Uncovering the hidden niche: incorporating microclimate temperature into species distribution models

Stef Haesen[1,2], Jonathan Lenoir[3], Eva Gril[3], Pieter De Frenne[4], Jonas J. Lembrechts[5], Martin Kopecký[6,7], Martin Macek[6], Matěj Man[6,8], Jan Wild[6,9], Koenraad Van Meerbeek[1,2]

**Corresponding author, OrcID = https://orcid.org/0000-0002-4491-4213, stef.haesen@kuleuven.be, +32 16 32 24 67*

[1]Department of Earth and Environmental Sciences, Celestijnenlaan 200E, 3001 Leuven, Belgium; [2]KU Leuven Plant Institute, KU Leuven, Leuven, Belgium; [3]UMR CNRS 7058 "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN), Université de Picardie Jules Verne, Amiens, France; [4]Forest & Nature Lab, Department of Environment, Ghent University, Geraardsbergsesteenweg 267, 9090 Melle-Gontrode, Belgium; [5]Research Group PLECO (Plants and Ecosystems), University of Antwerp, 2610 Wilrijk, Belgium; [6]Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-25243, Průhonice, Czech Republic; [7]Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, CZ-165 21, Prague 6 - Suchdol, Czech Republic; [8] Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-128 01 Prague 2, Czech Republic; [9] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, CZ, 165 21 Prague 6 - Suchdol, Czech Republic;

OrcIDs

Jonathan Lenoir: https://orcid.org/0000-0003-0638-9582
Eva Gril: https://orcid.org/0000-0002-7340-8264
Pieter De Frenne: https://orcid.org/0000-0002-8613-0943
Jonas J. Lembrechts: https://orcid.org/0000-0002-1933-0750
Martin Kopecký: https://orcid.org/0000-0002-1018-9316
Martin Macek: https://orcid.org/0000-0002-5609-5921
Matěj Man: https://orcid.org/0000-0002-4557-8768
Jan Wild: https://orcid.org/0000-0003-3007-4070
Koenraad Van Meerbeek: https://orcid.org/0000-0002-9260-3815

ABSTRACT

Species' environmental niches are conventionally modelled using coarse-grained macroclimate data. These data are known to deviate substantially from local, near-ground and proximal conditions (i.e., the microclimate), especially so below forest canopies. Here, we aimed to assess the impact of using gridded microclimate data instead of gridded macroclimate data on the performance of species distribution models (SDMs), as well as on the predicted geographical distribution and the derived species response curves of 140 forest specialist plant species across Europe over the 2000-2020 period. We performed a comparative study between SDMs constructed with different sets of bioclimatic predictors to separately test the effect of using (i) proximal climate data instead of conventional macroclimatic data and (ii) high-resolution proximal climate data rather than coarse-gridded macroclimatic data. Therefore, we challenged SDMs with: (1) a macroclimatic dataset at a spatial resolution of 1 km × 1 km; (2) an aggregated microclimatic dataset matching the same resolution of 1 km × 1 km; and (3) a microclimatic dataset at a much finer spatial resolution of 25 m × 25 m. We found significant differences in model performance, indicating that microclimate-based SDMs outperform both their macroclimatic and aggregated counterparts. Most importantly, this study makes clear that macroclimate-based SDMs tend to introduce a systematic bias into the perceived species response curves. Additionally, macroclimatic data is unable to identify warm and cold refugia beyond the range edges of species' distributions. We thus conclude that microclimate-based SDMs are a crucial tool to gain peculiar insights regarding biodiversity conservation, which is needed to align management actions and prioritize conservation efforts.

Keywords: climate change, species distribution modelling, MaxEnt, microclimate, ForestTemp, forest plant species, species response curves, understory temperatures

INTRODUCTION

Over the last decades, species distribution models (SDMs, also known as ecological niche models or habitat suitability models), have emerged as a central method to project the effects of changing environmental conditions on species' distributions in space and time (Elith & Leathwick, 2009; Guisan & Zimmermann, 2000; Zimmermann et al., 2010). Species distribution models are employed for a wide range of applications that are vital to support conservation decision making (Baker et al., 2021), ranging from quantifying the effects of contemporary climate change on biodiversity (Araújo et al., 2011; Pearce-Higgins et al., 2017) to the management of invasive species (Roy-Dufresne et al., 2019; Srivastava, 2019) and rewilding practices (Jarvie & Svenning, 2018).

SDMs commonly are correlative models that infer relationships between species occurrences and the environment using statistical or machine learning methods (Elith & Leathwick, 2009). Conventional SDM practices involve the incorporation of a standard set of bioclimatic variables with a maximal spatial resolution of 30 arc seconds (± 1 km² at the equator) such as WorldClim (Fick & Hijmans, 2017; 1 km²), CHELSA (Karger et al., 2017; 1 km²) or TerraClimate (Abatzoglou et al., 2018; 16 km²). However, these climatological data are derived from standardized meteorological stations at approximately 2 meters height above short vegetation, exposed to wind, and well away from trees and buildings to minimize any noise generated by microclimatic effects (Jarraud, 2008). Gridded macroclimatic data interpolate such weather stations' data and thus represent the free-air temperature conditions in open ecosystems. Although these data are sufficient to adequately capture changes in free-air temperatures, problems arise when using these data to model the response of species that live close to the ground in topographically heterogenous terrain and/or in ecosystems with trees and shrubs. For instance, within-pixel (1 km²) variability of mean annual temperatures can be as high as 6 °C in mountainous areas, and this might even increase when different land-use types are present within a single grid cell (i.e., forested and non-forested areas; Lenoir et al., 2013). This high thermal variability results from physical processes such as air flow and incoming solar radiation that interact with topographic factors such as slope aspect and surface roughness (Geiger, 1950). Additionally, vegetation cover is known to affect local microclimate temperature (De Frenne et al., 2019; Lenoir et al., 2017). Indeed, it is currently well acknowledged that forests harbour distinct microclimatic conditions owing to the structural complexity of the canopy, resulting in shading and evapotranspirative cooling (Geiger, 1950). Consequently, forest canopies are characterized by their buffering capacities of extreme temperatures in comparison to weather station data, with cooler sub-canopy maximum temperatures and warmer sub-canopy minimum temperatures (De Frenne et al., 2019). In European forests, this difference can add up to 9°C for mean monthly temperatures (Haesen et al., 2021).

It is clear that fine-scale microclimatic data should urgently be used within SDMs – and ecological research in general – as ignoring the mismatch between conventionally-used macroclimatic data and the apparent microclimatic conditions might lead to erroneous predictions, wrong ecological interpretations and, ultimately, questionable conservation decisions (Körner & Hiltbrunner, 2018). Especially under contemporary climate change, where species are shifting their distributions in accordance with the moving isotherms, accurate estimations of species distributions are needed. The correct identification of leading and trailing edges of moving species distributions is of particular interest for conservation since they coincide with the formation and disappearance of suitable habitat, respectively (Greiser et al., 2020). Furthermore, recent studies emphasize on the importance of so-called microrefugia, where species can find more stable climatic conditions in which they can persist for a longer amount of time (Finocchiaro et al., 2022; Nadeau et al., 2022). However, current SDM practices are unable to identify these microrefugia as conventional macroclimate data represents the overarching free-air temperatures rather than the local temperatures inside these microrefugia (Lenoir et al., 2017), which act at high-spatial resolution.

With recent advances in microclimate modelling (Gril et al., 2023; Maclean, 2019), there has been an increase in the usage of microclimate in SDMs (Graae et al., 2018; Greiser et al., 2020; Lenoir et al., 2017). Nevertheless, current studies have not yet been able to incorporate fine-scale microclimatic data over continental extents (Lembrechts, Nijs, & Lenoir, 2018). However, with the recent advent of sub-canopy microclimate layers for European forests at 25 m × 25 m resolution, a new avenue of species distribution modelling can be explored (Haesen et al., 2023).

Here, we compared SDMs constructed with different sets of bioclimatic variables to separately test the effect of using (i) proximal climate data instead of conventional macroclimatic data and (ii) high-resolution proximal climate data rather than coarse-gridded macroclimatic data. We therefore challenge SDMs with (1) a macroclimatic dataset at a spatial resolution of 1 km × 1 km; (2) an aggregated microclimatic dataset matching the resolution of the macroclimatic dataset, yet using proximal below-canopy temperatures; and (3) a microclimatic dataset at a spatial resolution of 25 m × 25 m, matching the resolution of understory vegetation communities and using the proximal below-canopy temperatures. Note that we did not opt to include a high-resolution macroclimatic dataset (i.e., 25 m × 25 m; topographically downscaled) within this comparative study as this would be representative for topoclimatic conditions, which are proximal as well. Additionally, topoclimate would not capture the influence of the canopy cover on below-canopy temperatures.

For 198 forest specialist plant species, we aimed to assess the impact of large-scale, gridded microclimate data on the performance of SDMs as well as on their predicted geographical distribution. Furthermore, we compared species thermal response curves constructed with the three methods for

the forest specialist plant species, and analysed their behaviour at their range edges. As forests are known to buffer temperatures, forest specialist plant species respond to warmer minimum temperatures and lower maximum temperatures as perceived by the free-air temperature data. Therefore, we hypothesize that (1) the actual thermal response curves of forest specialist species are narrower than the perceived thermal response curves (Figure 1). Intuitively, the incorrectly modelled thermal response curve would result in an overestimation of species range distributions when using macroclimate data. However, the resultant of the effects of the incorrectly modelled species response and the 'incorrect' macroclimate data on the predicted distribution may lead to deviation from this expectation. Therefore, our null hypothesis is that (2) the effect of the incorrectly modelled species response will be cancelled out by the 'incorrect' macroclimate data, so that distributions modelled by macroclimate and microclimate data will not differ. Finally, assuming that species are constrained by the maximum temperature at the trailing edge of their distribution and in the minimum temperature at the leading edge, we hypothesize that (3) populations of forest specialist species survive in local microrefugia, which are cooler than the surrounding area at the trailing edge and warmer than the surrounding area at the leading edge.

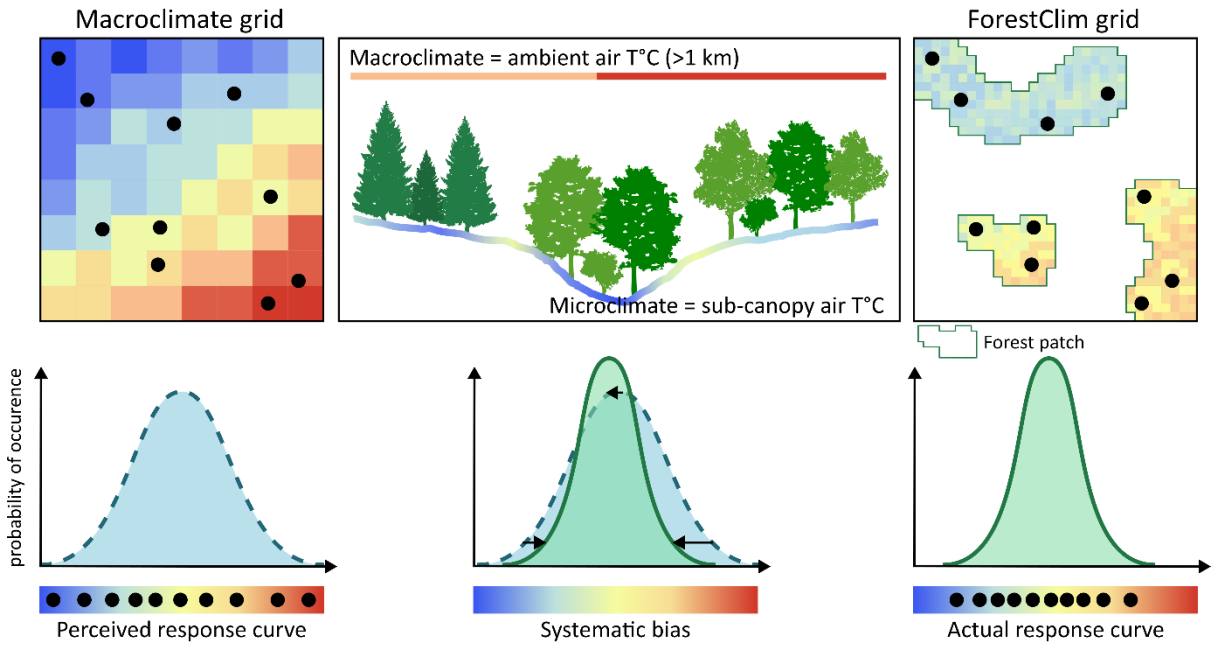


Figure 1: Design of this comparative study, where we compared species distribution models with different set-ups of climatic data. As forests are known to buffer sub-canopy temperatures, forest specialist plant species respond to warmer minimum temperatures and lower maximum temperatures as perceived by the free-air (i.e. macroclimate) temperature data. Therefore, we hypothesize that the actual thermal response curves of forest specialist species are narrower than the perceived thermal response curves. Black points indicate (simulated) species occurrences (adapted from Lenoir et al., 2017).

