https://helda.helsinki.fi

EU-Trees4F, a dataset on the future distribution of European tree species

Mauri, Achille

2022-02-03

scientific **data**

OPEN

DATA DESCRIPTOR

EU-Trees4F, a dataset on the future distribution of European tree species

Achille Mauri[1,2 ✉], Marco Girardello[2], Giovanni Strona[1], Pieter S. A. Beck[2], Giovanni Forzieri[2], Giovanni Caudullo [2], Federica Manca[1] & Alessandro Cescatti[2 ✉]

We present "*EU-Trees4F*", a dataset of current and future potential distributions of 67 tree species in Europe at 10 km spatial resolution. We provide both climatically suitable future areas of occupancy and the future distribution expected under a scenario of natural dispersal for two emission scenarios (RCP 4.5 and RCP 8.5) and three time steps (2035, 2065, and 2095). Also, we provide a version of the dataset where tree ranges are limited by future land use. These data-driven projections were made using an ensemble species distribution model calibrated using EU-Forest, a comprehensive dataset of tree species occurrences for Europe, and driven by seven bioclimatic parameters derived from EURO-CORDEX regional climate model simulations, and two soil parameters. "*EU-Trees4F*", can benefit various research fields, including forestry, biodiversity, ecosystem services, and bio-economy. Possible applications include the calibration or benchmarking of dynamic vegetation models, or informing forest adaptation strategies based on assisted tree migration. Given the multiple European policy initiatives related to forests, this dataset represents a timely and valuable resource to support policymaking.

Background & Summary

Covering 35% of EU land[1], forests play a fundamental economic and ecological role. Besides their obvious contribution to biodiversity and the provision of wood and non-wood products, forests maintain a wide range of ecosystem services, such as carbon storage and sequestration, habitat provision, and water regulation[2–4]. Nonetheless, forests are increasingly under threat from habitat fragmentation, the spread of invasive alien species, climate change, water scarcity, fires, storms, and pests[5,6]. By the end of the century, climate change alone will substantially alter the current distribution of climatically suitable areas for the majority of European trees species (Fig. 1), generating severe mismatches between species' niches and the local climatic conditions[7–9]. This might result in both the erosion of current species ranges and colonization of newly suitable areas.

Considering the relatively low dispersal ability of most European tree species, it is unlikely that natural dispersion will permit forests to compensate for range erosion by colonizing new territory[8,10–12]. Targeted forest management appears an obvious and realistic option to minimize the loss of local biodiversity and ensure the continued delivery of forest ecosystem services[13,14]. This is particularly so in Europe, where forests are very far from naturalness;[15] they have been managed for millennia through clearance to create croplands and pastures, and by intensive tree collection for fuelwood and construction materials[16,17]. Even now European forests are being managed rather intensively[1,18,19]. However, current management is mostly driven by economical considerations, while we strongly argue that ensuring a broad range of future forest ecosystem services in Europe urgently calls for a science-based change of direction to "design" forests capable to withstand environmental change while bringing economic, social and ecological benefits to humans and natural systems. Climate could change significantly in a given locality during the lifespan of an individual tree, therefore, forest management needs to consider not only the compatibility of target tree species with present climatic conditions but also with the climate expected in the near future[20–22]. Here we present a new dataset that, amongst others, might help forest managers to tackle these challenging issues. In particular, we provide current and future (to the end of the 21st century) distribution maps for 67 tree species in Europe under different modelling and climatic assumptions.

[1]Faculty of Biological and Environmental Sciences, Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland. [2]European Commission, Joint Research Centre, Ispra, Italy. ✉e-mail: achille.mauri@helsinki.fi; Alessandro.CESCATTI@ec.europa.eu
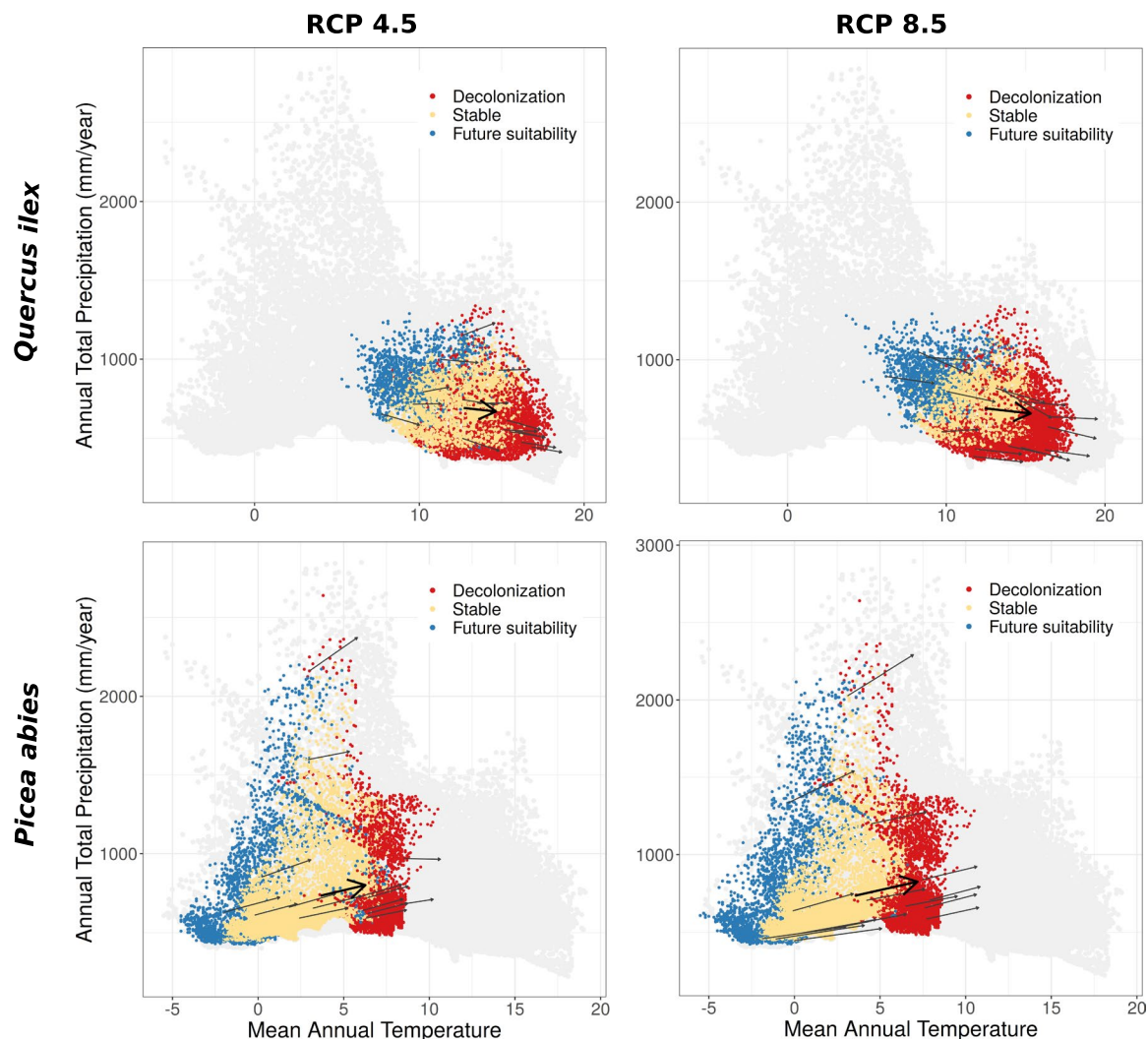
Fig. 1 Two examples of the expected mismatch between the present climatic conditions in which tree species currently grow, and the climate they will experience by the end of the century under two contrasting warming scenarios (RCP 4.5 and RCP 8.5). Black arrows represent the variations in mean annual temperature and annual precipitation for fifteen random climate bins while the thick black arrow represents the average variations for the entire climate envelope of the species. Climate bins are shown with different colours: yellow, for stable areas where by the end of the century the tree species will still have a suitable climate; red, for decolonized areas where the tree species will no longer be climatically suited; blue, for future suitability areas that are not currently colonized by the tree species but will become suitable in the future. Grey dots show all combinations of mean annual temperature and total precipitation for the present and the future and for all 67 species studied here. The targeted species (*Quercus ilex* and *Picea abies*) were arbitrarily chosen from the full set of tree species.

