Modelling and mapping beech forest distribution and site productivity under different climate change scenarios in the Cantabrian Range (North-western Spain)

Javier Castaño-Santamaría[1 2 3], Carlos A. López-Sánchez[1], José Ramón Obeso[2], Marcos Barrio-Anta[1*]

[1] GIS-Forest Research Group, BOS Department, Oviedo University, Mieres, Spain.

[2] UMIB-Research Unit of Biodiversity (UO, CSIC, PA), Oviedo University, Mieres, Spain.

[3] Directorate General for Cadastre, Ministry of Finance, Regional Office of Asturias, Oviedo, Spain.

* Corresponding author. Address: Polytechnic School of Mieres, E-33600 Mieres, Asturias, Spain. Tel.: +34 985 45 80 05; E-mail address: barriomarcos@uniovi.es (M. Barrio-Anta).

Declaration of interest: none.

Abstract

The beech forests in the Cantabrian Range occur at the southwestern limit of the distribution of the species and are very important for wildlife and biodiversity in the region. Climate change is expected to increase the frequency and severity of drought events over the next few decades in southwestern Europe, and establishing how this will alter the distribution, abundance and productivity of beech is fundamental for biodiversity conservation and management. In this study, we used spatially continuous environmental variables to develop spatial distribution and site-productivity models for beech forests in the Cantabrian Range and to project these models to different climate change scenarios. Two raster-based models of resolution 250 m were constructed to identify suitable habitat (species distribution model) and to estimate site index (productivity model) for beech in the Cantabrian Range. Of the 23 variables retained in the spatial distribution model, climate, soil and terrain were the most important (explaining respectively 51.2%, 34.2% and 10.1% of the variation). The productivity model retained only three variables (percentage of silt in soil, mean diurnal range of temperature and plan curvature of the terrain) but was able to explain 54% of the total variation. Future projections based on two emission scenarios suggest that suitable habitat will be drastically reduced by 2070 (loss of 40-90% of the area for the moderate and pessimistic scenarios, respectively). The reduction will probably also be accompanied by less favourable conditions for seedling establishment, higher mortality rates and a reduction in local density of populations. However, the projections do not imply current population removal. Productivity projections for suitable habitat suggest a large increase in the average site index (from current 15.19 to 18.18 m) in the moderate scenario and an increase of only 34 cm in the pessimistic scenario. The study findings provide basic information for conservation biology and could be used by decision-makers to develop and implement actions for mitigating the impact of climate change on beech forests.

Keywords

Fagus sylvatica L., Suitable habitat, Site Index, spatially-continuous environmental variables, Random Forest, modelling, Climate Change.

1. Introduction

Climate change is expected to increase the frequency and severity of drought events over the next few decades in Western Europe (Schar et al., 2004). These changes will have a significant influence on the distribution and abundance of plant species and on forest productivity (e.g. Monzón et al., 2011; Hacket-Pain et al., 2016), thus posing a great challenge to decision-makers. In fact, forest conservation strategies and plans may be unsuccessful if the projected changes are not taken into account (Noce et al., 2017).

The concept of site productivity, also known as site quality, refers to a quantitative estimate of the potential of a site (usually by designating and summarizing the local biophysical characteristics of a forest environment) to produce plant biomass (Bontemps and Bouriaud, 2014). Both terms (site quality and site productivity) can be considered equivalent when only biophysical site variables drive tree growth (i.e. absence of vegetation control, irrigation, drainage or severe effects of pest and/or diseases) (Skovsgaard and Vanclay, 2008) as, for instance, in unmanaged forests. The mean height growth of the dominant and codominant trees (dominant height) at a specific reference age, known as the site index (SI), is by far the most frequently used indicator of forest productivity and is related to stand structure, which greatly influences habitat and plant and animal species diversity (Pretzsch, 2009). This strong influence explains why this variable has also been used in studies concerning biodiversity (Potter and Woodall, 2014), forest structure (Larson et al., 2008) and forest disturbance (Wei et al., 2003), amongst others.

Although site quality and productivity can be predicted with reasonable accuracy in small areas by using measured SI values, the process is very costly (dominant height and stand age must be determined) and requires the species to be present. These drawbacks can be overcome by using indirect methods to estimate SI from on site environmental variables (also known as geocentric methods), applicable even when suitable trees are absent (Clutter et al., 1983).