METHODS

Study area & species selection

Our study area encompasses all 27 EU countries, plus Albania, Andorra, Bosnia and Herzegovina, Kosovo, Liechtenstein, Montenegro, North Macedonia, Norway, San Marino, Serbia, Switzerland and the United Kingdom. The Canary Islands and Azores, as well as Europe's overseas territories were excluded from the analysis.

Forest specialist species were selected based on the European forest vascular plant species list, which is based on vegetation databases, literature and expert knowledge (Heinken et al., 2022). From this list, we first selected shrub and herb species, which – unlike tree species – usually complete their entire life cycle within the forest understory layer, thus experiencing forest microclimate dynamics (Caron et al., 2021). Subsequently, we selected the species categorized as forest specialists (i.e. categories 1.1 and 1.2 in Heinken et al., 2022) throughout their entire range, meaning that these species occur only in closed-canopy forests, forest edges or forest openings. The final selection encompassed 198 forest specialist species (Table S1).

Environmental predictors

Three different sets of bioclimatic temperature-related variables (i.e., macroclimatic, aggregated microclimatic and microclimatic) were used to construct our SDMs, starting from the conventional set of eleven bioclimatic temperature variables. However, we excluded mean temperature of the wettest quarter (BIO8) and mean temperature of the driest quarter (BIO9) as these were recently criticized for their use within species distribution models (Booth, 2022). As the available CHELSA and WorldClim data are not fully covering our study period (2000-2020), we used TerraClimate to construct the 'macroclimatic dataset' at the typical spatial resolution of 1 km² as used in conventional SDMs. TerraClimate bioclimatic variables covering the 2000-2020 period are available at a spatial resolution of 16 km² and thus were spatially downscaled to a spatial resolution of 1 km². To do this, we first calculated, for each 1 km² grid cell, the difference between the bioclimatic variables of TerraClimate (1970-2000 period; 16 km²) and WorldClim (1970-2000 period; 1 km²) and then added these offset values or anomalies to the TerraClimate bioclimatic variables for the 2000-2020 period to come to a final macroclimate layer of 1 km² resolution. This method assumes that the offset values from the long-term average period of 1970-2000 are still valid for the period 2000-2020.

The 'microclimatic dataset' consists of the original bioclimatic variables provided within ForestClim, a new high-resolution dataset of forest understory temperature for all European forests at a spatial resolution of 25 m × 25 m derived from the ForestTemp model (Haesen et al., 2021, 2023). Briefly, ForestTemp was created by combining more than 1,200 time series of *in situ* near-surface

forest temperatures from across Europe with topographical, biological and macroclimatic predictors in a machine learning model. The 'aggregated dataset' was generated by aggregating (i.e. averaging) the ForestClim bioclimatic variables to a 1 km² resolution.

Each set of bioclimatic temperature variables was complemented with the conventional set of eight bioclimatic precipitation variables. Note that we omitted precipitation of the warmest quarter (BIO18) and precipitation of the coldest quarter (BIO19) for similar reasons discussed by Booth (2022). The six remaining variables were calculated from TerraClimate precipitation data for the 2000-2020 period and disaggregated to match the spatial resolution of each bioclimatic set. Finally, edaphic variables were added, since soil data often increase model performance (Hageer et al., 2017). Based on their effects on plant demography, we selected four soil variables: bulk density (bdod; cg/cm³), which reflects the soil porosity; soil clay content (clay; g/kg), which reflects the soil texture; pH H_2O (pH; unitless); and cation exchange capacity (cec; $mmol_c/kg$; Hageer et al., 2017). The soil raster layers were downloaded from the SoilGrids database (Poggio et al., 2021) at a resolution of 250 m for three different depths: 0-5 cm; 5-15 cm; and 15-30 cm. These three layers were averaged into one single layer representing the depth from 0 cm to 30 cm, with the exception of pH (i.e., a logarithmic scale), which was aggregated using the median value over the three layers.

To help reduce overfitting of SDMs, multicollinearity between the predictors was assessed using a pairwise Spearman correlation test (Figure S1). Highly correlated variables (Spearman correlation coefficients > 0.7) were removed from the analysis in order to reach the most parsimonious model (Dormann et al., 2013). When excluding one of the correlated covariate pair, we retained variables which are known to be more important for plant species distribution (Macek, Kopecký, & Wild, 2019) and which are important for our further analyses (e.g., BIO5 & BIO6). The final selection of covariates encompassed two temperature variables (maximum temperature of the warmest month (BIO5) and minimum temperature of the coldest month (BIO6)), two precipitation variables (mean annual precipitation, (BIO12) and precipitation seasonality (BIO15)) and two edaphic variables (cation exchange capacity and soil clay content). All covariate layers were projected in an equal-area projection (epsg:3035; ETRS89/LAEA).

Species occurrence data

Georeferenced occurrence data for 198 forest plant species were downloaded from the Global Biodiversity Information Facility on the 13th of September 2022 (https://doi.org/10.15468/dl.kf533a). To improve data quality for each species, the occurrence data were filtered in the following sequential steps: (1) only records of 'human observations' were selected; (2) records with an unknown coordinate uncertainty or coordinate uncertainty larger than 25 m (i.e., the pixel size) were excluded; (3) records located at country or capital centroids and biodiversity institutions (e.g., botanical gardens)

were omitted (Cheng et al., 2021); (4) duplicate records were removed; (5) records outside our study area were deleted; (6) only records observed during our climatic reference period (2000-2020) were selected; (7) records were spatially thinned to one random observation per 25 m × 25 m grid cell; and (8) we omitted species with less than 50 cleaned occurrence records, which has been postulated as a minimum standard to build robust species distribution models (van Proosdij et al., 2016; Wisz et al., 2008). Finally, we maintained occurrence data for 140 out of 198 species (Table S1). Note that exactly the same occurrence datasets are needed over the different climatic set-ups to have comparable model outputs. Here, we decided to work with occurrence datasets that underwent a cleaning protocol based upon the characteristics of the microclimatic dataset (i.e., maximum coordinate uncertainty of 25 m, and spatial thinning to a 25 m × 25 m grid cell).

Species distribution modelling

We used MaxEnt, a presence-background algorithm that combines species presence-only data with environmental predictors for the current climate to predict the environmental suitability of each study species across our study area (Phillips, et al., 2017; Phillips & Dudík, 2008). We did that for each of the three sets of bioclimatic variables (i.e., the macroclimatic set, the aggregated microclimatic set and the microclimatic set), thus generating three sets of habitat suitability maps for each study species. Background data were generated by sampling an equal amount of background points as occurrence points (i.e., so that species prevalence equals 50%) based on a 2D kernel-density estimate of the occurrence points (Venables & Ripley, 2002), meaning that the spatial density of the background points is proportional to the spatial density of occurrence points for a given species, thereby accounting for spatial bias in the occurrence points (Lake et al., 2020; Vollering et al., 2019).

Although widely-used in scientific research, MaxEnt could suffer from issues like spatial bias and bad model performance (Radosavljevic & Anderson, 2014). To deal with the problem of spatial bias, we conducted spatially independent evaluations in ENMeval2.0 (Kass et al., 2021; Muscarella et al., 2014) using block cross-validation and allocating 80% of our occurrence points to this cross-validation procedure (20% is kept for independent evaluation). Furthermore, model performance was improved by tuning the model settings in ENMeval2.0 rather than working with the default settings of MaxEnt. This was implemented by means of a grid search over the possible values of the two hyperparameters: feature classes (Linear, Quadratic, Product) and regularization multipliers (0.5, 1, 2, 3, 4 and 5). Linear, quadratic and product features were selected to allow for linear and quadratic relationships as well as interactions among predictors (Merow, Smith, & Silander, 2013). Regularization multipliers, on the other hand, control model complexity and overfitting. The larger these regularization multipliers, the smoother the model predictions.

Model performance & sensitivity

In order to customize the settings for the feature classes and the regularization multipliers, a total of 42 different models were run for every single species. The Akaike Information Criterion for small sample sizes (AIC_c) was used to select the best candidate models (Burnham & Anderson, 2004). Next, model performance was assessed using the Continuous Boyce Index (CBI), instead of the commonly-used area under the receiver-operating characteristic curve (AUC). The latter has recently been shown to be biased in presence-only models and should therefore be avoided (Jiménez & Soberón, 2020). The CBI is a threshold-independent metric that represents the relationship between predicted habitat suitability and the distribution of occurrence records (Hirzel et al., 2006). Additionally, we calculated the sensitivity enabling us to quantify how good our model is able at distinguishing true positives from false negatives. Both were calculated based on the independent 20% subset of the data.

Finally, we used Bayesian regression models (BRMs) in order to assess differences in model performance and sensitivity between SDMs constructed using the three different types of climate data. We opted for BRMs as they are able to account for data dependencies (i.e., values clustered within species), unequal variances among groups and skewed distributions. The final model structures are added in Table S2. Both sensitivity and CBI were modelled with a beta distribution. CBI was rescaled between 0 and 1 before analysis. Bayesian regression models were run using the *brms* package (Bürkner, 2021). All models were first run using standard priors and with 2 chains, 10,000 iterations and a warm-up of 1000 runs. When models did not converge, the flat priors were replaced by weakly informative priors (Table S2) and the models were run again with 4 chains. The final models converged with $\hat{R}$ values close to 1 (Gelman and Rubin's diagnostic) and all bulk and tail effective sample sizes of the means were greater than 2500. When the highest posterior density intervals (α = 0.05) of the contrasts, calculated using the *emmeans* package (Lenth, 2021), did not overlap with zero, contrasts are considered 'significant'.

Model predictions

Habitat suitability was predicted for each species and for each of the three sets of bioclimatic temperature variables (macroclimatic, aggregated microclimatic and microclimatic) for the 2000-2020 period. Furthermore, we transformed the logistic maps (i.e., probability values for habitat suitability) to binary (presence-absence) maps using the 10% training presence as a threshold, meaning that the suitable area contains 90% of the original occurrence records (Benito, Cayuela, & Albuquerque, 2013).

To compare between model predictions from SDMs constructed with different climate sources and resolutions, we calculated both the potential suitable area and the potential latitudinal range of each species. Note that we (i) disaggregated the binary maps derived from macroclimatic and

aggregated data (1 km × 1 km) to the finer resolution (25 m × 25 m), and (ii) masked out all non-forest pixels to make a valid comparison between the three climate types. First, the potential suitable area (km^2) was calculated as the sum of all pixels classified as potentially suitable under the binary maps. Second, the northern (i.e., leading or cold) and southern (i.e., trailing or warm) latitudinal limit of the predicted distributional ranges were defined as the 95% and 5% quantile in latitudinal position, respectively, of all pixels classified as potentially suitable. Next, we quantified species thermal response curves for mean annual temperature (BIO1), maximum temperature of the warmest month (BIO5) and minimum temperature of the coldest month (BIO6). Note that we randomly sampled 1,000,000 pixels over the potentially suitable area to optimize computational power. For each variable, we derived the cold edge (Q05), the optimum (mode), the warm edge (Q95), and the niche width (Q95 – Q05). Analogous to the model performance, we used BRMs with the same settings to assess differences in model predictions between the SDMs based on the three types of climate data (Table S2). Values of bioclimatic variables were standardized before the analysis to aid model convergence.

