Comparing temperature data sources for use in species distribution models : from in-situ logging to remote sensing

Comparing temperature data sources for use in species distribution models: from *in-situ* logging to remote sensing

Running header: temperature data for distribution models

Jonas J. Lembrechts[1,], Jonathan Lenoir[2], Nina Roth[3], Tarek Hattab[4,2], Ann Milbau[5], Sylvia Haider[6,7], Loïc Pellissier[8,9], Aníbal Pauchard[10,11], Amanda Ratier Backes[6,7], Romina. D. Dimarco[12], Martin A. Nuñez [13], Juha Aalto[14,15], Ivan Nijs[1]*

[1] Centre of Excellence Plants and Ecosystems (PLECO), University of Antwerp, 2610 Wilrijk, Belgium
[2] UR "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN, UMR 7058 CNRS-UPJV), Université de Picardie Jules Verne, 1 Rue des Louvels, 80037 Amiens Cedex 1, France
[3] Biogeography and Geomatics, Department of Physical Geography, Stockholm University, 10691 Stockholm, Sweden
[4] MARBEC (IRD, Ifremer, Université de Montpellier, CNRS), Sète Cedex, France
[5] Research Institute for Nature and Forest - INBO, Havenlaan 88, bus 73, 1000 Brussels, Belgium
[6] Institute of Biology / Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany
[7] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
[8] Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, CH-8092 Zürich, Switzerland
[9] Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland
[10] Laboratorio de Invasiones Biológicas, Facultad de Ciencias Forestales, Universidad de Concepción, Casilla 160-C, 4030000 Concepción, Chile
[11] Institute of Ecology and Biodiversity (IEB), 8320000 Santiago, Chile
[12] Grupo de Ecología de Poblaciones de Insectos, INTA-CONICET, Modesta Victoria 4450, 8400, Bariloche, Argentina
[13] Grupo de Ecología de Invasiones, INIBIOMA, CONICET-Universidad Nacional del Comahue, Av. de Los Pioneros 2350, 8400 Bariloche, Argentina
[14] Dept of Geosciences and Geography, Gustaf Hällströmin katu 2a, FIN–00014 Univ. of Helsinki, Finland
[15] Finnish Meteorological Inst., Finland

**Corresponding author: jonas.lembrechts@uantwerpen.be, +3232651727.*

Orcid ID JJL: orcid.org/0000-0002-1933-0750.
Orcid ID JL: orcid.org/0000-0003-0638-9582.
Orcid ID TH: orcid.org/0000-0002-1420-5758.
Orcid ID AM: orcid.org/0000-0003-3555-8883.
Orcid ID SH: orcid.org/0000-0002-2966-0534.
Orcid ID AP: orcid.org/0000-0003-1284-3163.

Orcid ID JA: orcid.org/0000-0001-6819-4911

Acknowledgements

The research leading to this publication has received funding from the Research Foundation-Flanders (FWO) through a personal grant to JJL, from the European INTERACT-program through a Transnational Access grant to JJL and through the Methusalem funding of the Flemish Community through the Research Council of the University of Antwerp. Computational resources and services were provided where needed by the HPC core facility CalcUA of the University of Antwerp, and VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government – department EWI. AP funded by CONICYT PFB-23 and Fondecyt 1180205. The authors declare no conflicts of interest.

Biosketch

This study is performed in the framework of (1) the Mountain Invasion Research Network (MIREN, www.mountaininvasions.org), a global consortium of plant ecologists focussing on species redistributions in mountain regions, and (2) SoilTemp (https://soiltemp.weebly.com), a global effort to create a database of in-situ soil temperature measurements for use in ecology.

Abstract

Aim: While species distribution models (SDMs) traditionally link species occurrences to free-air temperature data at coarse spatiotemporal resolution, the distribution of organisms might rather be driven by temperatures more proximal to their habitats. Several solutions are currently available, such as downscaled or interpolated coarse-grained free-air temperatures, satellite-measured land surface temperatures (LST) or *in-situ* measured soil temperatures. A comprehensive comparison of temperature data sources and their performance in SDMs is however currently lacking.

Location: Northern Scandinavia

Time period: 1970 - 2017

Major taxa studied: Higher plants

Methods: We evaluated different sources of temperature data (WorldClim, CHELSA, MODIS, E-OBS, topoclimate and soil temperature from miniature data loggers), differing in spatial resolution (1'' to 0.1°), measurement focus (free-air, ground-surface or soil temperature) and temporal extent (year-long vs. long-term averages), and use them to fit SDMs for 50 plant species with different growth forms in a high-latitudinal mountain region.

Results: Differences between these temperature data sources originating from measurement focus and temporal extent overshadow the effects of temporal climatic differences and spatiotemporal resolution, with elevational lapse rates ranging from -0.6 °C per 100 m for long-term free-air temperature data to -0.2 °C per 100 m for in-situ soil temperatures. Most importantly, we found that the performance of the temperature data in SDMs depended on species' growth forms. The use of in-situ soil temperatures improved the explanatory power

of our SDMS (R^2 on average +16%), especially for forbs and graminoids (R^2: +24% and +21% on average, respectively) compared to the other data sources.

Main conclusions: We suggest future studies using SDMs to use the temperature dataset that best reflects the species' ecology, rather than automatically using coarse-grained data from WorldClim or CHELSA.

Keywords: bioclimatic variables, climate change, growth forms, microclimate, mountains, land surface temperature, bioclimatic envelope modelling, soil temperature, species distribution modelling

Introduction

Species distribution models (SDMs) are widely used to describe and forecast the spatial distribution of species (Elith & Leathwick, 2009). SDMs relate species occurrence data with information about the environmental conditions at these locations (Guisan & Thuiller, 2007; Elith & Leathwick, 2009; Jiménez-Valverde *et al.*, 2011). The most common strategy is to work with long-term (e.g. 30 years) interpolated averages of a set of bioclimatic variables at 30'' resolution (ca. 1 × 1 km at the equator), e.g. WorldClim or CHELSA (Hijmans *et al.*, 2005; Warren *et al.*, 2008; Sears *et al.*, 2011; Slavich *et al.*, 2014; Gonzalez-Moreno *et al.*, 2015; Karger *et al.*, 2017). While such macroclimate data might be sufficient to capture the conditions on flat terrains, many environments host a heterogeneous topography (e.g. across steep elevational gradients in mountain regions) that make the microclimate near the ground vary noticeably over short distances (Gottfried *et al.*, 1999; Holden *et al.*, 2011; Scherrer & Körner, 2011; Sears *et al.*, 2011; Opedal *et al.*, 2015; Stewart *et al.*, 2018). In order to make realistic forecasts of species distributions and distribution shifts in such heterogeneous environments, it has been suggested that climate data at finer spatiotemporal resolutions are needed (Illan *et al.*, 2010; Scherrer & Körner, 2011; Graae *et al.*, 2012; Lenoir *et al.*, 2013; Opedal *et al.*, 2015; Graae *et al.*, 2018). Such new climate datasets including *in-situ* logging and remote sensing are now increasingly becoming available (Bramer *et al.*, 2018). Yet, an evaluation of their performance in species distribution models is necessary to provide guidance for future studies, in particular those predicting species responses to climate change (Stewart *et al.*, 2018).

In the high-latitude and high-elevation areas of northern Europe, local temperatures have been found to vary up to 6° C within 1 km² spatial units, reflecting the local topography (Lenoir *et al.*, 2013). This high temperature variation depends for instance on the interaction between temperature and snow distribution, and consequently affects the length of the local growing

120 season (Körner, 2003; Aalto *et al.*, 2018). Local temperatures also vary strongly between
121 seasons, and short-term extreme weather conditions have been shown to be more relevant for
122 species distributions than the average climatic conditions (Ashcroft & Gollan, 2012).
123 Including this variation into SDMs is likely to be crucial, for instance in the context of
124 stepping stones, holdouts or microrefugia (Dobrowski, 2011; Opedal *et al.*, 2015; Meineri &
125 Hylander, 2017). Stepping stones refer to areas with microclimates that facilitate species'
126 range shifts, e.g. upward or poleward movement during climate change or after non-native
127 species introductions (Pauchard *et al.*, 2009; Hannah *et al.*, 2014; Lembrechts *et al.*, 2017).
128 Holdouts and microrefugia on the other hand are areas with a relatively stable microclimate
129 where isolated populations can persist for a certain time (Ashcroft, 2010; Hannah *et al.*, 2014;
130 Lenoir *et al.*, 2017; Meineri & Hylander, 2017). Climatic variability within an area can indeed
131 considerably buffer climate warming effects (Lenoir *et al.*, 2013; Lenoir *et al.*, 2017), which
132 often remains undetected using macroclimate data, possibly leading to the overestimation of
133 rates of extinction and range expansion (Willis & Bhagwat, 2009).

134 Moreover, many organisms (particularly small-stature plants, certain types of insects and soil
135 microbes) experience temperatures at ground or sub-surface level, which can differ strongly
136 from ambient air temperatures that are usually measured at 2 m above the soil surface (Poorter
137 *et al.*, 2016; Aalto *et al.*, 2018; Körner & Hiltbrunner, 2018). Especially in high-latitude and
138 high-elevation regions, snow cover for example acts as an insulator, thereby strongly
139 decoupling soil and air temperatures (Pauli *et al.*, 2013; Poorter *et al.*, 2016; Thompson *et al.*,
140 2018), while biophysical processes due to vegetation cover may also decouple upper
141 atmospheric conditions from boundary layer conditions (Geiger, 1950).

142 In order to overcome this spatiotemporal mismatch between climate data and species ecology
143 and to improve predictions of species' current and future distributions, four main approaches
144 are commonly used: (i) to downscale existing coarse-grained (i.e. 1000 x 1000 m resolution)

145 climate data (McCullough *et al.*, 2016); (ii) to interpolate climate station data (Aalto *et al.*,
146 2017), (iii) to gather local climate data through field measurements (Potter *et al.*, 2013;
147 Slavich *et al.*, 2014; Lenoir *et al.*, 2017); or (iv) to monitor climatic conditions continuously
148 in space and time through remote sensing technologies (e.g. satellite-measured land surface
149 temperatures) (Wan, 2008; Metz *et al.*, 2014; Neteler *et al.*, 2014). In the first two approaches,
150 a high spatial resolution can be obtained using topographic variables derived from digital
151 elevation models which are available at much finer resolutions (e.g. 1'', which is about 30 ×
152 30 m at the equator). Such downscaled or interpolated climate data has been found to be a
153 significant improvement over macroclimatic variables for modelling species distributions
154 (Randin *et al.*, 2009b; Dobrowski, 2011; Slavich *et al.*, 2014; Meineri & Hylander, 2017).

155 In the third approach, one uses actual *in-situ* measurements to provide fine-grained climatic
156 conditions with high spatial accuracy (microclimate) (Opedal *et al.*, 2015; Meineri &
157 Hylander, 2017). Such field measurements can also be interpolated to the level of regional
158 climate using topographical information (Ashcroft *et al.*, 2008; Maclean *et al.*, 2017; Greiser
159 *et al.*, 2018), yet usually cover short temporal and small geographical extents only. In addition
160 to a fine spatial resolution, *in-situ* measurements provide the opportunity to adapt the
161 measurement focus to the ecology or life form of the species, e.g. by measuring near-surface
162 soil temperature instead of air temperature. Gathering *in-situ* temperature data, however,
163 requires considerably more resources than the previously mentioned downscaling approaches
164 (Opedal *et al.*, 2015; Meineri & Hylander, 2017). Increasing the spatiotemporal resolution and
165 extent of such field measurements generally refines the predictions, but also presents a
166 logistical challenge (Wundram *et al.*, 2010; Meineri & Hylander, 2017).