Many studies have attempted to relate SI to environmental factors by using parametric approaches (e.g. Fontes et al., 2003), nonparametric approaches (e.g. McKenney and Pedlar, 2003; Albert and Schmidt, 2010) or both (e.g. Aertsen et al., 2010). However, many of these studies include some nutrient soil variables, which require complicated and expensive analytical techniques for their determination, making the models of little practical use. Remotely sensed data and spatially interpolated surfaces provide spatially-continuous environmental information that was not available a few decades ago, thus solving the previously mentioned problems. Developing methods that provide spatially explicit estimates of SI as a function of this environmental information would enable development of a SI map without the need for fieldwork, thus making information about forest productivity available for further ecological study (Parresol et al., 2017). For this purpose, a species distribution model (SDM), understood as an empirical ecological model that relates species occurrence to environmental predictors (Guisan and Zimmermann, 2000), is the most suitable framework to depict this site productivity information.

SDMs have been widely used to estimate ecological requirements of particular species and to characterize and map the spatial distribution of habitats occupied by species at landscape scale (e.g. Li et al., 2016). However, in addition to describing the environmental needs of the populations, SDMs can also predict the potential temporal and spatial distribution in unsampled areas and future climatic conditions (Elith et al., 2006), which is fundamental for conservation biology (Johnston et al., 2015).

Fagus sylvatica L. (hereinafter "beech") is the most widely distributed of all *Fagus* species and the most abundant broad-leaved forest tree in Europe, with a geographical distribution spanning between southern Scandinavia and Sicily, across a wide range of environmental conditions (Fang and Lechowicz, 2006). In the southernmost part of the distribution range, where the climate is warmer and drier (e.g. Spain), beech populations are restricted to mountain slopes where there are fewer climatic constraints (Jump et al., 2006). In the Cantabrian Range (NW Spain), the climate is humid temperate and beech occurs as a climax species on slopes of elevation higher than 600 m above sea level (Gandullo et al., 2004), where the forests are characterized by natural regeneration and fast-growing stocks (Ruiz de la Torre, 2006). These forests can host diverse types of fauna, and they form part of the habitats of endangered and emblematic species such as the Cantabrian capercaillie (*Tetrao urogallus* sbsp. *cantabricus* J. Castroviejo) and the brown bear (*Ursus arctos* L.), leading to their inclusion in protected areas that are relatively unaffected by human influence. Indeed, the high ecological value, together with the complex topography of the area, has led to the stands being unmanaged and not harvested, except occasional cutting to produce local domestic firewood. Because of the low economic importance, no studies have been carried out to date to determine the site quality of beech stands in the Cantabrian Range.

The relationship between beech and climate has been widely investigated in Europe due to the socio-economic and ecological importance of the species (Dyderski et al., 2017). Numerous studies have shown decreasing trends of growth and productivity in beech forests, which have mainly been attributed to the impact of climate change in Northern Europe (e.g. Farahat and Linderholm, 2018), Central Europe (e.g. Scharnweber et al., 2011; with Bosela et al. (2016) as an exception) and Southern Europe (e.g. Jump et al., 2006; with Tegel et al. (2014) as an exception). Moreover, simulation studies suggest future changes in the current distribution and productivity of the species as a consequence of climate change (e.g. Geßler et al., 2007; Albert and Schmidt, 2010; Meier et al., 2011; Falk and Hempelmann, 2013; Brandl et al., 2018).

The Cantabrian Range has undergone a gradual increase in temperature and potential evapotranspiration, together with a decrease in precipitation in recent decades (Rubio-Cuadrado et al., 2018). According to recent findings, even more dramatic changes are expected to occur in the future (e.g. IPCC, 2013; EEA, 2017). It is therefore necessary to incorporate climate variables as predictors to model how these changes will affect productivity, to predict shifts in species distribution and to identify areas where the species will be able to persist.

