

Assisted tree migration can preserve the European forest carbon sink under climate change

In the format provided by the
authors and unedited

Supplementary Information

Assisted tree migration can preserve the European forest carbon sink under climate change

Table of Contents

SUPPLEMENTARY DATA	2
Provenance trial data	2
Description of climate data used to develop the Universal response functions (URFs) and species distribution models (SDMs)	6
SUPPLEMENTARY METHODS	9
Estimating annual carbon sink (CS) from range-wide provenance trial data	9
Description of species distribution models (SDMs) to identify the best-suited species	11
Selecting the climatically best-suited species from species distribution models	28
Identifying best-suited seed provenances – modeling provenance growth and CS	28
Demonstrating assisted migration – modeling seed sourcing scenarios	31
Analyzing uncertainty in identifying best-suited species and seed provenances	35
Spatial partitioning of the components of uncertainty in the estimated carbon sink	36
SUPPLEMENTARY RESULTS	37
Uncertainty in identifying best-suited species and seed provenances	37
Spatial patterns in the uncertainty of the estimation of CS	54
REFERENCES (SUPPLEMENTARY MATERIAL)	65

23 **SUPPLEMENTARY DATA**

24 Provenance trial data

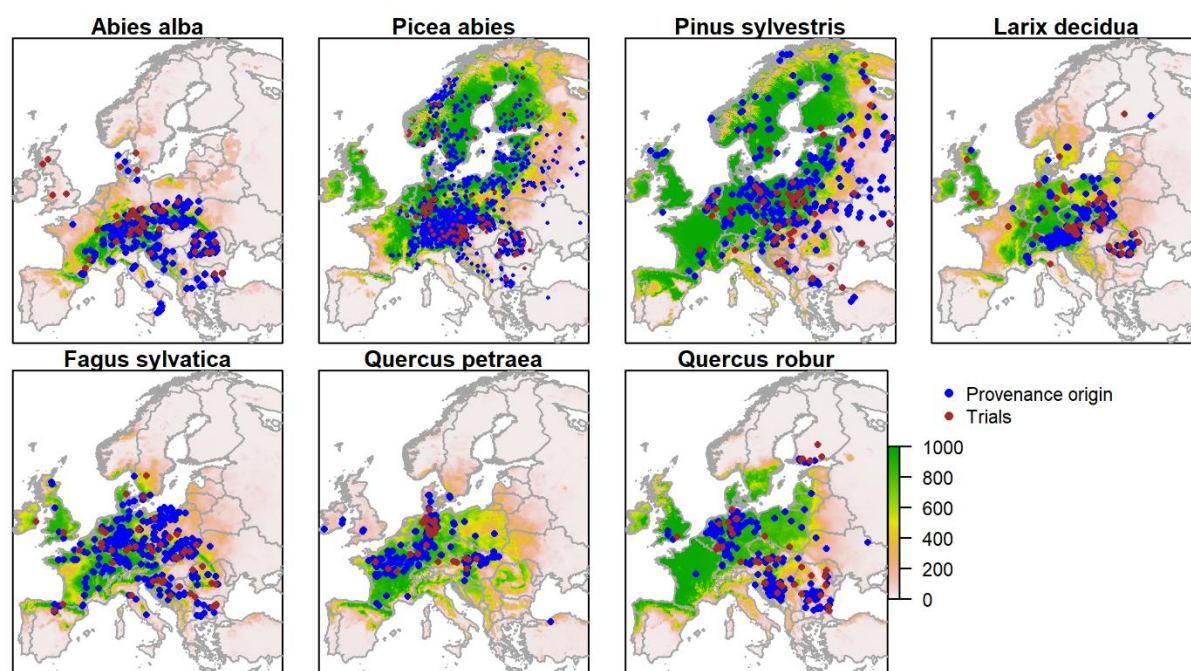
25 The local adaptation of forest tree populations has been quantified by data from range-wide
 26 provenance experiments. We collected provenance trial data for the seven European tree species, *A.*
 27 *alba*, *F. sylvatica*, *P. sylvestris*, *P. abies*, *L. decidua*, *Q. robur*, and *Q. petraea* (Supplementary Fig. 1 &
 28 Supplementary Table 1) within the Project INTERREG-CE SUSTREE

29 **Supplementary Table 1.** Overview of the provenance trials analyzed in this study. Number of provenance trial
 30 sites and provenances planted in each trial.

Tree species	Number of trial sites	Number of Provenances	Tree ages (years)
<i>Abies alba</i>	50	343	5- 53
<i>Picea abies</i>	236	1022	5-50
<i>Pinus sylvestris</i>	100	421	2-70
<i>Larix decidua</i>	50	247	5-64
<i>Fagus sylvatica</i>	74	491	2-40
<i>Quercus robur</i>	48	208	1-47
<i>Quercus petraea</i>	36	232	4-47
Total	594	2964	1-70

31

32



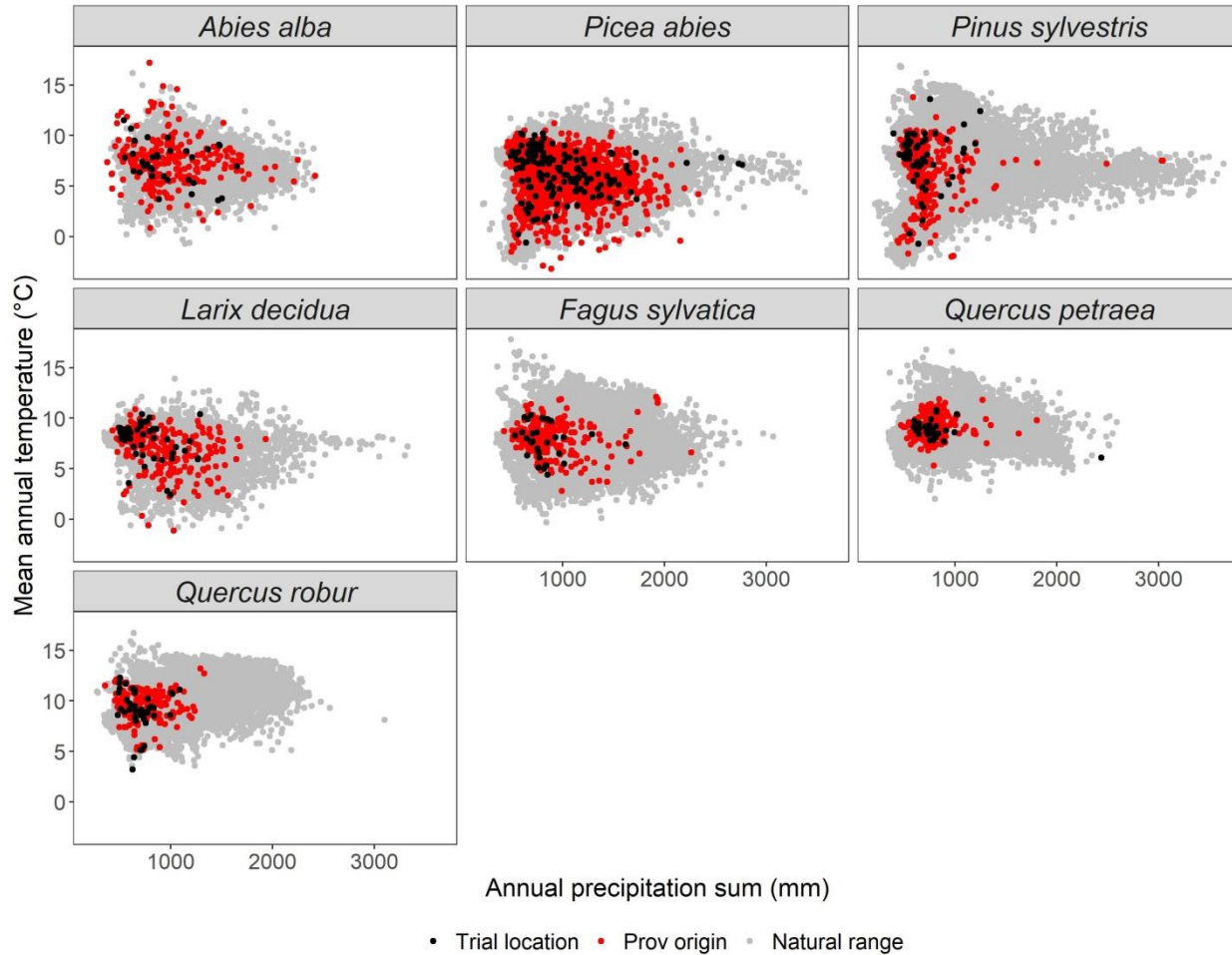
33

34 **Supplementary Fig. 1.** Geographic spread of the provenance trial data. Location of seed provenance origins
 35 (blue) and trial locations (brown), plotted over the predicted potential distribution, represented by climatic
 36 suitability (0= not suitable to 1000 most suitable) under the historical period (1961-1990).

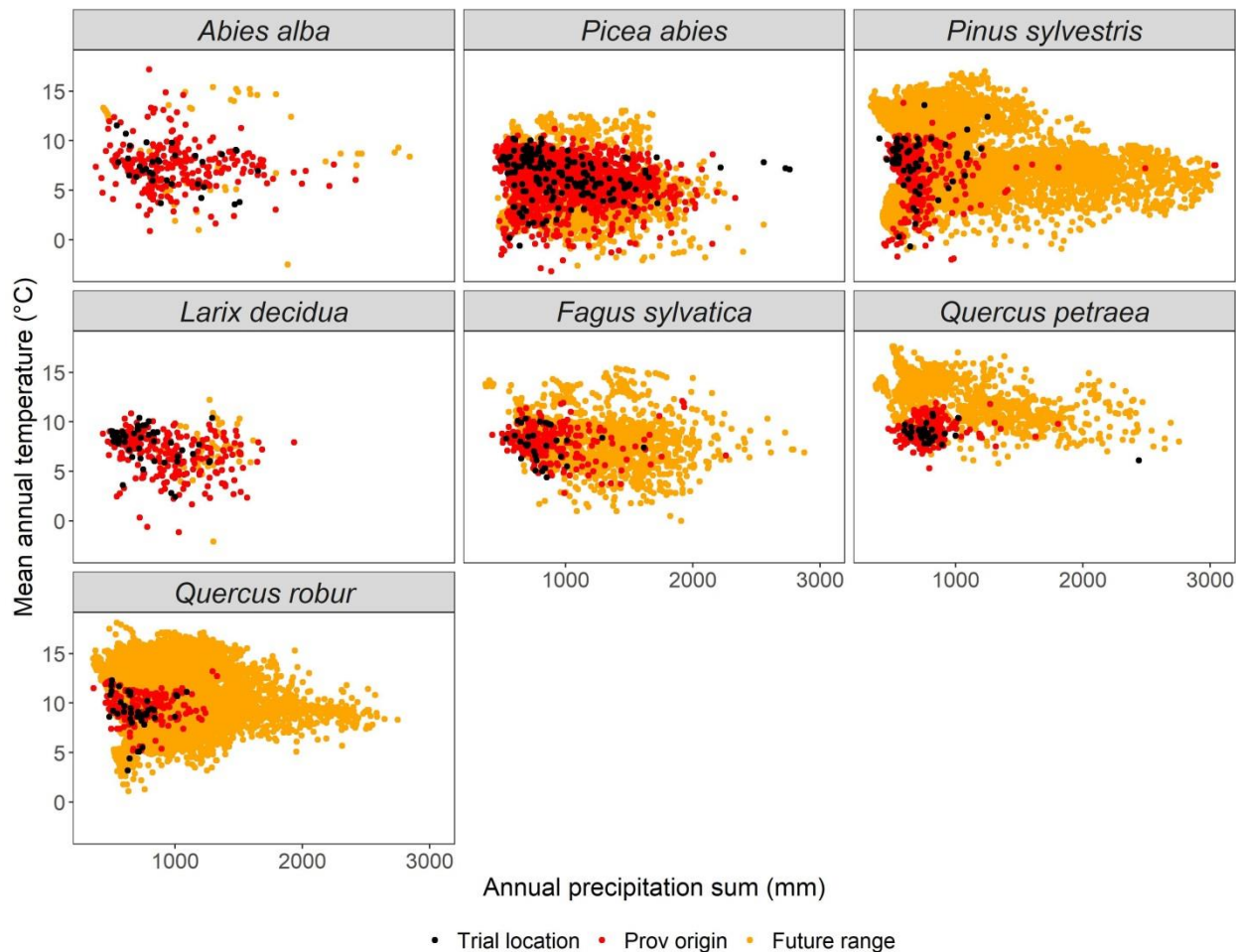
37

38 The dataset covers large parts of the climatic range of each species (Supplementary Fig. 2 & 3). This
 39 data collection includes traits such as tree age, height, dbh, volume, survival rate, and the trial mean
 40 values per provenance. In addition, information about the location (geographic coordinates and
 41 altitude) of trial sites and provenance origin were obtained. The database also contains information
 42 about the age and establishment (planting design, year of establishment) of trial sites and soil types, if
 43 available. The tree ages in the provenance trials range from 1-70 years (Supplementary Table 1). Data
 44 were collected from project partners of other European research organizations as well as collections
 45 from published data. The compiled data were checked for possible errors. Coordinates of provenances
 46 and trials were verified in a GIS framework. If possible, the error was rectified; otherwise, the
 47 erroneous data were removed. The dataset is available at:

48 <https://figshare.com/s/98e405d56bb789b08cb0>



Supplementary Fig. 2. Climate locations of seed provenance origins (red) and trial locations (black) of the provenance trial data plotted across the current natural distribution range (gray) of the seven tree species within a bioclimatic envelope of annual precipitation sum and mean annual temperature for the historical period (1961-1990). The natural distribution area has been obtained from⁹⁵.



Supplementary Fig. 3. Climate locations of seed provenance origins (red) and trial locations (black) of the provenance trial data plotted across predicted potential distribution (orange) of the seven tree species. The modeled predictions (orange) depict the future distribution under climate scenario RCP8.5 (mean of five Regional climate models (RCMs) in 2061-2080, see Supplementary Table 2) and their realized share if a single species with the highest suitability has been selected for the respective grid cell. The climate location of seed provenances and trial locations, however, represent observations for the historical period 1961-1990 to demonstrate that the universal response function (URF) models, which are built upon the provenance data, do not interpolate beyond the climate conditions of the observations.

67 Description of climate data used to develop the Universal response functions (URFs) and
68 species distribution models (SDMs)

69 Biologically relevant climate variables for developing the Universal response functions (URFs) and the
70 species distribution models (SDMs) were obtained from the ECLIPS 2.0 dataset⁹⁴. This dataset was
71 developed from dynamically downscaled and bias-corrected regional climate model results from the
72 EURO-CORDEX with a resolution of 30 arcsec. The dataset was tested against independent station data
73 from the European Climate Assessment (ECA).

74 With this approach, a gridded dataset for 80 bioclimate variables (Supplementary Table 3) for historic
75 climate (1961-90), contemporary climate (1991-2010), and three future time frames which include
76 averages of the periods 2041-2060, 2061-2080, and 2081-2100 for the two representative
77 concentration pathway RCP4.5 and RCP8.5⁹⁶ was obtained. RCP4.5 or the moderate scenario assumes
78 a 650 ppm atmospheric CO₂ concentration and a 1.0–2.6°C increase in annual temperature by 2100
79 whereas RCP8.5, a pessimistic scenario assumes a 1,350 ppm CO₂ and 2.6–4.8°C increase in annual
80 temperature by 2100⁹⁶. These RCP scenarios are complemented by the five Shared Socioeconomic
81 Pathways (SSPs)⁴⁰ provided in the recent IPCC 6th assessment report⁹⁶. The SSP pathways describe how
82 global societies might evolve according to population growth, socioeconomic developments, energy
83 use, technology changes, and environmental conditions. SSP1 assumes sustainable development, SSP2
84 refers to middle-of-the-road development, SSP3 refers to regional rivalry, SSP4 assumes social
85 inequality and SSP5 is based on fossil-fuelled development⁴⁰. RCP4.5 is roughly equivalent to the
86 Shared Socioeconomic Pathways, SSP2 is a middle-of-the-road or moderate scenario and RCP8.5
87 corresponds to SSP4, a fossil-fuel-based development or business-as-usual, scenario⁴⁰.

88 Climate variables for future scenarios were calculated for five regional climate models (RCM) to
89 account for uncertainty due to RCM projections (Supplementary Table 2). Five daily bias-corrected
90 RCM results out of nine projections available in the EURO-CORDEX database were used to calculate the
91 climate variables for future scenarios in the ECLIPS2.0 dataset (Supplementary Table 2). The criteria
92 used to select the five climate projections were: (i) representation of all available RCMs and GCMs; (ii)
93 two RCMs being nested in the same driving GCM; and (iii) one RCM being driven by two different
94 GCMs. Such criteria were adopted to ensure the representativeness of all combinations of RCMs and
95 GCMs available in the EURO-CORDEX database.

96 All our analysis was based on two-time frames, 1991-2010 and 2061-80.

97 To calibrate the models and to identify the best-suited seed provenances, the climate variables for trial
98 locations were obtained for the time when the provenance trials were inventoried, whereas climate
99 variables of seed provenance origin were restricted to the historical period (1961-1990) to match the
100 more stable past climate that promoted adaptation over generations.

101 The ECLIPS 2.0 dataset is available at <https://doi.org/10.5281/zenodo.3952159>.

102 **Supplementary Table 2** Bias-corrected EURO-CORDEX simulations selected for ECLIPS2.0⁹⁴.

Model ID	Institute	RCM	Driving GCM
M1	Climate Limited-area Modelling Community (CLM-Community)	CLMcom-CLM4-8-17	CNRM-CERFACS-CNRM- CM5
M2	Climate Limited-area Modelling Community (CLM-Community)	CLMcom-CLM4-8-17	MPI-M-MPI-ESM-LR
M3	Danish Meteorological Institute (DMI)	DMI-HIRHAM5	ICHEC-EC-EARTH
M4	Royal Netherlands Meteorological Institute (KNMI)	KNMI- RACMO22E	MOHC-HadGEM2-ES
M5	Helmholtz-Zentrum Geesthacht, Climate Service Center, Max Planck Institute for Meteorology (MPI-CSC)	MPI-CSC-REMO2009	MPI-M-MPI-ESM-LR

103

104

Climate variable	Variables	Unit
AHM	annual heat:moisture index $(MAT+10)/(MAP/1000)$	
bFFP	the Julian date on which FFP begins	
DDabove18	degree-days below 18°C, heating degree-days	
DDabove5	degree-days above 5°C, growing degree-days	
DDbelow0	degree-days below 0°C, chilling degree-days	
DDbelow18	degree-days below 18°C, heating degree-days	
eFFP	the Julian date on which FFP ends	
EMT	Extreme minimum temperature over 30 years	°C
FFP	frost-free period	days
MAP	mean annual precipitation (mm)	°C
MAT	mean annual temperature (°C)	°C
MCMT	Mean coldest month temperature (°C)	°C
MSP	mean summer (May to Sept.) precipitation (mm)	°C
MWMT	mean warmest month temperature (°C)	°C
NFFD	the number of frost-free days	days
PPT_at	autumn precipitation (mm)	mm
PPT_sm	summer precipitation (mm)	mm
PPT_sp	spring precipitation (mm)	mm
PPT_wt	winter precipitation (mm)	mm
PPT01 to 12	Precipitation month 01(Jan) to 12 (Dec)	mm
SHM	summer heat: moisture index $((MWMT)/(MSP/1000))$	
Tave_at	autumn (Sep. - Nov.) mean temperature (°C)	°C
Tave_sm	summer (Jun. - Aug.) mean temperature (°C)	°C
Tave_sp	spring (Mar. - May) mean temperature (°C)	°C
Tave_wt	winter (Dec.(prev. yr) - Feb.) mean temperature (°C)	°C
Tave01 to 12	Average temperature month 01(Jan) to 12 (Dec)	°C
TD	Temperature difference between MWMT and MCMT(°C)	°C
Tmax_an	Maximum yearly temperature	°C
Tmax_at	Maximum autumn temperature	°C
Tmax_sm	Maximum summer temperature	°C
Tmax_sp	Maximum spring temperature	°C
Tmax_wt	Maximum winter temperature	°C
Tmax 01 to 12	Maximum temperature 01(Jan) to 12 (Dec)	°C
Tmin_an	Minimum annual temperature	°C
Tmin_at	Minimum autumn temperature	°C
Tmin_sm	Minimum summer temperature	°C
Tmin_sp	Minimum spring temperature	°C
Tmin_wt	Minimum winter temperature	°C
Tmin01 to 12	Minimum temperature 01(Jan) to 12 (Dec)	°C

SUPPLEMENTARY METHODS

Estimating annual carbon sink (CS) from range-wide provenance trial data

Annual carbon sink (CS) was estimated from the provenance trial data by converting single tree above-ground living biomass (AGB) to carbon stock per hectare. Generalized biomass equations⁹⁸ were used to estimate AGB for the seven species. These biomass equations with the general form:

$$Eq\ 1: \quad l_n(AGB) = l_n\beta_0 + \beta_1 l_n(dbh) + \beta_2 l_n(Age) * CF$$

estimate AGB from a single tree diameter at breast height of 130 cm (dbh), and tree age (Age) of the provenance trial data. The correction factor (CF) was used to correct the biases if any, caused by transforming regional biomass equations into the generalized one (Eq 1)⁹⁸. The coefficients β_0 , β_1 , and β_2 are standardized coefficients that quantify how many standard deviations the dependent variable (AGB) will change per standard deviation increase in the independent variables in eq 1. The values of β_0 , β_1 , and β_2 are given in (Supplementary Table 4.).

Supplementary Table 4. Model parameter for equation 1. Parameter values of coefficients β_0 , β_1 , β_2 for Equation 1.

Species	β_0	β_1	β_2	CF
<i>Abies alba</i>	-2.3958	2.4497	0.00	1.000680438
<i>Fagus sylvatica</i>	-2.2208	2.3244	0.1749	1.091889992
<i>Larix decidua</i>	-5.0916	2.2557	1.0313	1.002566423
<i>Picea abies</i>	-2.6868	2.2288	0.2358	1.00325128
<i>Pinus sylvestris</i>	-2.7606	2.2336	0.2035	1.026162653
<i>Quercus petraea</i>	-2.3364	2.5068	0.00	0.96881138
<i>Quercus robur</i>	-2.6322	2.7535	-0.0843	1.024797498

AGB was converted to carbon stocks per tree $Cstock^{-tree}$ by multiplying it by 0.55 (Eq 2) according to IPCC-LULUCF guidelines^{99,100}:

$$Eq\ 2: \quad Cstock^{-tree} = AGB * 0.55$$

125 To estimate carbon stocks per ha, we calculated trees per ha using the species-specific competition-
 126 density rule¹⁰¹, where the number of trees per hectare (N) was calculated based on tree height and dbh
 127 (Eq 3). Here, d_g is the quadratic mean dbh, h is tree height and a_0 , a_1 , b_0 , are the species-specific
 128 coefficients of the competition-density rule given in (Supplementary Table 5).