Finally, we analyzed whether species are constrained to specific (relative) temperature conditions in their leading and trailing latitudinal limits, as this potentially has important implications for nature conservation. For the trailing and leading edge, we extracted the 5% most southern and northern occurrence records, respectively. Using paired two-sided t-tests ($\alpha = 0.05$), we compared the local temperature conditions of these occurrence points to the surrounding microclimatic conditions over a range of circular buffers (i.e., 100 m, 500 m, 1000 m, 2500 m, 5000 m; Figure S2) around each occurrence record.

All calculations were performed in R version 4.1.1 (R Core Team, 2021). The Tier-2 Genius cluster from the high-performance computing facilities of Flanders was used to make the predictions. In order to improve reproducibility, we followed the ODMAP (Overview, Data, Model, Assessment and Prediction) protocol to report on the SDMS in this study (Table S3; Zurell et al., 2020).

RESULTS

Model performance & sensitivity

We found significant differences (α = 0.05) in model performance between models constructed with (i) macroclimatic (mean CBI = 0.09; se = 0.04) and microclimatic (mean CBI = 0.67; se = 0.02) data, (ii) macroclimatic and aggregated microclimatic (mean CBI = 0.28; se = 0.04), and (iii) aggregated microclimatic and microclimatic data (Figure 2). Furthermore, there were no significant differences between any of the groups regarding the sensitivity of the models. We identified 11 species (i.e., *Anemone trifolia, Asarum europaeum, Clematis recta, Cyclamen purpurascens, Dictamnus albus, Gagea spathacea, Lathyrus vernus, Neottia nidus-avis, Polystichum aculeatum, Ribes spicatum,* and *Saxifraga hirsuta*) for which the incorporation of microclimatic data did not increase the performance of the SDMs in comparison to conventional SDM practices, contrarily to the other 129 species for which microclimate improved performance.

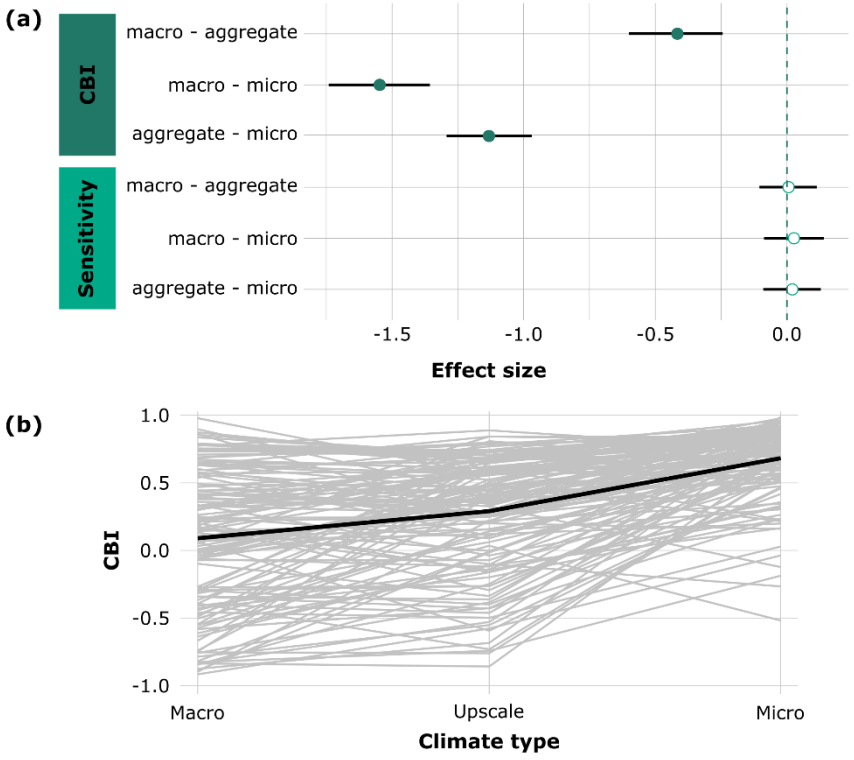


Figure 2: (a) Pairwise comparison of performance (CBI) and sensitivity between SDMs build with macroclimatic, aggregated microclimatic and microclimatic data. A positive effect size of the comparison reflects a higher model performance and sensitivity in SDMs built with the first group of climate data compared to SDMs built with the second group of climate data. Negative effect sizes reflect the opposite result. Points and associated black error bars correspond to posterior means and 95% highest posterior density intervals of the differences (of the scaled CBI and sensitivity). Significant differences are indicated by full dots whereas non-significant differences are indicated by transparent dots; (b) Parallel coordinate chart indicating the performance of each SDM per species over the three types of climate data (i.e., macroclimatic data, aggregated microclimatic data and microclimatic data). The thick black line shows the average CBI value over each of the three climate types. The GGally package was used to create the parallel coordinate plot (Schloerke et al., 2022).

Potential suitable area & latitudinal range

A first visual assessment of the binary maps showed clear differences in the potential suitable area and the potential latitudinal range covered by each species between models calibrated with macroclimatic data and models calibrated with microclimatic data at the native spatial resolution of 25 m × 25 m (non-aggregated data) (e.g., *Paris quadrifolia*; Figure S3). Indeed, the Bayesian regression models confirm these visual interpretations (Figure 3). Relative to the native microclimate-based SDMs, both the leading and trailing edge of the species' distributional ranges are significantly overestimated when using either macroclimatic or aggregated microclimatic data at 1-km^2. Consequently, species' potential latitudinal ranges are significantly smaller when using SDMs calibrated with microclimatic data (mean = 2,261 km; se = 42 km) in comparison with SDMs calibrated with aggregated microclimatic data (mean = 2,580 km; se = 43 km) or macroclimatic data (mean = 2,620 km; se = 49 km). Analogous, a species' potential suitable area is significantly smaller when using SDMs calibrated with microclimatic data (mean = 911,845 km²; se = 30,383 km²) in comparison with SDMs calibrated with aggregated microclimatic data (mean = 1,148,763 km²; se = 33,527 km²) or macroclimatic data (mean = 1,268,189 km²; se = 38,274 km²).

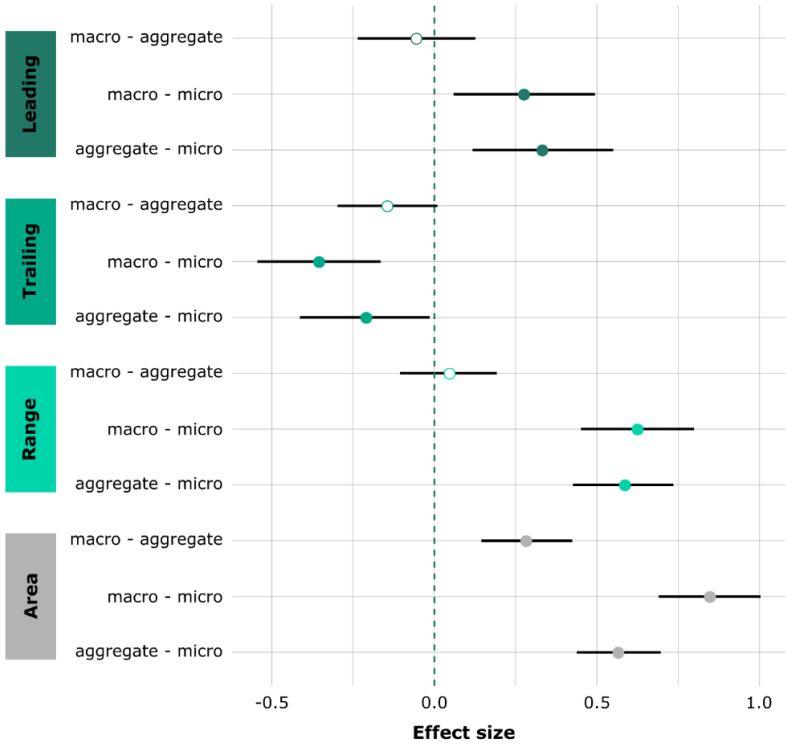


Figure 3: Pairwise comparison of the leading edge, trailing edge, latitudinal range, and potential suitable area, respectively between SDMs build with macroclimatic, aggregated microclimatic and microclimatic data. A positive effect size of the comparison reflects more northern leading edges, more northern trailing edges, higher latitudinal ranges and more potentially suitable area in SDMs built with the first group of climate data compared to SDMs built with the second group of climate data. Negative effect sizes reflect the opposite result. Points and associated black error bars correspond to posterior means and 95% highest posterior density intervals of the differences (of the standardized variables). Significant differences are indicated by full dots whereas insignificant differences are indicated by transparent dots.

Species response curves

A first visual assessment of the response curves showed that microclimate-based response curves of minimum temperature of the coldest month, mean annual temperature and maximum temperature of the warmest month have different optima, and narrower niches compared to macroclimate-based response curves (e.g., *Paris quadrifolia*; Figure 4).

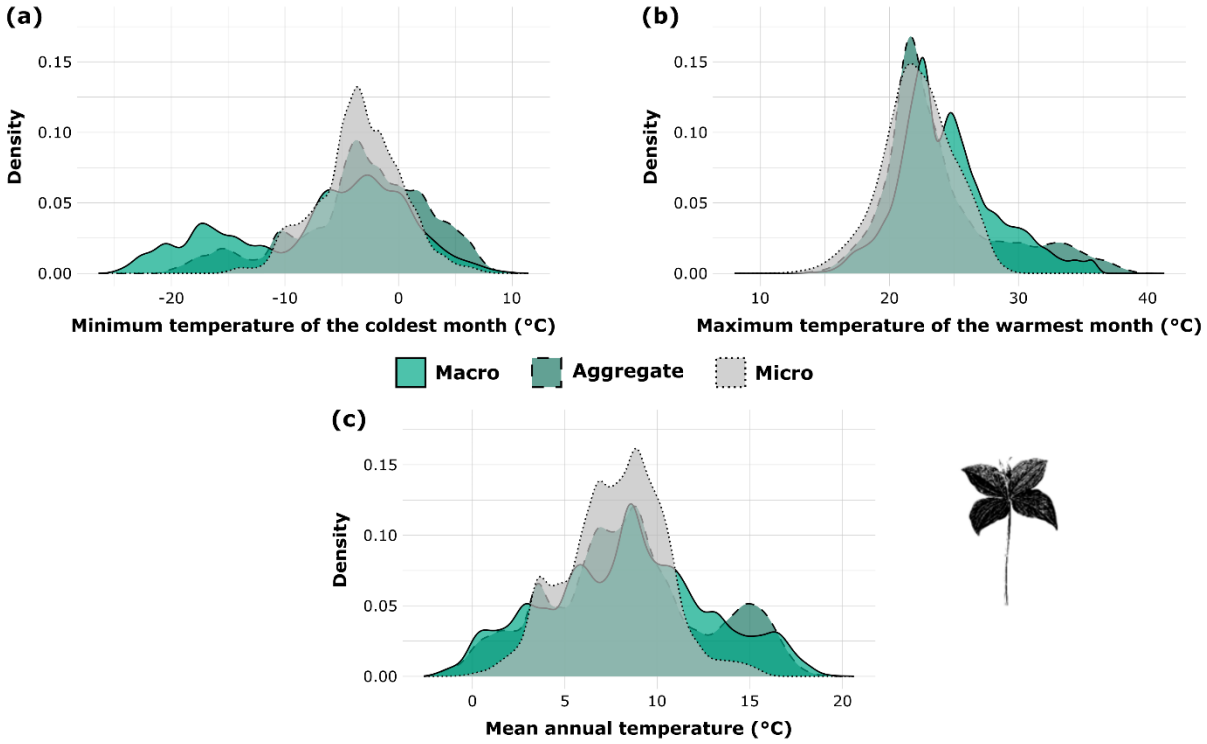


Figure 4: Response curves for (a) minimum temperature of the coldest month; (b) maximum temperature of the warmest month and (c) mean annual temperature for Paris quadrifolia.

Here, the Bayesian regression models partially supports these visual interpretations (Figure 5). Optima significantly differed between SDMs ran with microclimate and macroclimate data for minimum and maximum temperatures, with warmer optima in minimum temperature and cooler optima in maximum temperature for microclimate-based SDMs relative to macroclimate based SDMs. However, for mean temperature there are no significant differences in optima between the different climate types. Furthermore, the niche width was narrower in minimum and mean temperatures for microclimate-based SDMs relative to macroclimate based SDMs. Surprisingly, the niche width is significantly wider in maximum temperatures for microclimate-based SDMs relative to macroclimate based SDMs.

