macroclimate warming (Chen et al., 2021; Cooper, 2014; GISTEMP Team, 2021; Overland et al., 2014). As a result, mean annual temperatures in cold climate soils can be expected to warm slower than the corresponding macroclimate as offsets shrink with increasing macroclimate warming.

Contrastingly, predicted climate warming in hot and dry biomes could be amplified in the topsoil, where we show soils to become increasingly warmer than the air at higher temperatures. Similarly, changes in precipitation regimes—and thus soil moisture—can significantly alter the relationship between air and soil temperature, with critical implications for soil moisture-atmosphere feedbacks, especially in hot biomes (Zhou et al., 2021). Indeed, as precipitation decreases, offsets could turn more positive and soil temperatures might warm even faster than the observed macroclimate warming. Therefore, future research should not only use soil temperature data as provided here to study belowground ecological processes (De Frenne et al., 2013; Lembrechts et al., 2020), it should also urgently investigate future scenarios of soil climate warming in light of changing air temperature and precipitation, at ecologically relevant spatial and temporal resolutions to incorporate the non-linear relationships exposed so far (Lembrechts & Nijs, 2020).

4.4 | Within-pixel heterogeneity

We chose to use a 1-km² resolution spatial grid to model mismatches between soil and air temperature, aggregating all values from different microhabitats within the same 1-km² grid cell (e.g. sensors in forested versus open patches) as well as all daily and diurnal variation within a month. Additionally, we used coarse-grained free-air temperature rather than in situ measured air temperatures. We are aware that higher spatiotemporal resolutions would likely reveal the importance of locally heterogeneous variables. Finer-scale factors that affect the local radiation balance and wind (e.g. topography, snow and vegetation cover, urbanization) at the landscape to local scales and those that directly affect neighbouring locations (e.g. topographic shading and cold-air drainage, Ashcroft & Gollan, 2012; Lembrechts et al., 2020; Whitteman, 1982) would probably have emerged as more important drivers at regional scales and with higher spatiotemporal resolutions than those used here (Figure S12). The latter is illustrated by the multi-degree Celsius difference in mean annual temperature between forested and non-forested locations within the same biome (Table S7), as well as the lower accuracy obtained during winter months at high latitudes, where and when fine-scale spatial heterogeneity in snow cover and depth probably lowers models’ predictability at the 1-km² resolution. In situ measurements were largely from areas with a representative vegetation type, supporting the reliability of our predictions for the dominant habitat type within a pixel. However, improved accuracy at high latitudes will depend on the future development of high-resolution snow depth and/or snow water equivalent estimates (Luojus et al., 2010).

The SoilTemp database (Lembrechts et al., 2020) will facilitate the necessary steps towards mapping soil temperature at higher spatiotemporal resolutions in the future, with its georeferenced time series of in situ measured soil and near-surface temperature and associated metadata. Nevertheless, compared with existing soil temperature products such as those from ERA5L (Copernicus Climate Change Service (C3S), 2019), we...
emphasize that the increased resolution of our data products already provides a major technical advance, even though substantial finer within-pixel variation is still lost through spatio-temporal aggregation.

5 | CONCLUSIONS

The spatial (biome-specific) and temporal (seasonally variable) offsets between air and soil temperature quantified here likely bias predictions of current and future climate impacts on species and ecosystems (Bergstrom et al., 2021; Cooper, 2014; Graae et al., 2018; Kearney et al., 2009; Körner & Paulsen, 2004; Opedal et al., 2015; Zellweger et al., 2020). Temperature in the topsoil rather than in the air ultimately defines the distribution and performance of most terrestrial species, as well as many ecosystem functions at or below the soil surface (Gottschall et al., 2019; Hursh et al., 2017; Pleim & Gilliam, 2009; Portillo-Estrada et al., 2016). As many ecosystem functions are highly correlated with temperature (yet often non-linear, Johnston et al., 2021), soil temperature rather than air temperature should in those instances be the preferred predictor for estimating their rates and temperature thresholds (Coûteaux et al., 1995; Rosenberg et al., 1990; Schimel et al., 1996). Correcting for the non-linear relationship between air and soil temperature identified here is, thus, vital for all fields investigating abiotic and biotic processes relating to terrestrial environments (White et al., 2020). Indeed, soil temperature, macroclimate and land-use change will interact to define the future climate as experienced by organisms, and high-resolution soil temperature data are needed to tackle current and future challenges.