129 Eq 3:
$$N = \frac{d_g * a_0 * h^{a_1}}{b_0 * h^{\left(\frac{a_1}{-0.605}\right)}}$$

130

131 **Supplementary Table 5.** Model parameter for equation 3. Parameter values of coefficients a_0 , a_1 , b_0 for Equation
 132 3, were obtained from⁹⁷.

133

134

135

136

137

138

139

Species	a_0	a_1	b_0
<i>Abies alba</i>	$1.917 * 10^{-6}$	0.6670	0.9215
<i>Fagus sylvatica</i>	$1.806 * 10^{-6}$	0.6921	1.0364
<i>Larix decidua</i>	$2.526 * 10^{-6}$	0.6961	0.8278
<i>Picea abies</i>	$1.522 * 10^{-6}$	0.6493	0.8882
<i>Pinus sylvestris</i>	$2.682 * 10^{-6}$	0.5600	0.4717
<i>Quercus petraea</i>	$2.733 * 10^{-6}$	0.5922	0.6260
<i>Quercus robur</i>	$2.733 * 10^{-6}$	0.5922	0.6260

140 Carbon stocks per tree (Eq 2) and trees per hectare (Eq 3) for the respective tree age were the basis for
 141 estimating carbon stocks per hectare $Cstock^{-ha}$ (Eq 4).

142 Eq 4:
$$Cstock^{-ha} = Cstock^{-tree} * N$$

143

144

145 Description of species distribution models (SDMs) to identify the best-suited species

146

147 Climatically best-suited species for each 30 arc-sec grid cell of Europe were determined with species
148 distribution models (SDMs) developed with a multi-model ensemble approach implemented through
149 the R package biomod2¹⁰². These SDMs predicted potential distribution (probability of occurrence) for
150 each species in historic (1961-1990), contemporary (1991-2010), and future periods (2041-2060, 2061-
151 2080, 2081-2100) for the two climate change scenarios RCP4.5 and RCP8.5. All our analysis was based
152 on two time frames 1991-2010 and 2061-80.

153 The biomod2 offers a computational platform for multi-method modeling that generates models of
154 species' potential distribution for each species. The models include GLM (Generalized Linear Models),
155 GAM (Generalized Additive Models), GBM (Generalized Boosted regression Models), CTA (Classification
156 Tree Analysis), ANN (Artificial Neural Networks), SRE (Surface Range Envelop or BIOCLIM), FDA (Flexible
157 Discriminant Analysis), MARS (Multivariate Adaptive Regression Spline), RF (Random Forest for
158 classification and regression), and MAXENT. Tsuruoka.

159 Each model algorithm of the biomod2 predicted the probability of occurrence (potential distribution)
160 for the seven tree species. For contemporary climate, potential distributions of each of the seven
161 species were estimated from 10 SDM models of biomod2 and 1 ensemble SDM for each species
162 (Supplementary Table 9). In future climate scenarios of RCP4.5 and 8.5, the potential distributions
163 were estimated for combinations of 10 SDMs and 5 RCMs (Supplementary Tables 2 & 9) and 1
164 ensemble SDM for each species calibrated with a mean of 5 RCMs (Supplementary Tables 2& 9).

165 To ensure transparent reporting and reproducibility, we described the dataset according to the
166 ODMAP¹⁰³. ODMAP (Overview, Data, Model, Assessment, and Prediction), offers a standardized way of
167 communicating the results and outputs from SDMs by describing the objectives, model assumptions,
168 scaling issues, data sources, model workflow, model predictions, and uncertainty (Supplementary
169 Table 6).

170

171 **Supplementary Table 6** Description of the Species distribution models (SDMs) used to identify the best-suited
 172 species in each forested grid cell of Europe in this study according to the ODMAP protocol.

ODMAP elements	Contents
OVERVIEW	
Model objective	SDM Objective: Forecast/transfer Target output: probability of occurrence of target tree species
Taxon	Seven tree species of Europe: Abies alba, Fagus sylvatica, Larix decidua, Picea abies, Pinus sylvestris, Quercus petraea, Quercus robur
Location	Europe
Scale of analysis	Spatial extent (Lon/ Lat): Longitude: -32.65000 °E -69.44167 °E Latitude: 30.877982 °N -71.57893 °N Spatial resolution: 30 arcsec Temporal resolution: We modeled for historic climate (1961-1990), contemporary (1991-2010), and three future time frames which include averages of (2041-2060, 2061-2080, and 2081-2100). The predictions were done for two Representative Concentration RCP4.5 and RCP8.5. All our analysis for the paper was based on two time frames, 1991-2010 and 2061-80.
Biodiversity data overview	Observation type: standardized monitoring Response data type: presence/ absence data
Type of predictors	Climatic

Conceptual model/hypotheses	A large body of scientific studies indicate that climate is one of the major drivers of the distribution of tree species at the continental scale. We exploited this correlation between species' current occurrence and climate to develop SDMs that predict the potential distribution of the target tree species.
Assumptions	We assumed that species are at pseudo-equilibrium with the environment. The source of the presence/absence data⁹⁵ used in this study is largely from national forest inventories where tree individuals below a certain diameter at breast height are not recorded. We assume that this data collection procedure did not bias our occurrence data. Since our occurrence dataset covers the whole current distribution of the target species, which represents both the historic, and likely future climate of Europe, we assumed that the species retain their niches across space and time and the current occurrence~ climate correlation remains stable when predicting the models for future climate.
SDM algorithms	Algorithms: We selected 10 modeling algorithms: GLM (Generalized Linear Models), GAM (Generalized Additive Models), GBM (Generalized Boosted regression Models), CTA (Classification Tree Analysis), ANN (Artificial Neural Networks), SRE (Surface Range Envelop or BIOCLIM), FDA (Flexible Discriminant Analysis), MARS (Multivariate Adaptive Regression Spline), RF (Random Forest for classification and regression), and MAXENT. Tsuruoka. These model algorithms were implemented through biomod2¹⁰². Model complexity: The individual models were run using the standard default settings of biomod2, which are designed to balance model complexity and overfitting. See further details in the section “Model” detailed below. Ensembles: For each species, the prediction of 1) the 10 SDM models in contemporary climate and a combination of 10 SDMs and 5 RCMs in future scenarios (Supplementary Table 9) were ensembled through biomod2¹⁰².

	See further details in the section “Model” below.
Model workflow	The model workflow includes: 1. Data cleaning and presence-absences classification/assignment 2. Presence – absence data thinning within geographic and environmental space 3. Variable selection 4. Model calibration 5. Ensemble prediction 6. Identification of best-suited species for each forested grid cell of Europe
Software	Software: All analyses were conducted using R version 3.3.2 (R Core Team, 2016). R Packages used: biomod2 ¹⁰², Random Forest ¹⁰⁴, covsel ¹⁰⁵ Data availability: Presence absence data are available from⁹⁵ Climate data is available from⁹⁴ High-resolution gridded climate data for Europe based on bias-corrected EURO-CORDEX: the ECLIPS-2.0 dataset Zenodo: 10.5281/zenodo.3952159.

DATA	
Biodiversity data	Taxon names: <i>Abies alba</i>, <i>Fagus sylvatica</i>, <i>Larix decidua</i>, <i>Picea abies</i>, <i>Pinus sylvestris</i>, <i>Quercus petraea</i>, <i>Quercus robur</i> Ecological level: Species-level Data source: Species presence-absence data was obtained from the EU-Forest dataset⁹⁵. The dataset harmonizes European tree occurrence from National Forest inventories (NFI), Forest Focus and, Biosoil datasets¹⁰⁶. A major part of the data arises from the NFI data (96%) while 4% is contributed by Forest Focus, Biosoil datasets. Sampling design: The background data included in the EU-Forest⁹⁵ varied in their sampling intensity and design. This data was harmonized and aggregated to a spatial resolution of 1 square kilometer, in line with an INSPIREcompliant 1 km× 1 km grid Sample size The dataset includes a total of 588,983 occurrence records at a spatial resolution of 1x1km⁹⁵.
	Data filtering: Presence-Absence data and data thinning: In our case, the geographic locations of the target species in the EU-Forest dataset were assumed to be true-presences. Conversely, the remaining occurrences of all other species, including the remaining target species, were treated as true-absence locations. To reduce spatial autocorrelation the entire presence-absences data was thinned in the geographic and environmental space. To do so, the presence-absence locations were initially thinned, retaining only one occurrence (either presence or absence) within each cell 10 km apart. To ensure that absence locations are climatically distant from the observed presence locations, we utilized the presence data of each target species to generate a rectilinear Surface Range Envelope (SRE) considering all bioclimatic variables

	predictors (Supplementary Table 3). When running the SRE model, we excluded the top and bottom 2.5% of the most extreme values for each variable along the gradient¹⁰². Subsequently, we removed all absence data overlapping with the predicted surface range envelope while the remaining absence dataset was used for calibration. After thinning, 244,273 occurrence records (presence and absence) were retained for calibrating the SDMs with biomod2. The number of presences and absence records used to calibrate the SDMs for the seven target species. <table><tr><th colspan="4">Occurrence data</th></tr><tr><th>Species</th><th>Presence</th><th>Absence</th><th>Total</th></tr><tr><td>Abies alba</td><td>2279</td><td>34760</td><td>37039</td></tr><tr><td>Fagus sylvatica</td><td>7229</td><td>27723</td><td>34952</td></tr><tr><td>Larix decidua</td><td>2657</td><td>34244</td><td>36901</td></tr><tr><td>Picea abies</td><td>13062</td><td>19131</td><td>32193</td></tr><tr><td>Pinus sylvestris</td><td>14703</td><td>17638</td><td>32341</td></tr><tr><td>Quercus petraea</td><td>4653</td><td>31518</td><td>36171</td></tr><tr><td>Quercus robur</td><td>7981</td><td>26695</td><td>34676</td></tr></table>	Occurrence data				Species	Presence	Absence	Total	Abies alba	2279	34760	37039	Fagus sylvatica	7229	27723	34952	Larix decidua	2657	34244	36901	Picea abies	13062	19131	32193	Pinus sylvestris	14703	17638	32341	Quercus petraea	4653	31518	36171	Quercus robur	7981	26695	34676
Occurrence data																																					
Species	Presence	Absence	Total																																		
Abies alba	2279	34760	37039																																		
Fagus sylvatica	7229	27723	34952																																		
Larix decidua	2657	34244	36901																																		
Picea abies	13062	19131	32193																																		
Pinus sylvestris	14703	17638	32341																																		
Quercus petraea	4653	31518	36171																																		
Quercus robur	7981	26695	34676																																		
Data partitioning	The occurrence dataset for each target species was partitioned by splitting it into 75% for model training and 25% for model evaluation. The models were, however, trained and ensembled with the full dataset.																																				
Environmental predictors	Predictor variables Environmental predictors were the biologically relevant seasonal and annual climate variables (Supplementary Table 3). From this list, a small subset of potential predictor variables was selected for each target species during the variable selection process. Data sources: The spatial resolution of predictor data: 30 arcsec which is roughly equivalent to 1x1km or lower depending on latitude.																																				

	The temporal resolution of predictor variable: Contemporary (1991-2010), and future time frame which include averages of (2061-2080) for two Representative Concentration RCP4.5 and RCP8.5 were used for the SDM predictions. Geographic projection: WGS 84 (EPSG: 4326)
--	---

Model	
Variable selection and Multicollinearity	From the list of potential biologically relevant predictor variables (Supplementary Table 3), we selected those that explain most of the variation in the observed presence and absence of each species while accounting for multi-collinearity. We utilized the covsel R package¹⁰⁵ for covariate selection, employing a two-step process: collinearity filtering (Step A) and model-specific embedding (Step B). In Step A, less informative covariates among collinear pairs were iteratively removed based on univariate GLM p-values and the default correlation coefficient threshold of $r < 0.70$. In Step B, the covariates selected in Step A were used to fit models with embedded selection procedures, including GLM with elastic-net regularization, GAM with null-space penalization, and guided regularized RF. The retained covariates after regularization were ranked based on the absolute value of regularized regression coefficients for GLM, chi-square statistic for GAM, and Mean Decrease Gini index for RF. An overall ranking was determined by the sum of the ranks for each covariate, resulting in the identification of the most important predictors. All other parameters related to the number of covariates and the maximum number of covariates to be selected were maintained at their default settings.

Climate variables retained after variable selection were used to calibrate the SDMs for each target species.

Variable	Rank	Species
<i>SHM</i>	1	<i>Abies alba</i>
<i>Tmax_sm</i>	2	<i>Abies alba</i>
<i>TD</i>	3	<i>Abies alba</i>
<i>DDbelow0</i>	4	<i>Abies alba</i>
<i>PPT_sm</i>	5	<i>Abies alba</i>
<i>PPT_sp</i>	6	<i>Abies alba</i>
<i>bFFP</i>	1	<i>Fagus sylvatica</i>
<i>MAP</i>	2	<i>Fagus sylvatica</i>
<i>SHM</i>	3	<i>Fagus sylvatica</i>
<i>TD</i>	4	<i>Fagus sylvatica</i>
<i>PPT_sm</i>	5	<i>Fagus sylvatica</i>
<i>DDbelow0</i>	1	<i>Larix decidua</i>
<i>SHM</i>	2	<i>Larix decidua</i>
<i>TD</i>	3	<i>Larix decidua</i>
<i>PPT_sp</i>	4	<i>Larix decidua</i>
<i>Tmax_sm</i>	5	<i>Larix decidua</i>
<i>PPT_sm</i>	6	<i>Larix decidua</i>
<i>bFFP</i>	1	<i>Picea abies</i>
<i>AHM</i>	2	<i>Picea abies</i>
<i>PPT_sp</i>	3	<i>Picea abies</i>
<i>TD</i>	4	<i>Picea abies</i>
<i>bFFP</i>	1	<i>Pinus sylvestris</i>
<i>AHM</i>	2	<i>Pinus sylvestris</i>
<i>PPT_sm</i>	3	<i>Pinus sylvestris</i>
<i>PPT_sp</i>	4	<i>Pinus sylvestris</i>
<i>PPT_sm</i>	1	<i>Quercus petraea</i>
<i>DDbelow0</i>	2	<i>Quercus petraea</i>
<i>DDabove18</i>	3	<i>Quercus petraea</i>
<i>TD</i>	4	<i>Quercus petraea</i>
<i>PPT_sp</i>	5	<i>Quercus petraea</i>
<i>bFFP</i>	1	<i>Quercus robur</i>
<i>PPT_sm</i>	2	<i>Quercus robur</i>
<i>TD</i>	3	<i>Quercus robur</i>

	<i>AHM</i> 4 <i>Quercus robur</i> <i>MAP</i> 5 <i>Quercus robur</i>
Model settings	The models were run with the following settings of biomod2 ¹⁰² as listed below Parameters for model calibration i. models = all models in biomod2 were used such as GLM, GBM, GAM, CTA, ANN, SRE, FDA, MARS, RF, and MAXENT.Tsuruoka, ii. CV.strateg = cross-validation selection strategy was set to random with 10 repetitions iii. CV.nb.rep = 10 repetitions iv. CV.perc= 0.75 (25% for model evaluation only). For model calibration full dataset was used v. prevalence = 0.5, presences and absences will be weighted equally vi. metric.eval = Model evaluation metric ROC and TSS Parameter settings for individual model algorithms Generalised linear models (GLMs): GLMs were generated assuming a logistic link function and a binomial error distribution of the response variable without interactions. Generalised Boosted Models (GBMs): GBMs were fitted with distribution = 'bernoulli' and a maximum of 25000 trees fitted to the data and with an interaction depth of 4

	Generalised additive models (GAMs): GAMs were generated with default parameter in biomod2 with k (default 4): a smooth term cubic-smooth splines bounded by a degree of smoothness of four for each climatic predictor and family (binomial(link = 'logit')) Classification tree analysis (CTA): CTA was carried out using method = 'class' for presence-absence data, parms = 'default', and 10 fold cross validation Artificial neural networks (ANNs) with the following parameters - NbCV (default 5): nb of cross-validation to find best size and decay parameters; - size (default NULL): number of units in the hidden layer; - decay (default NULL): decay parameter will be optimized by cross-validation on model AUC i.e.. NbCv - rang (default 0.1) - maxit (default 200): maximum number of iterations. Surface range envelopes (SREs): quant (default 0.025): quantile of 'extreme environmental variable' removed for selection of species envelops Multivariate adaptive regression splines (MARS): MARS were fitted with default 0-level interactions between predictors. Random Forests (RFs): The number of trees grown were set to 500 with mtry (default 'default'), nodesize (default 5) and maxnodes (default NULL) MAXENT.Tsuruoka (Maxent) was used with the following default parameters in biomdo2 - l1_regularizer (default 0.0): A numeric turning on L1 regularization and setting the regularization parameter. A value of 0 will disable L1 regularization
--	--

	ii. l2_regularizer (default 0.0): A numeric turning on L2 regularization and setting the regularization parameter. A value of 0 will disable L2 regularization iii. use_sgd (default FALSE): A logical indication that SGD parameter estimation should be used. Defaults to FALSE iv. set_heldout (default 0): An integer specifying the number of documents to hold out. Sets a held-out subset of your data to test against and prevent overfitting v. verbose (default FALSE): A logical specifying whether to provide descriptive output about the training process
Model estimates	The models estimated probability of species occurrence and associated model uncertainty represented by the coefficient of variation from each model combination listed in Supplementary Table 9
Model ensemble	Predicted probabilities from the individual models for each target species were ensembled as a consensus model which combined the median probability over the selected models with the True Skill Statistics threshold (TSS > 0.7)^{107,108}. Parameters for model ensemble i. em.by= all ie. all models are combined into one ensemble ii. metric. select =a vector containing evaluation metric names (ROC, TSS) to be used together with metric. iii. select. thresh = 0.7 for both TSS and ROC Ensemble-models algorithms or the algorithms used for the ensemble were: i. EMmean : Mean of probabilities over the selected models. ii. EMmedian : Median of probabilities over the selected models. The median is less sensitive to outliers than the mean iii. EMcv : Coefficient of variation (sd/mean) of probabilities over the selected models. CV is a measure of uncertainty rather than a measure of the probability of occurrence. If the CV gets a high evaluation score, it means that the uncertainty is high where the

	species is observed																																																																																																																
Threshold selection	ROC (Receiver operating characteristic curve threshold >0.7) and True Skill Statistics threshold (TSS > 0.7), a commonly used threshold for SDMs ^{107,108} were used.																																																																																																																
Assessment																																																																																																																	
Model performance statistics	For each such model run as well as the final ensemble models for each target species, the model evaluation statistics were recorded. These statistics were true skill statistics (TSS) and area under the relative operating characteristic (ROC), model sensitivity (the ability of the model to predict true presences), and model specificity (the ability of the model to predict the true absences). TSS takes into account both omission and commission errors and ranges also from −1 to +1, not being affected by prevalence as KAPPA¹⁰⁸. TSS values ranging from 0.2 to 0.5 were considered poor, from 0.6 to 0.8 useful, and values larger than 0.8 were good to excellent (e.g.¹⁰⁷). Prediction accuracy is considered to be similar to random for ROC values lower than 0.5; poor, for values in the range 0.5–0.7; fair in the range 0.7–0.9; and excellent for values greater than 0.9¹⁰⁹. Statistics for evaluation for each of the models used to develop the ensemble SDM for the seven tree species.* MAXENT Tsuruoka <table><tr><th>Species</th><th>Criteria</th><th>Testing data</th><th>Evaluating data</th><th>Sensitivity</th><th>Specificity</th><th>Model</th></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.94</td><td>0.93</td><td>96.72</td><td>94.93</td><td>GLM</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>1.00</td><td>1.00</td><td>97.30</td><td>96.48</td><td>GLM</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.95</td><td>0.95</td><td>97.34</td><td>97.86</td><td>GBM</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>1.00</td><td>0.99</td><td>97.34</td><td>97.94</td><td>GBM</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.97</td><td>0.96</td><td>98.33</td><td>97.55</td><td>GAM</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>1.00</td><td>1.00</td><td>97.08</td><td>99.02</td><td>GAM</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.95</td><td>0.95</td><td>96.35</td><td>98.80</td><td>CTA</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>1.00</td><td>1.00</td><td>96.74</td><td>98.59</td><td>CTA</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.98</td><td>0.98</td><td>98.37</td><td>99.19</td><td>ANN</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>0.99</td><td>0.99</td><td>98.12</td><td>99.79</td><td>ANN</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.76</td><td>0.76</td><td>79.99</td><td>96.26</td><td>SRE</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>0.89</td><td>0.89</td><td>79.99</td><td>96.26</td><td>SRE</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.95</td><td>0.94</td><td>97.60</td><td>96.35</td><td>FDA</td></tr><tr><td><i>Abies alba</i></td><td>ROC</td><td>1.00</td><td>1.00</td><td>96.91</td><td>97.38</td><td>FDA</td></tr><tr><td><i>Abies alba</i></td><td>TSS</td><td>0.95</td><td>0.95</td><td>97.25</td><td>97.38</td><td>MARS</td></tr></table>	Species	Criteria	Testing data	Evaluating data	Sensitivity	Specificity	Model	<i>Abies alba</i>	TSS	0.94	0.93	96.72	94.93	GLM	<i>Abies alba</i>	ROC	1.00	1.00	97.30	96.48	GLM	<i>Abies alba</i>	TSS	0.95	0.95	97.34	97.86	GBM	<i>Abies alba</i>	ROC	1.00	0.99	97.34	97.94	GBM	<i>Abies alba</i>	TSS	0.97	0.96	98.33	97.55	GAM	<i>Abies alba</i>	ROC	1.00	1.00	97.08	99.02	GAM	<i>Abies alba</i>	TSS	0.95	0.95	96.35	98.80	CTA	<i>Abies alba</i>	ROC	1.00	1.00	96.74	98.59	CTA	<i>Abies alba</i>	TSS	0.98	0.98	98.37	99.19	ANN	<i>Abies alba</i>	ROC	0.99	0.99	98.12	99.79	ANN	<i>Abies alba</i>	TSS	0.76	0.76	79.99	96.26	SRE	<i>Abies alba</i>	ROC	0.89	0.89	79.99	96.26	SRE	<i>Abies alba</i>	TSS	0.95	0.94	97.60	96.35	FDA	<i>Abies alba</i>	ROC	1.00	1.00	96.91	97.38	FDA	<i>Abies alba</i>	TSS	0.95	0.95	97.25	97.38	MARS
Species	Criteria	Testing data	Evaluating data	Sensitivity	Specificity	Model																																																																																																											
<i>Abies alba</i>	TSS	0.94	0.93	96.72	94.93	GLM																																																																																																											
<i>Abies alba</i>	ROC	1.00	1.00	97.30	96.48	GLM																																																																																																											
<i>Abies alba</i>	TSS	0.95	0.95	97.34	97.86	GBM																																																																																																											
<i>Abies alba</i>	ROC	1.00	0.99	97.34	97.94	GBM																																																																																																											
<i>Abies alba</i>	TSS	0.97	0.96	98.33	97.55	GAM																																																																																																											
<i>Abies alba</i>	ROC	1.00	1.00	97.08	99.02	GAM																																																																																																											
<i>Abies alba</i>	TSS	0.95	0.95	96.35	98.80	CTA																																																																																																											
<i>Abies alba</i>	ROC	1.00	1.00	96.74	98.59	CTA																																																																																																											
<i>Abies alba</i>	TSS	0.98	0.98	98.37	99.19	ANN																																																																																																											
<i>Abies alba</i>	ROC	0.99	0.99	98.12	99.79	ANN																																																																																																											
<i>Abies alba</i>	TSS	0.76	0.76	79.99	96.26	SRE																																																																																																											
<i>Abies alba</i>	ROC	0.89	0.89	79.99	96.26	SRE																																																																																																											
<i>Abies alba</i>	TSS	0.95	0.94	97.60	96.35	FDA																																																																																																											
<i>Abies alba</i>	ROC	1.00	1.00	96.91	97.38	FDA																																																																																																											
<i>Abies alba</i>	TSS	0.95	0.95	97.25	97.38	MARS																																																																																																											