Among the available information on a particular species, occurrence, abundance, site productivity and stand structure (and the temporal and spatial variations in these) are of major interest for the purposes of biodiversity conservation. Spatially-continuous distribution and productivity models developed for different climate change scenarios will help decision-makers to develop and implement actions for mitigating the decline in biodiversity brought about by global warming. Thus, the overall aim of this study was to simulate the effects of climate change on suitable habitat and site productivity of the beech forests in the Cantabrian Range. The specific objectives were as follows: *i*) to develop a site index equation for *Fagus sylvatica* L. in the study region, *ii*) to investigate the environmental factors determining the distribution and productivity of the species and *iii*) to develop spatial distribution (SDM) and productivity (PM) models based on environmental variables, in order *iv*) to generate spatially-continuous maps, and *v*) to project the models and maps to different climate change scenarios.

2. Materials and Methods

2.1. Study area

The Cantabrian Range (NW Spain) constitutes an ecotone between the Eurosiberian and the Mediterranean biogeographic regions in Europe (Díaz and Fernández, 1987) with the main axis running in an east-west direction from the Galician Atlantic coast to the western extreme of the Pyrenees in the Basque Country. There is considerable asymmetry between the northern and southern sides of the Cantabrian Range. Thus, while the northern side always terminates at sea level, the southern side descends to the Northern Plateau (Douro River basin), with elevations rarely lower than 800 m. The terrain of this mountain range is complex, and the different combinations of topography and landform influence the type and vigour of the vegetation communities (Elena, 1997).

The present study was conducted in the Northwestern Cantabrian Range (42.82° to 43.51° N; -6.79° to -4.52° W), in the provinces of Asturias and León (Fig. 1). The climate of the region is Atlantic, with precipitation very uniformly distributed around the year (Rozas et al., 2015). Precipitation ranges from 1,217 to 1,855 mm, with an annual average of 1,568 mm, whereas the mean annual temperature varies from 6.7 to 10.5 °C, with an annual average of 8.1 °C. Geologically, ancient Paleozoic rocks (carboniferous limestone, slate, coal, conglomerates, quartzite, sandstone) predominate in the central axis, flanked by Mesozoic (limestone, dolomite, sandstone) and Tertiary rocks in the lower mountains of the eastern part of the Basque Country (IGME, 2015a).

2.2. Data collection

Four different types of data were considered in this study and used for different purposes: *i*) longitudinal tree height-age data, obtained by stem analysis in research plots, were used to develop a site quality system, *ii*) occurrence data obtained from the Third Spanish National Forest Inventory were used to develop the distribution model, *iii*) data of current spatial environmental variables were used to model and map distribution and site productivity; and *iv*) future climatic data projections under different emission scenarios were used to predict the impact of climate change.

2.2.1 Dominant tree height-age data

A total of 112 permanent sample plots, covering the existing range of stand densities and site qualities, were established in natural beech-dominated stands (90% or more of the standing basal area consisting of beech) throughout the Western Cantabrian Range in 2010 and 2011 (Fig. 1). The plots ranged in size from 400 to 3600 m^2, depending on stand density, in order to achieve a minimum of 30 trees per plot. These beech forests are located in environmentally protected areas, and we were not able to obtain permission to cut trees in all locations. We thus finally selected a sample of 30 plots to represent all site qualities. In each plot, two dominant trees were felled and destructively sampled. The felled trees were the first two dominant trees found outside the plots, but in the same stands, within ± 5% of the mean diameter at 1.3 m above ground level and mean height of the dominant trees (considered as the 100 largest-diameter trees per hectare). All of the trees (n=60) were cross-sectioned at stump height, at 0.50 m above ground, at breast height, and 1 m intervals thereafter along the stem. Each cross section was processed by electric brushing and sanding until the tree rings were clearly visible. The treated cross sections were scanned at 900 dpi (in an Epson Expression STD 1680 PLUS flatbed scanner) and the resulting data were analyzed using WinDENDRO image analysis software (Regent Instruments Canada Inc.) to produce the annual ring count. To reduce the bias when determining the height of each cross section at a given age, Carmean's algorithm, with the modification proposed by Newberry (1991), was applied. Summary statistics, including mean, maximum, minimum and standard deviation values for the main tree and stand variables are shown in Table 1.