By making our global soil temperature maps and the underlying monthly offset data openly available, we offer gridded soil temperature data for climate research, ecology, agronomy and other life and environmental sciences. Future research has the important task of further improving the spatial and temporal resolution of global microclimate products as microclimate operates at much higher temporal resolutions, with temporal variation over hours, days, seasons and years (Bütikofer et al., 2020; Potter et al., 2013), as well as to confirm accuracy of predictions in undersampled regions in the underlying maps (Lembrechts et al., 2021). However, we are convinced that the maps presented here bring us one step closer to having accessible climate data exactly where it matters most for many terrestrial organisms (Ashcroft et al., 2014; Kearney & Porter, 2009; Lembrechts & Lenoir, 2019; Niittynen & Luoto, 2018; Pincebourde et al., 2016). We, nevertheless, highlight that there is still a long way to go towards global soil microclimate data with an optimal spatiotemporal resolution. We, therefore, urge all scientists to submit their microclimate time series to the SoilTemp database to fill data gaps and help to increase the spatial resolution until it matches with the scale at which ecological processes take place (Bütikofer et al., 2020; Lembrechts et al., 2020).

ACKNOWLEDGEMENTS

JL received funding from the Research Foundation Flanders (grant nr. 12P1819N). The project received funding from the Research Foundation Flanders (grants nrs. G018919N, W001919N). JVDH and TWC received funding from DOB Ecology. JA received funding from the University of Helsinki, Faculty of Science (MICROCLIM, grant nr. 7510145) and Academy of Finland Flagship (grant no. 337552). PDF, CM and PV received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (ERC Starting Grant FORMICA 757833). JK received funding from the Arctic Interactions at the University of Oulu and Academy of Finland (318930, Profi 4), Maaja vesiteknikan tuki ry., Tiina and Antti Herlin Foundation, Nordenskiöld Samfundet and Societas pro Fauna et Flora Fennica. MK received funding from the Czech Science Foundation (grant nr. 20-28119S) and the Czech Academy of Sciences (grant nr. RVO 67985939). TWC received funding from National Geographic Society grant no. 9480-14 and WW-240R-17. MA received funding from CISSC (program ICRP (grant nr:2397) and INSF (grant nr: 96005914). The Royal Botanic Garden Edinburgh is supported by the Scottish Government’s Rural and Environment Science and Analytical Services Division. JMA received funding from the Funding Org. Qatar Petroleum (grant nr. QUEX-CAS-QP-RD-18/19). JMA received funding from the European Union’s Horizon 2020 research and innovation program (grant no. 678841) and from the Swiss National Science Foundation (grant no. 31003A_176044). AA was supported by research grants LTUSA19137 (program INTER-EXCELLENCE, subprogram INTER-ACTION) provided by Czech Ministry of Education, Youth and Sports and 20-05840Y of the Czech Science Foundation. AA was supported by the Ministry of Science and Higher Education of the Russian Federation (grant FSRZ-2020-0014). SN, UAT, JJA, and JvO received funding from the Independent Research Fund Denmark (7027-00133B). LvdB, KT, MYB and RC acknowledge funding from the German Research Foundation within the Priority Program SPP-1803 ‘EarthShape: Earth Surface Shaping by Biota’ (grant TI 338/14-1&2 and BA 3843/6-1). PB was supported by grant project VEGA of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences No. 2/0132/18. Forest Research received funding from the Forestry Commission (climate change research programme). JCB acknowledges the support of Universidad Javeriana. JLBA received funding from the Dirección General de Cambio Climático del Gobierno de Aragón; JLBA acknowledges fieldwork assistance by Ana Acín, the Ordesa y Monte Perdido National Park, and the Servicio de Medio Ambiente de Soria de la Junta de Castilla y León. RGB and MPB received funding from BECC - Biodiversity and Ecosystem services in a Changing Climate. MPB received funding from The European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 657627 and The Swedish Research Council FORMAS – future research leaders No. 2016-01187. JB received funding from the Czech Academy of Sciences (grant nr. RVO 67985939). NB received