	<i>Abies alba</i>	ROC	0.99	0.99	97.34	97.34	MARS
	<i>Abies alba</i>	TSS	1.00	1.00	99.15	99.62	RF
	<i>Abies alba</i>	ROC	1.00	1.00	99.67	99.32	RF
	<i>Abies alba</i>	TSS	0.90	0.91	95.19	95.14	MAXENT*
	<i>Abies alba</i>	ROC	1.00	0.99	96.44	94.19	MAXENT*
	<i>Fagus sylvatica</i>	TSS	0.82	0.82	94.06	88.12	GLM
	<i>Fagus sylvatica</i>	ROC	0.98	0.97	94.01	88.18	GLM
	<i>Fagus sylvatica</i>	TSS	0.84	0.83	93.92	89.69	GBM
	<i>Fagus sylvatica</i>	ROC	0.99	0.99	94.66	89.13	GBM
	<i>Fagus sylvatica</i>	TSS	0.84	0.84	95.32	89.27	GAM
	<i>Fagus sylvatica</i>	ROC	0.98	0.98	95.21	89.40	GAM
	<i>Fagus sylvatica</i>	TSS	0.87	0.86	94.48	92.36	CTA
	<i>Fagus sylvatica</i>	ROC	0.97	0.97	94.48	92.36	CTA
	<i>Fagus sylvatica</i>	TSS	0.86	0.86	94.93	91.49	ANN
	<i>Fagus sylvatica</i>	ROC	0.99	0.99	96.06	90.49	ANN
	<i>Fagus sylvatica</i>	TSS	0.61	0.60	76.06	84.17	SRE
	<i>Fagus sylvatica</i>	ROC	0.83	0.82	76.06	84.17	SRE
	<i>Fagus sylvatica</i>	TSS	0.83	0.82	91.81	89.28	FDA
	<i>Fagus sylvatica</i>	ROC	1.00	1.00	93.52	87.84	FDA
	<i>Fagus sylvatica</i>	TSS	0.84	0.84	93.72	88.95	MARS
	<i>Fagus sylvatica</i>	ROC	1.00	1.00	93.16	89.64	MARS
	<i>Fagus sylvatica</i>	TSS	0.99	0.97	98.25	97.15	RF
	<i>Fagus sylvatica</i>	ROC	1.00	0.99	98.16	97.31	RF
	<i>Fagus sylvatica</i>	TSS	0.79	0.78	94.55	82.34	MAXENT*
	<i>Fagus sylvatica</i>	ROC	0.96	0.96	94.77	82.19	MAXENT*
	<i>Larix decidua</i>	TSS	0.89	0.90	96.21	92.34	GLM
	<i>Larix decidua</i>	ROC	1.01	1.02	96.50	92.19	GLM
	<i>Larix decidua</i>	TSS	0.93	0.93	95.71	95.14	GBM
	<i>Larix decidua</i>	ROC	1.00	0.99	95.66	95.54	GBM
	<i>Larix decidua</i>	TSS	0.95	0.95	97.24	95.73	GAM
	<i>Larix decidua</i>	ROC	1.00	1.00	96.85	96.18	GAM
	<i>Larix decidua</i>	TSS	0.96	0.94	97.20	95.34	CTA
	<i>Larix decidua</i>	ROC	1.02	1.01	97.20	95.34	CTA
	<i>Larix decidua</i>	TSS	0.96	0.96	97.69	96.57	ANN
	<i>Larix decidua</i>	ROC	0.99	1.00	96.80	97.65	ANN
	<i>Larix decidua</i>	TSS	0.68	0.68	77.92	88.79	SRE
	<i>Larix decidua</i>	ROC	0.86	0.86	77.92	88.79	SRE
	<i>Larix decidua</i>	TSS	0.90	0.91	96.50	92.44	FDA
	<i>Larix decidua</i>	ROC	1.01	1.01	96.26	93.17	FDA
	<i>Larix decidua</i>	TSS	0.91	0.92	97.49	93.22	MARS
	<i>Larix decidua</i>	ROC	0.99	1.02	97.54	93.22	MARS
	<i>Larix decidua</i>	TSS	1.00	0.98	99.37	98.64	RF

<i>Larix decidua</i>	ROC	1.00	1.00	99.32	98.74	RF
<i>Larix decidua</i>	TSS	0.75	0.77	95.96	81.16	MAXENT*
<i>Larix decidua</i>	ROC	0.94	0.94	95.96	81.21	MAXENT*
<i>Picea abies</i>	TSS	0.85	0.85	92.55	92.75	GLM
<i>Picea abies</i>	ROC	1.00	0.98	92.85	92.50	GLM
<i>Picea abies</i>	TSS	0.92	0.91	97.64	94.02	GBM
<i>Picea abies</i>	ROC	1.00	1.00	97.58	94.12	GBM
<i>Picea abies</i>	TSS	0.92	0.91	96.09	95.54	GAM
<i>Picea abies</i>	ROC	1.00	1.00	95.64	96.12	GAM
<i>Picea abies</i>	TSS	0.95	0.93	98.61	94.97	CTA
<i>Picea abies</i>	ROC	1.01	0.99	98.61	94.97	CTA
<i>Picea abies</i>	TSS	0.90	0.88	97.36	91.82	ANN
<i>Picea abies</i>	ROC	0.98	0.98	97.45	91.80	ANN
<i>Picea abies</i>	TSS	0.66	0.66	76.34	89.55	SRE
<i>Picea abies</i>	ROC	0.84	0.84	76.34	89.55	SRE
<i>Picea abies</i>	TSS	0.85	0.85	92.47	92.56	FDA
<i>Picea abies</i>	ROC	0.99	0.99	94.27	91.01	FDA
<i>Picea abies</i>	TSS	0.88	0.88	94.35	94.00	MARS
<i>Picea abies</i>	ROC	1.00	1.00	95.29	93.15	MARS
<i>Picea abies</i>	TSS	1.00	0.99	99.98	99.96	RF
<i>Picea abies</i>	ROC	1.00	1.00	99.92	99.96	RF
<i>Picea abies</i>	TSS	0.67	0.66	94.96	67.03	MAXENT*
<i>Picea abies</i>	ROC	0.90	0.90	95.44	66.61	MAXENT*
<i>Pinus sylvestris</i>	TSS	0.80	0.80	92.19	88.51	GLM
<i>Pinus sylvestris</i>	ROC	0.98	0.98	92.66	88.15	GLM
<i>Pinus sylvestris</i>	TSS	0.89	0.88	97.97	90.24	GBM
<i>Pinus sylvestris</i>	ROC	1.00	0.99	95.81	92.73	GBM
<i>Pinus sylvestris</i>	TSS	0.87	0.88	95.47	92.28	GAM
<i>Pinus sylvestris</i>	ROC	1.00	1.00	94.72	93.18	GAM
<i>Pinus sylvestris</i>	TSS	0.93	0.92	98.71	93.17	CTA
<i>Pinus sylvestris</i>	ROC	0.99	0.98	98.66	93.29	CTA
<i>Pinus sylvestris</i>	TSS	0.90	0.90	97.04	93.46	ANN
<i>Pinus sylvestris</i>	ROC	0.99	0.98	96.75	93.80	ANN
<i>Pinus sylvestris</i>	TSS	0.58	0.59	76.97	82.50	SRE
<i>Pinus sylvestris</i>	ROC	0.80	0.80	76.97	82.50	SRE
<i>Pinus sylvestris</i>	TSS	0.82	0.82	92.55	89.42	FDA
<i>Pinus sylvestris</i>	ROC	0.98	0.98	93.94	88.39	FDA
<i>Pinus sylvestris</i>	TSS	0.85	0.85	96.13	88.80	MARS
<i>Pinus sylvestris</i>	ROC	0.99	0.99	95.57	89.47	MARS
<i>Pinus sylvestris</i>	TSS	1.00	0.99	99.91	99.45	RF
<i>Pinus sylvestris</i>	ROC	1.00	1.00	98.89	99.42	RF
<i>Pinus sylvestris</i>	TSS	0.51	0.52	94.03	58.05	MAXENT*

	<i>Pinus sylvestris</i>	ROC	0.84	0.83	94.35	57.82	MAXENT*
	<i>Quercus robur</i>	TSS	0.87	0.87	92.70	94.21	GLM
	<i>Quercus robur</i>	ROC	0.99	0.99	92.65	94.35	GLM
	<i>Quercus robur</i>	TSS	0.91	0.91	97.01	94.11	GBM
	<i>Quercus robur</i>	ROC	0.99	1.00	95.77	95.46	GBM
	<i>Quercus robur</i>	TSS	0.90	0.90	95.82	94.26	GAM
	<i>Quercus robur</i>	ROC	0.99	0.99	96.30	93.90	GAM
	<i>Quercus robur</i>	TSS	0.93	0.93	97.18	95.64	CTA
	<i>Quercus robur</i>	ROC	1.00	1.00	97.18	95.64	CTA
	<i>Quercus robur</i>	TSS	0.93	0.94	97.76	95.82	ANN
	<i>Quercus robur</i>	ROC	0.99	0.99	97.30	96.34	ANN
	<i>Quercus robur</i>	TSS	0.73	0.73	77.70	96.02	SRE
	<i>Quercus robur</i>	ROC	0.88	0.88	77.70	96.02	SRE
	<i>Quercus robur</i>	TSS	0.86	0.86	93.35	92.96	FDA
	<i>Quercus robur</i>	ROC	0.99	1.00	92.96	93.45	FDA
	<i>Quercus robur</i>	TSS	0.88	0.87	93.51	94.11	MARS
	<i>Quercus robur</i>	ROC	1.00	1.00	93.23	94.43	MARS
	<i>Quercus robur</i>	TSS	1.00	0.98	99.56	98.95	RF
	<i>Quercus robur</i>	ROC	1.00	1.00	99.07	99.54	RF
	<i>Quercus robur</i>	TSS	0.79	0.80	96.84	83.68	MAXENT*
	<i>Quercus robur</i>	ROC	0.96	0.97	96.73	83.80	MAXENT*
	<i>Quercus petraea</i>	TSS	0.77	0.76	95.63	80.82	GLM
	<i>Quercus petraea</i>	ROC	0.96	0.96	94.38	82.53	GLM
	<i>Quercus petraea</i>	TSS	0.80	0.80	90.83	89.88	GBM
	<i>Quercus petraea</i>	ROC	0.98	0.98	91.54	89.29	GBM
	<i>Quercus petraea</i>	TSS	0.82	0.82	92.78	89.45	GAM
	<i>Quercus petraea</i>	ROC	0.98	0.98	91.98	90.64	GAM
	<i>Quercus petraea</i>	TSS	0.87	0.85	94.93	90.35	CTA
	<i>Quercus petraea</i>	ROC	0.98	0.97	94.93	90.35	CTA
	<i>Quercus petraea</i>	TSS	0.85	0.85	93.89	91.42	ANN
	<i>Quercus petraea</i>	ROC	0.98	0.98	93.75	91.59	ANN
	<i>Quercus petraea</i>	TSS	0.66	0.67	80.47	86.99	SRE
	<i>Quercus petraea</i>	ROC	0.84	0.85	80.47	86.99	SRE
	<i>Quercus petraea</i>	TSS	0.78	0.78	92.97	85.11	FDA
	<i>Quercus petraea</i>	ROC	0.97	0.97	91.39	87.06	FDA
	<i>Quercus petraea</i>	TSS	0.80	0.80	96.41	83.71	MARS
	<i>Quercus petraea</i>	ROC	0.97	0.97	96.57	83.62	MARS
	<i>Quercus petraea</i>	TSS	0.99	0.92	96.39	95.76	RF
	<i>Quercus petraea</i>	ROC	1.00	1.00	96.46	95.76	RF
	<i>Quercus petraea</i>	TSS	0.71	0.70	93.16	77.05	MAXENT*
	<i>Quercus petraea</i>	ROC	0.91	0.92	95.07	75.36	MAXENT*

Statistics for evaluation for the ensemble SDM for the seven tree species.

Model evaluation					
Species	Criteria	Testing data	Evaluation data	Sensitivity	Specificity
<i>Abies alba</i>	ROC	0.99	0.99	94.74	96.67
<i>Fagus sylvatica</i>	ROC	0.96	0.96	92.09	88.19
<i>Larix decidua</i>	ROC	0.97	0.97	94.06	92.28
<i>Picea abies</i>	ROC	0.97	0.97	93.47	90.08
<i>Pinus sylvestris</i>	ROC	0.95	0.95	93.03	87.01
<i>Quercus petraea</i>	ROC	0.95	0.96	91.74	86.45
<i>Quercus robur</i>	ROC	0.98	0.98	92.96	93.36
<i>Abies alba</i>	TSS	0.93	0.92	94.78	96.44
<i>Fagus sylvatica</i>	TSS	0.82	0.81	91.79	88.40
<i>Larix decidua</i>	TSS	0.87	0.87	94.19	92.02
<i>Picea abies</i>	TSS	0.85	0.84	93.16	90.32
<i>Pinus sylvestris</i>	TSS	0.81	0.81	93.18	86.72
<i>Quercus petraea</i>	TSS	0.80	0.79	91.83	86.19
<i>Quercus robur</i>	TSS	0.87	0.87	93.21	93.04

In addition to internal evaluation, the model predictions were also tested against independent data from the European Forest Genetic Conservation Units (GCU)¹¹⁰. The ensemble models for the seven target species for the period 1961-1990 were used to predict their respective probability of occurrence at each GCU location. A GCU location was assumed to be correctly classified as a “presence location”, if its probability of occurrence is above the species-specific median of cut-offs used by biomod2 to convert the probability of SDM algorithms to binary values. biomod2 converts the continuous probability of occurrence to binary presence and absence by selecting a cutoff value that maximizes the TSS. The median cut-off was calculated from those models which constituted the ensemble prediction. The following table shows the number of GCUs correctly and incorrectly classified. The ensemble SDMs were able to correctly classify the GCU locations from 79% in the case of *Pinus sylvestris* up to 94 % for *Picea abies*.

	Number of correctly and incorrectly detected presence locations of GCUs <table><thead><tr><th colspan="6">Number of GCU</th></tr><tr><th>Species</th><th>Cut off</th><th>Correct</th><th>Incorrect</th><th>Total</th><th>% Correct</th></tr></thead><tbody><tr><td><i>Abies alba</i></td><td>468.94</td><td>214</td><td>13</td><td>227</td><td>94.27</td></tr><tr><td><i>Picea abies</i></td><td>571.05</td><td>348</td><td>21</td><td>369</td><td>94.31</td></tr><tr><td><i>Pinus sylvestris</i></td><td>637.41</td><td>166</td><td>45</td><td>211</td><td>78.67</td></tr><tr><td><i>Larix decidua</i></td><td>525.81</td><td>114</td><td>11</td><td>125</td><td>91.20</td></tr><tr><td><i>Fagus sylvatica</i></td><td>506.97</td><td>304</td><td>43</td><td>347</td><td>87.61</td></tr><tr><td><i>Quercus petraea</i></td><td>495.41</td><td>199</td><td>26</td><td>225</td><td>88.44</td></tr><tr><td><i>Quercus robur</i></td><td>487.83</td><td>136</td><td>37</td><td>173</td><td>78.61</td></tr></tbody></table>	Number of GCU						Species	Cut off	Correct	Incorrect	Total	% Correct	<i>Abies alba</i>	468.94	214	13	227	94.27	<i>Picea abies</i>	571.05	348	21	369	94.31	<i>Pinus sylvestris</i>	637.41	166	45	211	78.67	<i>Larix decidua</i>	525.81	114	11	125	91.20	<i>Fagus sylvatica</i>	506.97	304	43	347	87.61	<i>Quercus petraea</i>	495.41	199	26	225	88.44	<i>Quercus robur</i>	487.83	136	37	173	78.61
Number of GCU																																																							
Species	Cut off	Correct	Incorrect	Total	% Correct																																																		
<i>Abies alba</i>	468.94	214	13	227	94.27																																																		
<i>Picea abies</i>	571.05	348	21	369	94.31																																																		
<i>Pinus sylvestris</i>	637.41	166	45	211	78.67																																																		
<i>Larix decidua</i>	525.81	114	11	125	91.20																																																		
<i>Fagus sylvatica</i>	506.97	304	43	347	87.61																																																		
<i>Quercus petraea</i>	495.41	199	26	225	88.44																																																		
<i>Quercus robur</i>	487.83	136	37	173	78.61																																																		
Prediction																																																							
Prediction output	Predicted probabilities from the individual models and target species were ensembled as a consensus model which combined the median probability over the selected models with the True Skill Statistics threshold (TSS > 0.7)^{107,108}. The median was chosen because it is known to be less sensitive to outliers than the mean. The estimated ensemble model predictions were presented as GeoTIFF rasters.																																																						
Uncertainty quantification	Each model algorithm of the biomod2 predicted the probability of occurrence (potential distribution) for the seven tree species. For contemporary climate, potential distributions of each of the seven species were estimated from 10 SDM models of biomod2 and 1 ensemble SDM (Supplementary Table 9). In future climate scenarios of RCP4.5 and 8.5 the potential distributions were estimated for combinations of 10 SDMs and 5 RCMs and 1 ensemble SDM with a mean of 5 RCMs (Supplementary Table 9). For each such model combination parameters were estimated by summary statistics (minimum, mean, maximum), standard deviation, standard error, and coefficient of variation in the estimated potential distribution area of the seven species. For details see the section on uncertainty analysis below.																																																						

175

176

177 Selecting the climatically best-suited species from species distribution models

178 To select the climatically best-suited species for each 30 arc-sec grid cell of Europe, the potential
179 distribution range of the seven European tree species was first estimated. Thereafter, for each forested
180 grid cell of Europe, species with the highest probability of occurrence were selected and referred to as
181 “best-suited species” with the following steps:

182 Step1. The probability of occurrence for the 7 species was obtained for 10 SDMs and their ensemble in
183 the contemporary climate, and for all SDM-RCM combinations and their ensemble in future climate
184 scenarios.

185 Step2. For each forested grid cell of Europe, the species with the highest probability of occurrence was
186 identified to be the climatically best-suited species for the corresponding model combinations. In this
187 step, the untransformed probability of occurrence of each species was used. The best-suited species
188 were identified for all SDM-RCM model combinations described in Supplementary Table 9.

189 Identifying best-suited seed provenances – modeling provenance growth and CS
190

191 To quantify local adaptation and to identify best-suited seed provenances for future climate change, we
192 applied universal response functions (URF), where the climate of provenance origin, the climate of trial
193 locations, and the tree age of the respective provenance trial were integrated into a common model to
194 predict carbon stocks per hectare (C_{stock}^{-ha}) at age 40. Since the URFs have age as a co-variable, our
195 models allow the prediction of SPC performance at different tree ages. However, for concise reporting,
196 we chose the tree age 40 because; i) the majority of the provenance trials were inventoried until the
197 age of 40, ii) we avoid the uncertainty of predicting the models to ages that were not covered by the
198 experimental provenance trial data iii) it roughly corresponds to the middle of rotation of most of the
199 studied species. To fit the compiled provenance data of the seven tree species as the URF model, we
200 applied the machine learning approach Random Forests¹⁰⁴. From the list of 80 potential bioclimatic
201 variables (Supplementary Table 3), the set of variables to be used in the URFs was selected
202 (Supplementary Table 7) using a recursive feature elimination approach (RFE) implemented within the
203 Random Forest algorithm¹⁰⁴. Within the RFE approach, the variables were eliminated iteratively,

starting from the full set of potential predictors (Supplementary Table 3), and retaining only those variables that reduce the mean square error over random permutations of the same variable. Variables that were linearly correlated (Spearman's $r \geq 0.7$) with other variables and had a variance inflation factors $VIF > 5$ were identified and the ones with the lower value according to the Akaike Information Criteria (AIC) were retained for further model development. This subset of uncorrelated climate variables and tree age (Supplementary Table 7) was used as predictor variables for developing the URFs. The URFs explain around 91-95% of the variation in $Cstock^{-ha}$ (Supplementary Table 12). To account for uncertainty in the future, the CS was estimated for different URF and RCM combinations (Supplementary Table 9).

Supplementary Table 7. Selected bioclimatic variables used for universal response function(URF) development and model accuracy. Variance explained by each bioclimatic variable used in the URFs as the percentage increase in mean squared error (MSE). Variance explained depicts the variance explained in $Cstock^{-ha}$ by each species-specific URFs. Variables with suffix (_p) denote variables of provenance or seed origin and those without suffix _p denote variables of planting location.