2.2.2. Occurrence data

Information on beech occurrence was drawn from the Third Spanish National Forest Inventory (SNFI3) (DGCN, 2006). The plots of the SNFI3 are located at the nodes of a 1 km UTM square grid, comprising four concentric subplots of radius of 5, 10, 15 and 25 m, with a minimum diameter at 1.3 m above ground level threshold of 75, 125, 225 and 425 mm, respectively (DGCN, 2006). Presence was defined as the occurrence of one or more live beech trees in some of the subplots. A total of 1,877 plots falling within the study area with data on presence/absence of beech were available for analysis and were imported to a GIS database (ArcGIS 9.3, ESRI, Redlands, CA, USA). A minimum distance of 1 km between plots was considered in order to prevent the inclusion of spatially autocorrelated data.

2.2.3. Spatial environmental variables

Four types of environmental parameters were considered as possible predictors of the species distribution and site productivity: terrain, climate, soil and hydrographical variables. A total set of 48 variables was available for analysis (Table 2).

Terrain variables were based on a 5 m resolution digital elevation model (DEM) provided by the Spanish National Plan for Aerial Orthophotography (PNOA; www.pnoa.ign.es). We used the Automated Geoscientific Analyses Geographical Information System software v.3.0.0 (SAGA; Conrad et al., 2015) to calculate each of the terrain variables from the DEM. Seven topographic variables, plus three potential incoming solar radiation variables and one hydrographic variable were considered, excluding elevation, which is strongly correlated with climatic variables such as temperature and precipitation. Gridded data were obtained for all climate variables with a 30 arc-second resolution (approximately 800 m) from WorldClim (Hijmans et al., 2005). A total of 19 climatic variables were considered. Twelve soil variables were compiled from the SoilGrids250m (Hengl et al., 2017), which provides a collection of updatable soil property and world classification maps at 250 m spatial resolution, based on machine learning algorithms. Soil type and group were compiled from the European soil database (ESDB) v2.0. Lithostratigraphic type and permeability were obtained from the Spanish Stratigraphic Map (SSM) scale 1:200,000, and Geology from the Spanish Geological Map (SGM) scale 1:1,000,000 (IGME, 2015a; 2015b). All climate, soil and topography variable raster grids were resampled at 250 m resolution. To predict the future species distribution and site productivity under different climate change scenarios, we use the Global Climate Models (GCMs) for 2050 and 2070 based on the CMIP5 model of the IPCC 5th Assessment Report (http://www.worldclim.org/CMIP5).

2.3. Site quality system development and site index data

The algebraic difference approach (ADA; Bailey and Clutter, 1974) and its generalization (GADA; Cieszewski and Bailey, 2000) were used to develop the site quality system. We tested the six dynamic equations used by Barrio-Anta et al. (2008) for modelling both dominant height and basal area growth of I-214 poplar plantations in Spain. The dummy variables method (Cieszewski et al., 2000) considering a continuous-time autoregressive (CAR(x)) error structure for accounting for autocorrelation was used to estimate the model parameters. The dummy variables method and the CAR(x) error structure were programmed using the SAS/ETS® MODEL procedure (SAS Institute Inc., 2004), which allows dynamic updating of the residuals. We carried out a cross-validation, estimating the residuals in dominant height estimation for the two trees of one plot by fitting the model without the two trees from that plot. The root mean square error (RMSE) and the coefficient of determination for nonlinear regression (R^2) were calculated from the residuals obtained from cross-validation. The curves fitted over the dominant height trajectories were visually inspected to select the best model (Barrio-Anta et al., 2008). Determining SI at a particular location is generally a two-step process (McKenney and Pedlar, 2003): *i*) a dominant height-age equation for a particular species must first be developed on the basis of data obtained by destructive sampling of dominant trees, and *ii*) a site index is obtained for a site by measuring the height and age of several dominant and/or codominant trees and including the data in a previously developed equation.

2.4. Modelling species distribution and productivity

Various statistical approaches ranging from multiple linear regression to complex machine learning algorithms have been used to predict species occurrence (e.g. Falk and Hempelmann, 2013; Shirk et al., 2018) and forest productivity (e.g. Aertsen et al., 2010). However, simulating changes in vegetation characteristics relative to environmental variables can be extremely complex, posing significant challenges to traditional parametric regression analysis (Prasad et al., 2006). Thus, newly developed non-parametric methods have become more popular in recent decades.