The dataset has various features which make it stand out from available products. First, previous studies aiming at investigating the impact of climate change on tree species in Europe largely focused on few (<15) commercially important tree species[23–25] (but see Thuiller *et al.*[26] for a broader perspective). This might result in excluding tree species important for biodiversity, ecosystem functioning, wildlife habitat, and in turn ecosystem services. However, there is now increasing evidence that, in a climate change context, a greater number of plant species is vital to reduce vulnerability, guarantee ecosystem functioning and the future delivery of ecosystem services[27–29]. In fact, tree diversity is key to enhancing resilience of forest communities to climate-driven risks and disturbances, in particular when environmental conditions are rapidly changing[30,31]. Therefore, *EU-Trees4F* provides a more comprehensive view by mapping current and future ranges for a large number (67) of European tree species (Table S1).

Second, previous projections of future tree species ranges typically relied on bioclimatic parameters downscaled from coarse (grid cells greater than 100 km) simulations by one or more global climate models. Conversely, *EU-Trees4F* takes advantage of outputs of regional climate models (for Europe EURO-CORDEX)[32] at a higher spatial resolution which have now become available.

Third, most previous studies focused on the potential distribution of tree species regardless of dispersal constraints, i.e. they provide suitable future areas of occupancy without making hypotheses about future
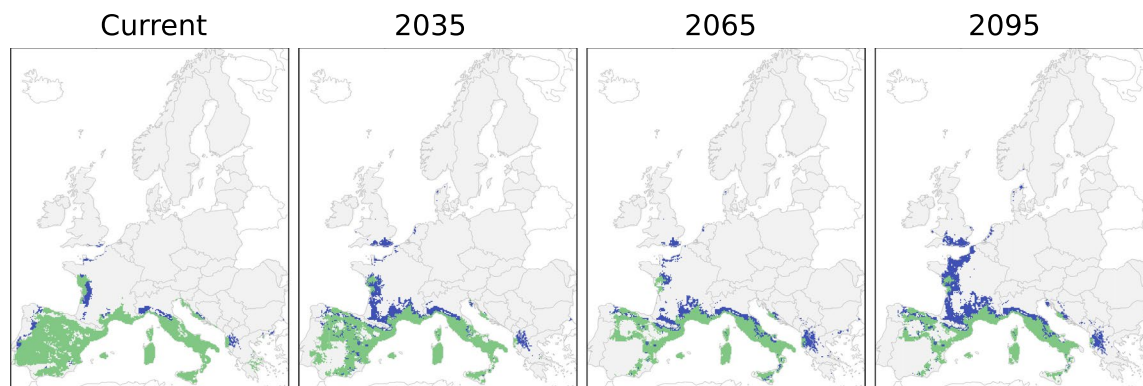
Current	2035	2065	2095

Fig. 2 Comparison between the *Querus ilex* potential suitable range (according to ensemble SDM projections, green + blue) and the expected distribution as simulated using a dispersal model (Migclim, green) for an RCP 4.5 emission scenario. The blue area represents the potential suitable range that is not occupied by *Quercus ilex* due to dispersal limitations. The targeted species (*Quercus ilex*) was arbitrarily chosen from the full set of tree species.

colonization patterns[23–26,33–37]. Here, we provide both the potential future suitable areas of occupancy as limited only by climate and soil, and the expected future tree species distribution under the assumption of natural dispersal. In addition, we provide a version of the dataset where tree ranges are limited to future modelled land use[38]. Combined, these two pieces of information provide a useful tool to simulate management scenarios regarding the future of European forests. In fact, the potential suitable area set the climate boundaries for targeted forest management, while the natural dispersal scenario defines the likely trajectory in case of no human intervention on species distribution. The difference between the two scenarios can therefore be interpreted as the room for manoeuvre of forest adaptation strategies based on the assisted migration of tree species.

The dataset includes future distribution maps corresponding to three 30-year periods, centred on 2035, 2065 and 2095, and modelled for two emission scenarios (RCP 4.5 and RCP 8.5) using an ensemble forecasting framework (Fig. 2 and Figure S1). Models were trained using the most comprehensive dataset of forest tree species occurrences in Europe currently available, EU-Forest[39,40]. Future projections were created using a set of 11 regional climate models (RCMs) from EURO-CORDEX[32], downscaled to 5 arc-minutes (~10 km) spatial resolution.

This dataset has a wide array of potential applications in various research fields, including forestry, bio-economy, biodiversity and ecosystem services. We envisage that *EU-Trees4F* will facilitate active forest management for climate adaptation (including assisted colonization strategies)[41] that addresses the balance between economic forest productivity (which currently hinges on rather few commercially exploited tree species) in the shorter or longer-term, the provision of non-economic ecosystem services, and the resilience of forest ecosystems to future environmental perturbations[22,30,31]. In addition, *EU-Trees4F* could contribute to biodiversity conservation and management to forecast changes in tree species richness through time[42]. In fact, the projected decline in tree species richness by the end of the 21st century, might provide spatial information to foresters and practitioners concerning the areas that may require assisted recolonization[19]. Another possible application could be in forest pest management by providing spatial and temporal distribution of host tree species for harmful pathogens[43]. In addition, *EU-Trees4F* could serve as a benchmarking dataset to calibrate and/or evaluate the output from dynamic vegetation models[37]. Finally, our dataset could be used to support policy-making, given the current European Commission's need to fulfil the European Green Deal's[44] objectives, the EU Biodiversity strategy 2030[45], the EU Bio-Economy strategy[46], and the new EU Forest Strategy[47].

Methods

We produced a dataset of tree species distribution maps using a framework for species distribution modeling (SDM, BIOMOD2[48]), driven by a large database of tree species occurrences available at pan-European scale[39]. We included in our analysis 67 tree species that cover a broad range of life histories and climatic tolerances, ecosystem functions, ranging from the Mediterranean to the Boreal region (Table S1). We built an ensemble of tree species distribution models based on a set of nine environmental parameters describing key features of climate and soils (Table S2). We projected them into the future using 11 regional climate models from EURO-CORDEX, for two emission scenarios (RCP 4.5 and RCP 8.5), downscaled to a spatial resolution of 5 arc-minutes (Fig. 3). Additionally, we provided a version of the dataset masked by future land use[38].

Tree occurrences. The core of our datasets is the EU-Forest data set that includes the best-quality data on forest tree species occurrences available in Europe[39,40]. We complemented EU-Forest with data from intensive monitoring plots (ICP Forests[49]) to fill in some geographical areas not well represented for a few species (Fig. 4). For Poland, which is an area with particular climate conditions not common elsewhere in our study domain, we enriched our dataset with occurrences obtained from Zając *et al.*[50].