funding from the SNF (grant numbers 40FA40_154245, 20FI21_148992, 20FI20_173691, 407340_172433) and from the EU (contract no. 774124). ICOS EU research infrastructure. EU FP7 NitroEurope. EU FP7 ECLAIRE. The authors from Biological Dynamics of Forest Fragments Project, PDBFF, Instituto Nacional de Pesquisas da Amazônia, Brazil were supported by the MCTI/ CNPq/FNDCT – Ação Transversal n°68/2013 – Programa de Grande Escala da Biosfera-Atmosfera na Amazônia – LBA; Project ‘Como as florestas da Amazônia Central respondem às variações climáticas? Efeitos sobre dinâmica florestal e sinergia com a fragmentação florestal’. This is the study 829 of the BDFFP Technical Series. To The EUCLIFLUX Cooperative Research Program and Forest Science and Research Institute-IPEF. NC acknowledges funding by Stelvio National Park. JC was funded by the Spanish government grant CGL2016-78093-R. ANID-FONDECYT 1181745 AND INSTITUTO ANTARTICO CHILENO (INACH FR-0418). SC received funding from the German Research Foundation (grant no. DFG- FZT 118, 202548816). The National Science Foundation, Poland (grant no. UMO-2017/27/B/ST10/02228), within the framework of the ‘Carbon dioxide uptake potential of sphagnum peatlands in the context of atmospheric optical parameters and climate changes’ (KUSCO2) project. SLC received funding from the South African National Research Foundation and the Australian Research Council. FM, MC, KU and MU received funding from Slovak Research and Development Agency (no. APVV-19-0319). Instituto Antartico Chileno (INACH_RT-48_16). Iniciativa Cientifica Milenio Núcleo Milenio de Salmóndidos Invasores INVASAL, Institute of Ecology and Biodiversity (IEB), CONICYT PIA APOYO CCTE AFB170008. PC is supported by NERC core funding to the BAS ‘Biodiversity, Evolution and Adaptation Team. EJCI received funding from the Norwegian Research Council (grant number 230970). GND was supported by NERC E3 doctoral training partnership grant (NE/L002558/1) at the University of Edinburgh and the Carnegie Trust for the Universities of Scotland. Monitoring stations on Livingston Island, Antarctica, were funded by different research projects of the Gobem of Spain (PERMAPLANET CTM2009-10165-E; ANTARPERMA CTM2011-15565-E; PERMASNOW CTM2014-52021-R), and the PERMATHERMAL arrangement between the University of Alcalá and the Spanish Polar Committee. GN received funding from the Autonomous Province of Bolzano (ITA). The infrastructure, part of the UK Environmental Change Network, was funded historically in part by ScotNature and NERC National Capability LTS-S: UK-SCAPE; NE/R016429/1). JD was supported by the Czech Science Foundation (GA17-19376S) and MSMT (LTAUSA18007). ED received funding from the Kempe Foundation (JCK-1112 and JCK-1822). The infrastructure was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme I (NPU I), grant number LO1415 and by the project for national infrastructure support CzeCOS/ICOS Reg. No. LM2015061. NE received funding from the German Research Foundation (DFG- FZT 118, 202548816). BE received funding from the GLORIA-EU project no EVK2-CT2000-00056, the Autonomous Province of Bolzano (ITA), from the Tiroler Wissenschaftsfonds and from the University of Innsbruck. RME was supported by funding to the SAFE Project from the Sime Darby Foundation. OF received funding from the German Research Foundation (DFG- FZT 118, 202548816). EPF was supported by the Jardín Botánico Atlántico (SV-20-GIJON-JBA). MF was funded by the German Federal Ministry of Education and Research (BMBF) in the context of The Future Okavango (Grant No. 01LL0912 and SASSCAL (01LLG1201M; 01LLG1201N) projects. EFL received funding from ANID PIA / BASAL FB210006. RAG received funding from Fondedct 11170516, CONICYT PIA AFB170008 and ANID PIA / BASAL FB210006. MBG received funding from National Parks (DYNBIO, #1656/2015) and The Spanish Research Agency (VULBIMON, #CGL2017-90040-R). MG received funding from the Swiss National Science Foundation (ICOS-CH Phase 2 20FI20_173691). FG received funding from the German Research Foundation (DFG- FZT 118, 202548816). KG and TS received funding from the UK Biotechnology and Biological Research Council (grant = 206/D16053). SG was supported by the Research Foundation Flanders (FWO) (project G0H1517N). KJ and PH received funding from the EU Horizon2020 INFRAIA project eLTER-PLUS (871128), the project LTER-CWN (FFG, F&E Infrastrukturförderung, project number 858024) and the Austrian Climate Research Program (ACRP7 – CentForCSink – KR14AC7K11960). SH and ARB received funding through iDiv funded by the German Research Foundation (DFG- FZT 118, 202548816). LH received funding from the Czech Science Foundation (grant nr. 20-28119S) and the Czech Academy of Sciences (grant nr. RVO 67985939). MH received funding from the Baden-Württemberg Ministry of Science, Research and Arts via the project DRieR (Drought impacts, processes and resilience: making the in-visible visible). LH received funding from International Polar Year, Weston Foundation, and ArcticNet. DH received funding from Natural Sciences and Engineering Council (Canada) (RGPIN-06691). TTH received funding from Independent Research Fund Denmark (grant no. 8021-00423B) and Villum Foundation (grant no. 17523). Ministry of Education, Youth and Sports of the Czech Republic (projects LM2015078, VAN2020/01 and CZ.02.1.01/0.0/0.0/16_013/0001708). KH, CG and CJJD received funding from Bolin Centre for Climate Research, Stockholm University and from the Swedish research council Formas [grant no. 2014-00590 to KH]. JJ received funding from the Funding Org. Swedish Forest Society Foundation (grant nr. 2018-485-Steg 2 2017) and Swedish Research Council FORMAS (grant nr. 2018-00792). AJ received funding from the German Federal Ministry of Education and Research BMF (Grant Nr. FKZ 031B0516C SUSALPS) and the Oberfrankenstiftung (Grant Nr. OFS FP00237). ISJ received funding from the Energy Research Fund (NÝR-11 -2019, NÝR-18 -2020). TJ was supported by a UK NERC Independent Research Fellowship (grant number: NE/S01537X/1). RJ received funding from National Science Centre of Poland (grant number: 2016/21/B/ST10/02271) and Polish National Centre for Research and Development (grant number: Pol-Nor/203258/31/2013). VK
received funding from the Czech Academy of Sciences (grant nr. RVO 67985939). AAK received funding from MoECC, Govt of India (AICOPTAX project F. No. 22018/12/2015/RE/Tax). NK received funding from FORMAS (grants nr. 2018-01781, 2018-02700, 2019-00836), VR, support from the research infrastructure ICOS-SE. BK received funding from the National Research, Development and Innovation Fund of Hungary (grant nr. K128441. Ministry of Education, Youth and Sports of the Czech Republic (projects LM20150708 and CZ.02.1.01/0.0/0.0/16_013/0001708). Project B1-RNM-163-UGR-18-Programa Operativo FEDER 2018, partially funded data collection. Norwegian Research Council (NORKLIMA grants #184912 and #244525) awarded to Vigdis Vandvik. MM received funding from the Czech Science Foundation (grant nr. 20-281195) and the Czech Academy of Sciences (grant nr. RVO 67985939). Project CONICYT-PAI 79170119 and ANID-MPG 190029 awarded to Roy Mackenzie. This work was partly funded by project MIUR PON Cluster OT4CLIMA. RM received funding from the SNF project number 407340_172433. FM received funding from the Stelvio National Park. PM received funding from AIAS-COFUND fellowship programme supported by the Marie Skłodowska-Curie actions under the European Union's Seventh Framework Programme for Research, Technological Development and Demonstration (grant agreement no 609033) and the Aarhus University Research Foundation, Denmark. RM received funding from the Ministry of Education, Youth and Sports of the Czech Republic (project LTT17033). SM and VM received funding from EU FP6 NitroEurope (grant nr. 17841), EU FP7 ÉCLAIRE (grant nr. 282910), the Ministry of Education and Science of Ukraine (projects