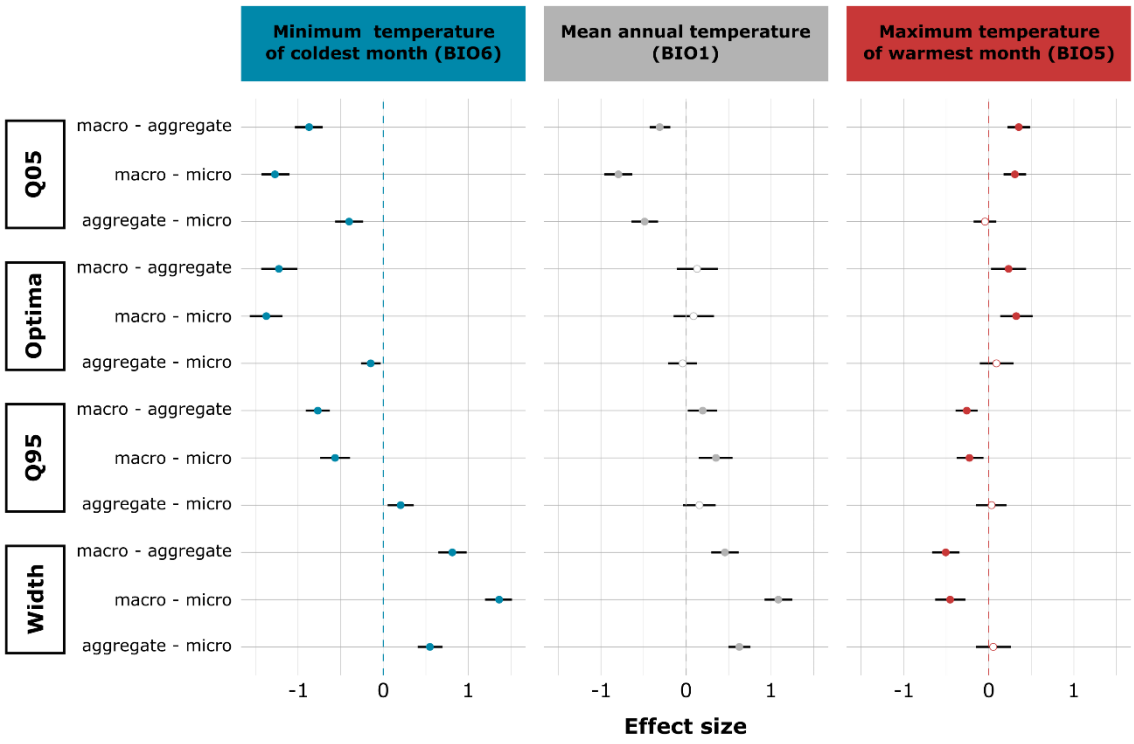


Figure 5: Pairwise comparison of the cold edge (Q05), optimum, warm edge (Q95), and niche width, respectively between SDMs build with macroclimatic, aggregated microclimatic and microclimatic data . Each of the comparisons is made for minimum temperature of the coldest month (BIO6), mean annual temperature (BIO1), and maximum temperature of the warmest month (BIO5), respectively. A positive effect size reflects warmer values for the cold edge, optima and warm edge as well as wider niche widths, respectively, in SDMs built with the first group of climate data compared to SDMs built with the second group of climate data. Negative effect sizes reflect the opposite result. Points and associated black error bars correspond to posterior means and 95% highest posterior density intervals of the differences (of the standardized variables). Significant differences are indicated by full dots whereas insignificant differences are indicated by transparent dots.

Microrefugia

For a buffer of 2500 m, we found that 66% of all the species are constrained in local microrefugia at their range edges. More specifically, 41% of the species found refuge in warm refugia, relative to the surrounding landscape, at the leading edge while 49% of the species have remnant populations in cool refugia, relative to the surrounding landscape, at the trailing edge (e.g., *Paris quadrifolia*; Figure 6).

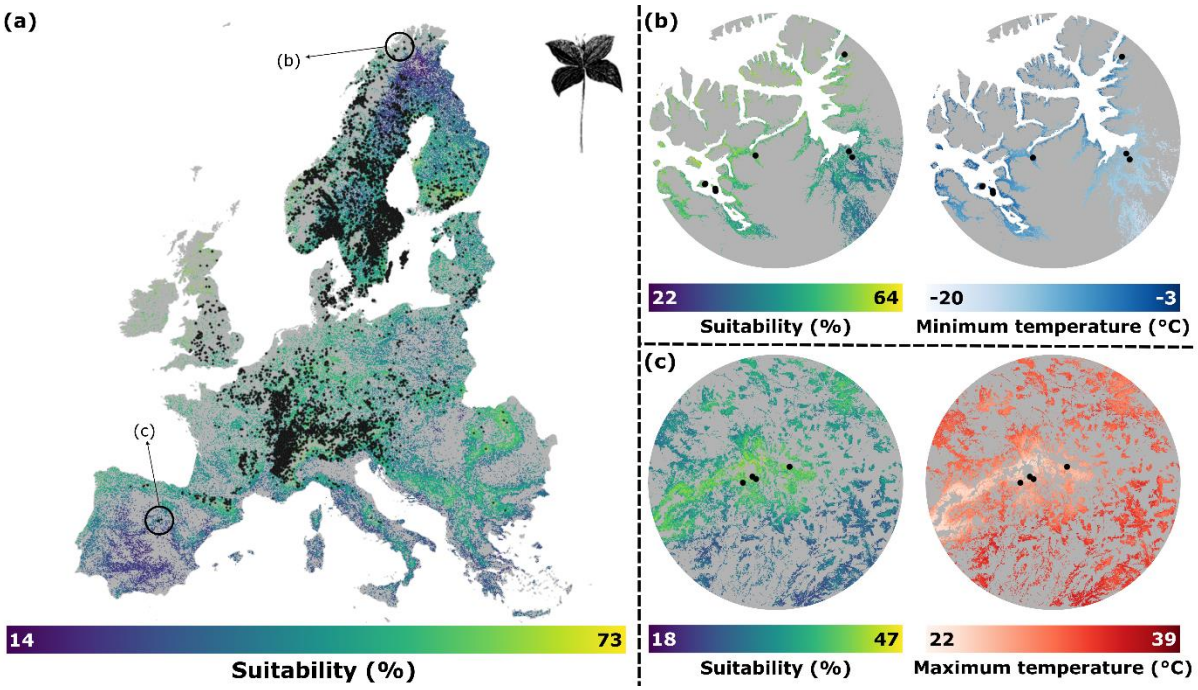


Figure 6: (a) Suitability map for Paris quadrifolia resulting from an SDM built with microclimatic data at 25 m × 25 m resolution. The black dots represent the occurrence points extracted from GBIF and used as an input to the SDMs. We see that the species is located in (b) warm refugia (i.e., higher minimum temperature values in the coldest month of the year) at their leading edge and in (c) cool refugia (i.e., lower maximum temperature values in the warmest month of the year) at their trailing edge. The grey background shows non-forest areas.

DISCUSSION

Here, we tested the effect of incorporating three different types of climatic data (i.e., macroclimate, aggregated microclimate matching the spatial resolution of macroclimate and microclimate at its fine spatial resolution) into species distribution models (SDMs) for 140 forest specialist plant species, allowing us to investigate the effect of microclimate as well as the effect of spatial resolution on the performance and predictions of SDMs. We found substantial differences in the model performance (based on the Continuous Boyce Index), indicating that microclimate-based SDMs significantly outperform their conventional (i.e., macroclimate) counterparts and that aggregating microclimate data at coarser spatial resolutions matching macroclimate leads to significant loss in performances. However, the use of aggregated proximal data is still a significant improvement to the use of conventional macroclimate data in SDMs. Furthermore, species response curves derived from the model predictions differed substantially between climate types. These findings highlight the importance of incorporating microclimate data within SDMs, as already postulated by Lembrechts et al. (2018). Certainly within the face of climate change, microclimate-based SDMs are a valuable tool as they allow to identify local refugia for biodiversity conservation. Indeed, up to 66% of the studied species have remnant populations in warm or cold refugia at the leading or trailing edge, respectively. Accordingly, management practices should be in line with the gained insights in order to preserve local microrefugia as they are able to preserve biodiversity despite macroclimatic warming.

The strength of microclimate-based species distribution models

Over the last years, microclimate research focused on improving our understanding of the drivers behind differences between microclimate and macroclimate temperatures (Zellweger et al., 2019) and predicting and mapping microclimate temperatures across space and time (Greiser et al., 2018; Kearney et al., 2019; Lembrechts et al., 2022). Although the drivers behind forest microclimates are relatively well understood, testing how microclimate layers perform within ecological applications such as SDMs has been limited, especially so across large (i.e., continental) spatial extents.

Previous research supports our findings on an increased performance of microclimate-based SDMs. For instance, Ashcroft et al. (2008) have already stated that explaining plant species distributions benefits from increasing the accuracy of local temperature, which is also confirmed by Slavich et al. (2014) who have shown that using topoclimate rather than coarse-gridded macroclimatic data leads to improved model performance. On the contrary, Stark & Fridley (2022) did not find significant differences in the performance of microclimate-based and macroclimate-based joint SDMs. Nevertheless, when studying plant species distributions, high-resolution environmental information has overall shown to significantly improve model performance and spatial predictions.

The increased performance of microclimate-based SDMs can be explained by several reasons, which are mainly related to the two main input sources of each SDM: the occurrence points and our predictors. First, each occurrence point is subjected to a certain amount of positional error (Wüest et al., 2020). In this study, we only used records with a very low coordinate uncertainty (< 25 m). However, putting such a threshold on the positional error might induce a loss of model power by reducing the sample size of the occurrence points (Guisan et al., 2007). Therefore, many studies often include occurrence records with higher positional uncertainties (e.g., Sanczuk et al., 2022). To deal with these errors within the analysis, it is conventionally suggested to increase the spatial resolution of the analysis to compensate for any positional errors in the occurrence points. However, SDMs are sensitive to changes in the spatial resolution (Chauvier et al., 2022; Manzoor et al., 2018). Decreasing spatial resolution inherently induces a loss of information as the data is smoothed (i.e., aggregated). This comes at the cost of model performance as shown by the CBI values from the models built with aggregated microclimatic data. Therefore, Gábor et al. (2022) strongly recommend to fit models as close as possible to the response grain of the species, meaning that it is recommended to calibrate SDMs with environmental data consistent with the biological scale of the system or organism under study (Randin et al., 2009). For instance, when modelling sessile species (i.e., species with low mobility) or organisms in systems characterized by high environmental heterogeneity, predictors with an increased spatial resolution will be needed to capture the details in their niches more accurately (Araújo et al., 2019; Elith & Leathwick, 2009). Consequently, we expect that the results from this study are not necessarily transferable when studying more mobile species (e.g., birds or mammals) in homogeneous systems (e.g., flat terrain, monoculture plantations). Nevertheless, aggregating proximal data still significantly improves the performance of SDMs relative to the use of conventional macroclimate data in SDMs. Especially when computational capacity is limited, aggregated microclimatic data could be used in order to improve models.

The climate that really matters

The increased availability of microclimatic data products over large spatial extents (Haesen et al., 2023; Lembrechts et al., 2022) opens new avenues within ecological research. With these data, we are able to unravel the hidden niches and describe the conditions that actually matter for species living close to the ground surface (e.g., tree seedlings and forest floor herbs) at very fine spatial resolutions. This study makes clear that species environmental niches derived from conventional macroclimatic data are much wider than one would expect given the buffering effect of forests (De Frenne et al., 2019). Furthermore, the niche estimations based on microclimatic data make it possible to more accurately pinpoint where species could come under pressure due to climate change, facilitating alignment of appropriate management actions (e.g., assisted migration). Indeed, an accurate

assessment of species distributional ranges will become vital as many forest specialist species are characterized by slow dispersal rates, up to several meters per year (Hermy et al., 1999; Svenning et al., 2008). It is very unlikely that these species will be able to follow contemporary macroclimate warming, where climate zones are shifting several kilometres each year (Burrows et al., 2011). In this respect, microclimate data and microclimate-based SDMs offer a solution as they will allow us to more accurately assess the velocity of microclimate warming experienced by organisms and its effect on species redistributions. Although outside the scope of this research, this could potentially reveal that the actual velocity of species redistributions does not have to be as high as predicted based on macroclimatic data.