Species		Variables	% increase in MSE	Total variance explained
<i>Abies alba</i>	AHM	Annual heat-moister index	20.89	92.7
	bFFP	Beginning of the frost-free period	21.88	
	TD	Continentality	27.90	
	TD_p	Continentality	9.36	
	SHM_p	Summer heat-moisture index	4.05	
	PPT_sp_p	Mean precipitation of spring	1.87	
	age	Tree age	16.23	
<i>Picea abies</i>	PPT_at	Mean precipitation of autumn	14.07	91.07
	MCMT	Mean coldest month temperature	9.34	
	FFP	Frost-free period	8.90	
	SHM	Summer heat-moisture index	14.42	
	TD	Continentality	15.51	
	PPT_at_p	Mean precipitation of autumn	33.05	
	PPT_sp_p	Mean precipitation of spring	37.89	
	MCMT_p	Mean coldest month temperature	24.65	
	bFFP_p	Beginning of the frost-free period	28.11	
	EMT_p	Extreme minimum temperature	29.30	
	TD_p	Continentality	32.75	
	age	Tree age	80.17	
<i>Pinus sylvestris</i>	PPT_at	Mean precipitation of autumn	16.87	94.91
	TD	Continentality	20.26	
	DDbelow0	Degree days-below 0°C	20.90	
	Tmin_wt	Minimum winter temperature	17.05	

Species		Variables	% increase in MSE	Total variance explained
	SHM	Summer heat-moisture index	19.80	
	TD_p	Continentality	10.06	
	MCMT_p	Mean coldest month temperature	7.45	
	DDbelow0_p	Degree days-below 0°C	6.08	
	NFFD_p	Number of frost free days	8.12	
	age	Tree age	46.86	
<i>Larix decidua</i>	Tmin_an	Minimum annual temperature	13.84	95.17
	Tave_sp	Average spring temperature	14.01	
	MCMT	Mean coldest month temperature	13.65	
	EMT	Extreme minimum temperature	13.39	
	DDbelow0	Degree days below 0°C	11.81	
	NFFD	Number of frost free days	10.93	
	AHM_p	Annual heat-moister index	13.80	
	PPT_sp_p	Mean precipitation of spring	10.23	
	SHM_p	Summer heat-moisture index	12.64	
	eFFP_p	End of frost-free period	7.95	
	age	Tree age	31.60	
<i>Fagus sylvatica</i>	SHM	Summer heat-moisture index	5.69	95.02
	TD	Continentality	5.86	
	DDabove5	Degree days above 5°C	7.97	
	Tave_sp	Average spring temperature	6.32	
	PPT_sp	Mean precipitation of spring	7.21	
	EMT	Extreme minimum temperature	9.73	
	MCMT	Mean coldest month temperature	7.55	
	NFFD_p	Number of frost free days	3.32	
	TD_p	Continentality	2.16	
	DDabove5_p	Degree days above 5°C	-0.13	
	SHM_p	Summer heat-moisture index	-0.63	
	PPT_sp_p	Mean precipitation of spring	0.14	
	age	Tree age	25.75	
<i>Quercus petraea</i>	EMT	Extreme minimum temperature	11.44	93.61
	Tmin_at	Minimum autumn temperature	9.52	
	DDbelow0	Degree days below 0°C	9.42	
	Tmin_sp	Minimum spring temperature	9.76	
	PPT_at	Mean precipitation of autumn	13.15	
	Tave_at_p	Average autumn temperature	3.65	
	FFP_p	Frost free period	3.48	
	MCMT_p	Mean coldest month temperature	3.78	
	SHM_p	Summer heat-moisture index	4.26	
	age	Tree age	44.29	
<i>Quercus robur</i>	bFFP	Beginning of frost-free period	13.93	95.43
	DDabove5	Degree days above 5°C	13.68	
	MWMT	Mean warmest month temperature	12.71	
	Tave_sm	Average summer temperature	11.99	

Species	Variables	% increase in MSE	Total variance explained
TD	Continentality	19.27	
Tmax_an	Annual maximum temperature	11.37	
AHM_p	Annual heat-moisture index	7.90	
DDabove5_p	Degree days above 5°C	9.38	
PPT_at_p	Mean precipitation of autumn	6.14	
bFFP_p	Beginning of frost-free period	6.93	
MWMT_p	Mean warmest month temperature	10.00	
age	Tree age	54.85	

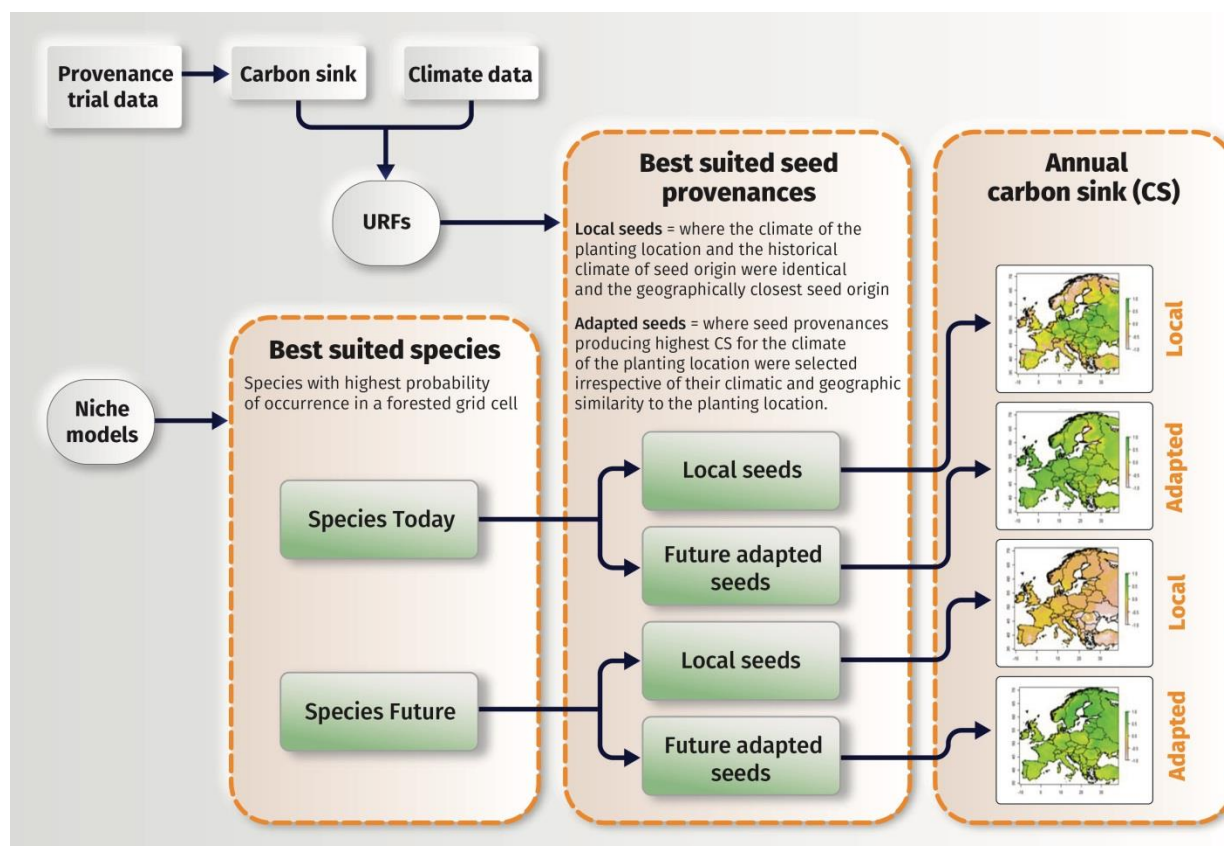
Demonstrating assisted migration – modeling seed sourcing scenarios

The URFs predicted Cstock^{-ha} for the tree age at 40 for each 30 arcsec resolution raster cell of Europe. We restricted our predictions only to those grid cells of Europe where at least 50% of the area is covered by forest according to European Forest cover data³⁶. We termed these grid cells the “forested grid cells” of Europe. We adopted this approach to avoid the uncertainty arising from predicting the URFs models beyond the current species’ ranges and tested provenances origin.

For concise reporting, the Cstock^{-ha} at age 40 predicted by the URFs for each species-specific SPC was converted to annual carbon sink (CS) by dividing the Cstock^{-ha} by 40 for contemporary climate (1991-2010) and the two climate change scenarios RCP4.5 and 8.5 for the periods 2021-40 and 2061-80.

For the URF predictions, the climate of the planting location corresponded to the climate of each forested grid cell at a given time frame, whereas the climate of seed origin was represented by historical (1961-90) mean values of species-specific seed provenance clusters (SPC) (Supplementary Table 8). These SPC represent climatically similar and geographically proximate groups of populations to correspond with range-wide local adaptation patterns. The SPC for each species was developed in two steps: 1) Principal component analysis of the climate variables used in the URFs, and 2) Partitioning around medoids of those principal components which cumulatively explain at least 95% of the variation in climate. The number of PAM clusters for each species was decided from the reducing error curve, whereby increasing the number of clusters does not further reduce the error in classification.

239 After identifying the best-suited species for each forested grid cell for each combination of SDM-URF
 240 and climate models, we demonstrated AM schemes for seven European tree species with two AM
 241 scenarios: 1) "local seeds " where the climate of the planting location and the historical climate of seed
 242 origin were identical and geographically proximate, and 2) "adapted seeds", where seed provenances
 243 producing highest CS for the climate of the planting location were selected irrespective of their climatic
 244 and geographic similarity to the planting location (Supplementary Fig.4).



245

246 **Supplementary Fig. 4.** Schematic diagram of the assisted migration (AM) scenarios: 1) "local seeds" where the
 247 climate of the planting location and the seed origin were identical and geographically proximate, and 2)
 248 "adapted seeds", where seed provenances producing the highest annual carbon sink (CS) and adapted to future
 249 conditions were selected.

250

251 **Supplementary Table 8. Climatic descriptions of the seed provenance clusters (SPC).** Mean values of the
252 bioclimatic variables defining SPC for the period 1961-1990.

Species	Cluster	Climatic variable of seed provenance clusters (SPC)							
		TD	SHM	PPT_sp					
A. alba	1	30.18	47.95	142.57					
	2	28.96	31.66	203.10					
	3	25.03	39.49	239.69					
	4	28.40	39.41	175.28					
	5	26.76	23.29	279.59					
	6	24.53	23.47	342.34					
	7	24.64	32.11	277.17					
	8	31.16	25.96	215.11					
	9	21.72	55.90	255.34					
Cluster		PPT_at	MCMT	FFP	SHM	TD	PPT_sp	bFFP	EMT
P. abies	1	244.06	-5.77	150.57	28.17	28.18	228.70	125.26	-31.87
	2	327.90	-5.93	123.22	19.68	26.92	290.84	143.52	-33.66
	3	155.37	-4.35	168.11	40.72	28.72	151.81	115.93	-28.67
	4	226.48	-3.00	174.85	39.19	26.05	202.61	114.68	-26.09
	5	401.14	-2.48	166.63	24.29	24.71	329.88	121.23	-27.15
	6	136.87	-3.36	184.75	50.95	28.50	128.77	107.74	-26.75
	7	221.84	-0.97	195.30	47.56	20.68	168.61	104.87	-22.58
	8	320.63	2.24	204.16	42.89	7.08	215.60	100.41	-18.09
	9	198.77	-6.17	161.23	51.42	26.59	120.87	118.68	-31.65
	10	171.15	-11.56	112.16	49.23	31.19	108.16	146.13	-39.87
	11	300.17	-7.43	125.99	30.75	20.88	195.09	139.55	-33.65
Cluster		TD	MCMT	DDbelow0	NFFD				
P. sylvestris	1	27.73	-2.70	305.95	251.74				
	2	27.46	-4.49	491.10	228.40				
	3	19.73	-2.96	338.43	219.99				
	4	31.63	-11.82	1352.73	161.01				
	5	28.80	-8.76	968.69	183.12				
	6	20.57	1.15	106.35	276.13				
	7	13.14	-0.90	206.98	247.55				
	8	31.85	-14.11	1761.66	139.28				

253

Supplementary Table 8 continued.

	Cluster	AHM	PPT_sp	SHM	eFFP	
<i>L. decidua</i>	1	13.52	262.75	22.68	267.03	
	2	24.50	179.76	47.89	297.46	
	3	13.51	290.07	24.62	276.60	
	4	29.53	134.27	49.51	290.99	
	5	20.10	190.73	34.34	279.67	
	6	24.79	150.93	42.13	285.33	
	7	19.58	216.59	41.43	302.68	
	Cluster	NFFD	TD	DDabove5	SHM	PPT_sp
<i>F. sylvatica</i>	1	223.15	24.07	1421.88	26.86	343.1
	2	244.65	27.09	1871.91	51.01	139.41
	3	192.70	27.82	1200.10	26.34	240.09
	4	221.33	29.18	1627.15	37.86	184.43
	5	259.78	24.49	1880.78	45.95	180
	6	262.58	24.01	1898.22	41.66	280.17
	7	295.79	11.42	1853.48	52.27	171.58
	8	303.11	18.73	2284.68	55.14	332.45
	Cluster	Tave_at	FFP	MCMT	SHM	
<i>Q. petraea</i>	1	7.48	163.13	-1.69	32.01	
	2	8.44	184.81	-3.31	53.05	
	3	8.11	178.34	-3.40	45.92	
	4	7.65	174.30	-3.71	41.55	
	5	8.85	191.95	-1.31	47.76	
	6	9.85	204.00	0.11	47.14	
	7	10.58	223.55	1.82	61.27	
	8	11.99	250.91	3.72	57.21	
	Cluster	AHM	DDabove5	PPT_at	bFFP	MWMT
<i>Q. robur</i>	1	29.09	1921.17	154.39	106.92	18.72
	2	24.51	1668.21	181.35	113.2	17.66
	3	22.32	1939.46	289.34	100.05	18.02
	4	32.87	2479.67	160.23	97.3	21.32
	5	14.96	1405.92	429.39	127.81	14.81
	6	26.74	2322.74	256.75	89.82	19.58
	7	24.27	2851.43	333.67	69.8	20.91
	8	20.29	2885.25	507.49	35.82	18.43

255 Analyzing uncertainty in identifying best-suited species and seed provenances

256

257 To account for uncertainty in the identified best-suited species and best-suited-provenance, we
 258 estimated potential species distributions from all possible combinations of SDMs and RCMs separately
 259 and then applied URFs for each such combination to identify the best-suited seed provenance cluster
 260 and to estimate the annual carbon sink of the identified forest regrowth. This resulted in 1666 model
 261 combinations. (Supplementary Table 9). For each such model combination, statistical parameters such
 262 as minimum, maximum, mean, standard deviation, standard error, and coefficient of variation in the
 263 area occupied by best-suited species and the estimated annual carbon sink (CS) were calculated.

264 **Supplementary Table 9.** Model combinations to analyze the uncertainty in identifying best-suited species and
 265 best-suited seed provenance cluster (SPC) under contemporary and future climate scenarios.

Climate scenario	Model type	Model combination	Total model combination
Contemporary climate (1991-2010)	Models estimating best-suited species	10 SDM x 7 species	70
		1 Ensemble SDM x 7 species	7
	Models estimating best SPC	10 SDMs x 7 species x1URFs	70
		1 Ensemble SDM x 7 species x1 URFs	7
RCP4.5 (2061-80)	Models estimating best-suited species	10 SDM x 7 species x 5 RCMs	350
		1 Ensemble SDM x 7 species x 1 mean RCM	7
	Models estimating best SPC	10 SDM x 7 species x 5 RCMs x 1 URFs	350
		1 Ensemble SDM x 7 species x7 URFs	49
RCP8.5 (2061-80)	Models estimating best-suited species	10 SDM x 7 speies x 5 RCMs	350
		1 Ensemble SDM x 7 species x 1 mean RCM	7
	Models estimating best SPC	10 SDM x 7 species x 5 RCMs x 7 URFs	350
		1 Ensemble SDM x 7 species x7 URFs	49
Sum			1666

267

268

269 Spatial partitioning of the components of uncertainty in the estimated carbon sink
270

271 In addition to quantifying the uncertainty in identifying the best-suited species by the SDMs and the
272 best suited SPCs and estimation of annual carbon sink (CS) by the URFs, we also quantified the relative
273 contribution or partitioning of the sources of such uncertainty and their geographic patterns. This was
274 done for two major sources of uncertainty i) uncertainty due to the modelling framework (U-MF) ii)
275 uncertainty due to the modelling framework and assisted migration scenario (U-AM). For both cases we
276 performed analysis of variance (ANOVA) at each of the 33,475,75 forested grid cells of Europe under
277 contemporary climate (1991-2010) and two climate change scenarios of RCP4.5 and 8.5 for the period
278 2061-2080.

279 For U-MF, we performed a three-way ANOVA ⁴¹ with interaction between all terms to partition the
280 uncertainty of CS estimates individually for both local and adapted seed sources. This analysis
281 considered the variation contributed by the 10 SDM algorithms of biomod2⁹ used to identify the best-
282 suited species, the five Regional Climate Models (RCMs) used to calibrate the SDM algorithms, and the
283 Universal Response Function (URF).

284 For U-AM, we examined the contribution of the AM scenario combined with the choice of models to
285 identify the best-suited species (SDM algorithms) and the URFs on the total variation of CS. For this
286 analysis, we considered the 10 SDM algorithms, the five RCMs used in SDM, the URF, and the seed
287 source recommendation (local or adapted) as additional factors in the ANOVA. In this case, a four-way
288 ANOVA without interaction between terms was employed. Although we initially considered an
289 interaction between terms, the percentage allocated to the unexplained variance component
290 (residuals) obtained from our initial runs was insignificant, leading us to discard the interaction terms.
291 Moreover, this allowed us to compute the *p-values* for each term of the ANOVA. For U-AM under the
292 contemporary climate, the variation in CS was explained by SDM algorithms and seed sources
293 (local/adapted).

294 In the RCP 4.5 and 8.5 scenarios, alongside seed source and the algorithms used in calibrating the
295 SDMs, we included the RCMs twice, as they were utilized for both SDM and URF calibration. This
296 expanded the dataset to 500 observations (10 x 2 x 5 x 5) for each of the 3,347,575 forested grid cells.

297 To assess the contribution of each component, we calculated the percentage of variance explained by
298 each component by dividing the sum of squares corresponding to each component by the total sum of
299 squares. Additionally, we extracted p-values at a significance level of 0.05. The percentages explained
300 by each component, derived from the ANOVA, were then used to create spatial representations.
301 Meanwhile, the percentages of significant and nonsignificant cells were presented as a histogram.

302 However, given that only the SDM algorithm could be considered for U-MF under contemporary
303 climate conditions, resulting in single values for each algorithm, we opted not to conduct this analysis
304 as it lacked meaningful interpretation

305

306 **SUPPLEMENTARY RESULTS**

307 Uncertainty in identifying best-suited species and seed provenances

308

309 Potential cultivation area varies between climate scenarios, SDMs, and the combination of SDMs and
310 RCMs used to estimate potential species distribution (Supplementary Table 10). The variation in
311 potential cultivation area represented by the coefficient of variation (CV) ranges from 4-15% under
312 contemporary climate, from 12-19% under RCP4.5, and 14-25% under RCP8.5 (Supplementary Table
313 10). In general, the variation in suitable area is higher in the broadleaved species compared to the
314 conifers in contemporary climate (Supplementary Table 10). Although the variation in potential
315 cultivation increases in future climate scenarios, it does not differ between broadleaved and conifer
316 species (Supplementary Table 10).

317 The variation in the suitable area under contemporary climate is due to the 10 SDMs (Supplementary
318 Fig. 5, Supplementary Table 10), while for future scenarios the variation is due to the combination of
319 10 SDMs and 5 RCMs climate models used to estimate potential species distribution (Supplementary
320 Fig.6, Supplementary Table 10). In both climate change scenarios, the variation in the suitable area
321 attributed to RCMs is comparatively higher than the variation attributed to the different SDMs
322 (Supplementary Fig.7). Despite such variation, the relative ranks of the species in terms of their total

323 suitable area for cultivation remains relatively constant irrespective of the used SDMs and RCMs
324 (Kruskal-Wallis test, $p>0.05$), (Supplementary Fig. 8).

325 The recommended cultivation area for each seed provenance clusters (SPC) varies between the climate
326 change scenarios due to the different URF-RCM combinations (Supplementary Fig. 9 &10). Despite the
327 variance in the suitability of the SPC, the relative ranks of the suitability of the SPC of each species
328 remains constant irrespective of the SDM-RCM combination (Supplementary Fig. 9) and does not differ
329 statistically (Kruskal-Wallis test, $p>0.05$) from the best-suited SPC identified by the URFs for the species
330 identified by the ensemble SDM (Fig.1, Supplementary Fig.11).

331 Estimated annual carbon sink varies due to the 10 SDM-URF combinations in contemporary climate
332 (Supplementary Fig.13, Supplementary Table 12) and 10 SDM-5RCM-URF combinations in future
333 scenarios (Supplementary Fig. 14, Supplementary Table 12). For example in the RCP8.5 scenario the
334 annual carbon sink from the seven species vary from 32.48 to 66.36TgCyr⁻¹ when adapted seed sources
335 and from 17.53 to 30.66 TgCyr⁻¹ when local seed sources are planted (Supplementary Table 6).
336 Irrespective of the choice of SDM, RCM, and URFs, the annual carbon sink of adapted seed sources is
337 significantly larger than the CS of local seed sources (Kruskal-Wallis test, $p<0.05$) (Supplementary Fig.
338 15).

339

340

341 **Supplementary Table 10. Variation in the potential cultivation areas in 1000 km², where each of the seven**
342 **species has the highest climatic suitability in contemporary climate (1991-2010) and projected climate change**
343 **scenarios RCP4.5 and 8.5 for the period (2061-80).** The variation in potential cultivation area is represented by
344 minimum, maximum, and mean values estimated by 10 species distribution models (SDMs) in contemporary
345 climate and a combination of 10 SDMs and 5 regional climate models (RCMs) from ECLIPS2.0 climate dataset⁹⁴ in
346 RCP4.5 and 8.5 scenarios. SD, CV, and SE represent standard deviation, coefficient of variation, and standard
347 error respectively. Potential cultivation areas have been assigned for each forested grid cell by selecting the
348 species with the highest probability of occurrence.