nr. 505, 550, 574, 602), GEF-UNEP funded “Toward INMS” project (grant nr. NEC05348) and ENI CBC BSB PONTOS (grant nr. BSB 889). The authors from Biological Dynamics of Forest Fragments Project, PDBFF, Instituto Nacional de Pesquisas da Amazônia, Brazil were supported by the MCTI/CNPq/FNDCT – Ação Transversal n°68/2013 – Programa de Grande Escala da Biosfera-Atmosfera na Amazônia – LBA; Project ‘Cómo as florestas da Amazônia Central respondem às variações climáticas? Efeitos sobre dinâmica florestal e sinergia com a fragmentação florestal’. FJRM was financially supported by the RFBR grant #18-05-60203-Arkтика. RPu received funding from the Polish National Science Centre (grant project number 2017/27/B/NZ8/00316), ODYSSEE project (ANR-13-ISV7-0004, PN-II-ID-JRP-RO-FR-2012). KR was supported through an Australian Government Research Training Program Scholarship. Fieldwork was supported by the Global Challenges program at the University of Wollongong, the ARC the Australian Antarctic Division and INACH. DR was funded by the project SUBANTECO IPEV 136 (French Polar Institute Paul-Emile Victor), Zone Atelier CNRS Antarctique et Terres Australes, SAD Région Bretagne (Project INFLECT), BiodivERsA 2019-2020 BioDivClim call ‘ASICS’ (ANR-20-EB15-0004), SAR received funding from the Australian Research Council. NSF grant #11556772 to the University of Notre Dame. Pavia University (Italy). OR received funding from EU-LEAP-Agri (RAMSES II), EU-DESIRA (CASSECS), EU-H2020 (SustainSaHEL), AGROPOLIS and TOTAL Foundations (DSCATT), CGIAR (GLDC). AR was supported by the Russian Science Foundation (Grant 18-74-10048).
received funding from Vetenskapsrådet grant nr (No: 2014-04270), ALTER-net multi-site grant, River LIFE project (LIFE08 NAT/S/000266), Flexpeil. Helmholtz Association long-term research program TERENO (Terrestrial Environmental Observatories). PS received funding from the Polish Ministry of Science and Higher Education (grant nr. N N305 304840). AS acknowledges funding by ETH Zürich project FEVER ETH-27 19-1. LSC received funding from NSERC Canada Graduate Scholarship (Doctoral) Program; LSC was also supported by ArcticNet-NCE (insert grant #). Conselho Nacional de Desenvolvimento Científico e Tecnológico (141513/2017-9); Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (E26/200.84/2019). ZS received funding from the SRDA (grants nos. APVV-16-0325 and APVV-20-0365) and from the ERDF (grant no. ITMS 3130115735, CE LignoSilva). JS, MB and CA received funding from core budget of ETH Zurich. State excellence Program M-V \"WETSCAPES\". AfricanBioServices project funded by the EU Horizon 2020 grant number 641918. The authors from KIT/IMK-IFU acknowledge the funding received within the German Terrestrial Environmental Observatories (TERENO) research program of the Helmholtz Association and from the Bavarian Ministry of the Environment and Public Health (UGVo6080204000). Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 192628686, in the framework of the collaborative German-Indonesian research project CRC 990 (SFB): \‘EFForTS, Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (Sumatra, Indonesia)\’. MS received funding from the Ministry of Education, Youth and Sports of the Czech Republic (grant nr. INTER-TRANSFER LTT19018). TT received funding from the Swedish National Space Board (SNSB Dnr 95/16) and the CASSECS project supported by the European Union. HJDT received funding from the UK Natural Environment Research Council (NERC doctoral training partner - the European Union. HJDT received funding from the Swiss National Science Foundation (grant nr. 172198 and 193645). PZ received funding from the Funding Org. Knut and Alice Wallenberg Foundation (grant no. 2015.0047). JL received funding from (i) the Agence Nationale de la Recherche (ANR), under the framework of the young investigators (UCJC) funding instrument (ANR JCJC Grant project N°ANR-19-CE32-0005-01: IMPRINT) (ii) the Centre National de la Recherche Scientifique (CNRS) (Défi INFINITI 2018: MORFO); and the Structure Fédérative de Recherche (SFR) Condorcet (FR CNRS 3417: CREUSE).