167 Finally, the fourth approach, i.e. using remotely sensed data, is now more frequently used in
168 SDMs (Pottier *et al.*, 2014), for instance through remotely sensed snow cover data or by using
169 the normalized difference vegetation index (NDVI) (Yannic *et al.*, 2014). One such remotely

sensed source of data of which the spatiotemporal resolution, extent and accuracy is rapidly improving is satellite-based land-surface temperatures (LST) (Wan, 2008; Wan *et al.*, 2015). Remotely-sensed LST are now freely available at the global scale at the vegetation canopy or land surface level, with a temporal resolution of days over a period of decades and with a spatial resolution ranging from 30'' (ca. 1000 × 1000 m at the equator) to as fine as 1'' (ca. 30 x 30 m) (Cook, 2014). This type of data does have the advantage over free-air temperature datasets like WorldClim or CHELSA of being a direct and contiguous measurement in space and time, as opposed to data interpolation and temporal averaging from a network of weather stations, yet might be strongly affected by land surface characteristics and cloud cover in the area (Zellweger *et al.*, 2019). Thanks to the increasing availability of these long-term and accurate time series, such satellite-based LST-datasets offer very promising research avenues to fill the gap between local temperature measurements and global-scale climatic datasets.

These different approaches to obtain suitable climate data have been extensively explored and applied in SDMs (Bramer *et al.*, 2018), yet a comparative study of all of these (downscaled and interpolated macroclimate data, field measurements, and satellite-based LST) together – both concerning their inherent characteristics and their role in SDMs - has up till now been missing. Such a comparison is nevertheless urgently needed in order to quantify the progress that can be made by replacing the traditional global climate models with other temperature data sources. We hypothesize in that regard that the best result depends in large on two critical factors: a) the climatic characteristics of the study region, and b), the growth forms of the study organisms. Here, we use a case study along steep climatic gradients in the Northern Scandes, a mountain range in northern Scandinavia, to assess both factors and to provide guidelines for the use of temperature data in SDMs in topographically challenging regions. We compare the characteristics of different temperature datasets within the region, as well as the descriptive and predictive power of SDMs for 50 plant species with different

195 growth forms: forbs, graminoids, (dwarf) shrubs and trees. We compare global climate
196 datasets (i.e. WorldClim and CHELSA) with datasets of remotely-sensed LST (MODIS), a
197 topographic downscaling and interpolation approach, and soil temperature obtained with
198 miniature data loggers, and use three widely applied and ecologically relevant (i.e.
199 bioclimatic) temperature variables: (i) mean annual temperature and mean temperature of the
200 (ii) warmest and (iii) coldest quarter. We hypothesize a significant effect of the spatial
201 resolution of the climate data, as well as of measurement focus (free-air, surface, or soil) and
202 temporal extent on temperature patterns across topographic gradients. Increasing
203 spatiotemporal accuracy of temperature data, especially through the use of *in-situ*
204 measurements, is expected to improve the descriptive and predictive power of the SDMs,
205 despite the associated loss in temporal extent. The optimal resolution, extent and
206 measurement focus are, however, likely to depend on the growth forms of the assessed
207 species, i.e. the spatiotemporal framework in which they operate.

208 **Methods**

209 ***Study region***

210 The study was conducted in the Northern Scandes mountain range in Norway and Sweden,
211 between N 67°46'23.5" / E 16°30'52.6" (south west) and N 68°40'33.6" / E 18°58'40.4"
212 (north east), covering an area of 100 × 100 km and an elevation range from 0 up to 2097 m
213 a.s.l. The area ranges from the Norwegian coast, with a relatively mild and wet climate
214 dominated by birch forests with heathland understory, to the significantly drier and colder
215 eastern side of the Northern Scandes, typically vegetated by subarctic, alpine dwarf shrub
216 vegetation (Lembrechts *et al.*, 2014). The region was chosen for its strong climatic gradient,
217 with large macro- and microclimatic variation due to a distinct topography and high latitude

location (Scherrer & Körner, 2011; Graae *et al.*, 2012; Lenoir *et al.*, 2013). In total, 106 temperature measurement locations were spread across the study area (Fig. 1).

Climate data

For this area, we obtained eight different types of climate data encompassing a wide range of measurement foci, spatiotemporal resolutions and temporal extents (Table 1). For each of these datasets, we extracted or calculated the mean annual temperature and mean temperature of the warmest and coldest quarter (bioclimatic variables Bio1, Bio10 and Bio11, following the definition of WorldClim, Hijmans *et al.*, 2005, hereafter called mean annual, summer and winter temperature, respectively). These ecologically relevant variables belong to the set of physiologically most pertinent bioclimatic determinants of spatial plant species distribution and are thus commonly used in SDMs (e.g. Austin & Van Niel, 2011; Cord & Rödder, 2011; Distler *et al.*, 2015), and they allow us to accurately take into account seasonal differences in climate. The different datasets are discussed in detail below.

a) WorldClim

The WorldClim database (Version 2.0) provides globally interpolated free-air temperature conditions over a 30-year time period (1970-2000) at a spatial resolution of 30'' (ca. 1000 × 1000 m at the equator) (Fick & Hijmans, 2017). The studied bioclimatic variables were directly downloaded from the website (www.worldclim.org).

b) CHELSA

The climatologies at high resolution for the earth's land surface areas (CHELSA, Version 1.2) is a global dataset based on quasi-mechanistical statistical downscaling of free-air temperatures from the ERA Interim (ECMWF) global circulation model (Dee *et al.*, 2011), over a period of 34 years (1979-2013) and with the same spatial resolution as WorldClim

241 (30'', ca. 1000 x 1000 m at the equator), yet for a more recent time period (Karger *et al.*,
242 2017). Bioclimatic variables were again downloaded directly from the website (www.chelsa-
243 climate.org).

244 c) Downscaled CHELSA-data (hereafter called 'downscaled')

245 We used the bioclimatic variables downloaded from CHELSA, at an original resolution of
246 30'' (ca. 1000 x 1000 m at the equator), and downscaled them statistically even further, to a
247 1'' (ca. 30 × 30 m at the equator) resolution based on topographic variation, using a
248 physiographically-informed model fitted with a geographically weighted regression (GWR)
249 technique (Fotheringham *et al.*, 2003). In short, GWR extends the traditional regression
250 approach by allowing estimated regression parameters to vary across space. Therefore, GWR
251 models are particularly relevant to explore the scale-dependent and spatial non-stationary
252 relationships between free-air temperatures and physiographic variables (here: elevation,
253 slope, eastness, northness, distance to the ocean and clear-sky solar radiation) (Su *et al.*,
254 2012). For more details, see Supplementary Material 1.

255 d) Topoclimate

256 Fine-resolution gridded climate data for the region was obtained from Aalto *et al.* (2017), who
257 included topography-driven small-scale climate heterogeneity in a topoclimatic interpolation
258 of weather station data across northern Scandinavia, using generalized additive modelling at a
259 resolution of 1'' (ca. 30 x 30 m at the equator). They modelled monthly average temperatures
260 from 1981 till 2010 using geographical location, elevation, water cover, solar radiation and
261 cold-air pooling. Bioclimatic variables were calculated based on these monthly averages.

262 e) MODIS LST

263 The moderate resolution imaging spectroradiometer (MODIS) satellite TERRA (Wan *et al.*,
264 2015) from the National Aeronatuics and Space Administration (USA) provides global land
265 surface temperature (LST). We extracted data from MOD11A2: 8-day averages based on the
266 clear sky day- and night-time records at a 30'' (ca. 1000 × 1000 m at the equator) resolution,
267 for a period of two years corresponding to the *in-situ* measurements (from August 2015 to
268 July 2017, see below). Mean annual temperature was calculated in ArcGIS by averaging the
269 temperature per pixel for 2015-2016 and 2016-2017, separately, from day of the year (DOY)
270 209 in year *n* (e.g. July 27th for 2015) till DOY 208 in year n+1 (e.g. July 26th for 2016),
271 which was the set of 8-day averages corresponding most closely with the period used for the
272 *in-situ* temperature measurements described below (see sub-section *h* on Soil temperatures).
273 Mean summer and winter temperatures were calculated similarly, yet for DOY 185 (e.g. July
274 3th in 2015) till 272 (September 28th in 2015) and from DOY 1 (e.g. January 1st in 2016) till
275 88 (March 28th in 2016), respectively.

276 f) EuroLST

277 The EuroLST dataset is a gap-filled dataset at the European scale of LST derived from
278 MODIS (see sub-section *e* focusing on MODIS LST) at a spatial resolution of 250 × 250 m
279 and averaged over a temporal extent of 10 years (Metz *et al.*, 2014). This dataset has been
280 created using a combination of weighted temporal averaging with statistical modelling and
281 spatial interpolation to fill in the gaps in the MODIS LST dataset, as well as to improve its
282 spatial resolution. Relevant bioclimatic variables were downloaded directly from the website
283 (courses.neteler.org/eurolst-seamless-gap-free-daily-european-maps-land-surface-
284 temperatures).

285 g) E-OBS

The E-OBS dataset (version 17.0) provides daily gridded climate data of free-air temperature for Europe at a 0.1° (ca. 10.000 × 10.000 m at the equator) spatial resolution, interpolated from weather stations (Haylock *et al.*, 2008), used here over the study period from August 2015 to July 2017 (as in sub-section *e* on MODIS LST). The gridded dataset is created by first interpolating the monthly mean temperature from the weather stations using three-dimensional thin-plate splines, interpolating the daily anomalies using a spatial kriging approach with an external drift for temperature, and then combining these monthly and daily estimates. Temperature data was downloaded directly from the website (https://www.ecad.eu/download/ensembles/download.php) and subsequently used to generate the three studied bioclimatic variables in R.

h) Soil temperatures

Near-surface soil temperatures were logged every 1.5 or 2 hours (iButtons: DS1922L or DS1921G, with 0.5 °C accuracy, www.maximintegrated.com, San José, CA, USA) at a depth of 3 cm below the soil surface in 106 locations along several elevation gradients in Norway and Sweden (Fig. 1, Table 2). Loggers were wrapped in parafilm and put in a small zipper bag to prevent water damage. The loggers were originally established for several different projects (Lembrechts *et al.*, 2014; Lembrechts *et al.*, 2016; Lembrechts *et al.*, 2017) along seven elevation gradients, together ranging from 0 to 1200 m a.s.l., of which three were in Norway and four in Sweden. The three bioclimatic variables were calculated in R (R Core Team, 2015) for each 106 locations and for each year (from 2015 till 2017, corresponding to the periods used in sub-section *e*) from daily averages. Based on these soil temperature data, we made predictions for each bioclimatic variable for the whole study area of 100 × 100 km for the period August 2016 till July 2017 using GWRs (as in sub-section *c* featuring the downscaling approach) based on the same physiographic variables (i.e. elevation, slope, eastness, northness, distance to the ocean and clear-sky solar radiation). The models were

311 used to predict the bioclimatic variables for every 1'' (ca. 30 x 30 m at the equator) pixel in
312 the study area. For more details on the interpolation approach, see Supplementary Material 1.

313 ***Plant species observations***

314 Plant species data were obtained during summer 2017 in the framework of the Mountain
315 Invasion Research Network (www.mountaininvasions.org) long-term monitoring effort, and
316 specifically as a follow-up of the survey of Lembrechts *et al.* (2014) in the Norwegian study
317 plots (59 out of the 106 plots with *in-situ* soil temperature measurements, see Fig. 1, Table 2).
318 Within the framework of this survey, three elevation gradients were selected (spanning on
319 average 700 m in elevation). The elevation range covered by each gradient was divided into
320 19 equally spaced elevation bands, resulting in 20 sampling sites per gradient. At each
321 elevation, presence/absence of all vascular plant species was recorded in plots of 2 × 50 m in
322 natural vegetation. At one end of each of these plots, the temperature logger (see dataset
323 described in sub-section *h* above) was buried. We used data for the 50 most common plant
324 species in the survey (i.e. at least 10 occurrences). Species were grouped based on their
325 growth forms (Table S1): forbs (N = 25); graminoids (N = 7); dwarf shrubs (N = 15); and
326 trees (N = 3). All species were native to the region.