We considered as trees those species that have a defined crown and a single main stem, as defined in Gschwantner *et al.*[51] for National Forest Inventory surveys in Europe. We included *Corylus avellana*, for its
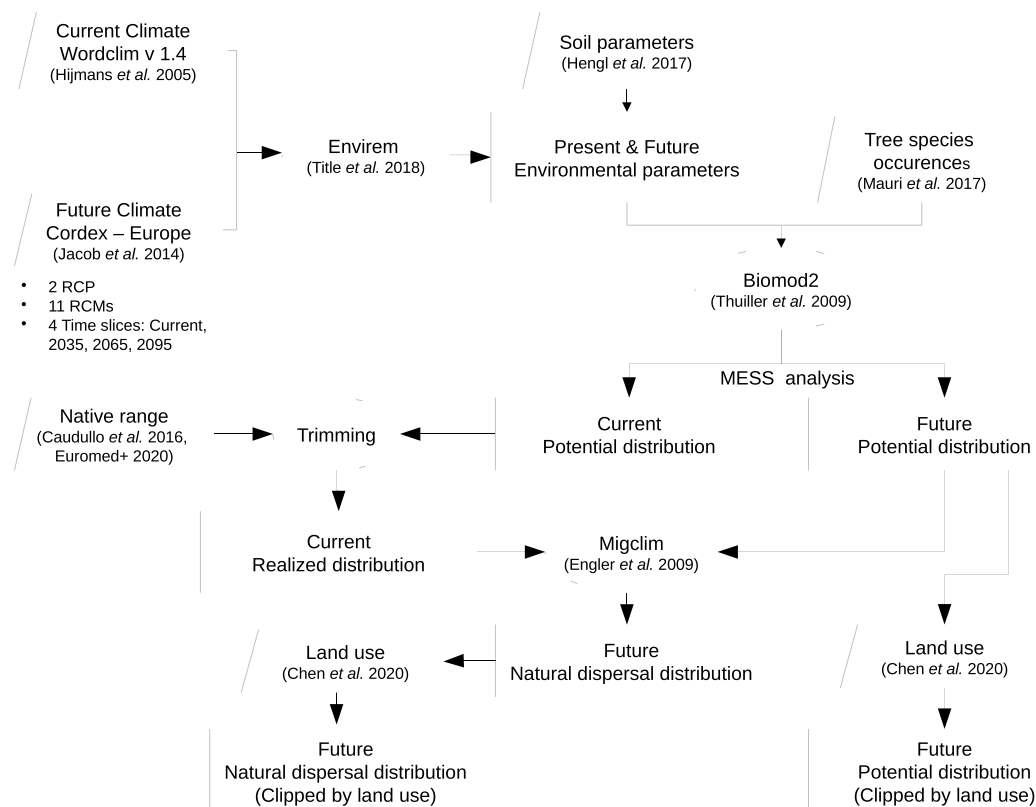
Fig. 3 Flowchart illustrating the various steps undertaken to produce the distribution maps.
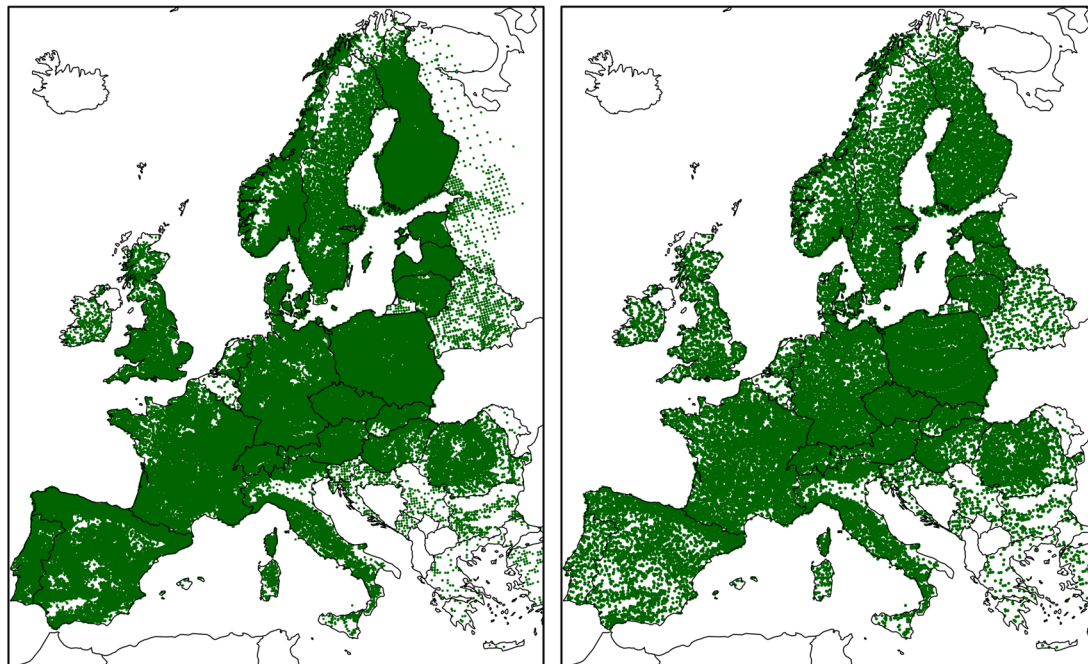

Fig. 4 Spatial distribution of all the tree species occurrences (left panel) and the thinned occurrences used to train the species distribution models (right panel).

importance in terms of ecosystem functioning and services. As a result, we retain a set of 67 tree species totalling 582,066 occurrences. We note that this set covers only a portion of the entire pool of native tree species available in the study domain (443)[52], but it includes all the tree species with significant commercial value as well as many

tree species known for their importance for ecosystem functioning. However, some rare endemic tree species such as *Abies nebrodensis*, which is found in only a few forest patches in Sicily, are not included in our species list, while other tree species that are not commercially exploited but have a broader distribution (e.g. *Aria edulis, Pinus brutia, Prunus padus, Quercus coccifera, Sorbus aucuparia, Taxus baccata*) are included.

Prior to the analyses, we implemented a data thinning procedure to reduce spatial bias within the species occurrence records. First, we removed all the duplicates of the same tree species falling in each 5 arc-min (~10 km) cell. Second, similarly to Dyderski *et al.*[23], we randomly selected a single occurrence in every 40 × 40 km grid cell and discarded the others. In this way, we overcame the problem of uneven sampling intensity in the occurrence datasets used[53], and ensured that observations were evenly distributed within the geographical area of interest (Fig. 4). It is important to note that the environmental parameters used to develop the species distribution models with these thinned points describe the local surroundings (~ 10 km cell) rather than the ~ 40 km cells used for the thinning.

Environmental data. We selected climate and soil parameters considered critical to plant physiological functioning and survival[54–56] and used in several earlier large-scale species distribution models[26,57–59]. These are winter and summer temperature (°C) and precipitation (mm/month), precipitation seasonality, mean annual temperature (°C), mean temperature of the coldest month (°C), total annual precipitation (mm/year), continentality, a humidity index (Alpha), growing degree days above 5 °C (GDD5), soil pH and organic carbon content (OCC) (Table S2). Four parameters (mean temperature of the coldest month, growing degree days above 5 °C, winter temperature, and the humidty index) were excluded as a result of multi collinearity tests made using the 'usdm' R package[60] (Figure S2).

For the current period, the climatic parameters were derived from the Worldclim climatology Version 1.4[61] with a spatial resolution of 5 arc-min (~10 km). It covers the period from 1961 to 1990, when climate conditions were more representative of the conditions experienced by the trees recorded in our occurrence dataset at the time of their establishment. GDD5, continentality and the humidity index (the ratio between actual and potential evapotranspiration) were calculated using the *envirem* library[62].

For the future, we used 11 regional climate models simulations (RCMs) sourced from the Coordinated Regional Downscaling Experiment (CORDEX) of the World Climate Research Programme (WCRP). The EURO-CORDEX[32] initiative, which is part of the CORDEX project, provides regional climate projections for Europe at ~12.5 km horizontal resolution (Table S3). These were downscaled to a 5 arc-minutes resolution (~10 km) using a change factor approach[63,64], a widely used method in climate change impact assessments[65–69]. First, we computed monthly values from daily temperature and precipitation. From these, we calculated climate anomalies between the future and the control period, by applying a multiplicative correction for precipitation and an additive correction for temperature. The control period (1961–1990) is matching the period on which the Worldclim Version 1.4 climatology was computed. Climate anomalies were then gridded to a finer spatial resolution of 5 arc-minutes using bilinear interpolation, before adding them back to baseline Worldclim climatology. In addition, a climatic ensemble mean of the output of the 11 RCMs was calculated and used to project the species distribution into the future. We focused our analysis on two representative concentration pathways (RCP 4.5 and RCP 8.5) depicting the greenhouse gas concentration by the end of the 21st century. The RCP 4.5 is an intermediate scenario corresponding to a projected change in global mean surface air temperature by the end of the century of + 1.8 °C relative to the reference period of 1986–2005[70] and with greenhouse gases concentration stabilizing shortly after 2100. Instead, the RCP 8.5 is a business-as-usual scenario with steady increases in greenhouse gas concentrations resulting in a global mean surface air temperature increase of 3.7 °C by the end of the century relative to 1986–2005. Organic carbon content (g per kg) and soil pH, in the first 15 cm of topsoil, were extracted from the 1- km spatial resolution SoilGrids dataset[71] and aggregated to a lower (5 arc-min) spatial resolution.