When taking a closer look at the trailing and leading edge, we found that 66% of the studied species already persist in cool or warm refugia, respectively (Figure 6; Figure S2). Current SDM practices are unable to identify these microrefugia as conventional macroclimate data represents the overarching free-air temperatures rather than the local temperatures inside these microrefugia (Lenoir et al., 2017), which act at high-spatial resolution. Microclimate-based SDMs thus enable the species-level identification of these microrefugia (Michalak et al., 2020). These insights will thus be crucial for conservation practices as the importance of microrefugia regarding the accumulation and conservation of biodiversity has been widely discussed in the recent scientific literature (Finocchiaro et al., 2022; Nadeau et al., 2022). Indeed, microclimate-based SDMs are a valuable tool to identify areas within the landscape where forest management practices can be aligned (i) to increase the capacity of species and communities to resist climate change (i.e., resistance strategy) or (ii) to facilitate the transformation of communities to species that are well adapted to the novel environmental conditions (Hylander et al., 2022).

We would like to note that the goal of this study was to compare different types of climate data for use in SDMs, meaning that we did not necessarily build the best possible model for each species. Therefore, the environmental niches and distributional ranges provided in this study should not be used to undertake conservation actions.

Future improvements

As gridded microclimatic data at 25 m × 25 m is currently only available for European forests, this multi-species study is limited to 140 forest specialist plant species. Here, we do not cover herbaceous plant species living in open habitats such as grasslands or heathlands, as relevant microclimate data at the relevant spatial resolution is not yet available for these habitats. This is mainly due to the fact that measurements from available microclimate sensors are strongly affected by incoming solar radiation when installed in open ecosystems (Maclean et al., 2021), which prevents the development of analogous microclimatic data for these systems. In order to assess whether the results from this

485 study are transferable to other species groups, accurate microclimate data over large spatial extents
486 in open systems is thus urgently needed. Furthermore, additional variables play an important role in
487 plant species distributions but are not yet available at high spatial resolution over large spatial extents.
488 For instance, soil moisture is known to be vital for plant survival, but up till now the topographic
489 wetness index has often been used as a proxy (Kopecký et al., 2021) as high-resolution soil moisture
490 products are hard to obtain. Finally, microclimatic data predicted under future shared socioeconomic
491 pathways (SSPs) will improve insights for future conservation efforts. For instance, knowledge on
492 species reshuffling under climate change or thermophilization of species communities could benefit
493 from such microclimate change predictions. However, forests are very dynamic systems and their
494 structural characteristics – known to influence the forest microclimate – cannot be assumed to remain
495 static over time, which hampers the development of such dynamic products up till now (De
496 Lombaerde et al., 2022). Especially within a warming world, disturbances affecting forest canopies
497 (e.g., drought, pests, storms) will become more frequent and pronounced (De Frenne et al., 2021;
498 Seidl et al., 2017).

CONCLUSIONS

In this study, we performed a comparative analysis between species distribution models (SDMs) with three different sets of climate data: (i) conventional macroclimatic data; (ii) aggregated microclimatic data matching the macroclimatic data in spatial resolution; and (iii) microclimatic data at the native fine spatial resolution that is relevant to the size of the studied organisms (i.e., understory forest specialists here). We conclude that the performance of SDMs for forest specialist species can be significantly improved by incorporating microclimatic data, although this might not necessarily be transferable to other species groups. We would like to emphasize the ability of microclimate-based SDMs to uncover the hidden niche of forest specialist plant species, which has implications for the predicted tolerance of these species at their warm and cold edge. Furthermore, this study makes clear that macroclimatic data is unable to identify warm and cold refugia beyond the range edges of species' distributions. Therefore, we conclude that microclimate-based SDMs are a crucial tool to gain peculiar insights regarding biodiversity conservation within the face of climate change, which is needed to align management actions and prioritize conservation efforts.

ACKNOWLEDGEMENTS

This work was funded by the European Union through the COST Action CA18201 – ConservePlants & Internal Funds of KU Leuven and by an FWO Research Network Grant to SoilTemp (W001919N). SH was supported by a FLOF fellowship (project nr. 3E190655) of the KU Leuven. JL received funding from: (i) the Agence Nationale de la Recherche (ANR) (project IMPRINT; https://microclimat.cnrs.fr; grant nr. ANR-19-CE32-0005-01); (ii) the Centre National de la Recherche Scientifique (CNRS) through the MITI interdisciplinary programs (Défi INFINITI 2018: MORFO); (iii) and the Structure Fédérative de Recherche (SFR) Condorcet (FR CNRS 3417: CREUSE). JJL is funded by the Research Foundation Flanders (12P1819N) and by the ASICS project (ANR-20-EBI5-0004, BiodivERsA, BiodivClim call 2019–2020). The study was further supported by the Czech Science Foundation (project GACR 20-28119S) and the Czech Academy of Sciences (project RVO 67985939).

REFERENCES

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, *5*(1), 170191. https://doi.org/10.1038/sdata.2017.191

Araújo, M. B., Alagador, D., Cabeza, M., Nogués-Bravo, D., & Thuiller, W. (2011). Climate change threatens European conservation areas. *Ecology Letters*, *14*(5), 484–492. https://doi.org/10.1111/j.1461-0248.2011.01610.x

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., … Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, *5*(1), eaat4858. https://doi.org/10.1126/sciadv.aat4858

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*(2), 211–225. https://doi.org/10.1007/s10980-007-9181-8

Baker, D. J., Maclean, I. M. D., Goodall, M., & Gaston, K. J. (2021). Species distribution modelling is needed to support ecological impact assessments. *Journal of Applied Ecology*, *58*(1), 21–26. https://doi.org/10.1111/1365-2664.13782

Benito, B. M., Cayuela, L., & Albuquerque, F. S. (2013). The impact of modelling choices in the predictive performance of richness maps derived from species-distribution models: Guidelines to build better diversity models. *Methods in Ecology and Evolution*, *4*(4), 327–335. https://doi.org/10.1111/2041-210x.12022

Booth, T. H. (2022). Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models. *Austral Ecology*, *47*(7), 1506–1514. https://doi.org/10.1111/aec.13234

Bürkner, P.-C. (2021). Bayesian Item Response Modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5). https://doi.org/10.18637/jss.v100.i05

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., … Richardson, A. J. (2011). The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science*, *334*(6056), 652–655. https://doi.org/10.1126/science.1210288

Caron, M. M., Zellweger, F., Verheyen, K., Baeten, L., Hédl, R., Bernhardt-Römermann, M., … De Frenne, P. (2021). Thermal differences between juveniles and adults increased over time in European forest trees. *Journal of Ecology*, (April), 1–14. https://doi.org/10.1111/1365-2745.13773

Chauvier, Y., Descombes, P., Guéguen, M., Boulangeat, L., Thuiller, W., & Zimmermann, N. E. (2022). Resolution in species distribution models shapes spatial patterns of plant multifaceted diversity. *Ecography*, *2022*(10), 1–13. https://doi.org/10.1111/ecog.05973

Cheng, Y., Tjaden, N. B., Jaeschke, A., Thomas, S. M., & Beierkuhnlein, C. (2021). Using centroids of spatial units in ecological niche modelling: Effects on model performance in the context of environmental data grain size. *Global Ecology and Biogeography*, *30*(3), 611–621. https://doi.org/10.1111/geb.13240

De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., … Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Global Change Biology*, (November 2020), gcb.15569. https://doi.org/10.1111/gcb.15569

De Frenne, P., Zellweger, F., Rodríguez-Sánchez, F., Scheffers, B. R., Hylander, K., Luoto, M., … Lenoir, J. (2019). Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution*, *3*(5), 744–749. https://doi.org/10.1038/s41559-019-0842-1

De Lombaerde, E., Vangansbeke, P., Lenoir, J., Van Meerbeek, K., Lembrechts, J., Rodríguez-Sánchez, F., … De Frenne, P. (2022). Maintaining forest cover to enhance temperature buffering under future climate change. *Science of The Total Environment*, *810*, 151338. https://doi.org/10.1016/j.scitotenv.2021.151338

575 Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013).
576 Collinearity: a review of methods to deal with it and a simulation study evaluating their
577 performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

578 Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction
579 Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697.
580 https://doi.org/10.1146/annurev.ecolsys.110308.120159

581 Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global
582 land areas. *International Journal of Climatology*, *37*(12), 4302–4315.
583 https://doi.org/10.1002/joc.5086

584 Finocchiaro, M., Médail, F., Saatkamp, A., Diadema, K., Pavon, D., & Meineri, E. (2022). Bridging the
585 gap between microclimate and microrefugia: A bottom-up approach reveals strong climatic and
586 biological offsets. *Global Change Biology*, (October), 1–13. https://doi.org/10.1111/gcb.16526

587 Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., … Moudrý, V. (2022). Positional errors
588 in species distribution modelling are not overcome by the coarser grains of analysis. *Methods in*
589 *Ecology and Evolution*, *2022*(June), 1–14. https://doi.org/10.1111/2041-210x.13956

590 Geiger, R. (1950). *The climate near the ground*. Cambridge, Mass.: Harvard University Press. 482p.
591 pages.

592 Graae, B. J., Vandvik, V., Armbruster, W. S., Eiserhardt, W. L., Svenning, J.-C., Hylander, K., … Lenoir, J.
593 (2018). Stay or go – how topographic complexity influences alpine plant population and
594 community responses to climate change. *Perspectives in Plant Ecology, Evolution and*
595 *Systematics*, *30*(September 2017), 41–50. https://doi.org/10.1016/j.ppees.2017.09.008

596 Greiser, C., Ehrlén, J., Meineri, E., & Hylander, K. (2020). Hiding from the climate: Characterizing
597 microrefugia for boreal forest understory species. *Global Change Biology*, *26*(2), 471–483.
598 https://doi.org/10.1111/gcb.14874

599 Greiser, C., Hylander, K., Meineri, E., Luoto, M., & Ehrlén, J. (2020). Climate limitation at the cold edge:
600 contrasting perspectives from species distribution modelling and a transplant experiment.
601 *Ecography*, *43*(5), 637–647. https://doi.org/10.1111/ecog.04490

602 Greiser, C., Meineri, E., Luoto, M., Ehrlén, J., & Hylander, K. (2018). Monthly microclimate models in a
603 managed boreal forest landscape. *Agricultural and Forest Meteorology*, *250–251*(December
604 2017), 147–158. https://doi.org/10.1016/j.agrformet.2017.12.252

605 Gril, E., Spicher, F., Greiser, C., Ashcroft, M. B., Pincebourde, S., Durrieu, S., … Lenoir, J. (2023). Slope

606 and equilibrium: A parsimonious and flexible approach to model microclimate. *Methods in*
607 *Ecology and Evolution*, *2022*(May), 1–13. https://doi.org/10.1111/2041-210X.14048

608 Guisan, A., Graham, C. H., Elith, J., & Huettmann, F. (2007). Sensitivity of predictive species distribution
609 models to change in grain size. *Diversity and Distributions*, *13*(3), 332–340.
610 https://doi.org/10.1111/j.1472-4642.2007.00342.x

611 Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological*
612 *Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

613 Haesen, S., Lembrechts, J. J., De Frenne, P., Lenoir, J., Aalto, J., Ashcroft, M. B., … Van Meerbeek, K.
614 (2021). ForestTemp – Sub-canopy microclimate temperatures of European forests. *Global*
615 *Change Biology*, *27*(23), 6307–6319. https://doi.org/10.1111/gcb.15892

616 Haesen, S., Lembrechts, J. J., De Frenne, P., Lenoir, J., Aalto, J., Ashcroft, M. B., … Van Meerbeek, K.
617 (2023). ForestClim—Bioclimatic variables for microclimate temperatures of European forests.
618 *Global Change Biology*, *29*(11), 2886–2892. https://doi.org/10.1111/gcb.16678

619 Hageer, Y., Esperón-Rodríguez, M., Baumgartner, J. B., & Beaumont, L. J. (2017). Climate, soil or both?
620 Which variables are better predictors of the distributions of Australian shrub species? *PeerJ*, *5*,
621 e3446. https://doi.org/10.7717/peerj.3446

622 Heinken, T., Diekmann, M., Liira, J., Orczewska, A., Schmidt, M., Brunet, J., … Vanneste, T. (2022). The
623 European Forest Plant Species List (EuForPlant): Concept and applications. *Journal of Vegetation*
624 *Science*, *33*(3). https://doi.org/10.1111/jvs.13132

625 Hermy, M., Honnay, O., Firbank, L., Grashof-Bokdam, C., & Lawesson, J. E. (1999). An ecological
626 comparison between ancient and other forest plant species of Europe, and the implications for
627 forest conservation. *Biological Conservation*, *91*(1), 9–22. https://doi.org/10.1016/S0006-
628 3207(99)00045-2

629 Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat
630 suitability models to predict species presences. *Ecological Modelling*, *199*(2), 142–152.
631 https://doi.org/10.1016/j.ecolmodel.2006.05.017

632 Hylander, K., Greiser, C., Christiansen, D. M., & Koelemeijer, I. A. (2022). Climate adaptation of
633 biodiversity conservation in managed forest landscapes. *Conservation Biology*, *36*(3), 1–9.
634 https://doi.org/10.1111/cobi.13847

635 Jarraud, M. (2008). *Guide to meteorological instruments and methods of observation (WMO-No. 8)*.
636 Geneva, Switzerland: World Meteorological Organisation.