327 ***Direct comparison of climatic variables***

328 1) Relation to elevation

329 To assess differences in the behaviour of the eight climate datasets along an elevation
330 gradient, the three bioclimatic variables derived from these climate datasets were plotted
331 separately against the elevation of the 106 locations of the *in-situ* soil temperature data
332 loggers. For the gridded climate datasets, we extracted a value for each bioclimatic variable
333 for each location. We used linear models (function *lm* in R, R Core Team, 2015) to assess the
334 lapse rate (i.e. the slope, °C per 100 m) of temperature decrease with elevation. For MODIS

LST, E-OBS and the soil temperature measurements, data was plotted and modelled separately for the two study years (2015-2016 and 2016-2017).

2) Paired comparisons

For each of the 106 studied locations, we compared the values for each climatic dataset (and each of the three bioclimatic variables) against the others, to investigate consistent temperature deviations between datasets. Trends for each bioclimatic variable and each dataset were visualised with general additive models (GAMs) with a cubic regression line and without pre-set smoothing value (function gam, R package mgcv, Wood, 2006), following procedures described in Zuur *et al.* (2009). GAMs were used as we did not want to make restrictive assumptions about the relationships of the datasets with each other.

3) Correlative dendrograms

For all 106 locations, we made correlative dendrograms (distance = $1 - \rho$, where ρ is the Pearson's product-moment correlation) to visualize correlations among and relationships between the different datasets, using the function *hclust* from the package *spatstat* (Baddeley *et al.*, 2015).

4) Regional climate predictions

We generated regional maps for the different climate datasets (see the Climate data section as well as Supplementary Material 1 for more details on how the maps were generated for the in-situ measurements), and calculated for each pixel the absolute temperature difference between the respective dataset and the regionally modelled soil temperature at a 1'' (ca. 30 × 30 m at the equator) spatial resolution.

5) Temporal correction

For a more formal comparison between the datasets with different temporal windows, we calculated for each climatic dataset its difference with the 'background climate', taken as temperatures for the window in question from the ERA Interim (ECMWF) 2 meter free-air temperature database (Dee *et al.*, 2011). This is a time series of monthly means of daily means from 1979 up till 2018 (hence covering the time period for all studied datasets except WorldClim), for which we calculated average Bio1, Bio10 and Bio11 over the whole 100 × 100 km study area (based on the original 0.75° × 0.75° resolution grid). We then re-ran the paired comparisons (see above) with the temperature off-set, i.e. the difference between the bioclimatic value (for each observation and for each dataset) and the average bioclimatic value from ERA Interim for the corresponding period, using paired t-tests to test for potential differences, e.g. differences between a) Bio1(soil temperature$_{(2016\text{-}2017)}$) – Bio1(ERA Interim$_{(2016\text{-}2017)}$) and b) Bio1(CHELSA$_{(1979\text{-}2013)}$) – Bio1(ERA Interim$_{(1979\text{-}2013)}$).

Using this off-set of temperatures from a standardized and common time series allowed to correct to some extent for differences in the temporal scope among the climatic datasets, and thus climate change and inter-annual weather variation. While this does not take into account possible decoupling of climate change between soil, surface and air temperature, it does allow to estimate the size of the temporal effect in the dataset, and thus quantify the difference between in-situ soil temperature and the other datasets more precisely.

Species distribution modelling

The regional distribution of the 50 plant species was modelled using species-specific generalized linear mixed-effect models (GLMMs) (function *glmer*, package *lme4* (Bates *et al.*, 2013), family = binomial) as a function of mean annual, summer and winter temperature, and their quadratic terms. Gradient (plant data were available from three different elevation gradients; Table 2) was used as a random intercept term in these models to account for

structural variation between gradients. This was repeated for each climate dataset (except for E-OBS, as due to the limited measured climate variation within the region, species distributions could not be modelled), resulting in a total of 350 SDMs (50 species × 7 datasets). For both MODIS LST and soil temperature, only the data from the measurement year prior to the species observations (2016-2017) were used, while the bioclimatic variables from 2015-2016 were highly correlated with those of 2016-2017 and thus excluded. The variance inflation factor (VIF, function vif, package car, Fox & Weisberg, 2011) was calculated for each of the climatic datasets to test the correlation between the different bioclimatic variables. As the VIF (a value between 0 and infinity) exceeded 5 (indicating a strong correlation) for some datasets (specifically those with long-term climatic averages), separate models including only Bio1 as explanatory variables were made, and results compared between both approaches.

The explained variance in the present distribution of the species (R^2 of the fixed effect, i.e. the marginal R^2, Nakagawa & Schielzeth, 2013) was then calculated for each model and compared across all species between the different datasets with an ANOVA and a post-hoc Tukey HSD test (R^2 ~ growth forms (factor with 4 levels), model assumptions were met). We also compared the increase in R^2 values obtained by using soil temperature versus the other climate datasets for the different growth forms (forbs, graminoids, shrubs and trees) separately.

Finally, we assessed the predictive power of the different SDMs using a leave-one-out method, each time calibrating the model with 59 data points (plots) and predicting for the remaining one. We calculated the area under the curve (AUC) of the receiver operation characteristic (ROC), using the function *performance* from the package *ROCR* (Sing *et al.*, 2005), as well as the sensitivity (presences correctly predicted as presences) and the specificity (absences correctly predicted as absences) metrics. A value of 0.5 was used to

406 binarize predictions. This was repeated for each species and for each climate dataset, and
407 differences in AUC, sensitivity and specificity between SDMs using the different climatic
408 datasets were again assessed with an ANOVA and a post-hoc Tukey HSD test. We also
409 compared the increase in AUC, sensitivity and specificity obtained by using soil temperature
410 versus the other climate datasets for the different growth forms separately. Note that this
411 predictive approach is limited for three reasons: First, the restricted dataset size likely
412 constrains the predictive power of the models. Secondly, for comparison purposes, our SDMs
413 are only calibrated using bioclimatic predictors, and thus predictive power (as estimated here
414 using AUC-values) will be relatively low. Thirdly , when using predictive modelling in small-
415 sized plots (i.e. 100 m^2 here, vs. 1 km^2 traditionally), one can expect a high accuracy in
416 correctly predicting presences as presence (i.e. if a species is observed, the model will also
417 predict its presence), yet low accuracy in predicting absences as absence (i.e. if a species is
418 absent, this could either be due to the plot falling outside its niche (correctly predicted
419 absence), or due to random absences due to the limited plot size, or microscale non-climatic
420 factors (incorrectly predicted absence)). Of course, incorrect absences can also be due to
421 observation bias, identification uncertainties and incomplete detection, further lowering
422 predictive power. We thus expect high sensitivity, yet relatively low specificity and AUC-
423 values, and encourage interpretation of these different evaluation metrics together to assess
424 the predictive power of the models (Jiménez-Valverde, 2012).

425 All analyses were performed in R (R Core Team, 2015).

426 **Results**

427 ***Direct comparison of climatic variables***

428 All three studied bioclimatic variables (Bio1 = mean annual, Bio10 = mean summer and
429 Bio11 = mean winter temperature) showed a consistent negative correlation with elevation in

almost all temperature datasets in the region, yet with large differences in lapse rate (Fig. 2). The latter ranged for mean annual temperature from around -0.6 °C per 100 m for CHELSA, downscaled CHELSA and Topoclimate, over around -0.4 °C per 100 m for WorldClim, EuroLST and MODIS LST to -0.2 °C per 100 m for soil temperature and -0.1 °C per 100 m for E-OBS. Mean annual temperatures were in both years consistently higher for the soil temperature than for all other datasets, i.e. both the long-term temperature data (WorldClim, CHELSA, downscaled CHELSA, Topoclimate and EuroLST, Fig. 3a-e) and the surface (MODIS LST, Fig. 3f) and free-air (E-OBS, Fig. 3g) temperature measurements from the same time period ($p < 0.001$ from a linear model), yet differences were larger at low than at high temperatures. Differences of 3 to 6 °C between soil temperature and all other datasets remained even after correcting for possible inter-annual and climate change effects (Table 3, Fig. S1a-f). Significant differences of up to 3 °C in mean annual temperature could also be observed between all other datasets (Table 4, Fig. S2).

Despite the higher mean annual temperature in the soil, mean summer soil temperature was in both years similar (compared to WorldClim, Topoclimate, EuroLST and E-OBS) or even lower (CHELSA, downscaled CHELSA and MODIS LST) than air and surface temperature (Fig. 3h-n). After correcting for inter-annual and climate change effects, differences between soil temperature and most other datasets (except MODIS LST) remained limited to around 1 to 1.5 °C (Table 4, Fig. S1g-l). Summer temperature recordings were highest in MODIS LST (Fig. 2n, Fig. S2i,k,l). The relationship with elevation was again the strongest for (downscaled) CHELSA (-0.6 °C per 100 m), and weakest for E-OBS and MODIS LST. Winter temperature showed the largest discrepancy between soil, free-air and surface temperatures (Fig. 3), with soil temperatures being close to 0 °C from sea level up to at least 900 m a.s.l., and as such driving the higher mean annual temperatures in the soil (Fig. 2x). Part of this variation was due having relatively warm winters with plenty of snow in the area

in the period 2015-2017, yet the difference remained as high as 4 °C to 11 °C after correcting for the temporal mismatch (Table 4, Fig. S1n-r). Surface temperatures were in addition colder than free-air temperatures (Fig. S2n-r) due to an extended frost period (Fig. S3). Temperature differences between years were relatively small, except for mean annual and mean summer surface temperatures from MODIS (Fig. 2f,n).

The above-mentioned differences along the elevation gradient, combined with additional effects from local topography, resulted in large regional differences between the different climate datasets in general (Fig. 4), and between interpolated soil temperature and the other datasets in particular (Fig. 5). The correlation analyses (Fig. 4) showed that the climate datasets were nested, with strongest relationships (across all bioclimatic variables) between the datasets with long-term averages: (downscaled) CHELSA, Topoclimate, WorldClim and EuroLST. The datasets with short-term measurements (in-situ soil, MODIS LST and free-air E-OBS) differed more from each other than from the long-term averages. Modelled mean annual temperature in the soil was, as expected, several degrees warmer than in all other datasets, especially at higher elevations (Fig. 5), while in summer soil temperature was warmer than CHELSA climate and MODIS LST at high elevations, yet colder at low elevations (Fig. 5). Winter temperature predictions were up to 17 °C higher in the soil than in the other datasets, except at the highest elevations. Due to the large local variation in snow cover, however, winter soil temperature predictions were unreliable (Fig. 5, Fig. S3), in contrast to annual and summer temperatures, for which the local R^2 (indicating the local spatial regression fit) of the regional interpolations was highly consistent across space, albeit only moderately high, i.e. on average 50% for Bio10 and 37% for Bio1.

Species distribution modelling

SDMs using soil temperatures explained on average 80% of variance (48% if only Bio1 was used), which was on average 18% (15% for models with Bio1 only) more than the models using other climate datasets (Fig. 6, significant differences with most datasets after correcting for multiple testing). Differences in explained variance among SDMs based on these other datasets were much smaller. Differences in predictive power were not significant between models (highest for Euro-LST and downscaled CHELSA (AUC ≈ 0.70), and between 0.61 and 0.64 for the other datasets (Fig. S5). As expected, sensitivity was high (≈ 0.85), yet specificity was low (≈ 0.27) for all datasets. Predictive modelling was nearly impossible with models with Bio1 only (AUC ≈ 0.5, specificity ≈ 0.20), even though sensitivity was still high (≈ 0.81).

Model performances depended strongly on growth forms (i.e. forbs, graminoids, dwarf shrubs, trees, Fig. 6b-c). We observed a significant net improvement in marginal R^2 values (as an indicator of descriptive power of the models) for SDMs based on soil temperature in the case of forbs and graminoids compared to the other datasets (on average +24% and +21% for the full model, respectively, and 20% and 25% for the model with Bio1 only), and moderately so for shrubs (full model: +8%, Bio1: +25%). Yet there was no such net increase for trees (+2% and 8% only). On the contrary, we observed a significant net decrease in predictive values for shrubs and trees when using soil temperature compared to most of the other datasets (AUC on average -0.12 and -0.11 respectively for both models; -0.06 and -0.08 for Se), yet not so for forbs and graminoids (Fig. S5b-c).