Fieldwork in the Arctic got facilitated by funding from the EU INTERACT program. SN, UAT, JJA and JvO would like to thank the field team of the Vegetation Dynamics group for their efforts and hard work. We acknowledge Dominique Tristan for letting access to the field. For the logistic support the crew of INACH and Gabriel de Castilla Station team on Deception Island. We thank the Inuvialuit and Kluane First Nations for the opportunity to work on their land. MADP acknowledges fieldwork assistance and logistics support to Unidad de Tecnología Marina CSIC, and the crew of Juan Carlos I and Gabriel de Castilla Spanish Antarctic Staitons, as well as to the different colleagues from UAH that helped on the instrument maintenance. ERF acknowledges fieldwork assistance by Martin Heggli. MBG acknowledges fieldwork and technical assistance by P Abadía, C Benedé, P Bravo, J Gómez, M Grasa, R Jimenez, H Miranda, B Ponz, J Revilla and P Tejero and the Ordesa and Monte Perdido National Park staff. LH acknowledges field assistance by John Jacobs, Andrew Trant, Robert Way, Darroch Whitaker; we acknowledge the Inuit of Nunatsiavut, and the Co-management Board of Torngat Mountains National Park for their support of this project and acknowledge that the field research was conducted on their traditional lands. We thank our many bear guides, especially Boonie, Eli, Herman, John and Maria Merkuratsuk. AAK acknowledges field support of Akhtar Malik, Rameez Ahmad. Part of micromagnetic records from Saxony was funded by the Saxox Switzerland National Park Administration. Tyson Research Center. JP acknowledges field support of Emmanuel Malet (Edytem) and Rangers of Reserves Naturelles de Haute-Savoie (ASTERS). Practical help: Roel H. Janssen, N. Huig, E. Bakker, Schools in the tepåseförsöket, Forskare fredag, Erik Herberg. The support by the Bavarian Forest National Park administration is highly appreciated. LvdB acknowledges CONAF and onsite support from the park rangers from PN Pan de Azucar, PN La Campana, PN Nahuellbuta and from community agricola Quebrada de Talca. JL and FS acknowledge Manuel Nicolas and all forest officers from the Office National des Forêts (ONF) who are in charge of the RENECOFOR network and who provided help and local support for the installation and maintenance of temperature loggers in the field.
CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
Jonas J. Lembrechts, Johan van den Hoogen and Jonathan Lenoir conceptualized the project; Jonas J. Lembrechts, Johan van den Hoogen, Juha Aalto, Michael B. Ashcroft, Pieter De Frenne, Martin Kopecký, Miska Luoto, Ilya M. D. Maclean, Tom W. Crowther, Ivan Nijs and Jonathan Lenoir designed the paper; the SoilTemp consortium acquired the data; Jonas J. Lembrechts, Johan van den Hoogen, Julia Kempfijn, Pekka Niitnyfen, Jonathan Lenoir analyzed the data; Jonas J. Lembrechts, Johan van den Hoogen, Juha Aalto, Michael B. Ashcroft, Pieter De Frenne, Julia Kempfijn, Martin Kopecký, Miska Luoto, Ilya M. D. Maclean, Tom W. Crowther, Joseph J. Bailey, Stef Haesen, David H. Klinges, Pekka Niitnyfen, Brett R. Scheffers, Koenraad Van Meerbeek, Ivan Nijs and Jonathan Lenoir interpreted the analyses. All authors significantly revised the manuscript and approved it for submission.