637 Jarvie, S., & Svenning, J.-C. (2018). Using species distribution modelling to determine opportunities for
638 trophic rewilding under future scenarios of climate change. *Philosophical Transactions of the*
639 *Royal Society B: Biological Sciences*, *373*(1761), 20170446.
640 https://doi.org/10.1098/rstb.2017.0446

641 Jiménez, L., & Soberón, J. (2020). Leaving the area under the receiving operating characteristic curve
642 behind: An evaluation method for species distribution modelling applications based on presence-
643 only data. *Methods in Ecology and Evolution*, *11*(12), 1571–1586. https://doi.org/10.1111/2041-
644 210X.13479

645 Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2017).
646 Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*(1), 170122.
647 https://doi.org/10.1038/sdata.2017.122

648 Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla-Buitrago, G. E., Boria, R. A., … Anderson, R.
649 P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species'
650 niches and distributions. *Methods in Ecology and Evolution*, *2021*(January), 2041-210X.13628.
651 https://doi.org/10.1111/2041-210X.13628

652 Kearney, M. R., Gillingham, P. K., Bramer, I., Duffy, J. P., & Maclean, I. M. D. (2019). A method for
653 computing hourly, historical, terrain-corrected microclimate anywhere on Earth. *Methods in*
654 *Ecology and Evolution*, (1), 2041-210X.13330. https://doi.org/10.1111/2041-210X.13330

655 Kopecký, M., Macek, M., & Wild, J. (2021). Topographic Wetness Index calculation guidelines based
656 on measured soil moisture and plant species composition. *Science of The Total Environment*, *757*,
657 143785. https://doi.org/10.1016/j.scitotenv.2020.143785

658 Körner, C., & Hiltbrunner, E. (2018). The 90 ways to describe plant temperature. *Perspectives in Plant*
659 *Ecology, Evolution and Systematics*, *30*(December 2016), 16–21.
660 https://doi.org/10.1016/j.ppees.2017.04.004

661 Lake, T. A., Briscoe Runquist, R. D., & Moeller, D. A. (2020). Predicting range expansion of invasive
662 species: Pitfalls and best practices for obtaining biologically realistic projections. *Diversity and*
663 *Distributions*, *26*(12), 1767–1779. https://doi.org/10.1111/ddi.13161

664 Lembrechts, J. J., Hoogen, J., Aalto, J., Ashcroft, M. B., De Frenne, P., Kemppinen, J., … Lenoir, J. (2022).
665 Global maps of soil temperature. *Global Change Biology*, *28*(9), 3110–3144.
666 https://doi.org/10.1111/gcb.16060

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

models. *Ecography*, 1–13. https://doi.org/10.1111/ecog.03947

Lenoir, J., Graae, B. J., Aarrestad, P. A., Alsos, I. G., Armbruster, W. S., Austrheim, G., … Svenning, J.-C. (2013). Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across Northern Europe. *Global Change Biology*, *19*(5), 1470–1481. https://doi.org/10.1111/gcb.12129

Lenoir, J., Hattab, T., & Pierre, G. (2017). Climatic microrefugia under anthropogenic climate change: implications for species redistribution. *Ecography*, *40*(2), 253–266. https://doi.org/10.1111/ecog.02788

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. Retrieved from https://cran.r-project.org/package=emmeans

Macek, M., Kopecký, M., & Wild, J. (2019). Maximum air temperature controlled by landscape topography affects plant species composition in temperate forests. *Landscape Ecology*, *34*(11), 2541–2556. https://doi.org/10.1007/s10980-019-00903-x

Maclean, I. M. D. (2019). Predicting future climate at high spatial and temporal resolution. *Global Change Biology*, (August), gcb.14876. https://doi.org/10.1111/gcb.14876

Maclean, I. M. D., Duffy, J. P., Haesen, S., Govaert, S., De Frenne, P., Vanneste, T., … Van Meerbeek, K. (2021). On the measurement of microclimate. *Methods in Ecology and Evolution*, 2041-210X.13627. https://doi.org/10.1111/2041-210X.13627

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size – finer may not always be better. *Scientific Reports*, *8*(1), 7168. https://doi.org/10.1038/s41598-018-25437-1

Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, *36*(10), 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Michalak, J. L., Stralberg, D., Cartwright, J. M., & Lawler, J. J. (2020). Combining physical and species-based approaches improves refugia identification. *Frontiers in Ecology and the Environment*, *18*(5), 254–260. https://doi.org/10.1002/fee.2207

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, *5*(11), 1198–1205. https://doi.org/10.1111/2041-210X.12261

699 Nadeau, C. P., Giacomazzo, A., & Urban, M. C. (2022). Cool microrefugia accumulate and conserve
700 biodiversity under climate change. *Global Change Biology*, *28*(10), 3222–3235.
701 https://doi.org/10.1111/gcb.16143

702 Pearce-Higgins, J. W., Beale, C. M., Oliver, T. H., August, T. A., Carroll, M., Massimino, D., … Crick, H. Q.
703 P. (2017). A national-scale assessment of climate change impacts on species: Assessing the
704 balance of risks and opportunities for multiple taxa. *Biological Conservation*, *213*(February), 124–
705 134. https://doi.org/10.1016/j.biocon.2017.06.035

706 Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box:
707 an open-source release of Maxent. *Ecography*, *40*(7), 887–893.
708 https://doi.org/10.1111/ecog.03049

709 Phillips, S. J., & Dudík, M. (2008). Modeling of species distribution with Maxent: new extensions and a
710 comprehensive evalutation. *Ecograpy*, *31*(December 2007), 161–175.
711 https://doi.org/10.1111/j.2007.0906-7590.05203.x

712 Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D.
713 (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty.
714 *SOIL*, *7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

715 R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Retrieved from
716 https://www.r-project.org/

717 Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions:
718 Complexity, overfitting and evaluation. *Journal of Biogeography*, *41*(4), 629–643.
719 https://doi.org/10.1111/jbi.12227

720 Randin, C. F., Engler, R., Normand, S., Zappa, M., Zimmermann, N. E., Pearman, P. B., … Guisan, A.
721 (2009). Climate change and plant distribution: local models predict high-elevation persistence.
722 *Global Change Biology*, *15*(6), 1557–1569. https://doi.org/10.1111/j.1365-2486.2008.01766.x

723 Roy-Dufresne, E., Saltré, F., Cooke, B. D., Mellin, C., Mutze, G., Cox, T., & Fordham, D. A. (2019).
724 Modeling the distribution of a wide-ranging invasive species using the sampling efforts of expert
725 and citizen scientists. *Ecology and Evolution*, *9*(19), 11053–11063.
726 https://doi.org/10.1002/ece3.5609

727 Sanczuk, P., De Lombaerde, E., Haesen, S., Van Meerbeek, K., Luoto, M., Van der Veken, B., … De
728 Frenne, P. (2022). Competition mediates understorey species range shifts under climate change.
729 *Journal of Ecology*, *110*(8), 1813–1825. https://doi.org/10.1111/1365-2745.13907

730 Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., … Crowley, J. (2022). *GGally:*
731 *Extension to "ggplot2."*

732 Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., … Reyer, C. P. O. (2017).
733 Forest disturbances under climate change. *Nature Climate Change*, *7*(6), 395–402.
734 https://doi.org/10.1038/nclimate3303

735 Slavich, E., Warton, D. I., Ashcroft, M. B., Gollan, J. R., & Ramp, D. (2014). Topoclimate versus
736 macroclimate: how does climate mapping methodology affect species distribution models and
737 climate change projections? *Diversity and Distributions*, *20*(8), 952–963.
738 https://doi.org/10.1111/ddi.12216

739 Srivastava, V. (2019). Species distribution models (SDM): applications, benefits and challenges in
740 invasive species management. *CAB Reviews: Perspectives in Agriculture, Veterinary Science,*
741 *Nutrition and Natural Resources*, *14*(020). https://doi.org/10.1079/PAVSNNR201914020

742 Stark, J. R., Fridley, J. D., & Gill, J. (2022). Microclimate-based species distribution models in complex
743 forested terrain indicate widespread cryptic refugia under climate change. *Global Ecology and*
744 *Biogeography*, *31*(3), 562–575. https://doi.org/10.1111/geb.13447

745 Svenning, J.-C., Normand, S., & Skov, F. (2008). Postglacial dispersal limitation of widespread forest
746 plant species in nemoral Europe. *Ecography*, *31*(3), 316–326. https://doi.org/10.1111/j.0906-
747 7590.2008.05206.x

748 van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J., & Raes, N. (2016). Minimum required number of
749 specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542–552.
750 https://doi.org/10.1111/ecog.01509

751 Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S, Fourth edition*. Springer, New
752 York.

753 Vollering, J., Halvorsen, R., Auestad, I., & Rydgren, K. (2019). Bunching up the background betters bias
754 in species distribution models. *Ecography*, *42*(10), 1717–1727.
755 https://doi.org/10.1111/ecog.04503

756 Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample
757 size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763–
758 773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

759 Wüest, R. O., Zimmermann, N. E., Zurell, D., Alexander, J. M., Fritz, S. A., Hof, C., … Karger, D. N. (2020).
760 Macroecology in the age of Big Data – Where to go from here? *Journal of Biogeography*, *47*(1),

761 1–12. https://doi.org/10.1111/jbi.13633

762 Zellweger, F., Coomes, D., Lenoir, J., Depauw, L., Maes, S. L., Wulf, M., … De Frenne, P. (2019). Seasonal
763 drivers of understorey temperature buffering in temperate deciduous forests across Europe.
764 *Global Ecology and Biogeography*, *28*(12), 1774–1786. https://doi.org/10.1111/geb.12991

765 Zimmermann, N. E., Edwards, T. C., Graham, C. H., Pearman, P. B., & Svenning, J.-C. (2010). New trends
766 in species distribution modelling. *Ecography*, *33*(6), 985–989. https://doi.org/10.1111/j.1600-
767 0587.2010.06953.x

768 Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., … Merow, C. (2020). A standard
769 protocol for reporting species distribution models. *Ecography*, *43*(9), 1261–1277.
770 https://doi.org/10.1111/ecog.04960

771

Supplementary material

Table S1: List of the study species, selected from the European forest vascular plant species list (Heinken et al., 2022). We report the amount of occurrence records used within the SDMs after data cleaning.