Discussion

Our comparison of different climate datasets highlights that the use of a specific source of climate data is species- and region-specific and can have strong repercussions on the outcome of SDMs, as exemplified here for the distributions of 50 plant species along steep climatic

gradients in a cold-climate region. Our data indeed revealed a strong sensitivity of SDMs to the used climate dataset depending on the growth form of the species. In general, the use of *in-situ* soil temperature instead of surface or free-air temperature did improve the explanatory power of our SDMs. It did so much more for forbs and graminoids, to a lesser degree for shrubs, yet not for trees (Fig. 6). This outcome confirms recent studies arguing for the use of more local climate variables in distribution modelling (e.g. Ashcroft *et al.*, 2008; Pradervand *et al.*, 2014; Slavich *et al.*, 2014; Opedal *et al.*, 2015; Meineri & Hylander, 2017) and proofs the validity of this concept across a whole range of possible temperature data sources. Yet, our results also indicate that an increased accuracy of climate data does not necessarily improve distribution models for all species or in all circumstances (Bennie *et al.*, 2014; Pradervand *et al.*, 2014), as it will depend on the growth forms of the species and perhaps also the regional climate characteristics. The differences in SDMs' explanatory power could result from differences in measurement focus and spatiotemporal resolution or extent, related to the different spatiotemporal framework in which different species groups operate, as discussed below.

Measurement focus

The most critical differences observed between the climate datasets in this study were likely driven by measurement focus (free-air, land surface or soil), with consistently higher average annual temperatures observed in the soil resulting to a large extent from differences in winter temperatures (Bio11). Even though free-air temperature predictions (WorldClim, CHELSA, E-OBS) for winter temperature easily dropped below -7 °C, and surface temperature measurements (EuroLST, MODIS LST) were even lower, winter temperatures just below the soil surface were close to 0 °C along most of the elevation gradient (Fig. 2). Only in those locations where global climate models predicted an average winter temperature below -10 °C, measured soil temperatures dropped below 0 °C (Fig. 2). These differences remained even

after correcting for the temporal mismatch in the different datasets (Table 4, Fig. S1). While some of the earliest studies on soil temperature reported a strong relationship with air temperature across all seasons (Shanks, 1956), it is clear that both a dense vegetation cover and a thick snowpack can provide effective insulation and protection against freezing events in the subnivium (Geiger, 1950; Dorrepaal *et al.*, 2004; Pauli *et al.*, 2013; Aalto *et al.*, 2017; Thompson *et al.*, 2018), and that snow in the Arctic is a crucial explanatory variable for the distribution of plant species (Randin *et al.*, 2009a; Niittynen & Luoto, 2017). In northern Norway, especially, the relatively mild climate and humid air from the ocean result in thick winter snow packs that can provide a significant decoupling between air, surface and soil temperature (Pauli *et al.*, 2013; Thompson *et al.*, 2018). Such an insulating snow pack can affect plant life in several ways, through its effects on overwintering survival, productivity, reproductive success and nutrient and water availability (Niittynen & Luoto, 2017), with both positive (e.g. less frost events) and negative effects (e.g. limited growing season) observed. For many species in the region, especially low-growing forbs and graminoids, we have shown that using near-surface soil temperatures instead of free-air temperatures, which allows incorporating these snow cover effects, is crucial to accurately describe the distribution of small-stature plants (Randin *et al.*, 2009a; Niittynen & Luoto, 2017). For trees, however, the absence of model improvement through the use of soil temperature might result from a stronger correlation with air than with soil temperature due to higher maximum canopy heights, at least in later life stages. In winter and early spring, trees are likely to be much more affected by air temperatures and freezing events affecting their buds above the snow than by temperatures in the soil (Körner, 2003).

These results also indicate that the relative importance of using soil temperature in SDMs will depend on the topography and large-scale climate of the region. Most importantly, the amount of fresh snow in winter will define the strength of the discrepancy between winter (and thus

552 indirectly annual) mean temperatures in the soil and in the air (Cohen, 1994; Zhang, 2005).
553 The mismatch is in our study indeed significantly larger in the warmer but snowier
554 (Norwegian) plots at low elevations than in the colder yet drier (Swedish) plots at high
555 elevations (Fig. 3). For summer temperature, our data overall showed a more consistent
556 match between the different datasets, although with minor buffering effects of the vegetation.
557 Even though the discrepancy between measurement foci is thus region-specific (and likely
558 even more different in e.g. tropical regions), we suggest that the use of climate data in close
559 proximity to the study species is always recommended. Importantly, however, the use of soil
560 temperature does not fully resolve this measurement mismatch, as only part of the plants are
561 belowground. Although our data demonstrates a significant improvement in the use of soil
562 temperature over free-air temperature data for species groups entirely covered by snow in
563 winter, an optimal approach would incorporate in-situ climate measurements both above and
564 below the soil surface. The latter can for example be achieved with the temperature and soil
565 moisture plant simulator sensors as described in Wild *et al.* (2019), measuring temperatures
566 at, above and below the surface.

567 Despite the clear benefits of using soil temperature data in SDMs, a major drawback (next to
568 the cost associated with obtaining *in-situ* soil temperature measurements) lies in the increased
569 local-scale heterogeneity, especially in winter. The soil temperatures were in our study indeed
570 hard to predict accurately using a 50 × 50 m DEM-based interpolation approach. More *in-situ*
571 temperature measurements, as well as the inclusion of other microclimate-related variables
572 like snow cover maps, might be needed to improve interpolations of microclimate at fine
573 spatial resolution. This is also a prerequisite for better SDMs' predictive performances.
574 Follow-up studies with larger datasets and in-situ measurements of more environmental
575 variables (e.g. soil moisture, air temperature, precipitation, or snow cover) are thus
576 recommended to investigate this further.

While satellite-measured land surface temperature data (MODIS LST and EuroLST) resulted in mean annual temperatures within the same range as those obtained with free-air temperature measurements, the land surface temperatures were, throughout the measurement period, significantly higher in summer and lower in winter, thus resulting in an increased overall annual temperature range (Fig. 2, Fig. S1, Table 4). These extremes were however smoothed out when using the EuroLST temperature averages over a ten-year period. While the use of satellite-based land surface temperature for SDMs has until now been largely underexplored, our study adds to the growing list of recent studies indicating the potential of these untapped data resources for accurately predicting species distributions (see e.g. Cord & Rödder, 2011; Bisrat *et al.*, 2012; Neteler *et al.*, 2013). We expect that LST-timeseries with an even higher spatial resolution, such as Landsat (Cook, 2014) will as such turn out the crucial link between local-scale temperature measurements and global climate models. Our results however indicate that smoothed, long-term averages like EuroLST are preferable above short-term measurements, especially for predictive modelling. Similar to the issue of spatial heterogeneity for *in-situ* soil temperature data, averages over long-term time series are, by nature, more likely to increase the predictive performances of SDMs compared with more erratic fluctuations based on short-term data.

Temporal - extent

Differences between the used climate datasets could also be attributed to variation in temporal extent, with the datasets building on long-term historic averages (WorldClim, CHELSA, Topoclimate and EuroLST) showing the strongest correlation with each other (Fig. 4). Correlations were however weakest for the three datasets with only two years of data, yet with different measurement foci as described above (MODIS LST, E-OBS and soil temperature). While patterns over time for these datasets were relatively consistent between measurement years (Fig. 2), they did reveal more variation between air and surface temperature than

between EuroLST and the other datasets with long-term climatic averages. The discrepancy in temporal extents might also explain why the performance of our predictive models decreased in some cases for shrubs and trees when using short-term soil (or surface) temperatures (Fig. S4 and 5). These long-lived species are indeed likely to be relatively inert towards short-term changes in their environment (Körner, 2003), which might make it harder to predict their distribution based on locally-measured short-term temperatures (Ashcroft *et al.*, 2008). Long-lived organisms like most arctic-alpine species in the study region could also persist outside their niche for considerable parts of their life (Bond & Midgley, 2001), adding to the complexity of predicting their distribution using short-term temperature data.

Spatial resolution

Our comparative approach indicates that the downscaling or interpolation of climate data – as applied here respectively to global datasets like CHELSA and the *in-situ* soil temperature data and topoclimatic dataset from Aalto *et al.* (2017) – was rather successful. Downscaling of CHELSA from 1000 × 1000 m to 30 × 30 m based on the physiography worked well, as indicated by the high local R^2-values (0.90 ± 0.06 for Bio1 and Bio10, 0.89 ± 0.06 for Bio11, Fig. S4), yet nevertheless only resulted in minor improvements of the regional SDMs compared to coarse-grained CHELSA-data (3.7% and 0.035 for the R^2 and AUC values, respectively). This lack of improvement is in disagreement with several other studies (e.g. Gillingham *et al.*, 2012; Slavich *et al.*, 2014). Part of this could be due to the inherent limitations in the original CHELSA dataset: unlike elevation, small-scale topographic variables like slope and aspect are not taken into account into the original CHELSA model, and their inclusion in the downscaling approach is thus unlikely to have major effects. Small-scale topographic effects on microclimate are more correctly taken into account in the topoclimatic dataset from Aalto *et al.* (2017), however, making the latter approach recommendable above the former. The fact that the topoclimatic dataset did not perform

627 significantly better in the SDMs than CHELSA either (ΔR^2 = -7% and +5%, and AUC = -0.01
628 and +0.06, depending on the model), might suggest again that an increased level of detail is
629 not better by default, yet depends on the context of the study (Bennie *et al.*, 2014). The most
630 likely explanation for this lack of improvement in model performance in this case thus is that
631 the distribution of the studied alpine species might be less driven by small-scale topoclimatic
632 variation in air temperature than by snow-cover induced variation in soil temperature.

633 Interpolation of the soil temperature data worked well across the whole study region, except
634 for winter temperature, where probably the strong local variation and the highly non-linear
635 correlation with elevation resulted in inaccurate predictions (Figure 4, Fig. S3, Ashcroft *et al.*,
636 2008). The large differences in winter temperatures between measurement locations – and the
637 low predictability of soil winter temperature in the region – thus suggest that caution is
638 needed, as in many regions winter temperatures are likely crucial for the distribution of
639 species (Williams *et al.*, 2015). A larger dataset and more accurate predictor variables, e.g.
640 related to snow cover duration (Niittynen & Luoto, 2017), might be needed to improve these
641 interpolation efforts.

642 ***Implications***

643 The observed differences in the climate datasets and SDMs at the regional scale advocate for
644 a careful selection of the climate data source when modelling species distributions, based on a
645 priori ecological assumptions about the relationship of the studied organism with the regional
646 environment, and the comparison – or joint use – of different datasets (Buermann *et al.*, 2008;
647 Rebaudo *et al.*, 2016). Measurement focus, temporal extent and spatiotemporal resolution
648 should all be taken into account with regard to the studied species and area: is the species
649 affected by snow cover; is it an annual or a perennial species; is the focal species mobile or
650 sessile; does the study area reach above the treeline; is it in topographically challenging

terrain, etc. Our study highlights the importance of growth forms: soil temperature was highly important for forbs and graminoids, and to a certain extent for shrubs, yet not so for trees. Only when making ecologically meaningful a priori decisions, and when comparing the performance of different datasets – and perhaps their interactions, one can be sure that the observed trends relate to the actual (micro)climate experienced by the study species or species group(s) in the study region. Understanding these processes in the current climate is a crucial step before model projections can be improved under climate change as well. In order to advance towards this goal, there is an urgent need for large-scale datasets of microclimate data; ecologists and climatologists should consider in-depth on-the-ground, long-term microclimate monitoring along climatic gradients to be able to improve our microclimatic models for use in SDMs (Lembrechts *et al.*, 2018). Nevertheless, our case study suggests that SDMs can be relatively robust to several characteristics of different types of climate datasets, like spatial and temporal resolution, especially in the relatively stable slow-reacting vegetation types of high-latitudinal mountains. Additionally, there is a need to improve our abilities to forecast microclimate data itself in the future, as climate change is likely to affect soil, surface and air temperatures differently (Ashcroft & Gollan, 2013; De Frenne *et al.*, 2019). Significant progress has been made in this regard, for example by integrating microclimatic dynamics and processes like microclimatic buffering in predictions (Keppel *et al.*, 2015; Lenoir *et al.*, 2017; Wason *et al.*, 2017), yet there is still a need for improvement before the same diversity and quality of climate datasets will be available for SDM projections into future climate as we have now for current climate.