Species distribution modelling. We modelled the potential distribution for each tree species using a well-accepted platform for ensemble species distribution modelling (BIOMOD2[48]) that is widely used for investigating the impact of climate change on forests[26,34,35,37,72]. We used six species distribution models: Generalized Linear Models (GLM), Generalized Additive Models (GAM), Generalized Boosting Models (GBM) or usually called Boosted Regression Trees, Multiple Adaptive Regression Splines (MARS), Maximum Entropy (Maxent) and Random Forest (RF). We used default settings except in MAXENT, where we sought to avoid overparameterization by setting product, threshold and hinge equal to false, as suggested in Merow *et al.*[73]. For each tree species, we calibrated the model on the thinned current occurrences and we predicted the potential distribution as a probability map, which values were converted for memory savings into integers ranging between zero and one thousand. Finally, for each tree species, we computed a consensus model by averaging the individual model's projections (Fig. 5). Presences and absences were weighted equally by setting the prevalence parameter of BIOMOD2 to 0.5. We selected for each species 10,000 pseudoabsence points outside of its suitable area, as estimated from a surface range envelop model derived from the climatic predictors[74]. We projected the models (calibrated for current conditions) into the future using two emission scenarios (RCP 4.5 and RCP 8.5) and three 30-year time periods (centred on the years 2035, 2065, and 2095). Projections into current and future climates were made using two approaches: 1) A climatic ensemble mean that projects the consensus model into current and future conditions using the average output of the 11 RCMs, and 2) an 'SDM' ensemble mean that projects the consensus model for every single RCM, and a posteriori average of the outputs of the 11 SDMs. We calculated the 'SDM' ensemble mean for each grid cell, when at least 8 out of the 11 values were not missing data[75] (Fig. 5). In addition, we translated probabilistic ensemble forecasts into binary maps using the same True Skill Statistics score as calculated for current predictions.
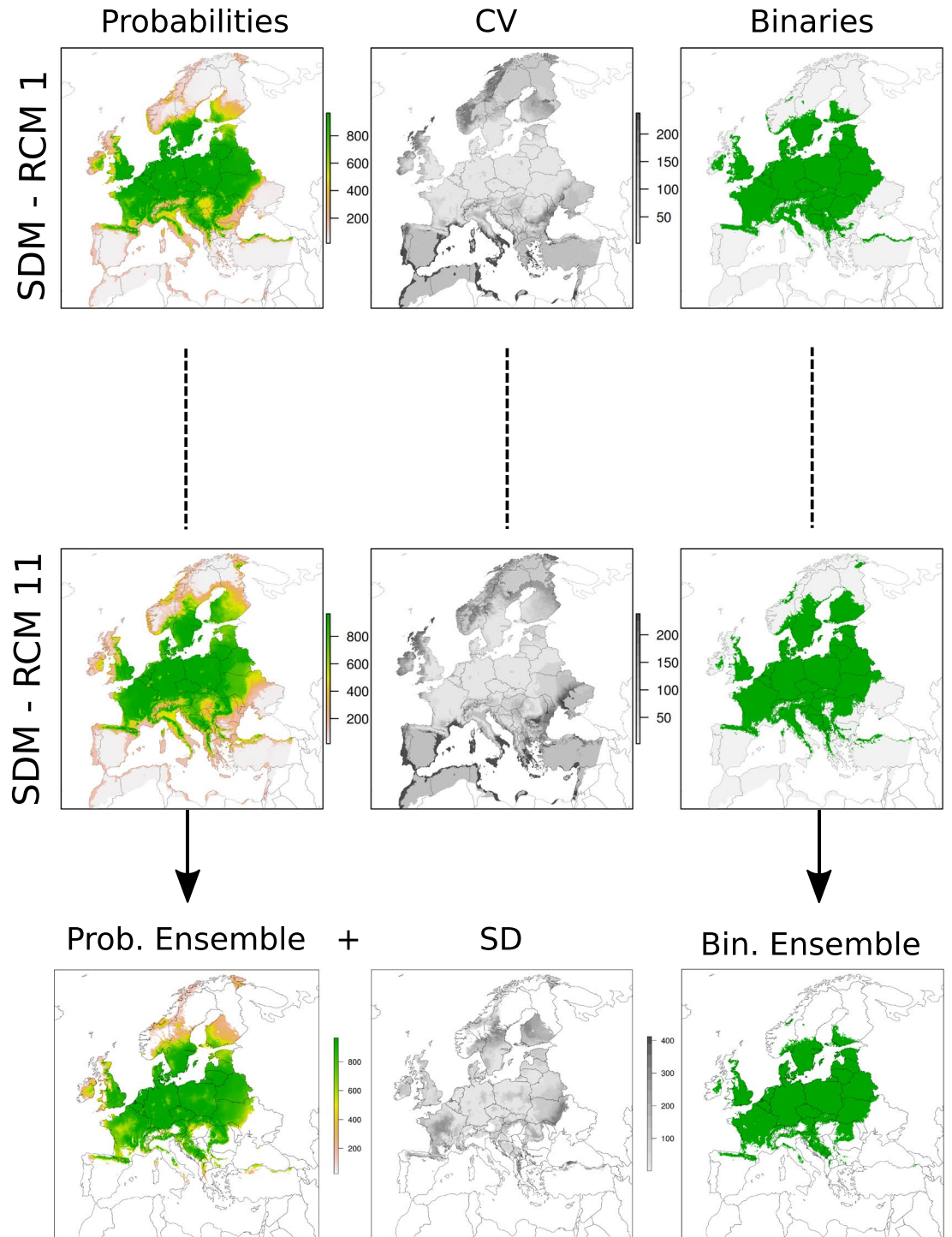
Fig. 5 Illustrative example of the type of data produced in our modelling workflow. The results are shown for a single tree species *"Acer campestre"*, for a single time slice (2095) and for the emission scenario RCP 8.5. The upper and central panels show the output from BIOMOD2 run using the 11 regional climate models (RCMs). This includes the projection as probability maps (*Probabilities*) from the consensus model (average of the seven species distribution models implemented in BIOMOD2), the associated coefficient of variation (*CV*) among the seven SDMs, and the binary maps (*Binaries*). The lower panel shows the average value of the 11 consensus projections with associated standard deviation (*SD*) and ensemble binary map. In addition (not illustrated here), we ran BIOMOD2 using the average output of the 11 RCMs.