Species	Nr. of records	Species	Nr. of records
Aconitum lycoctonum subsp. lasiostomum	0	Hymenophyllum tunbrigense	24
Actaea erythrocarpa	18	Hypericum androsaemum	2595
Actaea spicata	9158	Hypopitys monotropa	3921
Adoxa moschatellina	2365	Impatiens noli-tangere	4548
Aegonychon purpurocaeruleum	422	Impatiens parviflora	12917
Allium ursinum	8970	Inula helvetica	149
Androsace chaixii	14	Knautia drymeia	117
Anemone ranunculoides	2818	Lamium galeobdolon	0
Anemone trifolia	154	Lamium galeobdolon	22345
Arabis turrita	558	Laser trilobum	43
Aremonia agrimonoides	22	Laserpitium nestleri	30
Arenaria procera	0	Lathraea squamaria	3812
Arum italicum	8219	Lathyrus cirrhosus	1
Arum maculatum	7331	Lathyrus niger	1653
Asarum europaeum	1892	Lathyrus venetus	55
Asperula taurina	69	Lathyrus vernus	8234
Brachypodium sylvaticum	12233	Limodorum abortivum	888
Bromopsis benekenii	1632	Lonicera alpigena	698
Bromopsis ramosa	1094	Lonicera nigra	1241
Calamagrostis chalybaea	378	Lunaria rediviva	1365
Calypso bulbosa	7	Luzula forsteri	1103
Campanula latifolia	1393	Luzula luzulina	84
Cardamine bulbifera	4186	Luzula nivea	1032
Cardamine enneaphyllos	392	Luzula pedemontana	19
Cardamine glanduligera	65	Maianthemum bifolium	14196
Cardamine heptaphylla	1768	Melampyrum bohemicum	23
Cardamine pentaphyllos	664	Melica picta	5
Cardamine trifolia	244	Melica uniflora	8231
Carex alba	615	Melittis melissophyllum	4174
Carex digitata	10190	Mercurialis ovata	28
Carex disperma	304	Milium effusum	11644
Carex elongata	2403	Moehringia muscosa	648
Carex fritschii	0	Moneses uniflora	3624
Carex loliacea	562	Neottia cordata	5301
Carex pendula	3397	Neottia nidus-avis	4685
Carex pilosa	150	Neottianthe cucullata	0
Carex remota	7905	Omphalodes scorpioides	16
Carex strigosa	367	Onoclea struthiopteris	0
Cephalanthera damasonium	1508	Orchis spitzelii	6
Cerastium sylvaticum	4	Orobanche hederae	384
Chimaphila umbellata	1514	Orobanche lucorum	5
Chrysosplenium oppositifolium	2490	Orobanche salviae	14

Cinna latifolia	898
Circaea alpina	1744
Circaea lutetiana	9695
Circaea x intermedia	275
Cirsium carniolicum	8
Clematis recta	102
Clinopodium menthifolium	486
Coptidium lapponicum	1156
Coronilla coronata	45
Corydalis cava	2087
Cyclamen hederifolium	1805
Cyclamen purpurascens	752
Cypripedium calceolus	710
Cystopteris sudetica	19
Dactylis glomerata subsp. lobata	999
Daphne laureola	4501
Daphne mezereum	12362
Dictamnus albus	524
Digitalis purpurea	22045
Diplazium sibiricum	96
Drymochloa drymeja	8
Drymochloa sylvatica	2544
Dryopteris remota	13
Epipactis albensis	0
Epipactis bugacensis	1
Epipactis fageticola	0
Epipactis greuteri	0
Epipactis leptochila	29
Epipactis microphylla	69
Epipactis nordeniorum	0
Epipactis phyllanthes	34
Epipactis placentina	0
Epipactis pontica	0
Epipactis purpurata	210
Epipactis tallosii	0
Epipogium aphyllum	144
Euonymus latifolius	245
Euphorbia amygdaloides	7039
Festuca flavescens	8
Gagea spathacea	263
Galium aristatum	45
Galium intermedium	3
Galium laevigatum	19
Galium odoratum	18818
Galium rotundifolium	1155
Galium sylvaticum	590
Galium triflorum	391

Orthilia secunda	7210
Oxalis acetosella	27704
Paeonia mascula	29
Paris quadrifolia	15002
Pilosella hybrida	0
Poa remota	778
Poa stiriaca	0
Polygonatum multiflorum	7309
Polystichum aculeatum	1935
Polystichum braunii	561
Polystichum setiferum	1662
Prenanthes purpurea	2892
Prunus lusitanica	322
Pulmonaria collina	0
Pulmonaria longifolia	573
Pulmonaria obscura	2081
Pulmonaria officinalis	1965
Pulmonaria saccharata	103
Pyrola chlorantha	3752
Pyrola media	1304
Pyrola minor	3735
Ranunculus cassubicus	187
Rhamnus alpina	313
Ribes spicatum	1528
Rumex sanguineus	3069
Ruscus aculeatus	10151
Sanicula europaea	5624
Saxifraga geranioides	14
Saxifraga hirsuta	211
Saxifraga umbrosa	92
Schedonorus giganteus	2742
Scilla lilio-hyacinthus	126
Scrophularia alpestris	115
Scrophularia peregrina	38
Scutellaria altissima	73
Sedum cepaea	41
Senecio hercynicus	169
Senecio nemorensis subsp. jacquinianus	8
Silene viridiflora	10
Soldanella montana	42
Spiraea japonica	234
Stachys sylvatica	13096
Staphylea pinnata	243
Stellaria longifolia	795
Stellaria nemorum	587
Symphytum cordatum	4
Trifolium rubens	254

Geranium lanuginosum	56	Trochiscanthes nodiflora	14
Geum sylvaticum	181	Valeriana pyrenaica	56
Goodyera repens	21442	Veronica montana	2714
Gymnocarpium dryopteris	9594	Veronica urticifolia	892
Hacquetia epipactis	103	Vicia dumetorum	136
Helleborus niger	957	Vicia pisiformis	126
Helleborus viridis	1162	Viola jordanii	1
Hepatica nobilis	28312	Viola mirabilis	2449
Hesperis sylvestris	3	Viola pseudomirabilis	0
Hordelymus europaeus	2768	Viola reichenbachiana	6726

775

Table S2: Final model structures of the Bayesian regression models indicating which transformations, distributions, priors and formulas were used (modeled with brms package). Phi, sigma and alpha are distributional parameters representing the precision (beta distribution), standard deviation (gaussian and skewed normal distribution) and skewness (skewed normal distribution) respectively. The adapt_delta argument is used to adapt sampling speeds (default value = 0.8), with higher values corresponding to slower sampling speeds and more robust to posterior distributions. Q05, optimum and Q95 refer to cold edge, optimum and warm edge of the bioclimatic variables.

RESPONSE		Transformation	Distribution	PRIORS	Formula
CBI		rescaled	beta	standard	CBI ~ type + (1 | species) phi ~ type
SENSITIVITY		/	beta	standard	sensitivity ~ type + (1 | species) phi ~ type
AREA		standardized	skewed normal	user-defined*	area ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
EDGES	Leading	standardized	skewed normal	user-defined*	leading ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
	Trailing	standardized	skewed normal	standard	trailing ~ type + (1 | species) sigma ~ type alpha ~ type
	Range	standardized	skewed normal	user-defined*	range ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
BIO1	Q05	standardized	skewed normal	standard	BIO1_Q05 ~ type + (1 | species) sigma ~ type alpha ~ type
	Optimum	standardized	skewed normal	standard	BIO1_opt ~ type + (1 | species) sigma ~ type alpha ~ type
	Q95	standardized	skewed normal	standard	BIO1_Q95 ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
	Width	standardized	skewed normal	standard	BIO1_width ~ type + (1 | species) sigma ~ type alpha ~ type
BIO5	Q05	standardized	gaussian	standard	BIO5_Q05 ~ type + (1 | species)
	Optimum	standardized	skewed normal	standard	BIO5_opt ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.99
	Q95	standardized	skewed normal	standard	BIO5_Q95 ~ type + (1 | species) sigma ~ type alpha ~ type
	Width	standardized	skewed normal	user-defined*	BIO5_width ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
BIO6	Q05	standardized	skewed normal	user-defined*	BIO6_Q05 ~ type + (1 | species)

					sigma ~ type alpha ~ type adapt_delta = 0.99
	Optimum	standardized	skewed normal	standard	BIO5_opt ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9
	Q95	standardized	skewed normal	standard	BIO6_Q95 ~ type + (1 | species) sigma ~ type alpha ~ type
	Width	standardized	skewed normal	standard	BIO6_width ~ type + (1 | species) sigma ~ type alpha ~ type adapt_delta = 0.9

*User-defined priors: c(set_prior("normal(0,3)", class = "b"), set_prior("normal(0,5)", class = "b", par="alpha"), set_prior("student_t(3, 0, 2.5)", class = "b", dpar="sigma"))

785 *Table S3: ODMAP protocol for Haesen et al.*

OVERVIEW	
Authorship	• **Authors:** Stef Haesen, Jonathan Lenoir, Eva Gril, Pieter De Frenne, Jonas J. Lembrechts, Martin Kopecký, Martin Macek, Matěj Man, Jan Wild, Koenraad Van Meerbeek • **Contact email:** stef.haesen@kuleuven.be • **Title:** Uncovering the hidden niche: incorporating microclimate temperature into species distribution models • **DOI:** Not applicable
Model objective	• **Objective:** Mapping / interpolation • **Target outputs:** logistic maps (i.e, continuous habitat suitability index) & binary maps (i.e., suitable vs. unsuitable habitat)
Taxon	forest plant specialist species
Location	Europe
Scale of analysis	• **Spatial extent (Longitude / Latitude):** 8°E - 50°W ; 33°N – 67°N • **Spatial resolution:** 25 m × 25 m (microclimatic data) & 1 km × 1 km (macroclimatic data & aggregated microclimatic data) • **Temporal extent / time period:** 2000 - 2020 • **Type of extent boundary:** rectangular
Biodiversity data	• **Observation type:** human observations (GBIF) • **Response / Data type:** presence-only
Type of predictors	climatic, edaphic
Conceptual model	• **Hypotheses about species-environment relationships:** As forests are known to buffer temperatures, we hypothesize that forest specialist plant species respond to warmer minimum temperatures and lower maximum temperatures as perceived by the free-air temperature data.
Assumptions	**We assumed that:** o Species are at (pseudo-) equilibrium with their environment. o The offset values from the long-term average period of 1970-2000 are still valid for the period 2000-2020. These are used to spatially-downscale the 4 km × 4 km TerraClimate data to 1 km × 1 km.
SDM algorithms	• **Algorithms:** We fitted MaxEnt models, which were chosen due to the presence-only character of the occurrence records. • **Model complexity:** MaxEnt models were built with linear, quadratic and product features. Regularization multipliers ranges from 0.5 to 5. • **Ensembles:** Not applicable
Model workflow	Only weakly correlated predictors were retained in the analysis. We

	performed parameter tuning for the features and regularization multipliers in MaxEnt models based on AICc, using the ENMeval 2.0 package. Model performance was assessed using the Continuous Boyce Index (CBI)
Software	• **Software:** Analyses were conducted in R version 4.1.1 and models were constructed using ENMeval v2.0.0 with maxnet package v0.1.4. • **Data availability:** The raw biodiversity data is available through https://doi.org/10.15468/dl.kf533a. ForestClim is freely-available through https://doi.org/10.6084/m9.figshare.22059125.
DATA	
Biodiversity data	• **Taxon names:** All species are listed in Table S1 • **Ecological level:** individual point data • **Data source:** GBIF (https://doi.org/10.15468/dl.kf533a) • **Sampling design:** random • **Sample size:** Amount of records per species are listed in Table S1 • **Regional mask:** We clipped the data to the boundary of the study area • **Data cleaning / filtering:** The occurrence data were filtered in the following sequential steps: (1) only records of 'human observations' were selected; (2) records with an unknown coordinate uncertainty or coordinate uncertainty larger than 25 m were excluded; (3) records located at country or capital centroids and biodiversity institutions were omitted; (4) duplicate records were removed; (5) only records observed during our climatic reference period (2000-2020) were selected; (6) records were spatially thinned to one random observation per 25 m × 25 m grid cell; and (7) we omitted species with less than 50 cleaned occurrence records. • **Background data:** Background data were generated by sampling an equal amount of background points as occurrence points based on a 2D kernel-density estimate of the occurrence points. • **Errors and biases:** We performed spatial thinning of occurrence points to account for sampling bias. Furthermore, background data was sampled according to a 2D-kernel density function, which introduces an equal amount of sampling bias within the background data as in the presence-only data. We only used records with a very low coordinate uncertainty (< 25 m) in order to minimize positional error on the occurrence points.
Data partitioning	We allocated 80% of the occurrence points to a spatial block cross-validation procedure, whereas 20% is kept for independent evaluation.
Predictor variables	• **Predictor variables:** ○ **Climate:** maximum temperature of the warmest month