References

Aalto, J., Riihimäki, H., Meineri, E., Hylander, K. & Luoto, M. (2017) Revealing topoclimatic heterogeneity using meteorological station data. *International Journal of Climatology*, **37**, 544-556.

Aalto, J., Scherrer, D., Lenoir, J., Guisan, A. & Luoto, M. (2018) Biogeophysical controls on soil-atmosphere thermal differences: implications on warming Arctic ecosystems. *Environmental Research Letters*, **13**, 074003.

Ashcroft, M.B. (2010) Identifying refugia from climate change. *Journal of Biogeography*, **37**, 1407-1413.

Ashcroft, M.B. & Gollan, J.R. (2012) Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200 x 300 km) and diverse region. *International Journal of Climatology*, **32**, 2134-2148.

Ashcroft, M.B. & Gollan, J.R. (2013) Moisture, thermal inertia, and the spatial distributions of near-surface soil and air temperatures: Understanding factors that promote microrefugia. *Agricultural and Forest Meteorology*, **176**, 77-89.

Ashcroft, M.B., Chisholm, L.A. & French, K.O. (2008) The effect of exposure on landscape scale soil surface temperatures and species distribution models. *Landscape Ecology*, **23**, 211-225.

Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1-8.

Baddeley, A., Rubak, E. & Turner, R. (2015) *Spatial Point Patterns: Methodology and Applications with R.*

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2013) lme4: linear mixed-effects models using Eigen and S4. In, R package version 1.0-5.

Bennie, J., Wilson, R.J., Maclean, I.M.D. & Suggitt, A.J. (2014) Seeing the woods for the trees - when is microclimate important in species distribution models? *Global Change Biology*, **20**, 2699-2700.

Bisrat, S.A., White, M.A., Beard, K.H. & Richard Cutler, D. (2012) Predicting the distribution potential of an invasive frog using remotely sensed data in Hawaii. *Diversity and Distributions*, **18**, 648-660.

Bond, W.J. & Midgley, J.J. (2001) Ecology of sprouting in woody plants: the persistence niche. *Trends in ecology & evolution*, **16**, 45-51.

Bramer, I., Anderson, B., Bennie, J., Bladon, A., De Frenne, P., Hemming, D., Hill, R.A., Kearney, M.R., Körner, C., Korstjens, A.H., Lenoir, J., Maclean, I.M.D., Marsh, C.D., Morecroft, M.D., Ohlemüller, R., Slater, H.D., Suggitt, A.J., Zellweger, F. & Gillingham, P.K. (2018) Advances in monitoring and modelling climate at ecologically relevant scales. *Advances in Ecological Research*,

Buermann, W., Saatchi, S., Smith, T.B., Zutta, B.R., Chaves, J.A., Milá, B. & Graham, C.H. (2008) Predicting species distributions across the Amazonian and Andean regions using remote sensing data. *Journal of Biogeography*, **35**, 1160-1176.

Cohen, J. (1994) Snow cover and climate. *Weather*, **49**, 150-156.

Cook, M.J. (2014) Atmospheric Compensation for a Landsat Land Surface Temperature Product.

Cord, A. & Rödder, D. (2011) Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? *Ecological Applications*, **21**, 3285-3298.

De Frenne, P., Zellweger, F., Rodríguez-Sánchez, F., Scheffers, B.R., Hylander, K., Luoto, M., Vellend, M., Verheyen, K. & Lenoir, J. (2019) Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution*, 1.

Dee, D.P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G. & Bauer, d.P. (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, **137**, 553-597.

Distler, T., Schuetz, J.G., Velásquez-Tibatá, J. & Langham, G.M. (2015) Stacked species distribution models and macroecological models provide congruent projections of avian species richness under climate change. *Journal of Biogeography*, **42**, 976-988.

Dobrowski, S.Z. (2011) A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology*, **17**, 1022-1035.

727 Dorrepaal, E., Aerts, R., Cornelissen, J.H., Callaghan, T.V. & Van Logtestijn, R.S. (2004) Summer
728 warming and increased winter snow cover affect Sphagnum fuscum growth, structure and
729 production in a sub-arctic bog. *Global Change Biology*, **10**, 93-104.
730 Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: ecological explanation and prediction
731 across space and time. *Annual Review of Ecology Evolution and Systematics*, pp. 677-697.
732 Fick, S.E. & Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global
733 land areas. *International Journal of Climatology*, **37**, 4302-4315.
734 Fotheringham, A., Brunsdon, C. & Charlton, M. (2003) *Geographically weighted regression: the*
735 *analysis of spatially varying relationships*. John Wiley & Sons, Hoboken, USA.
736 Fox, J. & Weisberg, S. (2011) *An {R} Companion to Applied Regression, Second Edition*.
737 Geiger, R. (1950) *The climate near the ground*. Harvard University Press, Cambridge, Massachusets,
738 USA.
739 Gillingham, P.K., Huntley, B., Kunin, W.E. & Thomas, C.D. (2012) The effect of spatial resolution on
740 projected responses to climate warming. *Diversity and Distributions*, **18**, 990-1000.
741 Gonzalez-Moreno, P., Diez, J.M., Richardson, D.M. & Vila, M. (2015) Beyond climate: disturbance
742 niche shifts in invasive species. *Global Ecology and Biogeography*, **24**, 360-370.
743 Gottfried, M., Pauli, H., Reiter, K. & Grabherr, G. (1999) A fine-scaled predictive model for changes in
744 species distribution patterns of high mountain plants induced by climate warming. *Diversity*
745 *and Distributions*, **5**, 241-251.
746 Graae, B.J., De Frenne, P., Kolb, A., Brunet, J., Chabrerie, O., Verheyen, K., Pepin, N., Heinken, T.,
747 Zobel, M., Shevtsova, A., Nijs, I. & Milbau, A. (2012) On the use of weather data in ecological
748 studies along altitudinal and latitudinal gradients. *Oikos*, **121**, 3-19.
749 Graae, B.J., Vandvik, V., Armbruster, W.S., Eiserhardt, W.L., Svenning, J.-C., Hylander, K., Ehrlén, J.,
750 Speed, J.D., Klanderud, K., Bråthen, K.A., Milbau, A., Opedal, O.H., Alsos, I.G., Ejrnaes, R.,
751 Bruun, H.H., Birks, H.J.B., Westergaard, K.B., Birks, H.H. & Lenoir, J. (2018) Stay or go–how
752 topographic complexity influences alpine plant population and community responses to
753 climate change. *Perspectives in Plant Ecology, Evolution and Systematics*, **30**, 41-50.
754 Greiser, C., Meineri, E., Luoto, M., Ehrlén, J. & Hylander, K. (2018) Monthly microclimate models in a
755 managed boreal forest landscape. *Agricultural and Forest Meteorology*, **250**, 147-158.
756 Guisan, A. & Thuiller, W. (2007) Predicting species distribution: offering more than simple habitat
757 models (vol 8, pg 993, 2005). *Ecology Letters*, **10**, 435-435.
758 Hannah, L., Flint, L., Syphard, A.D., Moritz, M.A., Buckley, L.B. & McCullough, I.M. (2014) Fine-grain
759 modeling of species' response to climate change: holdouts, stepping-stones, and
760 microrefugia. *Trends in Ecology & Evolution*, **29**, 390-397.
761 Haylock, M., Hofstra, N., Klein Tank, A., Klok, E., Jones, P. & New, M. (2008) A European daily high-
762 resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal*
763 *of Geophysical Research: Atmospheres*, **113**
764 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution
765 interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**,
766 1965-1978.
767 Holden, Z.A., Abatzoglou, J.T., Luce, C.H. & Baggett, L.S. (2011) Empirical downscaling of daily
768 minimum air temperature at very fine resolutions in complex terrain. *Agricultural and Forest*
769 *Meteorology*, **151**, 1066-1073.
770 Illan, J.G., Gutierrez, D. & Wilson, R.J. (2010) The contributions of topoclimate and land cover to
771 species distributions and abundance: fine-resolution tests for a mountain butterfly fauna.
772 *Global Ecology and Biogeography*, **19**, 159-173.
773 Jiménez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. & Lobo, J.M. (2011) Use of
774 niche models in invasive species risk assessments. *Biological Invasions*, **13**, 2785-2797.
775 Jiménez-Valverde, A. (2012) Insights into the area under the receiver operating characteristic curve
776 (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and*
777 *Biogeography*, **21**, 498-507.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. & Kessler, M. (2017) Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, **4**, 170122.

Keppel, G., Mokany, K., Wardell-Johnson, G.W., Phillips, B.L., Welbergen, J.A. & Reside, A.E. (2015) The capacity of refugia for conservation planning under climate change. *Frontiers in Ecology and the Environment*, **13**, 106-112.

Körner, C. (2003) *Alpine plant life: functional plant ecology of high mountain ecosystems*. Springer, Berlin Heidelberg, Germany.

Körner, C. & Hiltbrunner, E. (2018) The 90 ways to describe plant temperature. *Perspectives in Plant Ecology, Evolution and Systematics*, **30**, 16-21.

Lembrechts, J., Nijs, I. & Lenoir, J. (2018) Incorporating microclimate into species distribution models. *Ecography*,

Lembrechts, J.J., Milbau, A. & Nijs, I. (2014) Alien roadside species more easily invade alpine than lowland plant communities in a subarctic mountain ecosystem. *PLoS One*, **9**, e89664.

Lembrechts, J.J., Lenoir, J., Nuñez, M.A., Pauchard, A., Geron, C., Bussé, G., Milbau, A. & Nijs, I. (2017) Microclimate variability in alpine ecosystems as stepping stones for non-native plant establishment above their current elevational limit. *Ecography*, **40**, 001-009.

Lembrechts, J.J., Pauchard, A., Lenoir, J., Nuñez, M.A., Geron, C., Ven, A., Bravo-Monasterio, P., Teneb, E., Nijs, I. & Milbau, A. (2016) Disturbance is the key to plant invasions in cold environments. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 14061-14066.

Lenoir, J., Hattab, T. & Pierre, G. (2017) Climatic microrefugia under anthropogenic climate change: implications for species redistribution. *Ecography*, **40**, 253-266.

Lenoir, J., Graae, B.J., Aarrestad, P.A., Alsos, I.G., Armbruster, W.S., Austrheim, G., Bergendorff, C., Birks, H.J.B., Brathen, K.A., Brunet, J., Bruun, H.H., Dahlberg, C.J., Decocq, G., Diekmann, M., Dynesius, M., Ejrnaes, R., Grytnes, J.A., Hylander, K., Klanderud, K., Luoto, M., Milbau, A., Moora, M., Nygaard, B., Odland, A., Ravolainen, V.T., Reinhardt, S., Sandvik, S.M., Schei, F.H., Speed, J.D.M., Tveraabak, L.U., Vandvik, V., Velle, L.G., Virtanen, R., Zobel, M. & Svenning, J.C. (2013) Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across Northern Europe. *Global Change Biology*, **19**, 1470-1481.