To the model projections in future climates, we applied the same 30-year time difference, as we had for the current projection, which was simulated with a climatology centred on the year 1975 but was referring to forest
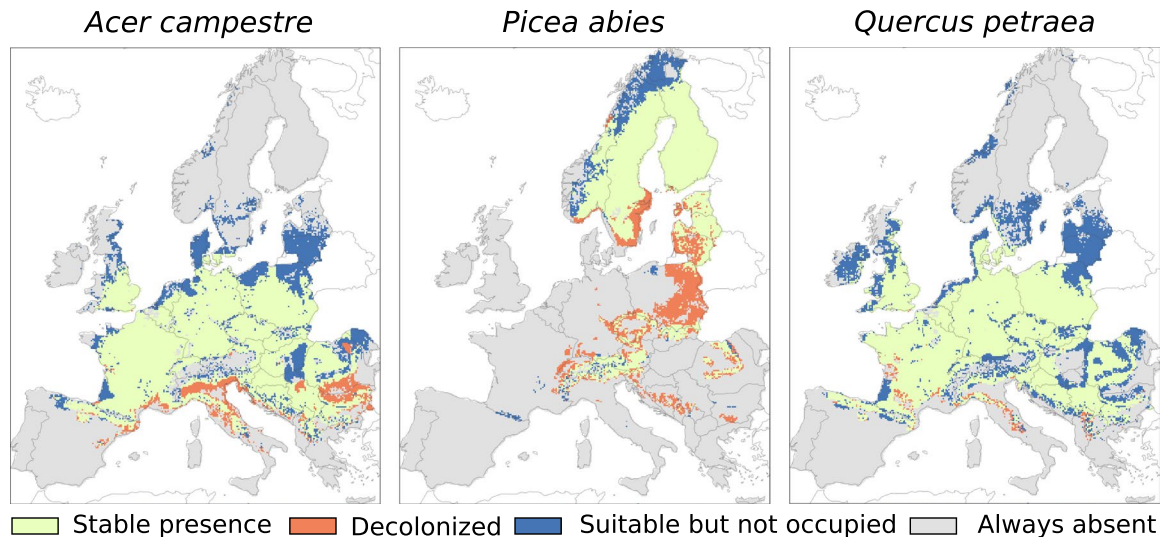
Fig. 6 Illustrative example of Migclim output for three species modelled until the end of the century (2100) and for an emission scenario RCP 4.5. 'Stable presence' denotes areas that will remain suitable habitats from the present until the end of the century. 'Decolonized', indicates areas that will become climatically unsuitable by the end of the century. 'Suitable but not occupied', represents areas that will become climatically suitable by the end of the century but will not be naturally colonized due to dispersal limitations. The targeted species (*Acer Campestre*, *Picea abies* and *Quercus petraea*) were arbitrarily chosen from the full set of tree species.

occurrences collected around the year 2005. The rationale behind this is that the tree occurrences in the calibration data set were recorded in the landscape around the year 2005 but mostly established themselves under the climatic conditions of the previous decades. Therefore, we assumed that the distribution of a species around the years 2035, 2065, and 2095 is predominantly constrained by the climatic conditions of the previous 30 years too.

In summary, for each tree species, we computed six SDMs and one consensus ensemble model. These were projected into current and future (three time periods) using two emission scenarios (RCP 4.5 and RCP 8.5). Future projections were made for 11 RCMs, either using a single climatic ensemble (average of the 11 RCMs) or by averaging the 11 consensus SDM output run using the single RCM output. Results produced from single climate model realization are provided in *EU-Trees4F* as well.

Realistic dispersal scenario. We considered a realistic dispersal scenario that allows species to move from their initial position into climatically suitable areas according to their natural dispersal capacity (Fig. 6). We implemented the scenario using MigClim, a cellular automation dispersal model (MigClim[76,77]) that simulates dispersal, colonization, growth, and local extinction. MigClim used, for each tree species, the current and future binary distribution, as derived from BIOMOD2, for three time periods centred on the years 2035, 2065 and 2095 and a dispersal step of 30 years, which approximates the age at which most tree species become reproductive. As in Merow *et al.*[78], we incorporated a spatial prior into our models, by trimming tree species occurrences by their native range distribution developed based on expert knowledge[79,80]. These constrained the predictions for the present in order to account for factors that were not included in the model covariates (e.g. biotic interactions or dispersal limitations).

Species dispersal was modelled using dispersal kernels driven by mean dispersal distances in combination with propagule production potential from the time a cell became colonized. We used a dispersal kernel based on a negative exponential function as implemented in other studies investigating plant species distribution in a changing climate[77,81–83]. Since species-specific mean dispersal distances are sparse in the literature, and not available for the entire set of tree species, we estimated them from maximum dispersal distances (MDD) using the following formula (Mean dispersal distance = $10 \wedge (\log 10 \text{ MDD} - 0.795)/0.984$) derived from Tamme *et al.*[84] and Thompson *et al.*[85]. We computed maximum distance dispersal using the dispeRsal algorithm (version 0.2)[84] and for a few tree species we estimated it on the basis of authors' knowledge of the species' reproductive traits. The algorithm computes maximum dispersal distances using a linear mixed-effects model on the following functional traits downloaded from the TRY database[86]: species dispersal syndrome (in our case: wind, animal or no particular mechanism), growth form (tree), seed mass, realizing height (approximated to tree height) and terminal velocity when available (Table S1). Before the analysis, we $\log_{10}$ transformed seed mass, maximum dispersal distance and maximum plant height data. The maturity age was set to 1 since we were using a time step of 30 years, which is about the maturity age for most tree species. Propagule production was set to 1, which assumes that all mature trees will produce propagules. To assess variability due to randomness, for each species, we produced, and averaged, 30 replicates of the dispersal model.

Land use datasets. We trimmed by land use the maps of future climatic suitability (potential occupancy) and the maps of future distribution expected under a scenario of natural dispersal. This was done for two emission scenarios (RCP 4.5 and RCP 8.5) and three time steps (2035, 2065, and 2095). We used a global land use

	Pse	Occ	pH	Pw	Ptot	Ps	Ci	Ts	Mat
Mean	0.13	0.17	0.21	0.25	0.28	0.35	0.46	0.49	0.55

Table 1. Relative importance of climatic and soil parameters averaged for the entire pool of tree species. Pse = precipitation seasonality; Occ = Organic carbon content; pH = soil pH; Pw = winter mean precipitation; Ptot = total annual precipitation (mm/year); Ps = summer mean precipitation; Ci = continentality; Ts = summer temperature (°C); Mat = mean annual temperature (°C).

dataset[38] with a spatial resolution (3 arc-min), which was aggregated at the spatial resolution of our dataset (5 arc-min). We masked our dataset by the forest layer, which was derived by aggregating eight land use types as in Chen *et al.*[38], specifically, needleleaf evergreen and deciduous trees, and broadleaf evergreen and deciduous trees from temperate, boreal and tropical regions. As these layers are presented as a cover fraction, we binarized them using a threshold of 40%, above which we consider to be forest. For the RCP 4.5 we used the SSP2 Socioeconomic Pathway scenario that represents a world where trends broadly follow their historical patterns with medium challenges to mitigation and adaptation ("a middle of the road").

Data Records

EU-Trees4F is available in GeoTIFF format at a resolution of 5 arc-minutes in LAEA (EPSG:3035) coordinate reference system. The files are freely accessible through *Figshare*[87] (https://doi.org/10.6084/m9.figshare.c.5525688) and https://forest.jrc.ec.europa.eu.

As illustrated in the main text and in Fig. 5, projections into future climates were made using two approaches: 1) A climatic ensemble mean that projects the consensus model from Biomod2 into future conditions using the average output of the 11 RCMs. These maps are stored in a binary format in the folder entitled "ens_clim"; 2) An 'SDM' ensemble mean that projects the consensus model for every single RCM, and a posteriori averages of the output of the 11 SDMs. These maps are stored in a binary format in the folder entitled "ens_sdms". For the 'SDM' ensemble mean, we provided as well probability maps as GeoTIFF rasters in a WGS84 (EPSG:4326) reference system, together with the associated standard deviation calculated from the 11 consensus models for every single RCM.

In a third directory, entitled "single_models" we included GeoTIFF rasters in a WGS84 (EPSG:4326) reference system for the single climate species distribution model realizations. Here, there are three subdirectories: 1) "bin", representing binaries distributions maps, 2) "prob" and 3) "CV", representing respectively probabilities distribution maps and associated coefficient of variation maps. The coefficient of variation is relative to the six species distribution models implemented to produce the consensus projection from BIOMOD2. As an example, for a single tree species there are 66 GeoTIFF rasters maps in each subdirectory. These are based on 11 consensus SDM realizations (one for each RCM), two emission scenarios (RCP 4.5 and RCP 8.5), and three periods in the future (2035, 2065, 2095).

Finally, in the fourth folder entitled "datasets", there are all the datasets needed to reproduce *EU-Trees4F*. In the "species occurrences" directory there are two subdirectories entitled "p_ICP", which includes data from ICP-Forests, and "p_Poland" which includes additional occurrences from Poland. We merged these two datasets to the EU-Forest[39,40] and we trimmed the merged occurrences by the species native ranges[79,80] as described in the main text. The resulting dataset, in the subdirectory "input_sdm" was used as input to BIOMOD2 together with the environmental parameters included in the folder "climate" and "soil".