DATA AVAILABILITY STATEMENT
All source codes are available at https://doi.org/10.5281/zenodo.4558663.

CODE AVAILABILITY
All source codes are available at https://doi.org/10.5281/zenodo.4558663.

DATA AVAILABILITY STATEMENT
All gridded maps are available at https://zenodo.org/record/4558732#.Ydv0Hf7MJPY. Soil bioclimatic layers SBIO1-11 are also available on Google Earth Engine under projects/crowtherlab/soil_bioclim/SBIO_v1_0_5cm and projects/crowtherlab/soil_bioclim/SBIO_v1_5_15cm. Monthly soil temperature maps are available under projects/crowtherlab/soil_bioclim/soil_temp_monthly_0_5_cm and projects/crowtherlab/soil_temp_monthly_5_15_cm. All monthly data to train the models and reproduce the figures, sampled covariate data and models are available at https://doi.org/10.5281/zenodo.4558663.

ORCID
Jonas J. Lembrechts https://orcid.org/0000-0002-1933-0750
Johan van den Hoogen https://orcid.org/0000-0001-6624-8461
Juha Aalto https://orcid.org/0000-0001-6819-4911
Pieter De Frenne https://orcid.org/0000-0002-8613-0943
Julia Kempfijn https://orcid.org/0000-0001-7521-7229
Martin Kopecký https://orcid.org/0000-0002-1018-9316
Miska Luoto https://orcid.org/0000-0001-6203-5143
Ilya M. D. Maclean https://orcid.org/0000-0001-8030-9136
Thomas W. Crowther https://orcid.org/0000-0001-5674-8913
Joseph J. Bailey https://orcid.org/0000-0002-9526-7095
Stef Haesen https://orcid.org/0000-0002-4491-4213
David H. Klinges https://orcid.org/0000-0002-7900-9379
Pekka Niitnyfen https://orcid.org/0000-0002-7290-029X
Brett R. Scheffers https://orcid.org/0000-0003-2423-3821
Koenraad Van Meerbeek https://orcid.org/0000-0002-9260-3815
Peter Aartsma https://orcid.org/0000-0001-5086-856X
Otar Abdalaze https://orcid.org/0000-0001-8140-0900
Mehdi Abedi https://orcid.org/0000-0002-0051-1491
Rien Aerts https://orcid.org/0000-0001-6694-0669
Negah Ahmadian https://orcid.org/0000-0002-7427-7198
Antje Ahrends https://orcid.org/0000-0002-5083-7760
Juha M. Alatalo https://orcid.org/0000-0001-5084-850X
Jake M. Alexander https://orcid.org/0000-0003-2226-7913
Camille Nina Allonsius https://orcid.org/0000-0003-2599-9941
Jan Altman https://orcid.org/0000-0003-4879-5773
Christof Ammann https://orcid.org/0000-0002-0783-5444
Christian Andres https://orcid.org/0000-0003-0576-6446
Christopher Andrews https://orcid.org/0000-0003-2428-272X
Jonas Ardá https://orcid.org/0000-0002-9318-0973
Nicola Arriga https://orcid.org/0000-0001-5321-3497
Alberto Arzac https://orcid.org/0000-0002-3361-5349
Valeria Aschero https://orcid.org/0000-0003-3865-4133
Rafael L. Assis https://orcid.org/0000-0001-8468-6414
Jakob Johann Assmann https://orcid.org/0000-0002-3492-8419
Maaike Y. Bader https://orcid.org/0000-0003-4300-7598
Khadije Bahalke https://orcid.org/0000-0003-1485-0316
Peter Barančok https://orcid.org/0000-0003-1171-2524
Isabel C. Barrio https://orcid.org/0000-0002-8120-5248
Agustina Barros https://orcid.org/0000-0002-6810-2391
Edmund W. Basham https://orcid.org/0000-0002-0167-7908
Marijn Bauters https://orcid.org/0000-0003-0978-6639
Manuele Bazzichetto https://orcid.org/0000-0002-9874-5064
Luca Belelli Marchesini https://orcid.org/0000-0001-8408-4675
Michael C. Bell https://orcid.org/0000-0002-3401-7746
Juan C. Benavides https://orcid.org/0000-0002-9694-2195
José Luis Benito Alonso https://orcid.org/0000-0003-1086-8834
Bernd J. Berauer https://orcid.org/0000-0002-9472-1532
Jarle W. Bjerke https://orcid.org/0000-0003-2721-1492
Robert G. Björk https://orcid.org/0000-0001-7346-666X
Mats P. Björkman https://orcid.org/0000-0001-5768-1976
Katrin Björsndottir https://orcid.org/0000-0001-7421-9441
Benjamin Blonder https://orcid.org/0000-0002-5061-2385
Pascal Boeckx https://orcid.org/0000-0003-3998-0010
Julia Boike https://orcid.org/0000-0002-5875-2112
Stef Bokhorst https://orcid.org/0000-0003-0184-1162
Bárbara N. S. Brum https://orcid.org/0000-0002-8421-3200
Josef Brúna https://orcid.org/0000-0002-4839-4593
Nina Buchmann https://orcid.org/0000-0003-0826-2980
Josef Luis Camargo https://orcid.org/0000-0003-0370-9878
Otávio C. Campoe https://orcid.org/0000-0001-9810-8834
Onur Candan https://orcid.org/0000-0002-9254-4122
Rafaela Canessa https://orcid.org/0000-0002-6979-9880
Nicoletta Cannone https://orcid.org/0000-0002-3390-3965
Michele Carbognani https://orcid.org/0000-0001-7701-9859
Jofre Carmicer https://orcid.org/0000-0001-7454-8296
Ángelica Casanova-Katny https://orcid.org/0000-0003-3860-1445
Simone Cesarz https://orcid.org/0000-0003-2334-5119
Bagdan Chojinicki https://orcid.org/0000-0002-9012-4060
Philippe Choler https://orcid.org/0000-0002-9062-2721
Steven L. Chown https://orcid.org/0000-0001-6069-5105
and soil moisture logger for long-term microclimate measurement. *Agricultural and Forest Meteorology*, 268, 40–47. https://doi.org/10.1016/j.agrformet.2018.12.018