Time	Species	Min	Mean	Max	SD	CV	SE
1991-2010	<i>Abies alba</i>	169.7	173.03	200.3	11.1	0.06	3.5
	<i>Picea abies</i>	622.7	647.80	692.5	25.6	0.04	8.1
	<i>Pinus sylvestris</i>	490.4	528.88	571.4	26.5	0.05	8.4
	<i>Larix decidua</i>	19.2	21.48	25	2.1	0.10	0.7
	<i>Fagus sylvatica</i>	43.3	48.86	57.3	4.8	0.10	1.5
	<i>Quercus petraea</i>	59.5	72.27	93.9	11.0	0.15	3.5
	<i>Quercus robur</i>	70.6	77.77	95.9	8.0	0.10	2.5
RCP4.5 (2061-80)	<i>Abies alba</i>	4.1	5.33	7.5	0.9	0.17	0.1
	<i>Picea abies</i>	268.7	349.50	452	49.5	0.14	7
	<i>Pinus sylvestris</i>	469.6	610.45	756.4	71.0	0.12	10
	<i>Larix decidua</i>	3.4	4.92	7.2	0.9	0.18	0.1
	<i>Fagus sylvatica</i>	64.5	89.89	125	15.6	0.17	2.2
	<i>Quercus petraea</i>	5.4	8.07	11.9	1.5	0.19	0.2
	<i>Quercus robur</i>	362.8	501.94	631.4	65.4	0.13	9.2
RCP8.5 (2061-80)	<i>Abies alba</i>	0.7	1.25	2.2	0.3	0.24	0.01
	<i>Picea abies</i>	133.3	219.76	330.6	47.1	0.21	6.7
	<i>Pinus sylvestris</i>	477.7	630.84	869.2	90.4	0.14	12.8
	<i>Larix decidua</i>	1.9	3.77	5.6	0.9	0.24	0.1
	<i>Fagus sylvatica</i>	49.4	84.88	131.5	20.9	0.25	3.0
	<i>Quercus petraea</i>	8.3	14.27	23.7	3.6	0.25	0.5
	<i>Quercus robur</i>	436.2	615.34	739	89.9	0.15	12.7
% Change in cultivation area from 1991-2010 - RCP4.5 (2061-80)	<i>Abies alba</i>	-97.58	-96.92	-96.26	0.13	0.00	0.02
	<i>Picea abies</i>	-56.85	-46.05	-34.73	3.21	-0.07	0.45
	<i>Pinus sylvestris</i>	-4.24	15.42	32.38	7.19	0.47	1.02
	<i>Larix decidua</i>	-82.29	-77.12	-71.20	1.54	-0.02	0.22
	<i>Fagus sylvatica</i>	48.96	83.96	118.15	20.12	0.24	2.85
	<i>Quercus petraea</i>	-90.92	-88.84	-87.33	2.50	-0.03	0.35
	<i>Quercus robur</i>	413.88	545.39	558.39	77.93	0.14	11.02
% Change in cultivation area from 1991-2010 - RCP8.5 (2061-80)	<i>Abies alba</i>	-99.59	-99.28	-98.90	0.15	0.00	0.02
	<i>Picea abies</i>	-78.59	-66.08	-52.26	2.02	-0.03	0.28
	<i>Pinus sylvestris</i>	-2.59	19.28	52.12	15.13	0.78	2.14
	<i>Larix decidua</i>	-90.10	-82.47	-77.60	1.78	-0.02	0.25
	<i>Fagus sylvatica</i>	14.09	73.71	129.49	39.31	0.53	5.56
	<i>Quercus petraea</i>	-86.05	-80.26	-74.76	12.01	-0.15	1.70
	<i>Quercus robur</i>	517.85	691.19	670.59	104.38	0.15	14.76

Supplementary Table 11. Potential cultivation areas of the best-suited tree species and preferred seed provenance selection. The total suitable forest area (given in 1000 km² ± standard deviation) for each species and the two seed source scenarios (local seed provenances and adapted seed provenances), while the share of this area for local and adapted seed provenances is given in percent ± standard deviation. SDM refers to species distribution models used to identify best-suited species in a forested grid cell of Europe and SDM+URF refers to the combination of SDM and Universal Response Functions (URFs) to identify best-suited seed provenance according to the highest CS for the species with the highest climatic suitability in that grid cell. Standard deviations are estimated from the combination of SDMs, URFs, and regional climate models (RCMs).

Time	Models	Suitability	<i>Abies alba</i>	<i>Picea abies</i>	<i>Pinus sylvestris</i>	<i>Larix decidua</i>	<i>Fagus sylvatica</i>	<i>Quercus petraea</i>	<i>Quercus robur</i>
1991-2010	SDM	Total area (1000 km ²)	173.03 ± 11.1	647.8 ± 25.6	528.88 ± 26.5	21.48 ± 2.1	48.86 ± 4.8	72.27 ± 11.0	77.77 ± 8.0
		Area suitable for local (1000 km ²)	31.15 ± 2.26	27.93 ± 1.69	17.82 ± 1.44	3.27 ± 0.31	24.92 ± 0.55	12.02 ± 0.43	6.38 ± 0.2
		Local (% of total area)	18.00 ± 1.21	4.31 ± 0.15	3.37 ± 0.29	15.23 ± 1.85	50.99 ± 0.59	16.63 ± 0.85	8.21 ± 0.14
	SDM+ URF	Area suitable for adapted (1000 km ²)	141.88 ± 12.19	619.87 ± 17.00	511.06 ± 12.12	18.21 ± 1.48	23.95 ± 1.60	60.25 ± 0.31	71.39 ± 1.23
		Adapted (% of total area)	82.00 ± 2.17	95.69 ± 1.35	96.63 ± 1.36	84.77 ± 2.96	49.01 ± 1.09	86.37 ± 2.07	91.79 ± 0.63
RCP4.5 (2061-80)	SDM	Total area (1000 Km ²)	5.33 ± 0.9	349.5 ± 49.5	610.45 ± 71	4.92 ± 0.9	89.89 ± 15.6	8.07 ± 1.5	501.94 ± 65.4
		Area suitable for local (1000 km ²)	0.88 ± 0.19	9.38 ± 1.23	19.1 ± 2.1	0.78 ± 0.15	8.91 ± 11.99	0.83 ± 0.08	89.69 ± 10.52
		Local (% of total area)	16.6 ± 1.23	2.7 ± 1.21	3.1 ± 0.42	15.8 ± 1.21	9.9 ± 0.12	10.03 ± 0.09	17.9 ± 1.22
	SDM+ URF	Area suitable for adapted (1000 km ²)	4.45 ± 0.26	340.12 ± 19.26	591.36 ± 39.00	4.14 ± 0.03	80.97 ± 13.6	7.23 ± 0.21	412.25 ± 22.12
		Adapted (% of total area)	83.44 ± 2.12	97.32 ± 5.12	96.76 ± 6.17	84.17 ± 1.21	90.08 ± 22.85	89.70 ± 0.31	82.13 ± 5.12
RCP8.5 (2061-80)	SDM	Total area (1000 km ²)	1.25 ± 0.3	219.76 ± 47.1	630.84 ± 90.4	3.77 ± 0.9	84.88 ± 20.9	14.27 ± 3.6	615.34 ± 89.9
		Area suitable for local (1000 km ²)	0.24 ± 0.05	4.37 ± 0.91	22.94 ± 3.52	0.59 ± 0.14	4.02 ± 1.03	1.21 ± 0.26	94.4 ± 23
		Local (% of total area)	19.17 ± 1.23	1.99 ± 0.28	3.64 ± 1.32	15.57 ± 2.11	4.73 ± 1.02	8.49 ± 2.00	15.34 ± 0.05
	SDM+ URF	Area suitable for adapted (1000 km ²)	1.01 ± 0.21	215.39 ± 28.25	607.90 ± 52.39	3.18 ± 1.02	80.86 ± 2.36	13.05 ± 0.89	520.94 ± 82.05
		Adapted (% of total area)	80.83 ± 3.25	98.01 ± 24.02	96.36 ± 15.02	84.43 ± 1.02	95.27 ± 1.28	91.51 ± 1.78	84.66 ± 9.98

359 **Supplementary Table 12. Variation in annual carbon sink (TgCyr⁻¹) and CO₂ removal (TgCO₂yr⁻¹) in Europe**
360 **when best-suited species are planted with local or adapted seed provenances.** This variation in annual carbon
361 sink and CO₂ removal is due to the combination of species distribution models (SDMs), regional climate models
362 (RCMs), and universal response functions (URFs) used to identify the best-suited species and seed provenances.

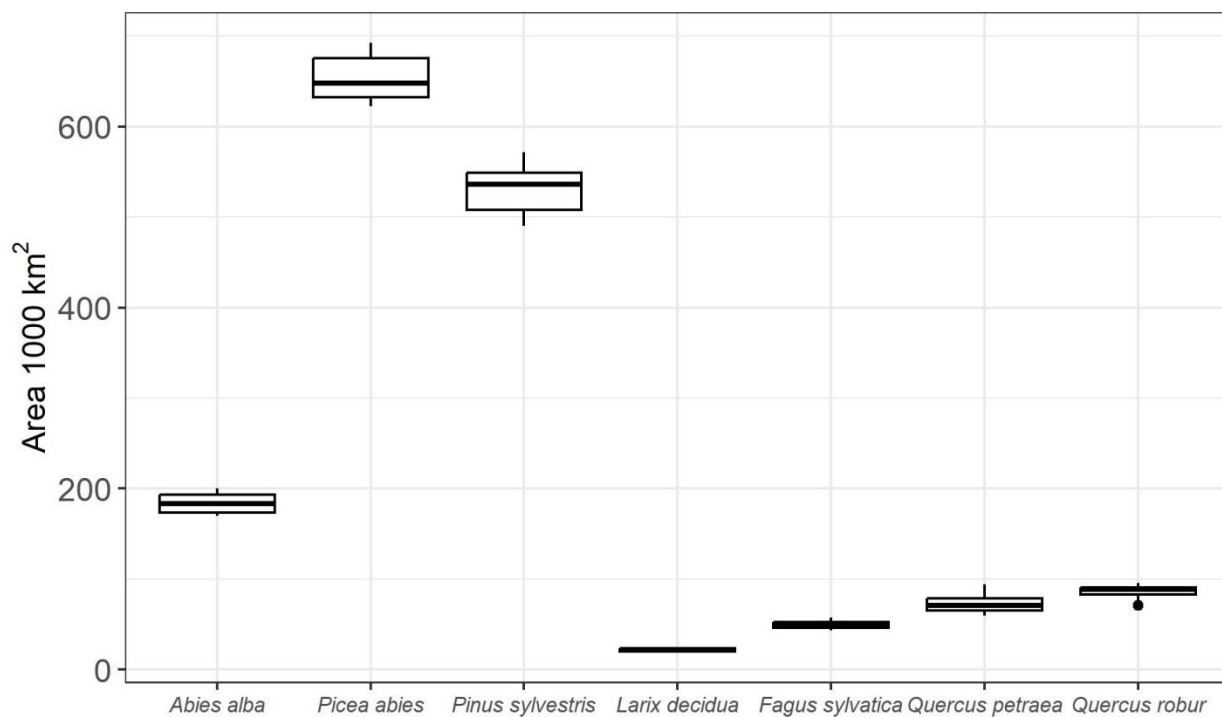
363

Time	Species	Seed source	Min	Mean	Max	SD	CV	SE
1991-2010	<i>Abies alba</i>	Adapted	3.87	3.94	4.56	0.23	0.06	0.07
	<i>Picea abies</i>		21.81	22.69	24.26	0.80	0.04	0.25
	<i>Pinus sylvestris</i>		17.70	19.09	20.63	0.85	0.04	0.27
	<i>Larix decidua</i>		4.00	4.48	5.21	0.38	0.09	0.12
	<i>Fagus sylvatica</i>		8.35	9.42	11.05	1.00	0.11	0.31
	<i>Quercus petraea</i>		0.08	0.10	0.13	0.01	0.10	0.00
	<i>Quercus robur</i>		6.41	7.06	8.71	0.78	0.11	0.25
	Annual carbon sink (TgCyr⁻¹)		62.22	66.78	74.53	1.72	0.03	0.54
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		228.13	244.85	273.29	6.30	0.03	1.99
RCP4.5	<i>Abies alba</i>	Adapted	3.71	4.83	6.79	0.89	0.18	0.28
	<i>Picea abies</i>		13.82	17.98	23.25	2.37	0.13	0.75
	<i>Pinus sylvestris</i>		9.87	12.83	15.90	1.42	0.11	0.45
	<i>Larix decidua</i>		0.19	0.28	0.41	0.05	0.19	0.02
	<i>Fagus sylvatica</i>		5.25	7.32	10.18	1.26	0.17	0.40
	<i>Quercus petraea</i>		0.29	0.44	0.65	0.09	0.20	0.03
	<i>Quercus robur</i>		11.52	15.93	20.04	2.26	0.14	0.72
	Annual carbon sink (TgCyr⁻¹)		44.66	59.61	77.22	2.36	0.04	0.75
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		163.75	218.57	283.13	8.65	0.04	2.73
RCP8.5	<i>Abies alba</i>	Adapted	0.29	0.52	0.91	0.18	0.34	0.06
	<i>Picea abies</i>		7.35	12.12	18.24	2.93	0.24	0.93
	<i>Pinus sylvestris</i>		8.81	11.63	16.02	1.83	0.16	0.58
	<i>Larix decidua</i>		0.08	0.15	0.22	0.04	0.24	0.01
	<i>Fagus sylvatica</i>		3.98	6.83	10.59	1.67	0.24	0.53
	<i>Quercus petraea</i>		0.07	0.13	0.21	0.04	0.28	0.01
	<i>Quercus robur</i>		11.91	16.79	20.17	1.90	0.11	0.60
	Annual carbon sink (TgCyr⁻¹)		32.48	48.18	66.36	3.52	0.07	1.11
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		119.10	176.64	243.33	9.52	0.05	3.01
1991-2010	<i>Abies alba</i>	Local	2.38	2.42	2.81	0.16	0.07	0.05
	<i>Picea abies</i>		12.94	13.46	14.39	0.52	0.04	0.16
	<i>Pinus sylvestris</i>		7.18	7.75	8.37	0.40	0.05	0.13
	<i>Larix decidua</i>		2.77	3.10	3.61	0.27	0.09	0.09
	<i>Fagus sylvatica</i>		6.76	7.62	8.94	0.73	0.10	0.23
	<i>Quercus petraea</i>		0.03	0.04	0.05	0.00	0.08	0.00

	<i>Quercus robur</i>		4.86	5.36	6.61	0.54	0.10	0.17
	Annual carbon sink (TgCyr⁻¹)		36.93	39.76	44.78	1.14	0.03	0.36
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		135.41	145.79	164.19	4.15	0.03	1.31
RCP4.5	<i>Abies alba</i>		1.43	1.85	2.61	0.30	0.16	0.10
	<i>Picea abies</i>		5.44	7.07	9.14	0.89	0.13	0.28
	<i>Pinus sylvestris</i>		3.27	4.25	5.26	0.55	0.13	0.17
	<i>Larix decidua</i>		0.08	0.11	0.16	0.02	0.19	0.01
	<i>Fagus sylvatica</i>	Local	2.28	3.17	4.41	0.53	0.17	0.17
	<i>Quercus petraea</i>		0.11	0.16	0.24	0.03	0.18	0.01
	<i>Quercus robur</i>		4.95	6.84	8.61	0.95	0.14	0.30
	Annual carbon sink (TgCyr⁻¹)		17.54	23.46	30.44	0.98	0.04	0.31
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		64.31	86.03	111.61	3.58	0.04	1.14
RCP8.5	<i>Abies alba</i>		0.16	0.29	0.52	0.10	0.33	0.03
	<i>Picea abies</i>		4.32	3.21	5.10	1.88	0.26	0.60
	<i>Pinus sylvestris</i>		3.58	4.02	6.51	0.79	0.17	0.25
	<i>Larix decidua</i>		0.08	0.15	0.22	0.02	0.14	0.01
	<i>Fagus sylvatica</i>	Local	2.44	4.19	6.49	1.02	0.24	0.32
	<i>Quercus petraea</i>		0.04	0.06	0.11	0.02	0.29	0.01
	<i>Quercus robur</i>		6.91	9.75	11.71	1.40	0.14	0.44
	Annual carbon sink (TgCyr⁻¹)		17.53	21.67	30.66	2.54	0.10	0.80
	Annual CO₂ removal (Tg CO₂ yr⁻¹)		64.28	79.46	112.42	9.33	0.10	2.95

364

365

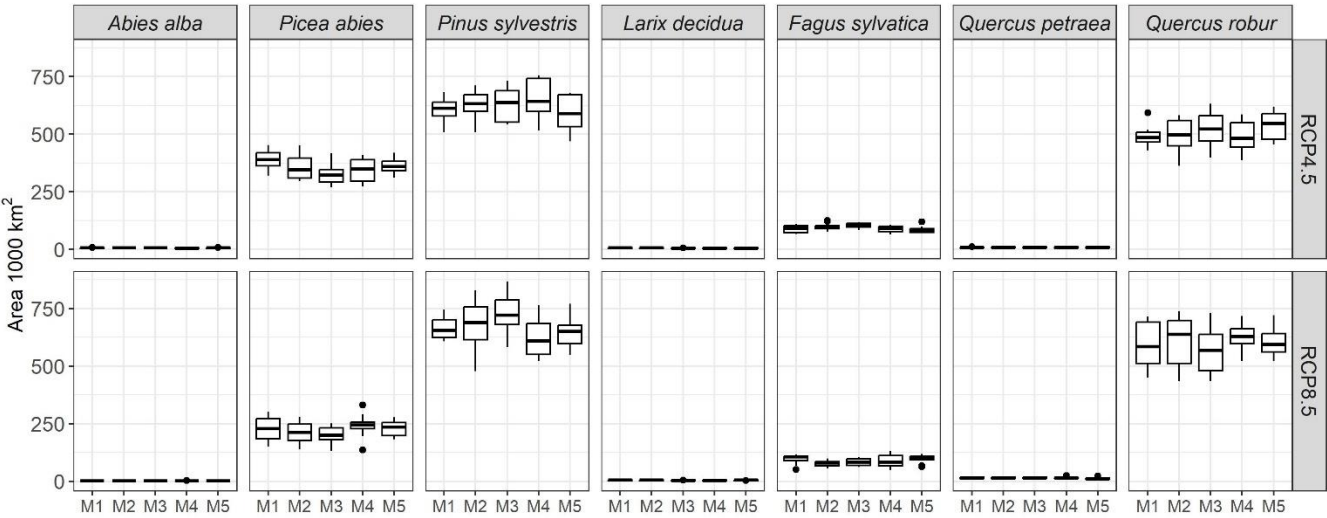


366

367 **Supplementary Fig. 5.** Size of the potential cultivation areas of the seven target species, where for each grid cell
 368 the species with the highest climatic suitability in contemporary climate (1991-2010) has been selected. The
 369 variation in the potential cultivation area shown by the box and whiskers is due to the 10 species distribution
 370 models (SDMs) used to estimate potential species distribution. Boxplot centre, hinges and whiskers denote
 371 median, upper/lower quartiles and upper/lower deciles, respectively.

372

373



374

375

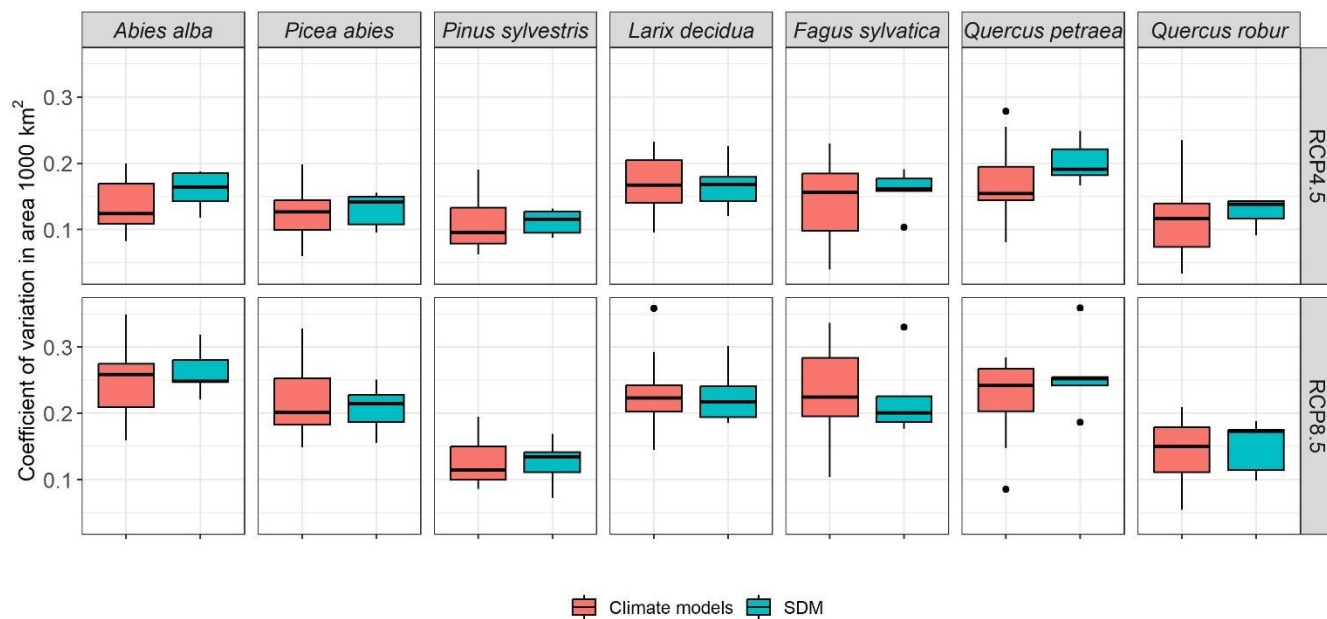
376 **Supplementary Fig. 6.** Size of the potential cultivation areas of the seven target species, where for each grid cell
377 the species with the highest climatic suitability under expected future climate (2061-80) for the scenarios RCP4.5
378 and 8.5 has been selected. The variation in the potential cultivation area shown by box and whiskers is due to
379 the combination of 10 species distribution models (SDMs) and M1 to M5 representing the 5 regional climate
380 models (RCMs) (Supplementary Table 2) used to estimate potential species distribution. Boxplot centre, hinges
381 and whiskers denote median, upper/lower quartiles and upper/lower deciles, respectively.