Technical validation

Tree occurrences. We excluded from our analyses alien tree species, with the exception of *Robinia pseudoacacia* which is a highly naturalized tree species in Europe[88]. We further excluded tree species that had fewer than 30 occurrences as well as those occurrences located in areas that are climatically suitable for the species to survive but not to reproduce. For instance, *Quercus ilex* is present and surviving in Scotland because it has been planted there, but the climate doesn't allow it to reproduce. To keep such cases from skewing our projections, we filtered out occurrences outside a species' natural range as documented in chorological maps[79]. For the few species without detailed chorologies, we used the native distribution information taken from the Euro + Med PlantBase[80] at the country level (accessed date: 11–11–2019). For *Robinia pseudoacacia*, we considered all the occurrences in the dataset. Overall we retained a set of 67 tree species totalling 589,862 occurrences. We further filtered the dataset by applying a thinning procedure aimed at removing spatial biases in the dataset. Altogether we retained 50,453 thinned occurrences that were used to train the species distribution model (Fig. 4).

Environmental parameters assessment. We assessed the relative importance of individual environmental parameters to predict the distribution of each tree species. The assessment relied on a permutation procedure, as described in Thuiller *et al.*[48]. We calculated the Pearson correlation coefficient (r) between the model predictions and a prediction generated when the climatic variable in question was randomly permutated. We repeated the procedure ten times and kept the mean of all the r values. A high r (little difference between the two predictions) means that the permutated variable is not important for the model. Conversely, a low value reflects a significant difference between the predictions and therefore a high importance of the variable. The importance is therefore expressed as 1- r. Table 1 presents the relative importance of the individual environmental parameters averaged across the entire set of tree species, while Fig. 7 presents the results for all the species modelled.

SDM evaluation and uncertainties. We evaluated the models' predictive performances through a block-cross validation[89], which is a particularly suited validation method to assess model transferability in
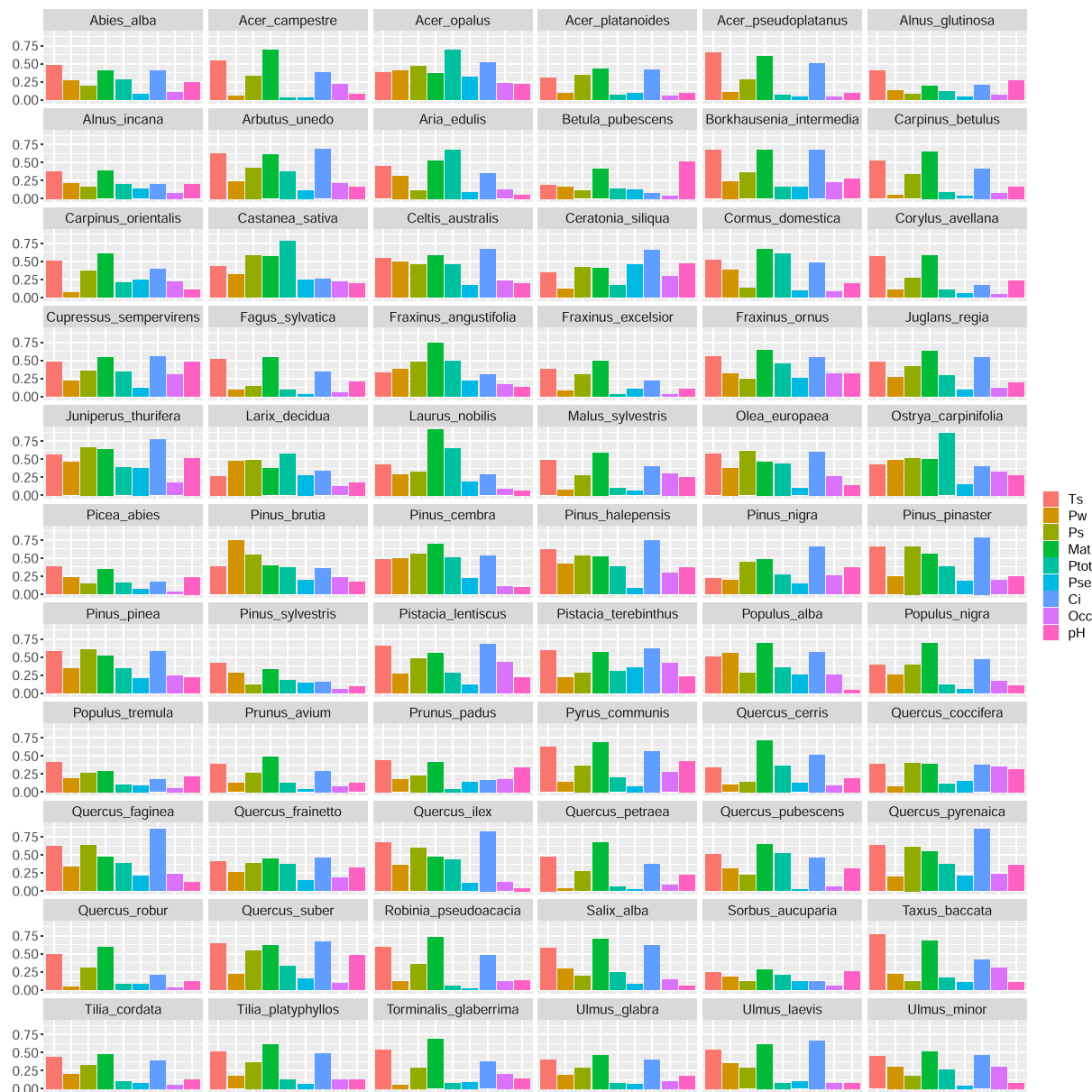
Fig. 7 Relative importance of environmental parameters in the distribution models as calculated using the BIOMOD2 package. Ts = summer temperature; Pw = winter precipitation; Ps = summer precipitation; Mat = mean annual temperature; Ptot = total annual precipitation; Pse = precipitation seasonality; Ci = continentality; Occ = Organic carbon content; pH = soil pH.

	GAM	GBM	GLM	MARS	ME	RF	MEAN
TSS	0.73	0.77	0.74	0.73	0.75	0.75	**0.74**

Table 2. Overall model evaluation using the True Skill Statistics (TSS), averaged over the full set of tree species. Generalized Additive Models (GAM); Generalized Boosting Models (GBM); Generalized Linear Models (GLM); Multiple Adaptive Regression Splines (MARS); Maximum Entropy (Maxent), Random Forest (RM); Mean of all models (MEAN).

geographic and environmental space where non-analog climatic conditions might be present[89]. Model validation results are reported using the true skill statistic[90] (TSS), which takes into account both omission and commission errors, and, unlike other statistical measures such as KAPPA, is not affected by prevalence[90]. For each tree species, we performed the block-cross validation, and then averaged the results. Table 2 shows the results of model evaluation as an average by model for the entire pool of tree species. In addition, in Figure S3, we present the results for the entire pool of tree species.

In addition, we computed a consensus model by averaging single model predictions for each tree species. In doing so, we retained only predictions from models with TSS > 0.7, to avoid working with poorly calibrated models. We projected the consensus model into present and future climates and we computed the coefficient of variation (CV) from the single SDM projections. The higher (lower) the CV, the higher (lower) the uncertainty.

Finally, to avoid extrapolation to non-analog climates[91], we excluded the areas of the projections where the future climatic conditions are unlike any currently observed in Europe. The exclusions were based on multivariate environmental similarity surface[92] (MESS), which was created using the "mess" function from the "Dismo" package of the R library.