Maclean, I.M.D., Suggitt, A.J., Wilson, R.J., Duffy, J.P. & Bennie, J.J. (2017) Fine-scale climate change: modelling spatial variation in biologically meaningful rates of warming. *Global Change Biology*, **23**, 256-268.

McCullough, I.M., Davis, F.W., Dingman, J.R., Flint, L.E., Flint, A.L., Serra-Diaz, J.M., Syphard, A.D., Moritz, M.A., Hannah, L. & Franklin, J. (2016) High and dry: high elevations disproportionately exposed to regional climate change in Mediterranean-climate landscapes. *Landscape ecology*, **31**, 1063-1075.

Meineri, E. & Hylander, K. (2017) Fine-grain, large-domain climate models based on climate station and comprehensive topographic information improve microrefugia detection. *Ecography*, **40**, 1003-1013.

Metz, M., Rocchini, D. & Neteler, M. (2014) Surface temperatures at the continental scale: tracking changes with remote sensing at unprecedented detail. *Remote Sensing*, **6**, 3822.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133-142.

Neteler, M., Rocchini, D., Delucchi, L. & Metz, M. (2014) Massive data processing in GRASS GIS 7: A new gap-filled MODIS Land Surface Temperature time series data set. *FOSS4G-Europe 2014* (ed by.

Neteler, M., Metz, M., Rocchini, D., Rizzoli, A., Flacio, E., Engeler, L., Guidi, V., Lüthy, P. & Tonolla, M. (2013) Is Switzerland suitable for the invasion of Aedes albopictus? *PLoS One*, **8**, e82090.

Niittynen, P. & Luoto, M. (2017) The importance of snow in species distribution models of arctic vegetation. *Ecography*,

Opedal, O.H., Armbruster, W.S. & Graae, B.J. (2015) Linking small-scale topography with microclimate, plant species diversity and intra-specific trait variation in an alpine landscape. *Plant Ecology & Diversity*, **8**, 305-315.
Pauchard, A., Kueffer, C., Dietz, H., Daehler, C.C., Alexander, J., Edwards, P.J., Arévalo, J.R., Cavieres, L.A., Guisan, A., Haider, S., Jakobs, G., McDougall, K., Millar, C.I., Naylor, B.J., Parks, C.G., Rew, L.J. & Seipel, T. (2009) Ain't no mountain high enough: plant invasions reaching new elevations. *Frontiers in Ecology and the Environment*, **7**, 479-486.
Pauli, J.N., Zuckerberg, B., Whiteman, J.P. & Porter, W. (2013) The subnivium: a deteriorating seasonal refugium. *Frontiers in Ecology and the Environment*, **11**, 260-267.
Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W.H., Kleyer, M., Schurr, U. & Postma, J. (2016) Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist*, **212**, 838-855.
Potter, K.A., Woods, H.A. & Pincebourde, S. (2013) Microclimatic challenges in global change biology. *Global Change Biology*, **19**, 2932-2939.
Pottier, J., Malenovský, Z., Psomas, A., Homolová, L., Schaepman, M.E., Choler, P., Thuiller, W., Guisan, A. & Zimmermann, N.E. (2014) Modelling plant species distribution in alpine grasslands using airborne imaging spectroscopy. *Biology letters*, **10**, 20140347.
Pradervand, J.-N., Dubuis, A., Pellissier, L., Guisan, A. & Randin, C. (2014) Very high resolution environmental predictors in species distribution models: Moving beyond topography? *Progress in Physical Geography*, **38**, 79-96.
R Core Team (2015) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing.
Randin, C.F., Vuissoz, G., Liston, G.E., Vittoz, P. & Guisan, A. (2009a) Introduction of snow and geomorphic disturbance variables into predictive models of alpine plant distribution in the Western Swiss Alps. *Arctic, Antarctic, and Alpine Research*, **41**, 347-361.
Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B., Vittoz, P., Thuiller, W. & Guisan, A. (2009b) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557-1569.
Rebaudo, F., Faye, E. & Dangles, O. (2016) Microclimate data improve predictions of insect abundance models based on calibrated spatiotemporal temperatures. *Frontiers in Physiology*, **7**, 139.
Scherrer, D. & Körner, C. (2011) Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography*, **38**, 406-416.
Sears, M.W., Raskin, E. & Angilletta, M.J. (2011) The world is not flat: defining relevant thermal landscapes in the context of climate change. *Integrative and Comparative Biology*, **51**, 666-675.
Shanks, R.E. (1956) Altitudinal and microclimatic relationships of soil temperature under natural vegetation. *Ecology*, **37**, 1-7.
Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. (2005) ROCR: visualizing classifier performance in R. *Bioinformatics*, **21**, 7781.
Slavich, E., Warton, D.I., Ashcroft, M.B., Gollan, J.R. & Ramp, D. (2014) Topoclimate versus macroclimate: how does climate mapping methodology affect species distribution models and climate change projections? *Diversity and Distributions*, **20**, 952-963.
Stewart, L., Simonsen, C.E., Svenning, J.C., Schmidt, N.M. & Pellissier, L. (2018) Forecasted homogenization of high Arctic vegetation communities under climate change. *Journal of biogeography*, **45**, 2576-2587.
Su, Y.F., Foody, G.M. & Cheng, K.S. (2012) Spatial non-stationarity in the relationships between land cover and surface temperature in an urban heat island and its impacts on thermally sensitive populations. *Landscape and Urban Planning*, **107**, 172-180.
Thompson, K.L., Zuckerberg, B., Porter, W.P. & Pauli, J.N. (2018) The phenology of the subnivium. *Environmental Research Letters*, **13**, 064037.

Wan, Z., Hook, S. & Hulley, G. (2015) MOD11C2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 0.05Deg CMG V006 [Data set]. In, NASA EOSDIS LP DAAC.
Wan, Z.M. (2008) New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sensing of Environment*, **112**, 59-74.
Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.
Wason, J.W., Bevilacqua, E. & Dovciak, M. (2017) Climates on the move: Implications of climate warming for species distributions in mountains of the northeastern United States. *Agricultural and Forest Meteorology*, **246**, 272-280.
Wild, J., Kopecký, M., Macek, M., Šanda, M., Jankovec, J. & Haase, T. (2019) Climate at ecologically relevant scales: A new temperature and soil moisture logger for long-term microclimate measurement. *Agricultural and Forest Meteorology*, **268**, 40-47.
Williams, C.M., Henry, H.A.L. & Sinclair, B.J. (2015) Cold truths: how winter drives responses of terrestrial organisms to climate change. *Biological Reviews*, **90**, 214-235.
Willis, K.J. & Bhagwat, S.A. (2009) Biodiversity and climate change. *Science*, **326**, 806-807.
Wood, S. (2006) *Generalized Additive Models: An introduction with R*. Chapman and Hall/CRC.
Wundram, D., Pape, R. & Loffler, J. (2010) Alpine soil temperature variability at multiple scales. *Arctic Antarctic and Alpine Research*, **42**, 117-128.
Yannic, G., Pellissier, L., Le Corre, M., Dussault, C., Bernatchez, L. & Côté, S.D. (2014) Temporally dynamic habitat suitability predicts genetic relatedness among caribou. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20140502.
Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D. & Coomes, D. (2019) Advances in microclimate ecology arising from remote sensing. *Trends in ecology & evolution*,
Zhang, T. (2005) Influence of the seasonal snow cover on the ground thermal regime: An overview. *Reviews of Geophysics*, **43**
Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York, USA.

Data Accessibility Statement

Most used climate datasets are freely available (see Methods section). In-situ soil temperature and species distribution data will be published in an open access data repository.

Tables

Table 1: The eight studied climate datasets *and their geographical and temporal extent, spatial resolution and measurement focus.*

Dataset	**Initial source**	**Geographical extent**	**Spatial resolution**	**Measurement focus**	**Temporal coverage**
a) WorldClim	WorldClim	Global	30''	Free-air	1970-2000
b) CHELSA	CHELSA	Global	30''	Free-air	1979-2013
c) Downscaled	CHELSA	10000 km²	1''	Free-air	1979-2013
d) Topoclimate	Aalto et al. (2018)	10000 km²	1''	Free-air	1981-2010
e) MODIS LST	MODIS	Global	30''	Surface	2015-2017
f) EuroLST	MODIS	Europe	~7.5"	Surface	2001-2011
g) E-OBS	E-OBS	Europe	0.1°	Free-air	2015-2017
h) Soil temperature	iButtons	10000 km²	1''	Soil	2015-2017

914 ***Table 2: Overview of* in-situ *soil temperature measurement plots in Sweden and Norway (n=106).*** *For each gradient (numbers from 1) to 4)*
915 *refer to the map in Fig. 1), we present the number of elevation gradients (i.e. different mountains monitored), sites and plots (with more plots*
916 *than sites indicating repeated temperature measurements in a < 20 × 20 m area), as well as the temporal extent, the length of the elevation*
917 *gradient, and if species data is available to run species distribution models (SDMs).*

Region	# of gradients	Sites	Plots	Surface area	Temporal extent	Elevation (m a.s.l.)	Species data
1) Norway	3	59	59	2 × 100 m	01/08/15-31/07/17	0-700	Yes
2) Sweden	2	4	23	0.6 × 1.2 m	01/08/15-31/07/16	900-1100	No
3) Sweden	2	6	11	0.6 × 1.2 m	01/08/16-31/07/17	400-900	No
4) Sweden	2	13	13	2 × 10 m	01/08/16-31/07/17	400-1200	No

918

Table 3: Differences in average temperature between the climatic datasets. *Two-by-two comparisons between the three studied bioclimatic variables (Bio1 = mean annual temperature, Bio10 = mean temperature of the warmest quarter, Bio11 = mean temperature of the coldest quarter) for the different climatic datasets (except WorldClim) after correcting for inter-annual and climate change effects using ERA Interim (see methods for details). Analysis based on data from all 106 measurement locations, for MODIS LST, E-OBS and in-situ soil temperature, only the data from 2016-2017 is tested. Values show the differences in average temperature in °C between the two datasets, with positive values indicating higher temperatures in the variable in the column than in the row. Values in bold are significant at $p<0.05$ from paired t-tests. Relationships with in-situ soil temperature are visualised in Fig. S1, while some relationships among the other variables are visualised in Fig. S2.*

	CHELSA down	*Topo-climate*	*EuroLST*	*MODIS LST*	*E-OBS*	*In-situ soil*
Bio1						
CHELSA	*-0.03*	***-0.36***	***-3.19***	***-1.96***	***-1.11***	***2.67***
CHELSA down	-	***-0.33***	***-3.16***	***-1.92***	***-1.08***	***2.68***
Topoclimate	-	-	***-2.84***	***-1.59***	***-0.75***	***3.00***
EuroLST	-	-	-	***1.22***	***2.08***	***5.77***
MODIS LST	-	-	-	-	***0.91***	***4.53***
E-OBS	-	-	-	-	-	***3.53***
Bio10						
CHELSA	*-0.03*	***-2.86***	***-3.28***	***1.45***	***-2.85***	***-1.48***
CHELSA down	-	***-2.83***	***-3.25***	***1.49***	***-2.81***	***-1.48***
Topoclimate	-	-	***-0.42***	***4.30***	*0.01*	***1.24***
EuroLST	-	-	-	***4.70***	***0.43***	***1.67***
MODIS LST	-	-	-	-	***-4.23***	***-3.15***
E-OBS	-	-	-	-	-	***1.12***
Bio11						
CHELSA	*-0.03*	***2.60***	***-2.47***	***-4.82***	*0.02*	***6.30***
CHELSA down	-	***2.63***	***-2.44***	***-4.78***	*0.05*	***6.29***
Topoclimate	-	-	***-5.07***	***-7.39***	***-2.58***	***3.74***
EuroLST	-	-	-	***-2.35***	***2.49***	***8.72***
MODIS LST	-	-	-	-	***4.89***	***10.99***
E-OBS	-	-	-	-	-	***6.06***