382

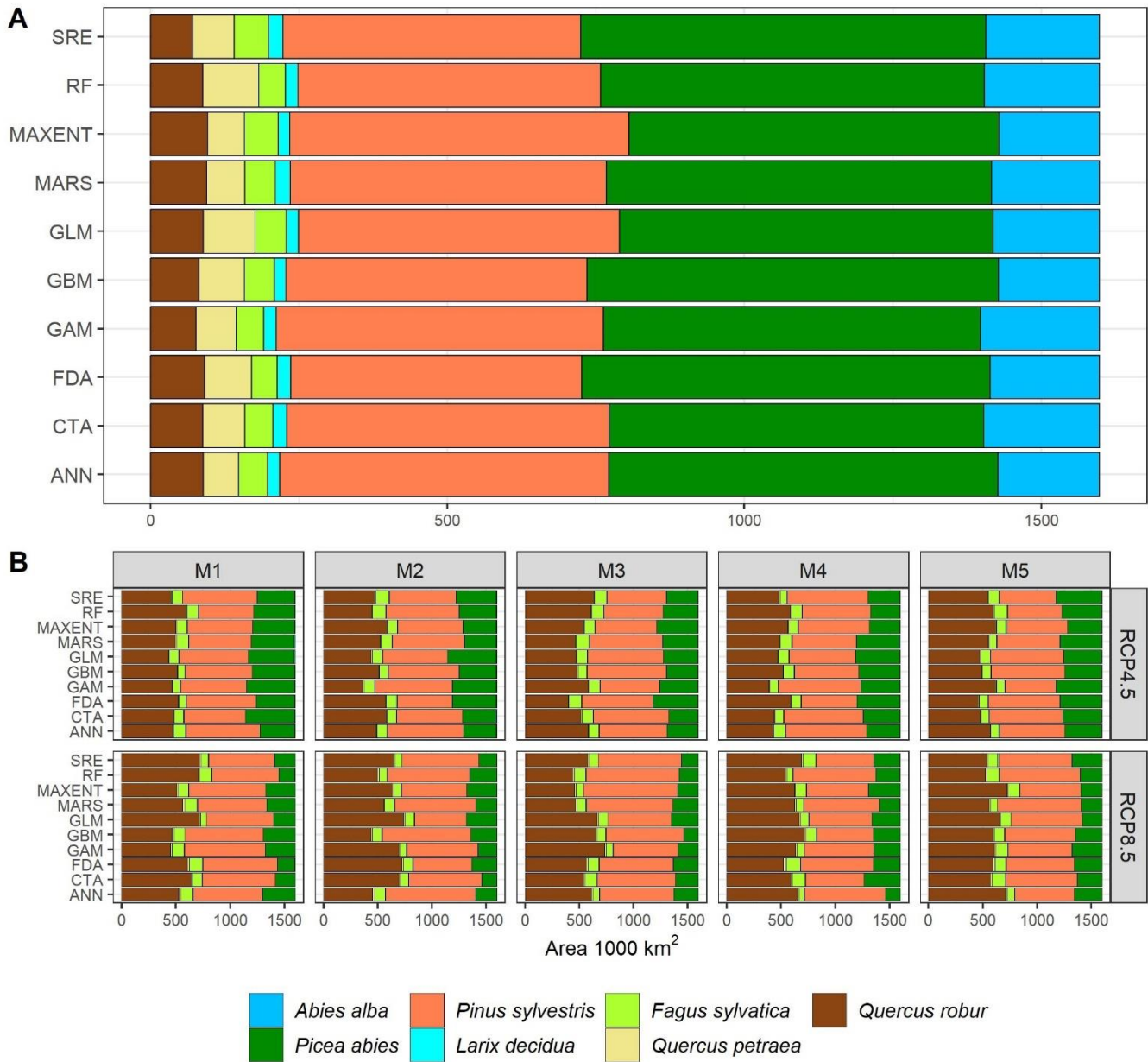
383

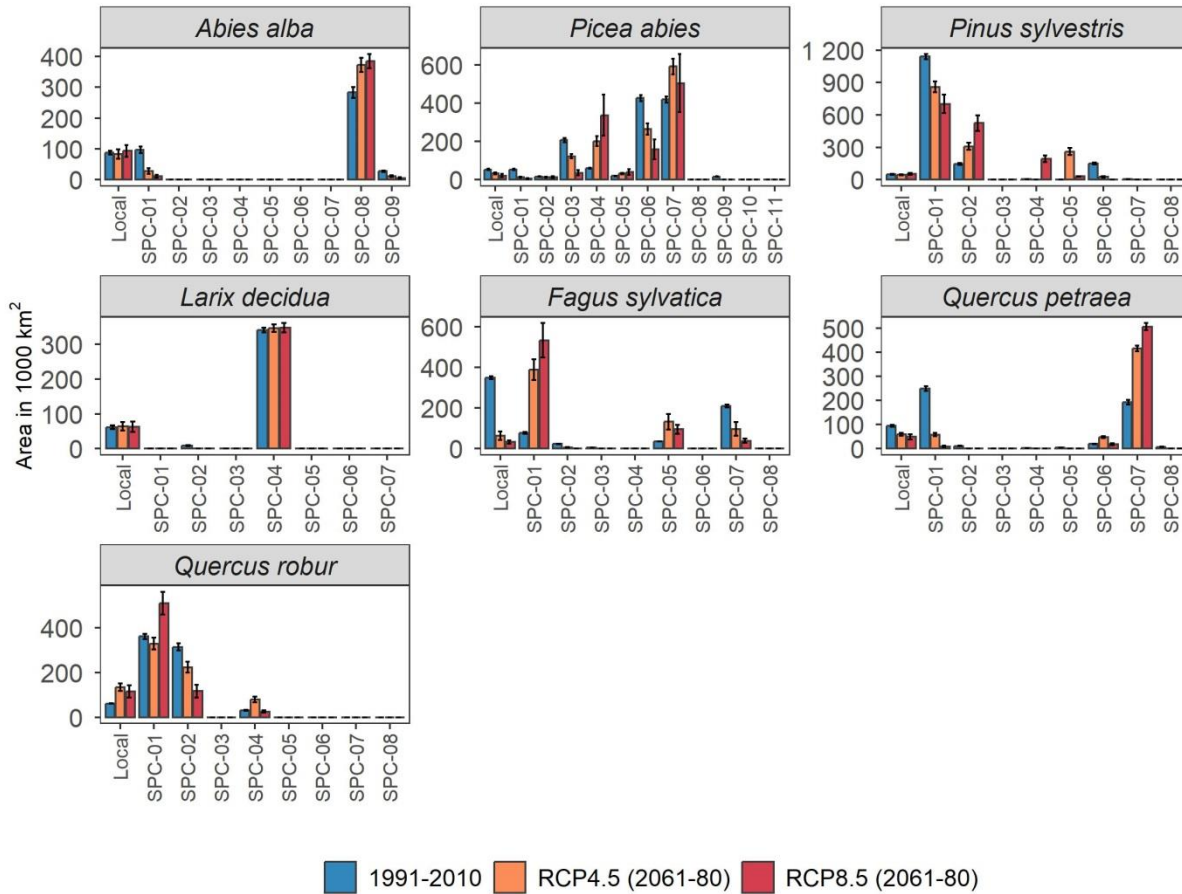
384

385



Supplementary Fig. 7. Coefficient of variation in potential cultivation areas under the climate change scenarios RCP4.5 and 8.5 for the period 2061-80 due to the 10 species distribution models (SDMs) and the 5 regional climate models (RCMs) used to estimate the potential distribution of the seven species. Boxplot centre, hinges and whiskers denote median, upper/lower quartiles and upper/lower deciles, respectively.

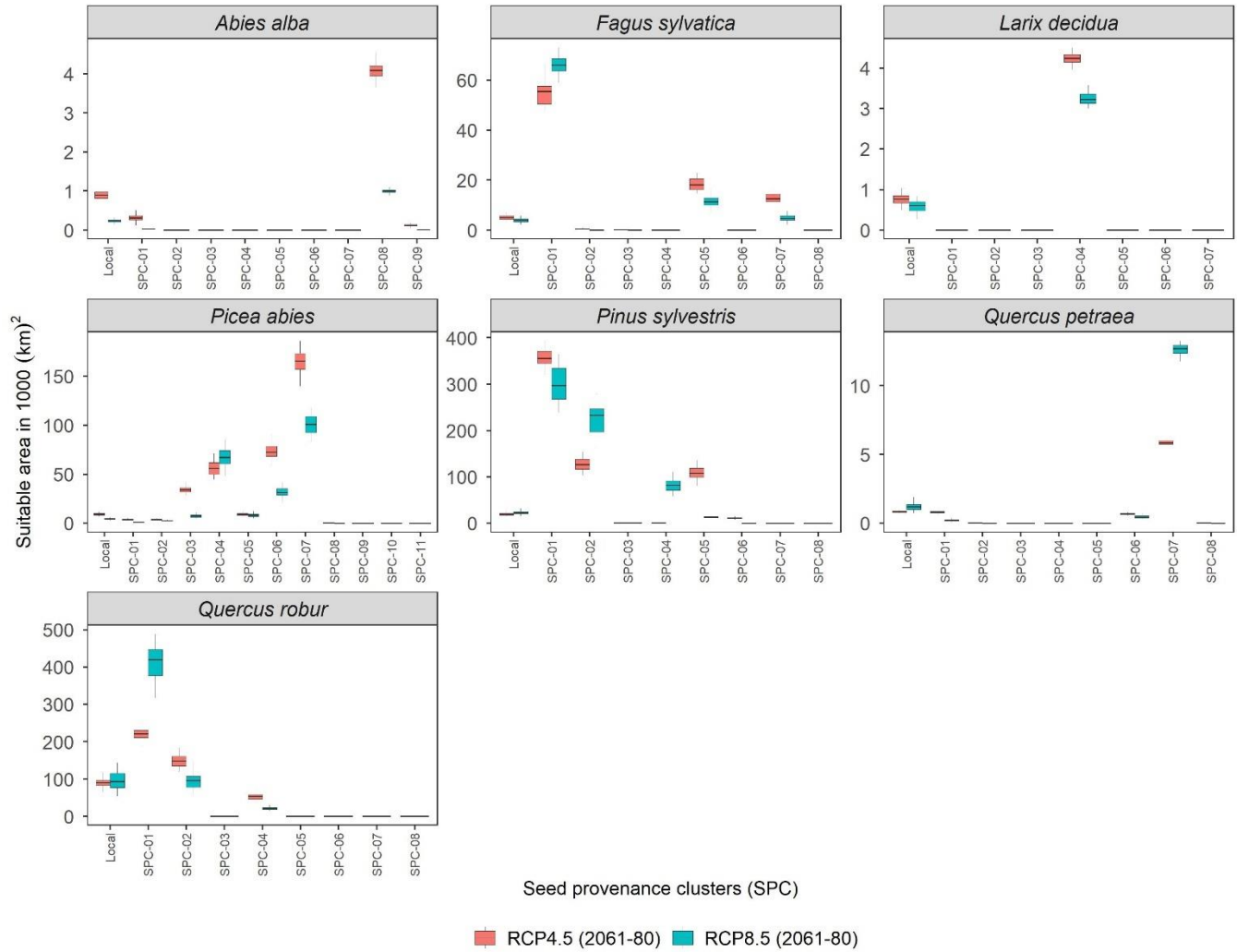




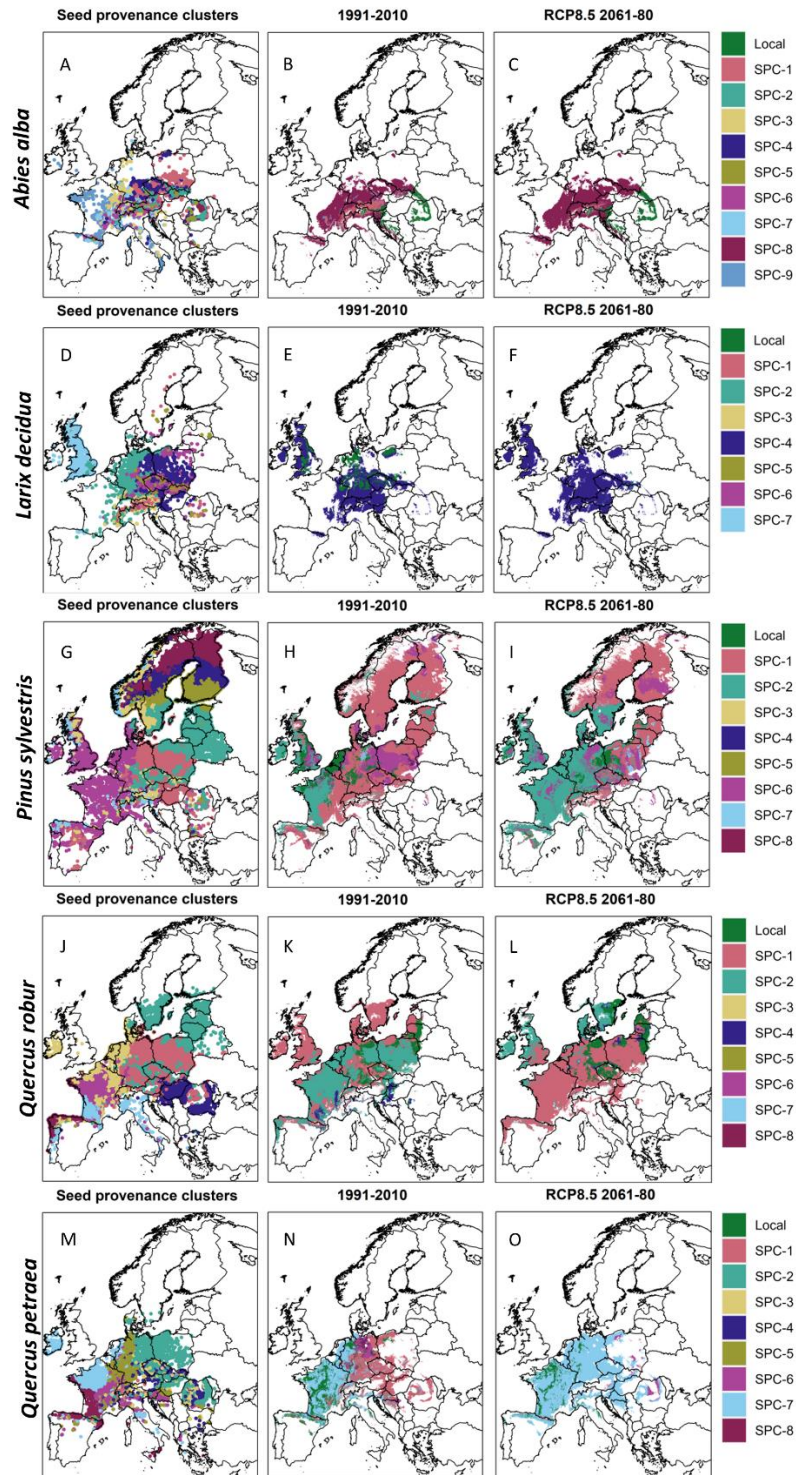
397

398 **Supplementary Fig. 9.** Selected seed provenance clusters of the seven tree species in contemporary and
 399 projected future climate. Size of the recommended cultivation area for the various seed provenance clusters
 400 (SPC) and local seed sources, given in 1000 km², according to the highest annual above-ground carbon sink (CS)
 401 that can be achieved with the respective SPC. Best-suited SPC under climate change scenarios was estimated by
 402 universal response functions (URFs) predicted for the mean climate of 5 regional climate models (RCMs)
 403 (Supplementary Table 2). Error bars show the standard deviation of the recommended cultivation area due to
 404 different combinations of SDMs, URFs, and RCMs. Different colors mark the climate change scenarios RCP4.5
 405 and RCP8.5 for the time period (2061-2080) and contemporary climate for comparison. For the estimated
 406 suitability of SPC under different URF-RCM combinations see Supplementary Fig. 5. The error bars represent the
 407 standard deviation of the size of the recommended cultivation area for the various seed provenance clusters
 408 (SPC) and local seed sources, given in 1000 km², according to the highest annual above-ground carbon sink.

409

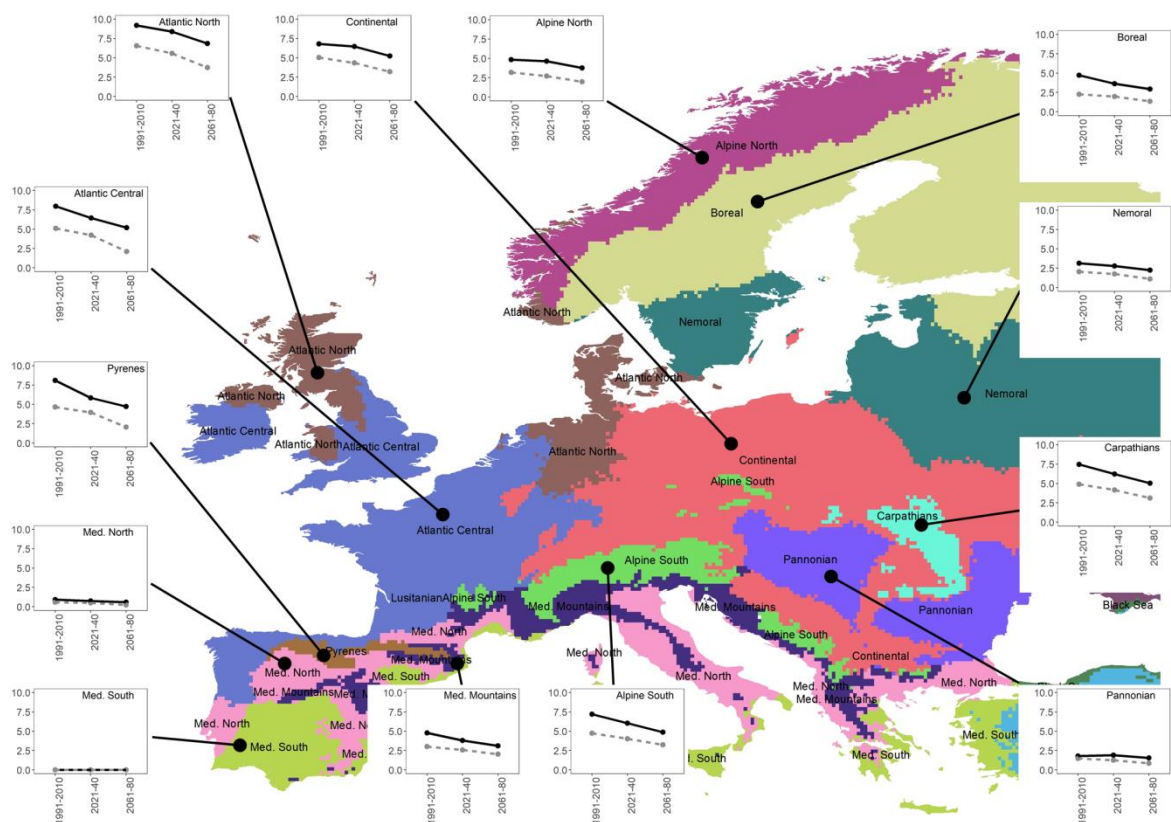


Supplementary Fig. 10. Variation in the size of the recommended cultivation area for the various seed provenance clusters (SPC) and local seed sources (in 1000 km²) according to the highest annual above-ground carbon sink (CS) that can be achieved with the respective SPC. The variation in the recommended cultivation area of each SPC is due to the five regional climate models (M1-M5) (Supplementary Table 2). Boxplot centre, hinges and whiskers denote median, upper/lower quartiles and upper/lower deciles, respectively.

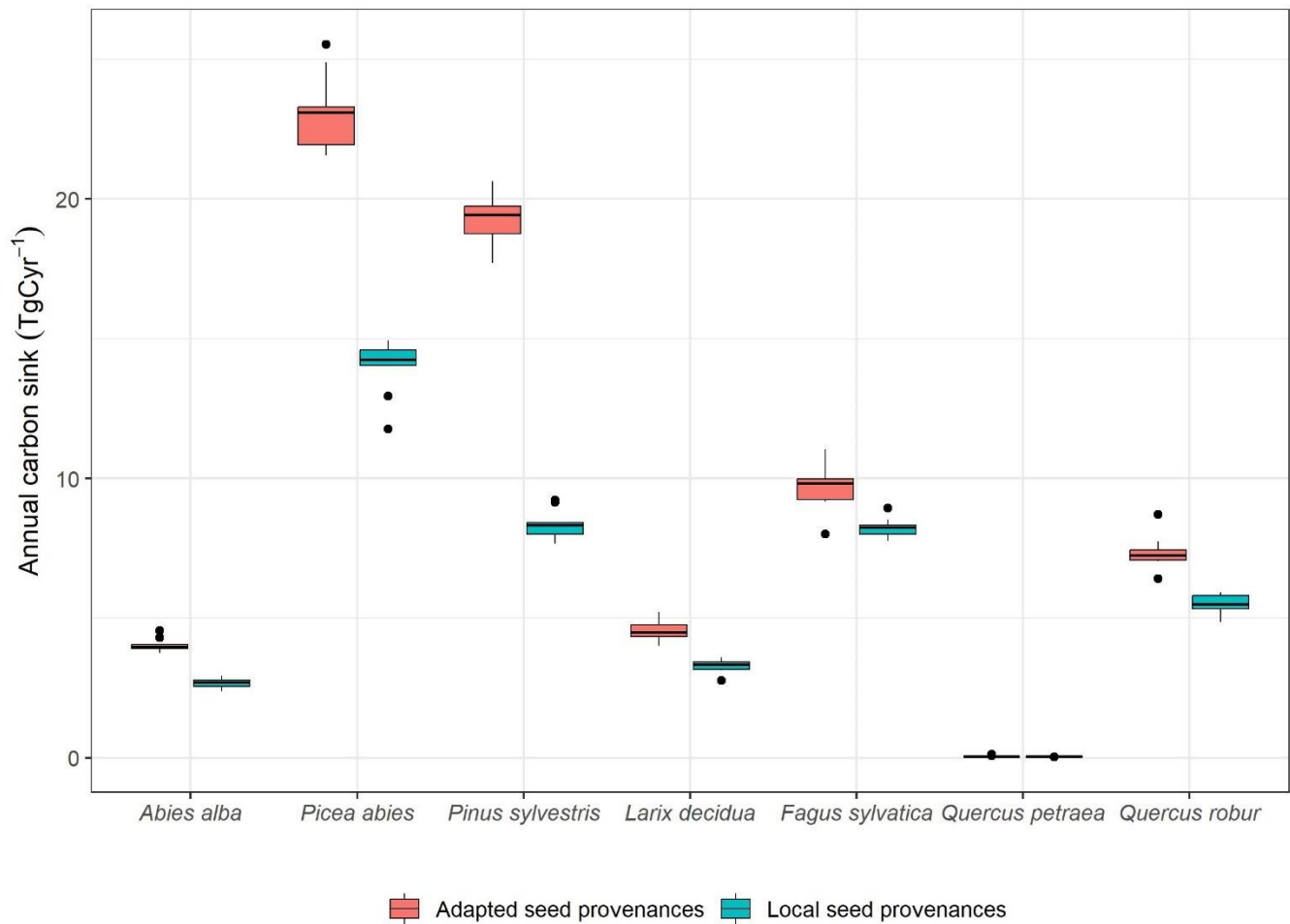


Supplementary Fig. 11. Seed provenance clusters and their application for assisted migration (AM). Geographic locations of the seed provenance clusters (SPC) and distribution maps of the tree species, where the SPC that is predicted to yield the highest annual carbon sink (CS) for each 1km grid cell are marked in the respective SPC

color. Green colors mark areas, where local seed sources are the most productive seed provenances. SPC selections are shown for the contemporary climate (1991-2010) and future (2061-2080) climate under the RCP8.5 scenario. Best-suited SPC in the RCP8.5 scenario were identified by URFs predicted for the mean climate of 5 RCMs (Supplementary Table 2). For better visualization, these maps present every 1km grid cell of the contemporary distribution of the species, while the aboveground carbon sink reported in the study is restricted only to forested grid cells of Europe.