Code availability

Three scripts are available in figshare[87]. "*1_EU-Trees4F_sdm_present_future.R*" runs BIOMOD2 to model tree species distributions until the end of the century. It uses tree species occurrences and current/future environmental parameters that are stored in the "datasets" directory. This script does the following tasks: a) projects BIOMOD2 model output into the current and future environmental conditions. The projection into the future is made using two approaches: 1) A climatic ensemble mean of the 11 RCMs, or 2) an 'SDM' ensemble mean that projects the consensus model for every single RCM, and a posteriori averages the outputs of the 11 SDMs; b) calculates multivariate similarity between current and future projections, to avoid extrapolation to non-analog climates, c) derives the realized niche for each tree species by trimming the potential distribution maps with their native ranges distributions. The other two scripts calculate dispersal into the future using output from the first script. "*2_EU-Trees4F_dispersal_migclim_ens_clim.R*" using a climatic ensemble mean, whereas "*3_EU-Trees4F_dispersal_migclim_ens_sdms.R*" uses the 'SDM' ensemble mean.

The directory "datasets" contains various files that are needed for the code to run, such as the environmental parameters and the species occurrences. All scripts were written and run in R software version 3.6.3[93] (2020–02–29).

References

1. FOREST EUROPE. *State of Europe's Forests* (Ministerial Conference on the Protection of Forests in Europe, Bratislava, 2020).
2. Gamfeldt, L. *et al.* Higher levels of multiple ecosystem services are found in forests with more tree species. *Nat. Commun.* **4**, 1–8 (2013).
3. Brockerhoff, E. G. *et al.* Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodiv. Conserv.* **26**, 3005–3035 (2017).
4. Mori, A. S., Lertzman, K. P. & Gustafsson, L. Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology. *J. Appl. Ecol.* **54**, 12–27 (2017).
5. Forzieri, G. *et al.* Emergent vulnerability to climate-driven disturbances in European forests. *Nat. Commun.* **12**, 1–12 (2021).
6. Senf, C. & Seidl, R. Mapping the forest disturbance regimes of Europe. *Nat. Sustain.* **4**, 63–70 (2021).
7. Talluto, M. V., Boulangeat, I., Vissault, S., Thuiller, W. & Gravel, D. Extinction debt and colonization credit delay range shifts of eastern North American trees. *Nat. Ecol. Evol.* **1**, 1–6 (2017).
8. Zhu, K., Woodall, C. W. & Clark, J. S. Failure to migrate: lack of tree range expansion in response to climate change. *Glob. Change Biol.* **18**, 1042–1052 (2012).
9. Williams, J. W., Ordonez, A. & Svenning, J.-C. A unifying framework for studying and managing climate-driven rates of ecological change. *Nat. Ecol. Evol.* **5**, 17–26 (2021).
10. Jump, A. S. & Penuelas, J. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol. Lett.* **8**, 1010–1020 (2005).
11. Saltré, F. *et al.* Climate or migration: what limited European beech post-glacial colonization? *Glob. Ecol. Biogeogr.* **22**, 1217–1227 (2013).
12. Svenning, J.-C. & Skov, F. Limited filling of the potential range in European tree species. *Ecol. Lett.* **7**, 565–573 (2004).
13. Pedlar, J. H. *et al.* Placing forestry in the assisted migration debate. *BioScience* **62**, 835–842 (2012).
14. Overpeck, J. T. & Breshears, D. D. The growing challenge of vegetation change. *Science* **372**, 786–787 (2021).
15. Strona, G. *et al.* Far from naturalness: How much does spatial ecological structure of European tree assemblages depart from potential natural vegetation? *Plos One* **11**, e0165178 (2016).
16. Giesecke, T. *et al.* Postglacial change of the floristic diversity gradient in Europe. *Nat. Commun.* **10**, 1–7 (2019).
17. Kaplan, J. O., Krumhardt, K. M. & Zimmermann, N. The prehistoric and preindustrial deforestation of Europe. *Quat. Sci. Rev.* **28**, 3016–3034 (2009).
18. Sabatini, F. M. *et al.* Where are Europe's last primary forests? *Divers. Distrib.* **24**, 1426–1439 (2018).
19. Nabuurs, G.-J. *et al.* Next-generation information to support a sustainable course for European forests. *Nat. Sustain.* **2**, 815–818 (2019).
20. Williams, J. W., Jackson, S. T. & Kutzbach, J. E. Projected distributions of novel and disappearing climates by 2100 AD. *Proc. Natl. Acad. Sci.* **104**, 5738–5742 (2007).
21. Hoegh-Guldberg, O. *et al.* Assisted colonization and rapid climate change. *Science* **321**, 345–346 (2008).
22. Jandl, R., Spathelf, P., Bolte, A. & Prescott, C. E. Forest adaptation to climate change - is non-management an option? *Ann. For. Sci.* **76**, 1–13 (2019).
23. Dyderski, M. K., Paź, S., Frelich, L. E. & Jagodziński, A. M. How much does climate change threaten European forest tree species distributions? *Glob. Change Biol.* **24**, 1150–1163 (2018).
24. Hanewinkel, M., Cullmann, D. A., Schelhaas, M.-J., Nabuurs, G.-J. & Zimmermann, N. E. Climate change may cause severe loss in the economic value of European forest land. *Nat. Clim. Change* **3**, 203–207 (2013).
25. Thurm, E. A. *et al.* Alternative tree species under climate warming in managed European forests. *For. Ecol. Manag.* **430**, 485–497 (2018).
26. Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T. & Prentice, I. C. Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci.* **102**, 8245–8250 (2005).
27. Isbell, F. *et al.* Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **526**, 574–577 (2015).
28. Morin, X. *et al.* Long-term response of forest productivity to climate change is mostly driven by change in tree species composition. *Sci. Rep.* **8**, 1–12 (2018).