Figures

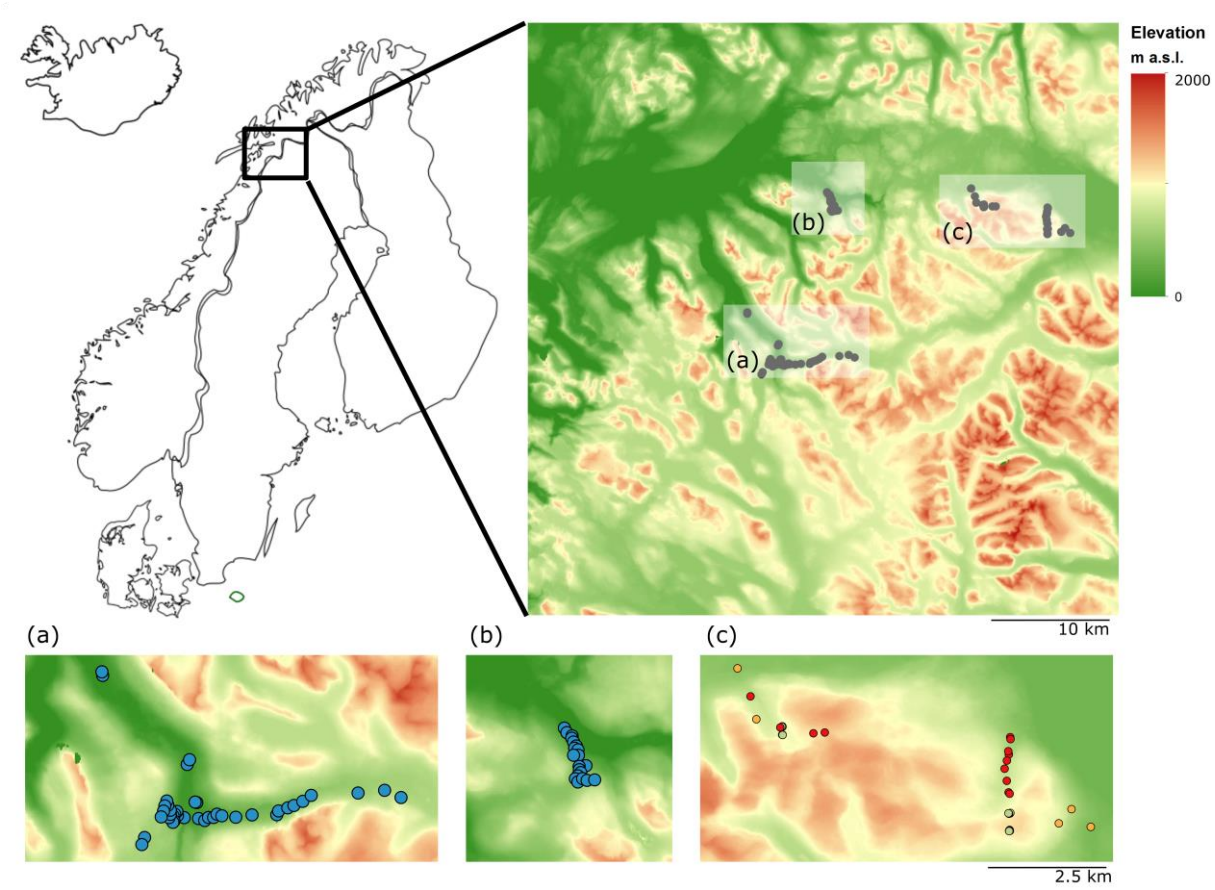
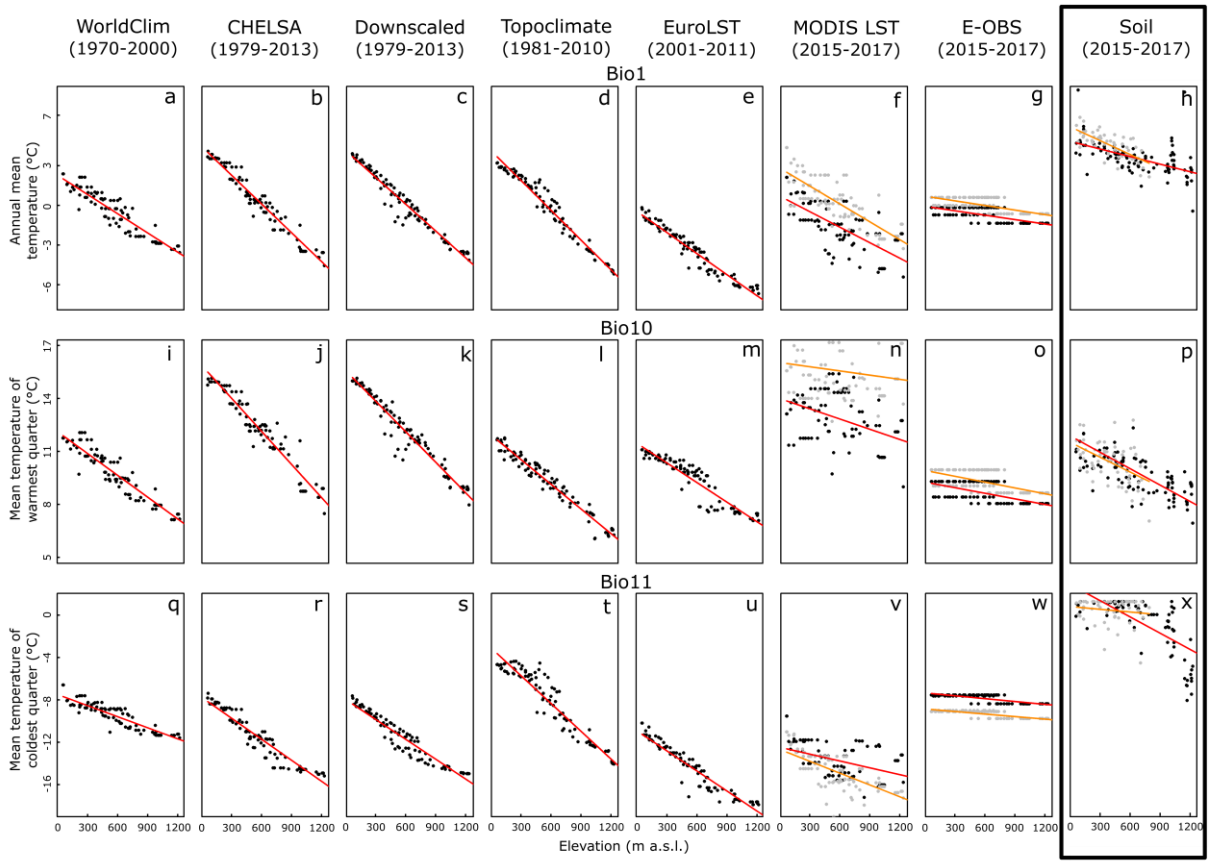


Figure 1: Study area and measurement locations. *Location of the study area in Scandinavia (left) and digital elevation model (DEM) at 1 arc-second resolution (ca. 30 x 30 m at the equator) across the study area (right). Dots on the DEM show locations of the 106 soil temperature measurements. Species data sampling was done in the locations marked with blue dots (a and b). See Table 2 for datasets: blue = 1), orange = 2), green = 3), red = 4). Elevational gradients ranging from 0 till 700 m a.s.l. (a and b) and from 400 to 1200 m a.s.l. (c).*



940

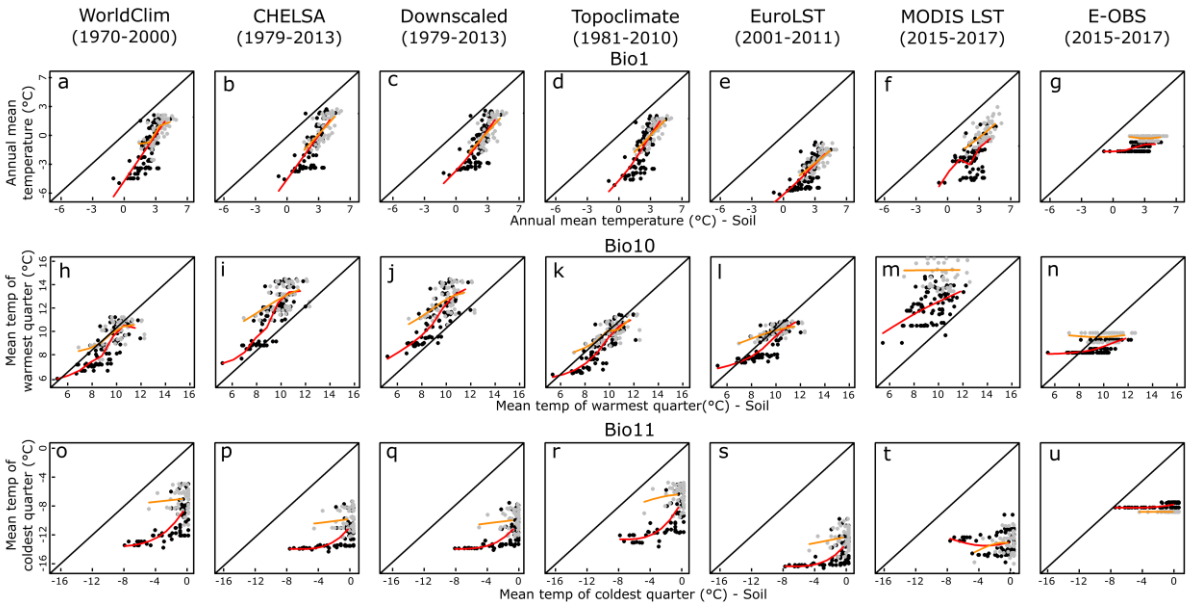
941 ***Figure 2: Temperature patterns against elevation for the different temperature data sets.***

942 *Average annual (Bio1, a-h), summer (Bio10, i-p) and winter (Bio11, q-x) temperature for the*

943 *eight climate datasets (columns, temporal extent between brackets) against elevation of the*

944 *106 measurement locations. Orange (2015-2016) and red (2016-2017) lines are fitted with*

945 *linear models.*



946

947 ***Figure 3: Plot-by-plot comparisons of soil temperature data against 7 other sources of***

948 ***temperature data.*** *Mean annual (Bio1, a-g), summer (Bio10, h-n) and winter (Bio11, o-u)*

949 *temperature, for all 106 measurement locations for 2015-2016 (orange lines, grey dots) and*

950 *2016-2017 (red lines, black dots). Black lines show first bisectors (a hypothetical perfect*

951 *match), red and orange lines are fitted with generalised additive models for each year of*