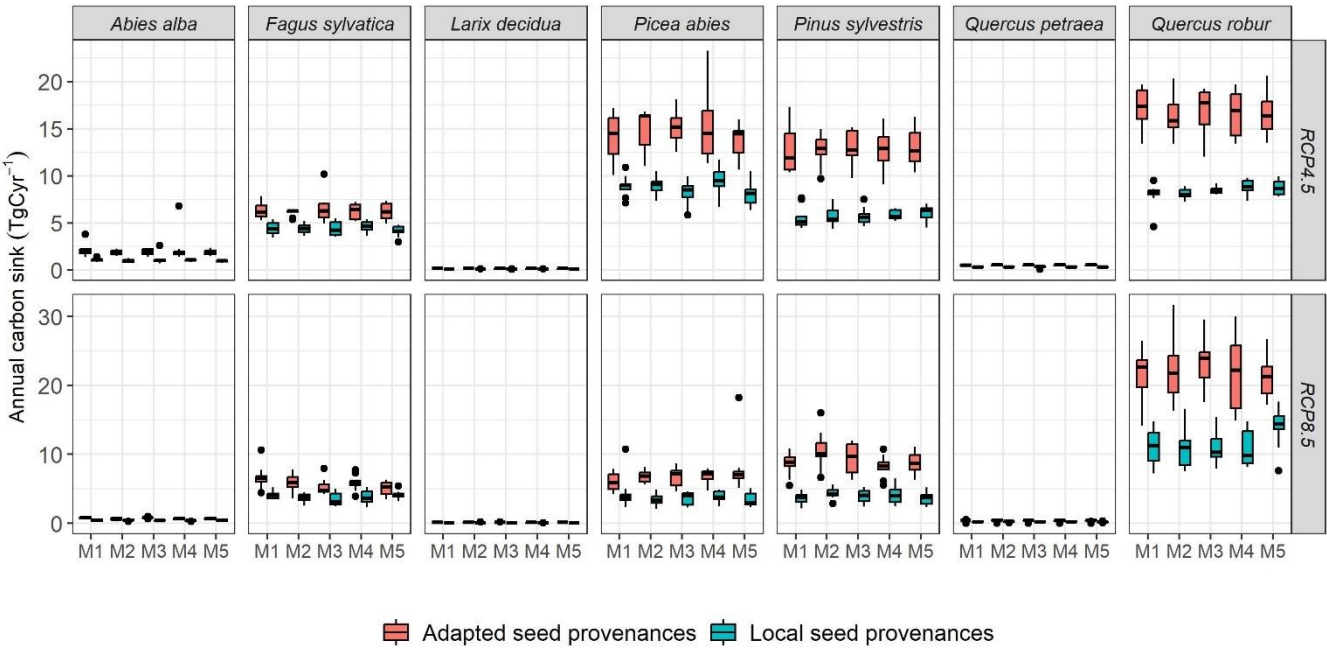


Supplementary Fig. 12. Regional variation of the annual carbon sink for the two seed source scenarios within all European bioclimatic zones. Total annual carbon sink (TgCyr-1) across the seven European tree species within bioclimatic zones for the assisted migration (AM) scenario with adapted seed provenances (solid line) and local seed provenances (dashed line) for contemporary climate (1991-2010) and two future periods 2021-40 and 2061-80 under the CC change scenario RCP8.5. The annual carbon sink of adapted and local seed sources were estimated from universal response functions (URFs) predicted for the mean climate of 5 regional climate models (RCMs) (Supplementary Table 2).



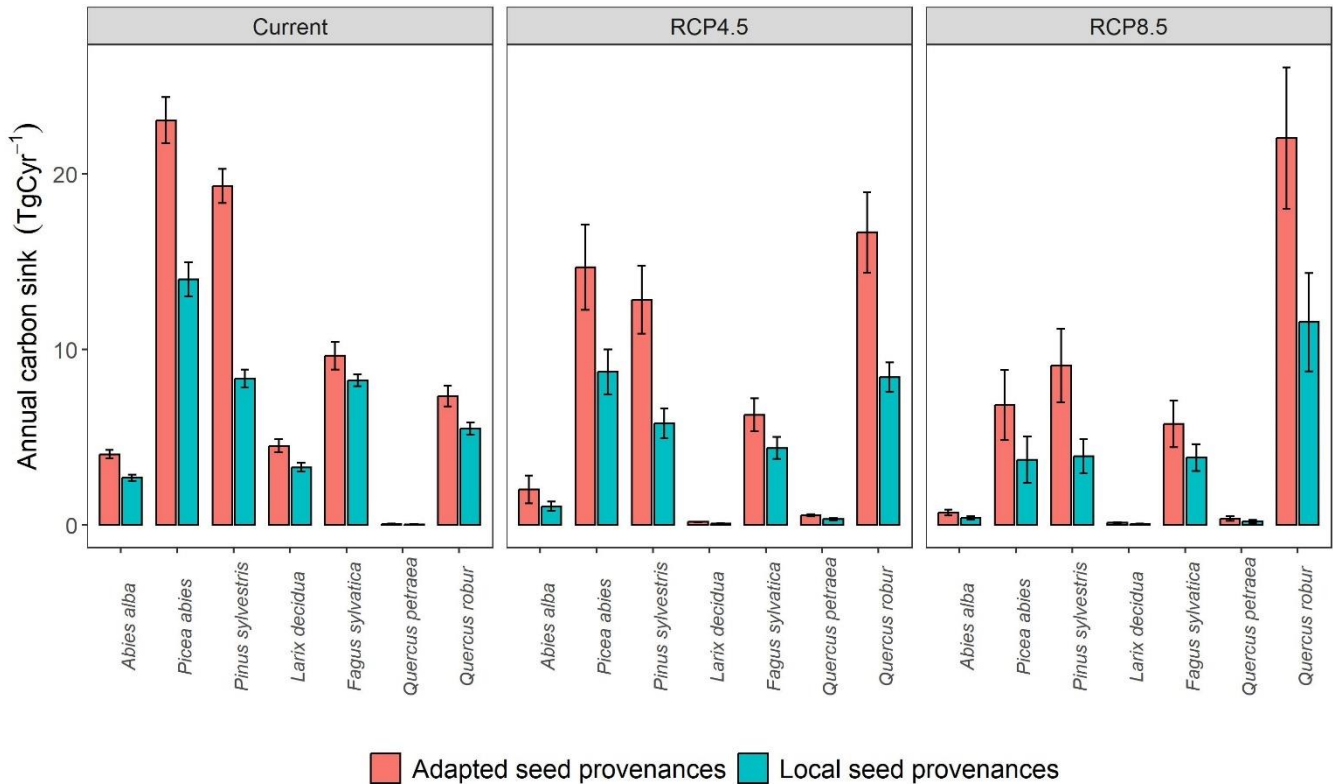
Supplementary Fig. 13. Variation in annual carbon sink of the seven species due to the 10 species distribution models (SDMs) used to identify potential species distribution and best-suited species in contemporary climate (1991-2010) shown for local seed sources and the adapted seed provenance clusters (SPC). Boxplot centre, hinges and whiskers denote median, upper/lower quartiles and upper/lower deciles, respectively.

446



447

448 **Supplementary Fig. 14.** Variation in annual carbon sink of the seven species due to the combination of 10
449 species distribution models (SDMs) and 5 regional climate models (RCMs) used to identify potential species
450 distribution and best-suited species for the climate scenarios RCP4.5 and RCP8.5 in 2061-80 shown for local seed
451 sources and the adapted seed provenance clusters. Boxplot centre, hinges and whiskers denote median,
452 upper/lower quartiles and upper/lower deciles, respectively.
453



Supplementary Fig. 15. Annual above-ground carbon sink (TgCyr^{-1}) for the seven European tree species under the two assisted migration scenarios “adapted” and “local” seed provenances. Local are those seed provenances that are geographically proximate to the planting locations and where the climate of the planting location and the historical climate of seed origin are identical. Adapted seeds are those seed provenance clusters producing the highest annual carbon sink (CS) for the climate of the planting location irrespective of their climatic similarity and geographic proximity to the planting location. CS from local and adapted seed sources in RCP4.5 & RCP8.5 scenarios were identified by URFs predicted for the mean climate of 5 RCMs (Supplementary Table 2). The error bars represent the standard deviation of the annual CS (TgCyr^{-1}) for each species due to different combinations of SDMs, URFs, and RCMs.

Spatial patterns in the uncertainty of the estimation of CS

The spatial analysis of uncertainty including the AM scenario (U-AM), revealed that the seed source component explains the largest share of variance in CS, confirming the significant role of seed-provenance selection compared to the SDM algorithms and RCMs. Across all climate scenarios, the seed source component explains on average more than 91% of the total variance, whereas the algorithms employed for SDM calibration account for only up to 1.46% (Supplementary Table 13).

Supplementary Table 13. Summary of variance analysis showing the variance explained by the different modelling components of the uncertainty analysis. Within the U-MF analysis, the AM scenarios were excluded, and the analysis undertaken individually for each AM scenario (local and adapted). In the U-AM analysis, effects of the species distribution model algorithm (SDM Algo), effects of the regional climate models (RCMs) and the effects of the assisted migration (AM) scenario were considered simultaneously. The SDM algorithm had 10 factors corresponding to each of the 10 SDMs of biomod2 used to identify best-suited species. RCM* used to calibrate the SDMs; URF-RCM** use to calibrate the URF. Resid. is the residual sum of squares.

Analysis	Time	Seed source	Component	Variance explained (%)				
				Mean	Median	Mode	Min	Max
U-AM	1991-2010		SDM Algo	1.46	0	0	0	98
U-AM	1991-2011		Seed source	98.25	100	100	1	100
U-AM	1991-2012		Resid.	0.29	0	0	0	79
U-AM	RCP 4.5		SDM Algo	1.43	0	0	0	49
U-AM	RCP 4.5		RCM*	0.24	0	0	0	16
U-AM	RCP 4.5		Seed source	93.21	98	100	3	100
U-AM	RCP 4.5		URF-RCM**	1.32	1	0	0	4
U-AM	RCP 4.5		Resid.	3.81	1	0	0	79
U-AM	RCP 8.5		SDM Algo	1.19	0	0	0	43
U-AM	RCP 8.5		RCM*	0.47	0	0	0	14
U-AM	RCP 8.5		Seed source	91.24	97	97	0	99
U-AM	RCP 8.5		URF-RCM**	1.83	1	1	0	100
U-AM	RCP 8.5		Resid.	5.26	1	0	0	89
U-MF	RCP 4.5	Local	SDM Algo	8.66	0	0	0	64
U-MF	RCP 4.5	Local	SDM Algo x RCM*	16.62	0	0	0	88
U-MF	RCP 4.5	Local	URF-RCM**	72.6	100	100	0	100
U-MF	RCP 4.5	Local	Other	2.12	0	0	0	39
U-MF	RCP 4.5	Adapted	SDM Algo	8.52	0	0	0	64
U-MF	RCP 4.5	Adapted	SDM Algo x RCM*	18.91	0	0	0	88
U-MF	RCP 4.5	Adapted	URF-RCM**	70.46	100	100	0	100
U-MF	RCP 4.5	Adapted	Other	2.11	0	0	0	43
U-MF	RCP 8.5	Local	SDM Algo	5.97	1	0	0	59
U-MF	RCP 8.5	Local	SDM Algo x RCM*	19.59	3	0	0	91
U-MF	RCP 8.5	Local	URF-RCM**	71.55	94	100	0	100
U-MF	RCP 8.5	Local	Other	2.89	1	0	0	30
U-MF	RCP 8.5	Adapted	SDM Algo	8.83	6	0	0	61
U-MF	RCP 8.5	Adapted	SDM Algo x RCM*	28.79	24	0	0	92
U-MF	RCP 8.5	Adapted	URF-RCM**	58.44	64	100	0	100
U-MF	RCP 8.5	Adapted	Other	3.94	3	0	0	29

482 The small contribution of the SDM algorithms to the variation of the modelled CS was also revealed by
483 the separate uncertainty analysis of the modelling framework (U-MF). Here, the SDM algorithms alone
484 explained up to 8.83% of the total variance across all scenarios, while the interaction with the RCMs
485 accounted for up to 28.79%. Furthermore, the most frequent (mode) of the percentage of variance
486 explained across all raster cells is 0 (Supplementary Table 13).

487 The small contribution of the SDMs to the modelled CS suggests, that in spite of variation in suitability
488 of the species, due to SDM algorithms and RCMs (Supplementary Fig.6), the species identified by each
489 algorithm as the best remains consistent, resulting in a low variation in the predicted carbon sink.
490 (Supplementary Table 13).

491 In the U-AM, the RCM models used to calibrate the SDM and URF models explain only up to 0.5% and
492 2% of the total variation, respectively, in both RCP scenarios, while in the U-MF, the RCMs used to
493 calibrate the URF emerges as the most important component of the variance explained. However,
494 when providing a seed source recommendation, the difference between the local and the adapted seed
495 sources is still higher than the variation of the 5 RCMs according to U-AM (Supplementary Table 13).

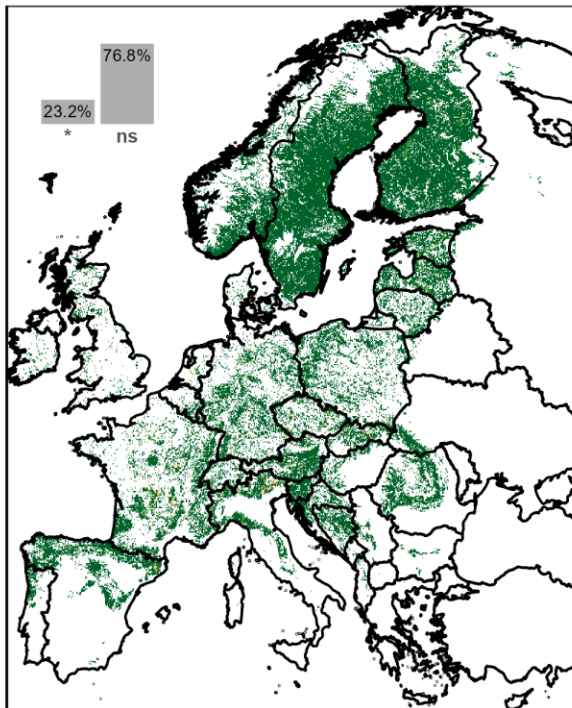
496 The geographical distribution of the variance explained by each component in the U-AM analysis
497 indicates no discernible spatial patterns observed for the SDM Algorithm, the RCM*, and URF-RCM**
498 components (see Supplementary Fig. 16, 17, and 18). However, the seed source component explained
499 notably lower parts of the variation in the Northern Europe and at high altitudes around the Alps and
500 the Carpathian mountains (see Supplementary Fig. 17c and 18c). Additionally, we observed changes in
501 the number of significant components across all analysed grid cells between the two climate change
502 scenarios: the number of cells where the SDM algorithm and the RCMs contributed significantly to the
503 variance increased from RCP4.5 to 8.5 (see Supplementary Fig. 17A, B and 18A, B), while the
504 contribution of the AM scenario was consistently high. This suggests that with more stronger climate
505 change, the variation of the SDM algorithms and the RCMs in predicting the best-suited species
506 increases. Although the URF-RCM** component explains a low percentage of the total variance, was it
507 was found to contribute significantly within a high number of grid cells within both climate change
508 scenarios.

509 In contrast to our findings from the U-AM analysis, the variance partitioning in the U-MF analysis
510 revealed pronounced spatial pattern in the uncertainty across all components (Supplementary Fig. 20 -
511 23). In central and southern regions of Europe, the primary source of variation is attributed to the 5
512 RCMs used to calibrate the URF(Supplementary Fig. 20-23C). However, in parts of Northern and
513 Northeastern Europe, the SDM algorithms and their interaction with the RCM models play a more
514 significant role (Supplementary Fig. 20-23A, B). While his pattern is evident for both AM scenarios, the
515 contribution of the interaction between SDM algorithms and RCM is more prominent for adapted seed
516 source, where it extend towards the nemoral and continental climatic zones of Europe (Supplementary
517 Fig. 21B, 23B).

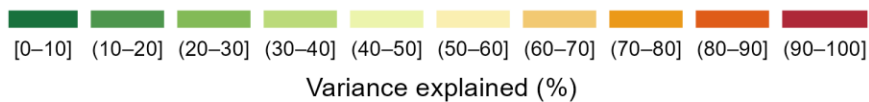
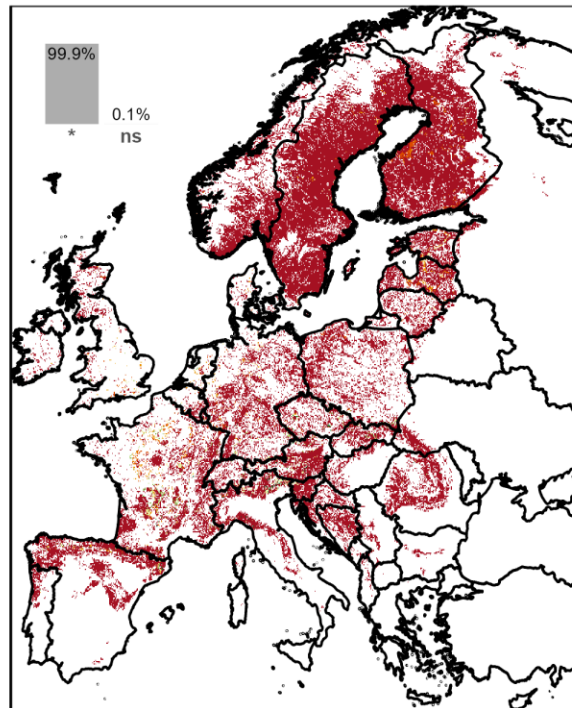
518 Altogether, the uncertainty analysis indicates significant levels of uncertainty in estimated CS which
519 can be attributed particularly to the SDM algorithms and which have notable prediction effects within
520 the boreal, nemoral, and north-continental climatic regions. In contrast, predictions for the
521 Mediterranean, Atlantic, and central continental climatic zones of Europe appear to be primarily
522 influenced by the RCMs for calibrating the URF models. While some variability in CS is to be expected
523 with changes in RCMs, the selection of algorithm should not drastically alter the results. Both, the U-
524 MF and U-AM analysis underline the limitations inherent in individual SDM algorithms or single RCMs
525 and advocates the need of ensemble modelling and detailed uncertainty analysis to build robust SDMs
526 and derived CS estimates

527

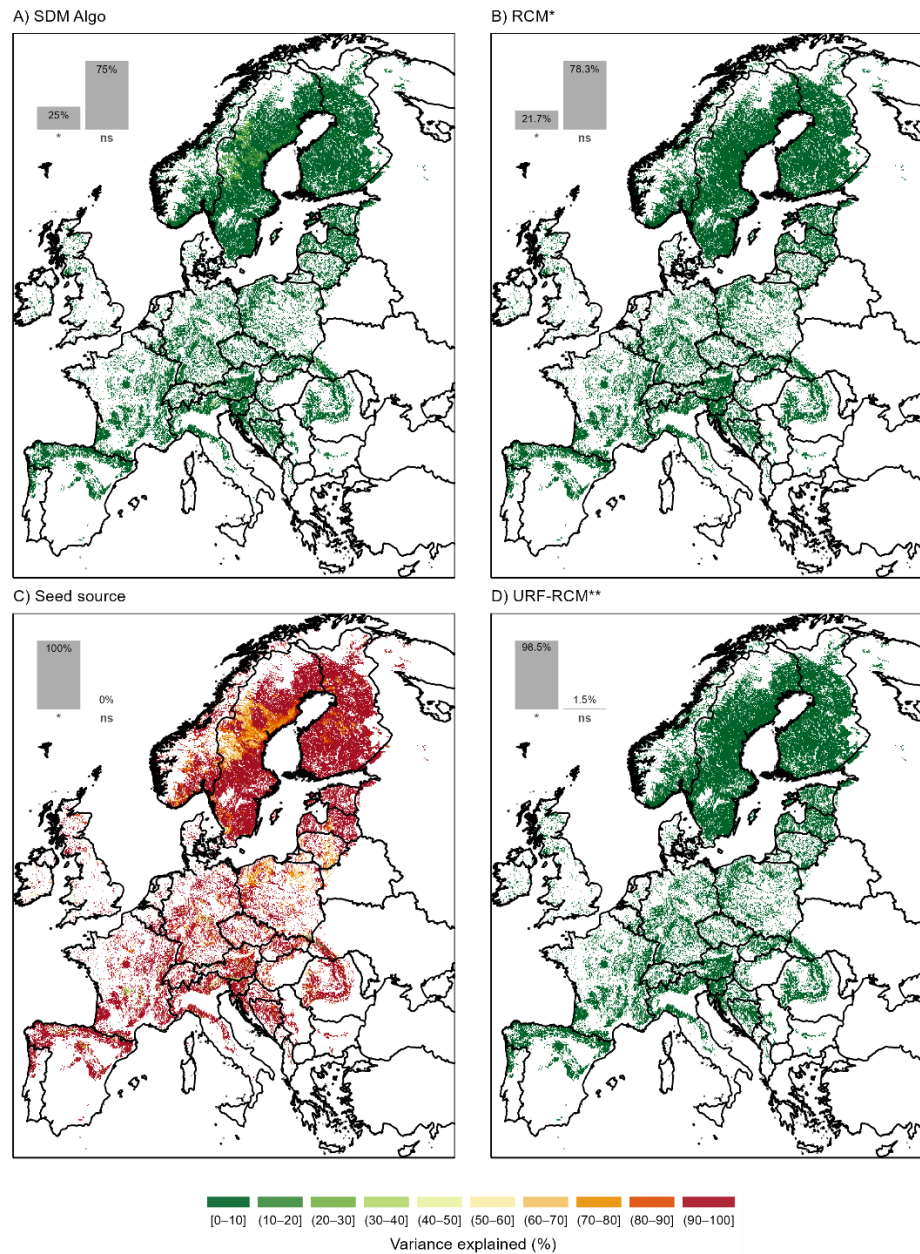
A) SDM Algo



B) Seed source

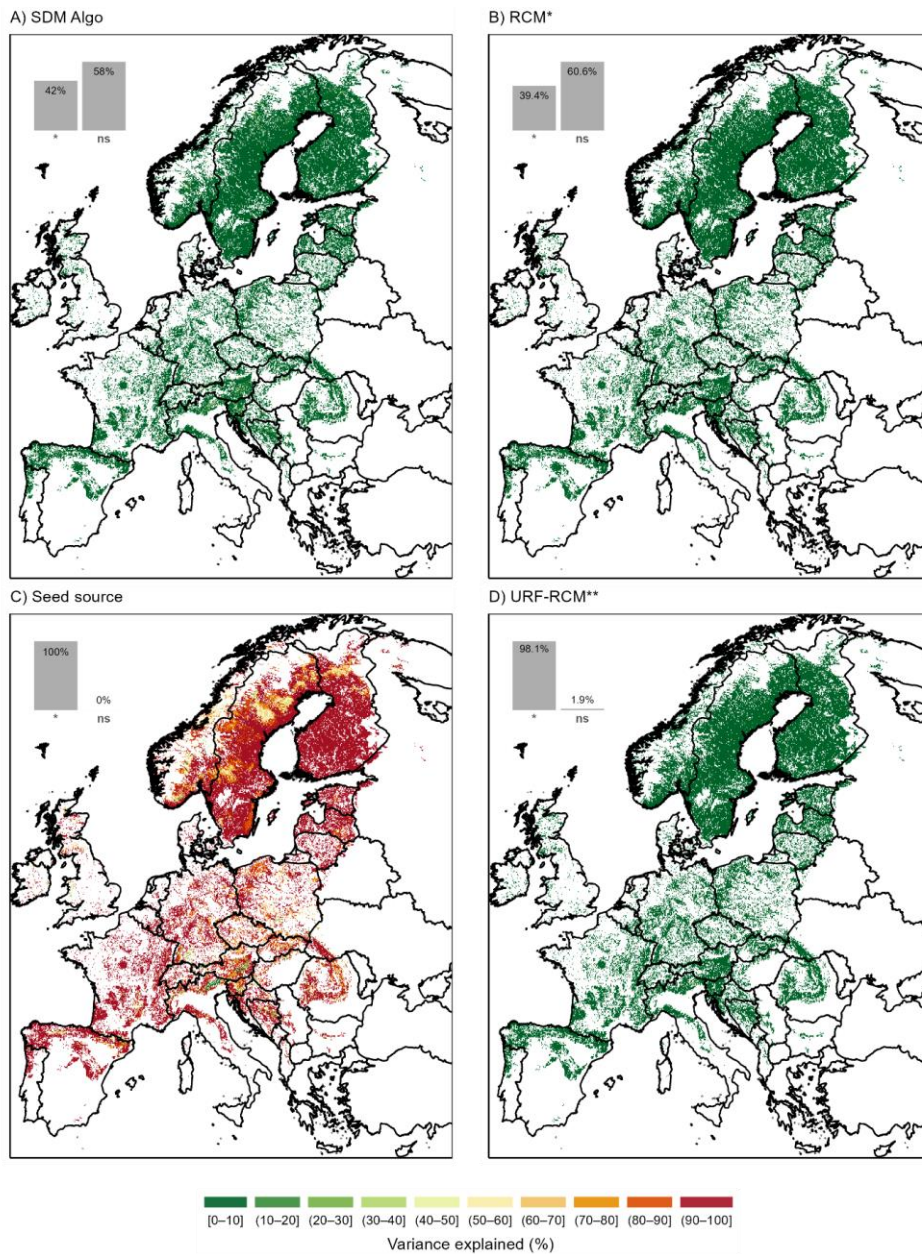


Supplementary Fig. 16. Variance partitioning of annual carbon sink under the contemporary climate (1991-2010) showing the uncertainty due to assisted migration scenario (U-AM). Variance explained as the proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms in biomod2, B) percentage of the total sum of squares explained by the two AM scenarios (local or adapted seed provenances). Within each map, the bar plots represent the percentage of total cells where the two variance components (SDM and Seed Source) were found significant. Asterisks (*) indicate a significant level of < 0.05, while "ns" denotes nonsignificant values.