29. Hisano, M., Searle, E. B. & Chen, H. Y. Biodiversity as a solution to mitigate climate change impacts on the functioning of forest ecosystems. *Biol. Rev.* **93**, 439–456 (2018).
30. Messier, C. *et al*. The functional complex network approach to foster forest resilience to global changes. *For. Ecosyst.* **6**, 1–16 (2019).
31. Di Sacco, A. *et al*. Ten golden rules for reforestation to optimize carbon sequestration, biodiversity recovery and livelihood benefits. *Glob. Change Biol.* **27**, 1328–1348 (2021).
32. Jacob, D. *et al*. EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg. Environ. Change* **14**, 563–578 (2014).
33. Buras, A. & Menzel, A. Projecting tree species composition changes of European forests for 2061–2090 under RCP 4.5 and RCP 8.5 scenarios. *Front. Plant Sci.* **9**, 1986 (2019).
34. Chakraborty, D., Móricz, N., Rasztovits, E., Dobor, L. & Schueler, S. Provisioning forest and conservation science with high-resolution maps of potential distribution of major European tree species under climate change. *Ann. For. Sci.* **78**, 1–18 (2021).
35. Noce, S., Collalti, A. & Santini, M. Likelihood of changes in forest species suitability, distribution, and diversity under future climate: The case of Southern Europe. *Ecol. Evol.* **7**, 9358–9375 (2017).
36. Hickler, T. *et al*. Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Glob. Ecol. Biogeogr.* **21**, 50–63 (2012).
37. Takolander, A., Hickler, T., Meller, L. & Cabeza, M. Comparing future shifts in tree species distributions across Europe projected by statistical and dynamic process-based models. *Reg. Environ. Change* **19**, 251–266 (2019).
38. Chen, M. *et al*. Global land use for 2015–2100 at 0.05 resolution under diverse socioeconomic and climate scenarios. *Sci. Data* **7**, 1–11 (2020).
39. Mauri, A., Strona, G. & San-Miguel-Ayanz, J. EU-Forest, a high-resolution tree occurrence dataset for Europe. *Sci. Data* **4**, 1–8 (2017).
40. Strona, G., Mauri, A. & San-Miguel-Ayanz, J. A high-resolution pan-European tree occurrence dataset. *Figshare* https://doi.org/10.6084/m9.figshare.c.3288407.v1 (2016).
41. Benito-Garzón, M. & Fernández-Manjarrés, J. F. Testing scenarios for assisted migration of forest trees in Europe. *New For.* **46**, 979–994 (2015).
42. Thuiller, W., Lavorel, S., Sykes, M. T. & Araújo, M. B. Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Divers. Distrib.* **12**, 49–60 (2006).
43. Robinet, C. *et al*. A suite of models to support the quantitative assessment of spread in pest risk analysis. *PLoS ONE* **7**, 10 (2012).
44. European Commission. *The European Green Deal*. (Publications office of the European Union, 2019).
45. European Commission. *EU Biodiversity Strategy for 2030, Bringing nature back into our lives*. (Publications office of the European Union, 2020).
46. European Commission. *A sustainable bioeconomy for Europe: strengthening the connection between economy, society and the environment*. (Publications office of the European Union, 2018).
47. European Commission. *New EU Forest Strategy for 2030*. (Publications office of the European Union, 2021).
48. Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M. B. BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).
49. ICP Forests. International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests. http://icp-forests.net/ (2019).
50. Zając, A., Zając, M., Tertil, R. & Harman, I. *Atlas rozmieszczenia roślin naczyniowych w Polsce–Distribution Atlas of Vascular Plants in Poland*. (Nakładem Pracowni Chorologii Komputerowej Instytutu Botaniki Uniwersytetu - Laboratory of Computer Corology - Institute of Botany - Jagiellonian University, 2001).
51. Gschwantner, T. *et al*. Common tree definitions for national forest inventories in Europe. *Silva Fennica* **43**, 303–321 (2009).
52. Rivers, M. *et al*. *European Red List of Trees*. (International Union for Conservation of Nature and Natural Resources, 2019).
53. Rocchini, D. *et al*. Anticipating species distributions: Handling sampling effort bias under a Bayesian framework. *Sci. Total Environ.* **584**, 282–290 (2017).
54. Bartlein, P. J., Prentice, I. C. & Webb III, T. Climatic response surfaces from pollen data for some eastern North American taxa. *J. Biogeogr*. 35–57 (1986).
55. Woodward, F. I. & Woodward, F. *Climate and plant distribution*. (Cambridge University Press, 1987).
56. Harrison, S. *et al*. Towards a global scheme of plant functional types for ecosystem modelling, palaeoecology and climate impact research. *J Veg Sci* **21**, 300–317 (2009).
57. Thuiller, W. BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* **9**, 1353–1362 (2003).
58. Prentice, I. C. *et al*. Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. *J. Biogeogr*. 117–134 (1992).
59. Pouteau, R. *et al*. Potential alien ranges of European plants will shrink in the future, but less so for already naturalized than for not yet naturalized species. *Divers. Distrib.* **27**, 2063–2076 (2021).
60. Naimi, B. USDM: Uncertainty analysis for species distribution models. https://www.rdocumentation.org/packages/usdm/versions/ (2015).
61. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol. J. R. Meteorol. Soc.* **25**, 1965–1978 (2005).
62. Title, P. O. & Bemmels, J. B. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* **41**, 291–307 (2018).
63. Teutschbein, C. & Seibert, J. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.* **456**, 12–29 (2012).
64. Ekström, M., Grose, M. R. & Whetton, P. H. An appraisal of downscaling methods used in climate change research. *Wiley Interdiscip. Rev. Clim. Change* **6**, 301–319 (2015).
65. Beck, H. E. *et al*. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 1–12 (2018).
66. Baker, B., Diaz, H., Hargrove, W. & Hoffman, F. Use of the Köppen–Trewartha climate classification to evaluate climatic refugia in statistically derived ecoregions for the People's Republic of China. *Clim. Change* **98**, 113–131 (2010).
67. Barredo, J. I., Caudullo, G. & Dosio, A. Mediterranean habitat loss under future climate conditions: Assessing impacts on the Natura 2000 protected area network. *Appl. Geogr.* **75**, 83–92 (2016).
68. Klausmeyer, K. R. & Shaw, M. R. Climate change, habitat loss, protected areas and the climate adaptation potential of species in Mediterranean ecosystems worldwide. *PloS One* **4**, e6392 (2009).
69. Tabor, K. & Williams, J. W. Globally downscaled climate projections for assessing the conservation impacts of climate change. *Ecol. Appl.* **20**, 554–565 (2010).
70. Collins, M. *et al*. Long-term climate change: projections, commitments and irreversibility. in *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* 1029–1136 (Cambridge University Press, 2013).
71. Hengl, T. *et al*. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* **12**, e0169748 (2017).
72. Zhang, L. *et al*. Consensus forecasting of species distributions: The effects of niche model performance and niche properties. *PloS One* **10**, e0120056 (2015).

73. Merow, C., Smith, M. J. & Silander, J. A. Jr. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069 (2013).
74. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol. Evol.* **3**, 327–338 (2012).
75. De Jong, R., Verbesselt, J., Zeileis, A. & Schaepman, M. E. Shifts in global vegetation activity trends. *Remote Sens.* **5**, 1117–1133 (2013).
76. Engler, R. & Guisan, A. MigClim: predicting plant distribution and dispersal in a changing climate. *Divers. Distrib.* **15**, 590–601 (2009).
77. Engler, R., Hordijk, W. & Guisan, A. The MIGCLIM R package–seamless integration of dispersal constraints into projections of species distribution models. *Ecography* **35**, 872–878 (2012).
78. Merow, C., Wilson, A. M. & Jetz, W. Integrating occurrence data and expert maps for improved species range predictions. *Glob. Ecol. Biogeogr.* **26**, 243–258 (2017).
79. Caudullo, G., Welk, E. & San-Miguel-Ayanz, J. Chorological maps for the main European woody species. *Data Brief* **12**, 662–666 (2017).
80. Euro+Med. Euro+Med PlantBase – the information resource for Euro-Mediterranean plant diversity. http://ww2.bgbm.org/EuroPlusMed/ (2019).
81. Summers, D. M., Bryan, B. A., Crossman, N. D. & Meyer, W. S. Species vulnerability to climate change: impacts on spatial conservation priorities and species representation. *Glob. Change Biol.* **18**, 2335–2348 (2012).
82. García-Valdés, R., Zavala, M. A., Araujo, M. B. & Purves, D. W. Chasing a moving target: Projecting climate change-induced shifts in non-equilibrial tree species distributions. *J. Ecol.* **101**, 441–453 (2013).
83. Lischke, H., Zimmermann, N. E., Bolliger, J., Rickebusch, S. & Löffler, T. J. TreeMig: a forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecol. Model.* **199**, 409–420 (2006).
84. Tamme, R. *et al*. Predicting species' maximum dispersal distances from simple plant traits. *Ecology* **95**, 505–513 (2014).
85. Thomson, F. J., Letten, A. D., Tamme, R., Edwards, W. & Moles, A. T. Can dispersal investment explain why tall plant species achieve longer dispersal distances than short plant species? *New Phytol.* **217**, 407–415 (2018).
86. Kattge, J. *et al*. TRY plant trait database–enhanced coverage and open access. *Glob. Change Biol.* **26**, 119–188 (2020).
87. Mauri, A., Girardello, M. & Strona, G. EU-Trees4F. A dataset on the future distribution of European tree species, *figshare*, https://doi.org/10.6084/m9.figshare.c.5525688 (2021).
88. Vítková, M., Müllerová, J., Sádlo, J., Pergl, J. & Pyšek, P. Black locust (Robinia pseudoacacia) beloved and despised: A story of an invasive tree in Central Europe. *For. Ecol. Manag.* **384**, 287–302 (2017).
89. Muscarella, R. *et al*. ENM eval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol. Evol.* **5**, 1198–1205 (2014).
90. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232 (2006).
91. Fitzpatrick, M. C. & Hargrove, W. W. The projection of species distribution models and the problem of non-analog climate. *Biodivers. Conserv.* **18**, 2255–2261 (2009).
92. Elith, J. *et al*. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **17**, 43–57 (2011).
93. R Core Team. *R: A language and environment for statistical computing.* (2020).

Acknowledgements

Author contributions

A.M. performed the analysis and wrote the paper. M.G. and G.S. assisted in modeling tasks. A.M., P.B., G.S., A.C. designed the study. All authors contributed to interpreting the results and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-022-01128-5.

Correspondence and requests for materials should be addressed to A.M. or A.C.

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

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