952 *temperature measurements separately. Measurement periods between brackets.*

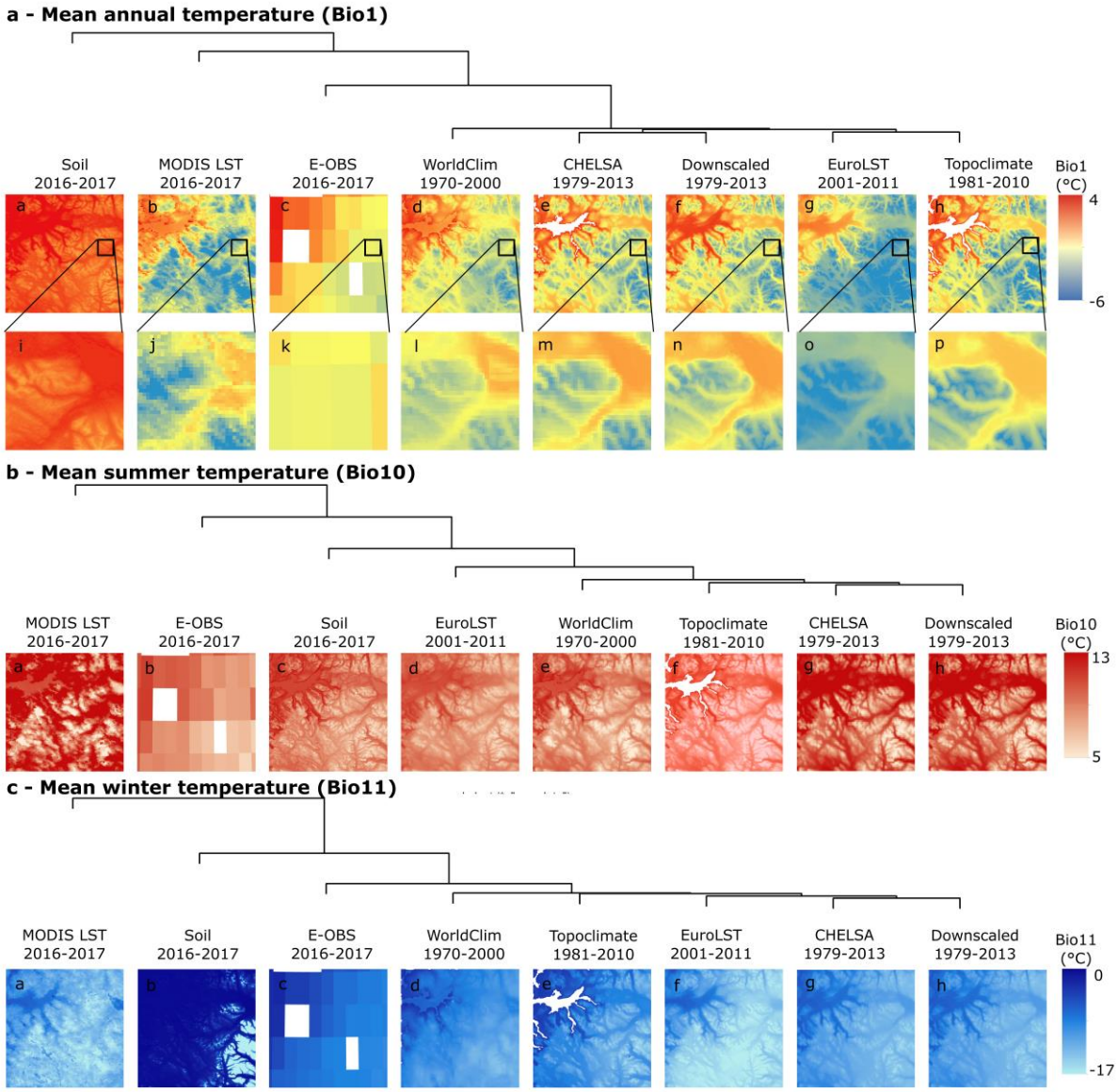


Figure 4: Dendrograms of collinearity between different temperature datasets. *Data from the 106 measurement locations for mean annual (a - Bio1), summer (b – Bio10) and winter (c – Bio11) temperature. Measurement periods between brackets. Maps show the regional (100 × 100 km) predictions for each dataset and bioclimatic variable. For Bio1, cut-outs of the maps are shown (location specified by black squares).*

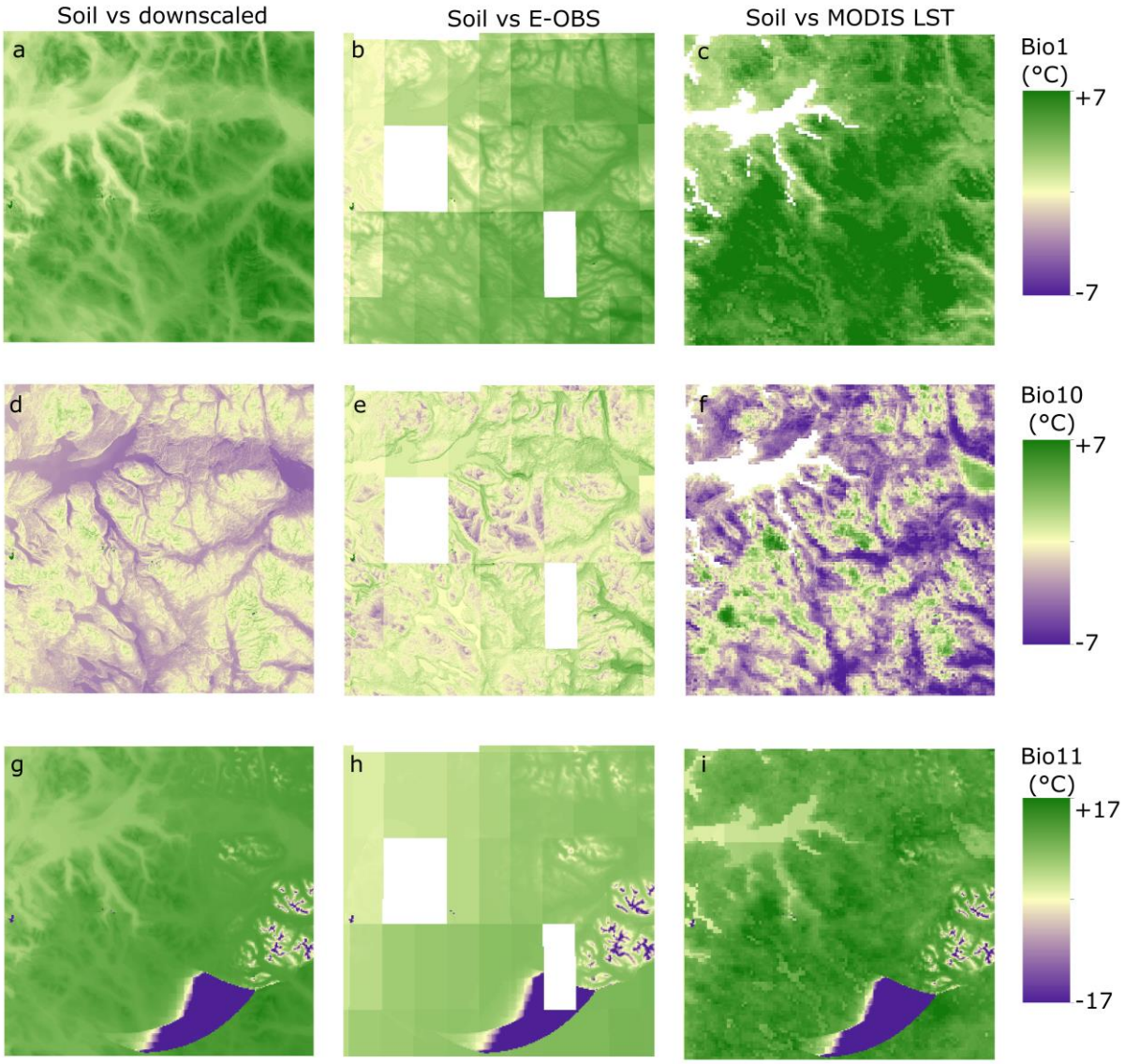


Figure 5: Differences (in °C) between regionally modelled soil temperature and other temperature data sources. *Differences in annual average temperature (Bio1), mean temperature of the warmest quarter (Bio10) and mean temperature of the coldest quarter (Bio11) are shown for soil temperature versus downscaled CHELSA (left), E-OBS (middle) and MODIS LST (right). Comparisons between soil temperature and CHELSA, WorldClim and EuroLST are not shown, as trends were similar. Values below zero indicate a lower value for the soil temperature compared with the other dataset; values above zero a higher value.*

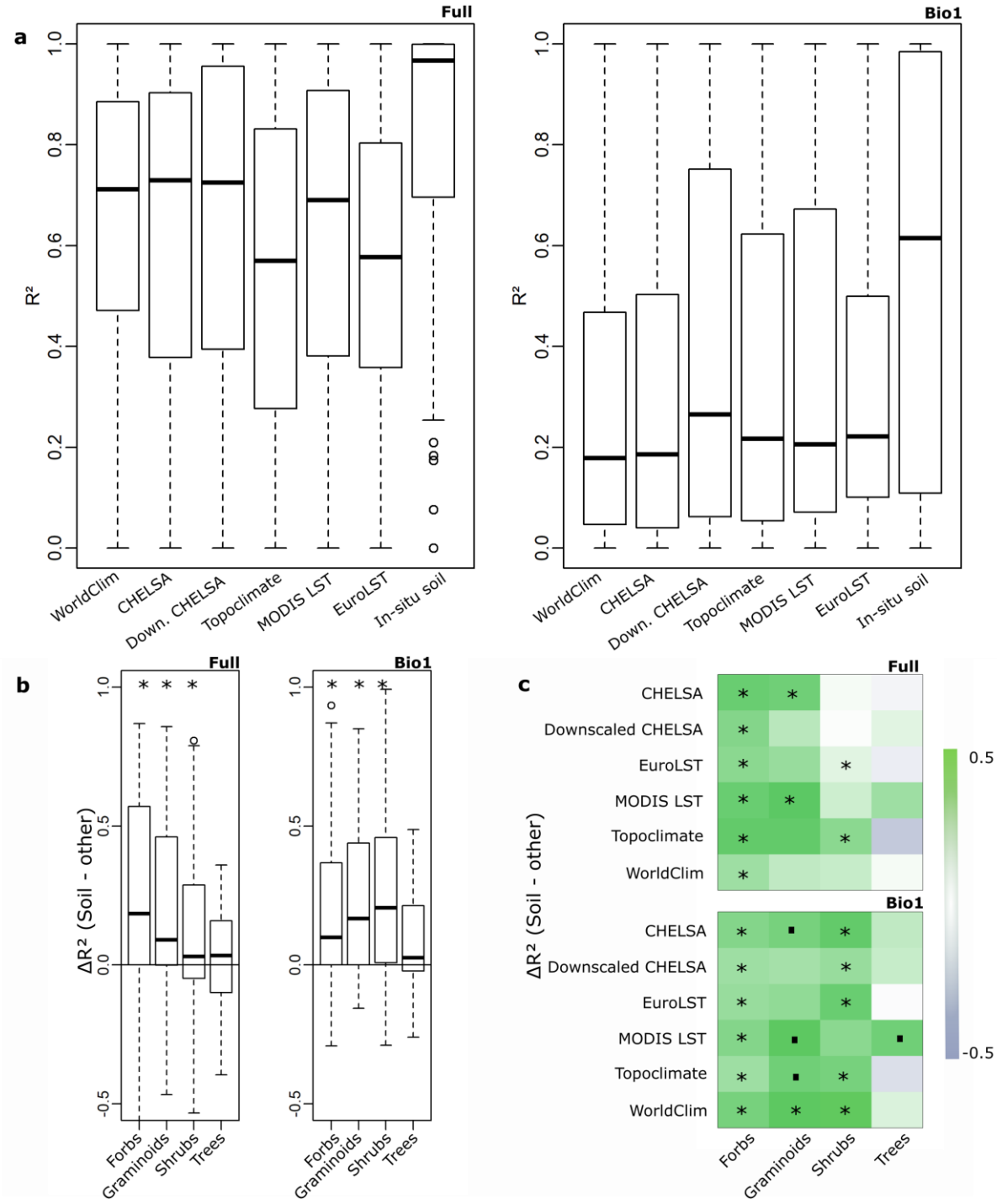


Figure 6: Proportion of explained variance (marginal R²) by species distribution models (SDMs) using the different temperature datasets. *(a) Boxplots of the marginal R² of distribution models for 50 plant species in a subset of 59 plots, based on binomial GLMMs built with the different temperature datasets: using Bio1, 10 and 11 together (left, 'Full') or Bio1 only (right, 'Bio1). (b) Differences in marginal R² between the models using soil*

temperature and all other datasets for forbs (N = 25), graminoids (N = 7), (dwarf) shrubs (N = 15) and trees (N = 3). (c) Heatmaps visualising the differences in marginal R^2 between the models using soil temperature and each of the other climatic datasets for the different growth forms. Green (positive values) indicates better performance of soil temperature models, blue a better performance of the other dataset in question. "" and "." respectively indicate significant ($p < 0.05$) and marginally significant ($0.05 < p < 0.1$) differences from zero as obtained with a two-sided t-test.*