537

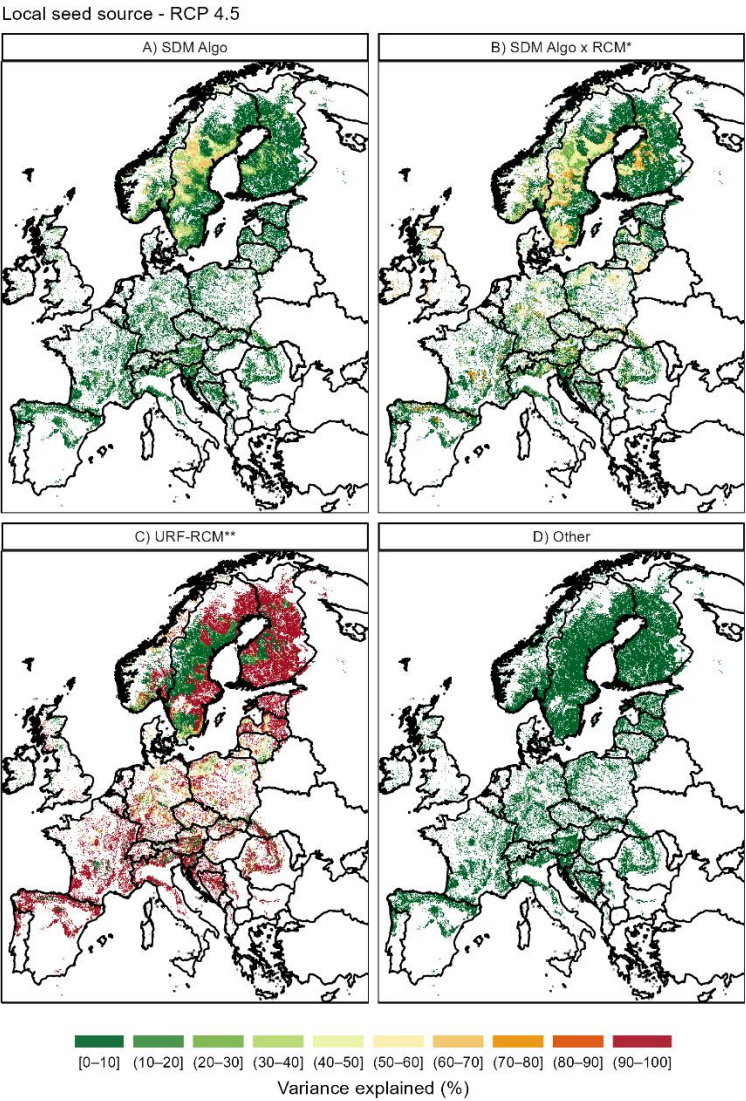
538 **Supplementary Fig. 17.** Variance partitioning of annual carbon sink under the climate change scenario RCP4.5
 539 (2061-80) showing the uncertainty due to assisted migration scenario (U-AM). Variance explained as the
 540 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms , B)
 541 variance due to 5 regional climate models (RCMs) used to calculate most-suited species, C) variance explained
 542 by the two AM scenarios (local or adapted seed provenances), & D) variance due to 5 RCMs used to identify
 543 most-suited seed provenance clusters with the URFs. Within each map, the bar plots illustrate the percentage of
 544 total cells where the variance components were found significant. Asterisks (*) indicate a significant level of <
 545 0.05, while "ns" denotes nonsignificant values.



546

547 **Supplementary Fig. 18.** Variance partitioning of annual carbon sink under the climate change scenario RCP8.5
 548 (2061-80) showing the uncertainty due to assisted migration scenario (U-AM). Variance explained as the
 549 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms, B)
 550 variance due to 5 regional climate models (RCMs) used to calculate most-suited species, C) variance explained
 551 by the two AM scenarios (local or adapted seed provenances) & D) variance due to 5 RCMs used to identify
 552 most-suited seed provenance clusters with the URFs. Within each map, the bar plots illustrate the percentage
 553 of total cells where the variance components were found significant. Asterisks (*) indicate a significant level of <
 554 0.05, while "ns" denotes nonsignificant values.

555



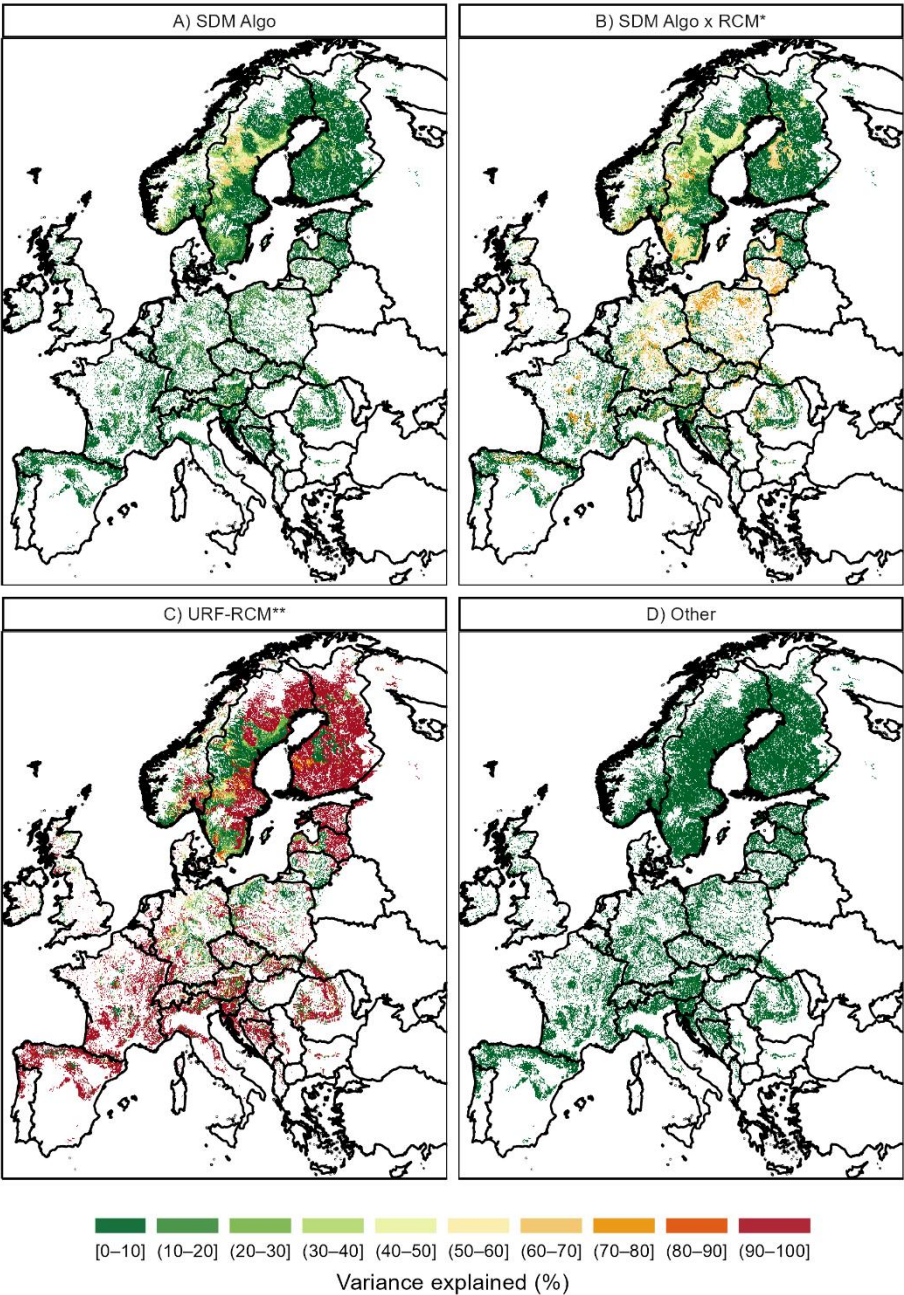
557

558

559 **Supplementary Fig. 19.** Variance partitioning of annual carbon sink for the local seed source predictions due to
560 modelling framework (U-MF) under the climate change scenario RCP4.5 (2061-80). Variance explained is the
561 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms in
562 biomod2, B) variance due to the interaction of the 10 SDMs with 5 regional climate models (RCMs) used to
563 predict most-suited species, C) variance due to 5 RCMs used to identify most-suited seed provenance clusters
564 with the URFs. & D) variance due to the other factors.

565

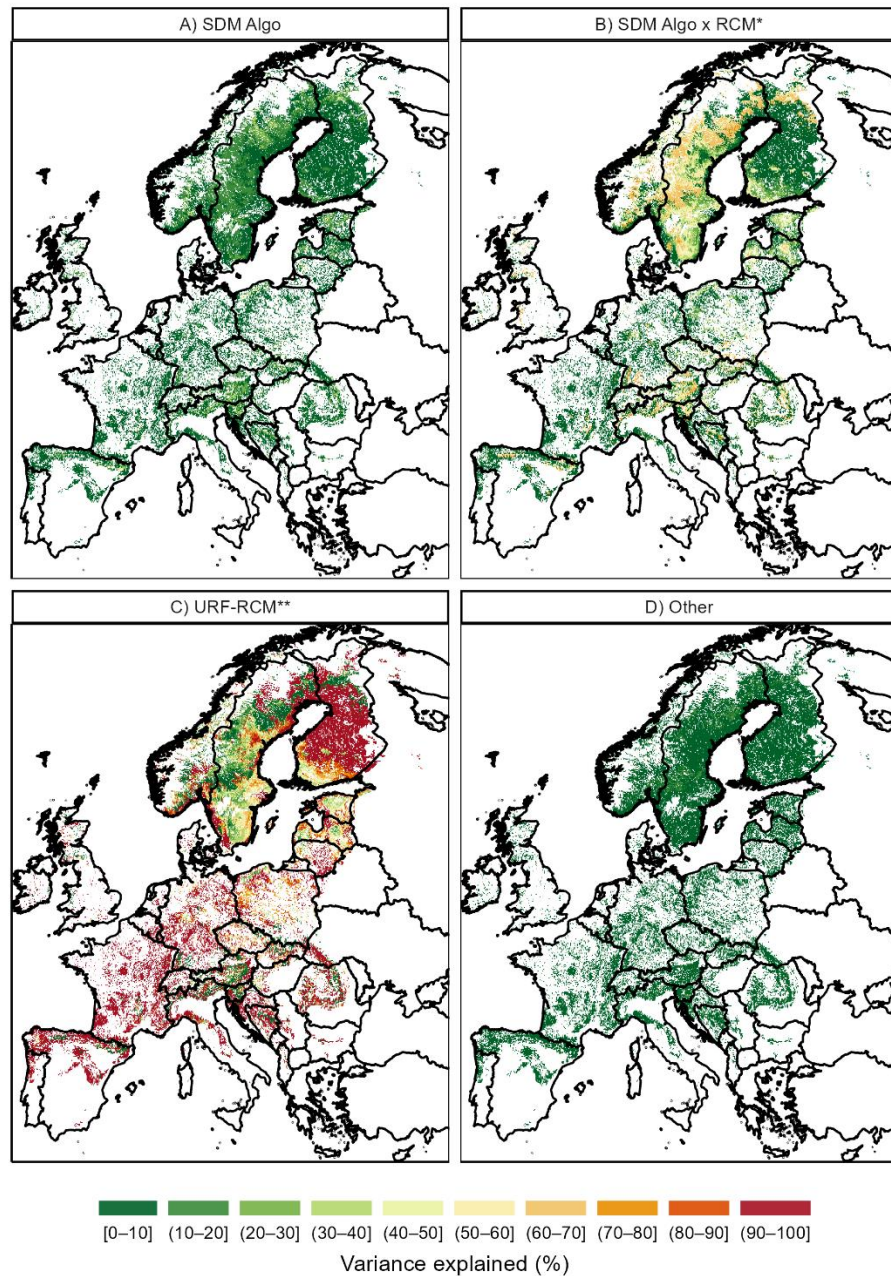
Adapted seed source - RCP 4.5



566

567 **Supplementary Fig. 20.** Variance partitioning of annual carbon sink for the adapted seed source predictions due
568 to modelling framework (U-MF) under the climate change scenario RCP4.5 (2061-80). Variance explained is the
569 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms in
570 biomod2, B) variance due to the interaction of the 10 SDMs with 5 regional climate models (RCMs) used to
571 predict most-suited species, C) variance due to 5 RCMs used to identify most-suited seed provenance clusters
572 with the URFs. & D) variance due to the other factors.

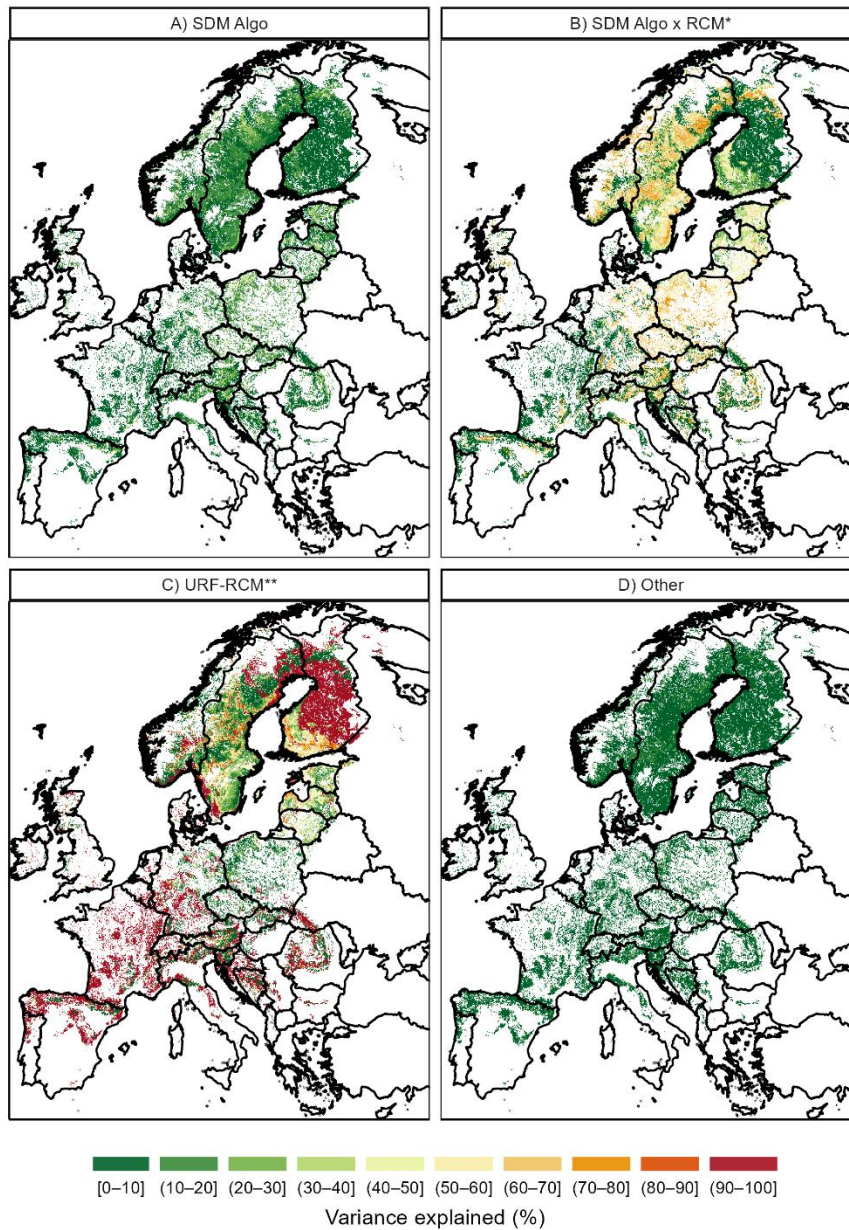
Local seed source - RCP 8.5



573

574 **Supplementary Fig. 21.** Variance partitioning of annual carbon sink for the local seed source predictions due to
 575 modelling framework (U-MF) under the climate change scenario RCP8.5 (2061-80). Variance explained is the
 576 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms in
 577 biomod2, B) variance due to the interaction of the 10 SDMs with 5 regional climate models (RCMs) used to
 578 predict most-suited species, C) variance due to 5 RCMs used to identify most-suited seed provenance clusters
 579 with the URFs. & D) variance due to the other factors.

Adapted seed source - RCP 8.5



580

581

582 **Supplementary Fig. 22.** Variance partitioning of annual carbon sink for the local seed source predictions due to
 583 modelling framework (U-MF) under the climate change scenario RCP8.5 (2061-80). Variance explained is the
 584 proportion of the total sum of squares explained by A) the 10 species distribution models (SDM) algorithms in
 585 biomod2, B) variance due to the interaction of the 10 SDMs with 5 regional climate models (RCMs) used to
 586 predict most-suited species, C) variance due to 5 RCMs used to identify most-suited seed provenance clusters
 587 with the URFs. & D) variance due to the other factors.

588

589 **REFERENCES (SUPPLEMENTARY MATERIAL)**

- 590 95. Mauri, A., Strona, G. & San-Miguel-Ayanz, J. EU-Forest, a high-resolution tree occurrence dataset for
591 Europe. *Sci. Data* (2017) doi:10.1038/sdata.2016.123.
- 592 96. van Vuuren, D. P. et al. The representative concentration pathways: An overview. *Clim. Change* 109, 5–31
593 (2011).
- 594 97. Masson-Delmotte, V. et al. IPCC, 2021: Climate Change 2021: The Physical Science Basis. Cambridge
595 University Press. In Press. (2021).
- 596 98. Forrester, D. I. et al. Generalized biomass and leaf area allometric equations for European tree species
597 incorporating stand structure, tree age and climate. *For. Ecol. Manage.* (2017)
598 doi:10.1016/j.foreco.2017.04.011.
- 599 99. Penman, J. et al. Good Practice Guidance for Land Use, Land-Use Change and Forestry. IPCC (2003).
- 600 100. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories Volume 4 Agriculture, Forestry and
601 Other Land Use Chapter 4 forest land 2006. Forestry (2006).
- 602 101. Vospernik, S. & Sterba, H. Do competition-density rule and self-thinning rule agree? *Ann. For. Sci.* (2015)
603 doi:10.1007/s13595-014-0433-x.
- 604 102. Thuiller, W., Georges, D. & Engler, R. biomod2: Ensemble platform for species distribution modeling. R
605 Packag. version 2, r560 (2013).
- 606 103. Zurell, D. et al. A standard protocol for reporting species distribution models. *Ecography (Cop.)*. (2020)
607 doi:10.1111/ecog.04960.
- 608 104. Breiman, L. Random forests. *Mach. Learn.* 45, 5–32 (2001).
- 609 105. Häggström, J., Persson, E., Waernbaum, I. & De Luna, X. Covsel: An R package for covariate selection
610 when estimating average causal effects. *J. Stat. Softw.* (2015) doi:10.18637/jss.v068.i01.
- 611 106. Hiederer, R., Houston Durrant, T. & Micheli, E. Evaluation of BioSoil demonstration project - Soil data
612 analysis. (2011).
- 613 107. Coetsee, B. W. T., Robertson, M. P., Erasmus, B. F. N., van Rensburg, B. J. & Thuiller, W. Ensemble models
614 predict important bird areas in southern Africa will become less effective for conserving endemic birds
615 under climate change. *Glob. Ecol. Biogeogr.* 18, 701–710 (2009).
- 616 108. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: Prevalence,
617 kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232 (2006).
- 618 109. Pontius, R. G. & Parmentier, B. Recommendations for using the relative operating characteristic (ROC).
619 *Landsc. Ecol.* 29, 367–382 (2014).
- 620 110. Lefèvre, F. et al. Dynamic Conservation of Forest Genetic Resources in 33 European Countries. *Conserv.*
621 *Biol.* (2013) doi:10.1111/j.1523-1739.2012.01961.x.