Wood, S. (2012). mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation. http://cran.r-project.org/web/packages/mgcv/index.html

World Meteorological Organization. (2008). *Guide to meteorological instruments and methods of observation*, Geneva, WMO-No. 8.

Xu, T., & Hutchinson, M. (2011). ANUCLIM version 6.1 user guide. The Australian National University, Fenner School of Environment and Society.

Xu, Y., Ramanathan, V., & Victor, D. G. (2018). Global warming will happen faster than we think. *Nature*. https://doi.org/10.1038/d41586-018-07586-5

Zellweger, F., De Frenne, P., Lenoir, J., Vangansbeke, P., Verheyen, K., Bernhardt-Römermann, M., Baeten, L., Hédil, R., Berki, I., Brunet, J., Van Calster, H., Chudomelová, M., Decocq, G., Dirnböck, T., Durak, T., Heinken, T., Jaroszewicz, B., Kopecký, M., Malis, F., ... Coomes, D. (2020). Forest microclimate dynamics drive plant responses to warming. *Science*, 368, 772–775. https://doi.org/10.1126/science.aba6880

Zhang, Y., Sherstiukov, A. B., Qian, B., Kokelj, S. V., & Lantz, T. C. (2018). Impacts of snow on soil temperature observed across the circumpolar north. *Environmental Research Letters*, 13, 044012. https://doi.org/10.1088/1748-9326/aab1e7

Zhang, Y., Wang, S., Barr, A. G., & Black, T. (2008). Impact of snow cover on soil temperature and its simulation in a boreal aspen forest. *Cold Regions Science and Technology*, 52, 355–370. https://doi.org/10.1016/j.coldregions.2007.07.001

Zhou, S., Williams, A. P., Lintner, B. R., Berg, A. M., Zhang, Y., Keenan, T. F., Cook, B. I., Hagemann, S., Seneviratne, S. I., & Gentine, P. (2021). Soil moisture–atmosphere feedbacks mitigate declining water availability in drylands. *Nature Climate Change*, 11(1), 38–44. https://doi.org/10.1038/s41558-020-00945-z

Zomer, R. J., Trabucco, A., Bossio, D. A., & Verchot, L. V. (2008). Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems and Environment*, 126, 67–80. https://doi.org/10.1016/j.agee.2008.01.014

**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

**How to cite this article:** Lembrechts, J. J., van den Hoogen, J., Aalto, J., Ashcroft, M. B., De Frenne, P., Kemppinen, J., Kopecký, M., Luoto, M., Maclean, I. M. D., Crowther, T. W., Bailey, J. J., Haesen, S., Klinges, D. H., Niittynen, P., Scheffers, B. R., Van Meerbeek, K., Aartsma, P., Abdalaze, O., Abedi, M., ... Lenoir, J. (2022). Global maps of soil temperature. *Global Change Biology*, 28, 3110–3144. https://doi.org/10.1111/gcb.16060