	(BIO5), minimum temperature of the coldest month (BIO6), mean annual precipitation (BIO12), and precipitation seasonality (BIO15). ○ **Edaphic variables:** cation exchange capacity and soil clay content • **Data sources:** Macroclimatic data was collected from the TerraClimate database, whereas microclimatic data was downloaded from ForestClim. SoilGrids was used for the edaphic variables. • **Spatial resolution:** 25 m × 25 m (microclimatic data) & 1 km × 1 km (macroclimatic data & aggregated microclimatic data) • **Extent:** 8°E - 50°W ; 33°N – 67°N • **Geographic projection:** ETRS89/LAEA • **Time period:** 2000-2020 • **Data processing:** Bilinear interpolation was used in case data needed to be spatially-downscaled, whereas data aggregation (i.e., averaging) was used when data needed to be spatially-upscaled. All layers were reprojected to ETRS89/LAEA, if needed.
MODEL	
Variables preselection	We started from the conventional set of nineteen bioclimatic variables. Based on Booth (2022), we excluded bioclimatic variables combining both temperature and precipitation data (i.e., BIO8 = mean temperature of the wettest quarter; BIO9 = mean temperature of the driest quarter; BIO18 = precipitation of the warmest quarter; BIO19 = precipitation of the coldest quarter)
Multicollinearity	Highly correlated variables (Spearman correlation coefficients > 0.7) were removed from the analysis in order to reach the most parsimonious model. When excluding one of the correlated covariate pair, we retained variables which are known to be more important for plant species distribution and which are important for our further analyses (e.g., BIO5 & BIO6)
Model settings	MaxEnt models were built with linear, quadratic and product features. Regularization multipliers ranges from 0.5 to 5
Non-independence	We accounted for spatial autocorrelation by implementing a spatial block-cross validation.
Threshold selection	We used a 10% training presence as a threshold, meaning that the suitable area contains 90% of the original occurrence records
ASSESSMENT	
Performance statistics	• **Performance statistics estimated on validation data:** The Akaike Information Criterion for small sample sizes (AICc) was used to select the best candidate models • **Performance statistics estimated on testing data:** Model

	performance was assessed using the Continuous Boyce Index (CBI)
Plausibility checks	Maps of modelled predictions were checked by experts for an ad-hoc subset of species.
PREDICTION	
Prediction output	• **Prediction unit:** logistic maps (i.e, continuous habitat suitability index) & binary maps (i.e., suitable vs. unsuitable habitat) • **Post-processing:** Non-forested areas where masked out

786

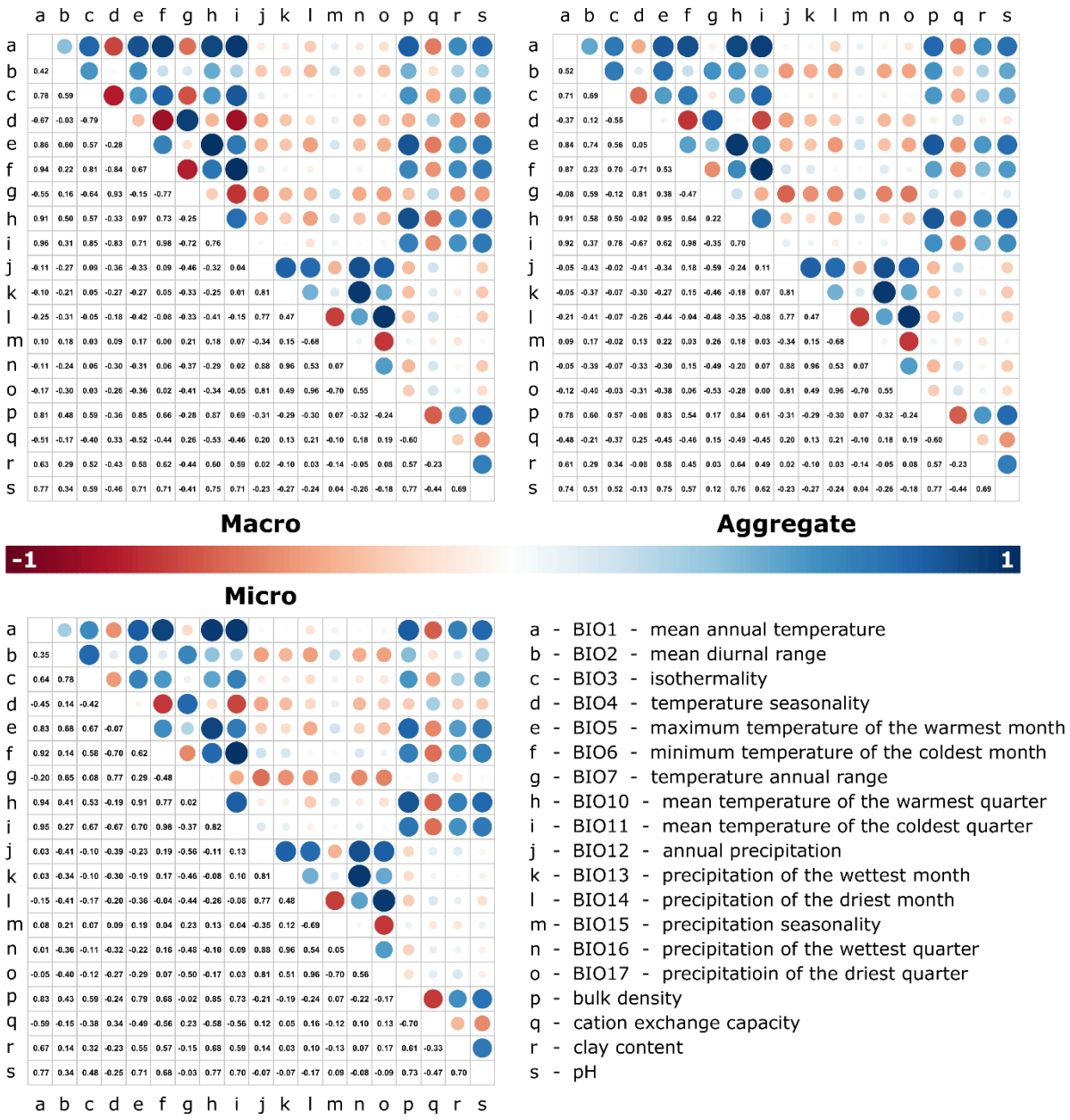
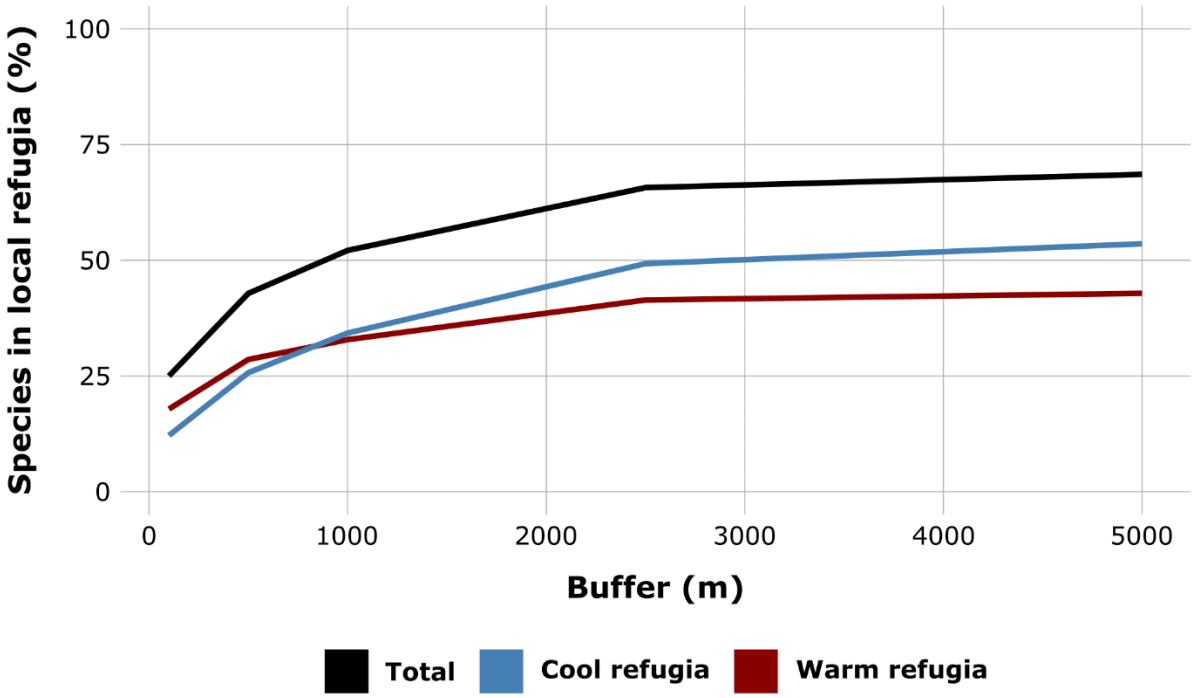


Figure S1: Correlation matrix containing all Spearman correlation coefficients (r^2) between all combinations of quantitative variables. Highly correlated variables (Spearman correlation coefficients > 0.7) were removed from the analysis (Dormann et al., 2013). The final selection of covariates encompassed two temperature variables (maximum temperature of the warmest month (BIO5) and minimum temperature of the coldest month (BIO6)), two precipitation variables (mean annual precipitation, (BIO12) and precipitation seasonality (BIO15)) and two edaphic variables (cation exchange capacity and soil clay content).



794

795 *Figure S2: Range of circular buffers (i.e., 100 m, 500 m, 1000 m, 2500 m, 5000 m) over which we quantified*
796 *whether or not a species was located in cool or warm refugia. For the trailing and leading edge, we extracted the*
797 *5% most southern and northern occurrence records, respectively. Using paired t-tests (α = 0.05), we compared*
798 *the local temperature conditions of these occurrence points to the surrounding microclimatic conditions over*
799 *these circular buffers around each occurrence record.*

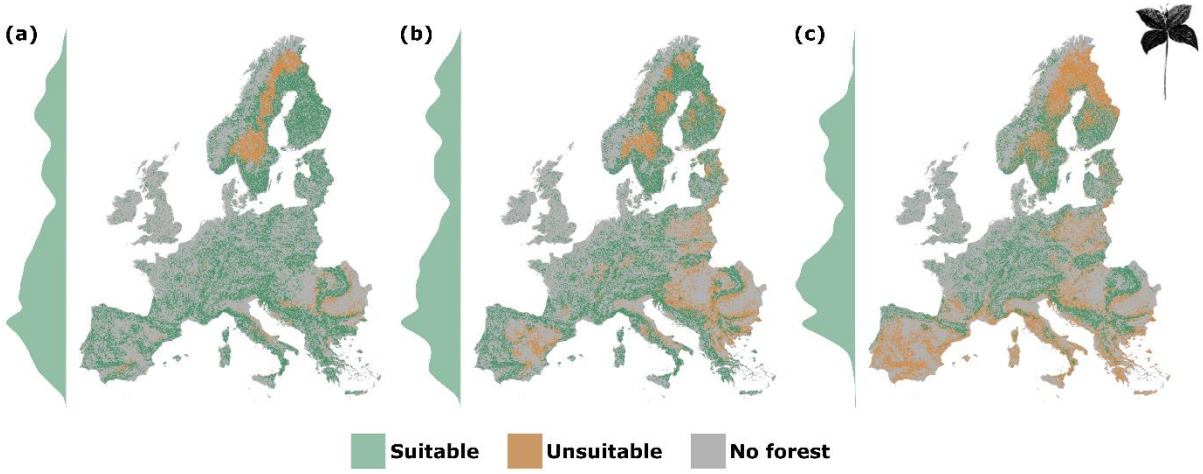


Figure S3: Binary maps indicating the potential suitable area for Paris quadrifolia. These maps are the result of species distribution modelled using (a) macroclimate data at a spatial resolution of 1 km × 1 km, (b) aggregated microclimatic data at a spatial resolution of 1 km × 1 km and (c) microclimatic data at the native spatial resolution of 25 m × 25 m over the 2000-2020 period. Density of suitable pixels along the latitudinal gradient is represented at the left side, respectively.