In this study, we used the Random Forest (RF) non-parametric ensemble learning method (Breiman, 2001). RF is a widely used non-parametric classification and regression approach that consists of building an ensemble of decision trees (Gislason et al., 2006). The success of this technique is based on the use of numerous trees and different independent variables that are randomly selected from the complete original set of features (e.g. Deschamps et al., 2012). In machine learning, spurious data features must be removed before a model is generated (Hall, 1999). Thus, the potentially most important variables are selected. For this purpose, WEKA open source software (Hall et al., 2009) was used to fit the RF algorithm by implementing a wrapper methodology to select the subsample of variables, which usually produces the best results (Zhiwei and Xinghua, 2010). This method selects the subsample of variables by using a learning algorithm as part of the evaluation function. Final fitted models were applied to environmental spatial variables resampled at a 250m x 250m resolution to generate spatially continuous maps.

2.5. Model assessment and analysis

Several approaches can be used to test the accuracy of supervised learning algorithms. We used the common method of k-fold cross validation. In this process the data set is divided into k subsets. Each time the model is applied, one of the subsets is used as the test set and the other k-1 subsets form the training set. This provides a good indication of how well the classifier will perform on unseen data. We used k=10 and applied the RF algorithm several times and computed various standard metrics to assess the model performance. In order to assess the accuracy of SDM predictions, we used a confusion matrix that reflects the four possible ways that a sample point can be classified and observed (Fielding and Bell, 1997). The values of this matrix were used to calculate several metrics. Some of the metrics commonly used include the following (Shirk et al., 2018): *i*) the overall accuracy (OA), *ii*) sensitivity (SN), *iii*) specificity (SP), *iv*) the True Skill Statistic (TSS), *v*) Cohen's Kappa; and *vi*) the area under the ROC curve, (AUC). A binary model, which was required to calculate Cohen's Kappa and OA, was created on the basis of a threshold probability where sensitivity equals specificity, with equally weighted errors of omission and commission. All modelling methods report a probability of presence (PoP) for each species as an output variable. We selected a threshold PoP for converting all other PoP data to binary presence–absence outputs. To select a threshold for presence–absence delineation from the PoP data, we used the average result of two methods: (1) the PoP that maximized the sum of sensitivity and specificity, and (2) the PoP that minimized the difference between the absolute values of sensitivity and specificity.

To evaluate the productivity model, we used the coefficient of determination for nonlinear regression (R^2) (Ryan, 1997), the absolute and relative values of respectively the mean absolute error (MAE) and the root mean squared error (RMSE). RF has an embedded feature ranking technique called the variable importance measure (VIM), which was used to guide selection of predictors for the final model. These scores were determined as follows. The out-of-bag Mean Squared Error (MSE) was first stored in each tree of the RF. Each of the predictors was then permuted in turn (values are randomly reassigned among the set of out-of-bag samples) and the difference in MSE, usually an increase, was computed and averaged over all the trees. Finally, normalization was carried out by standard deviation of the differences and the output represented as a percentage increase in MSE. Thus, the potentially most important variables were selected by RF according to the VIM. To ensure values of variable importance were expressed in comparable scales, the VIM values were normalized so that they summed to a unitary value (normalized importance) and they were also expressed in relative terms (relative importance = $(VIM-VIM_{min})/(VIM_{max}-VIM_{min})$). The marginal response curves were then constructed in order to explore the relationships between the response and each of important predictor variables. These curves represent the predicted probability of presence of the species or the site productivity prediction value (*y-axis*) as function of a single environmental variable (*x-axis*), when all other explanatory variables are held constant at their mean values.

2.6. Current and future predictions of models

Two raster databases of resolution 250 x 250 m were obtained, resulting in fitting SDM and PM models to the current environmental variables to enable generation of a current map of suitable habitat and site productivity for beech.

We also projected the fitted models onto spatial projections at 250 m resolution of the environmental variables reflecting two climate change scenarios (moderate and pessimistic) for 2050 and 2070 under different emissions pathways. These scenarios are expressed by the Representative Concentration Pathways (RCP), using values comparing the level of radiative forcing between the preindustrial era and 2100. The moderate scenario (RCP 4.5) assumes that climate policies limit greenhouse-related emissions and total radiative forcing is stabilized at 4.5 W m^{-2} in 2100 without ever exceeding that value in prior years with a CO_2 concentration of 650 ppm and 1.0–2.6°C increase by 2100 (Thomson et al., 2011). The pessimistic scenario (RCP 8.5) assumes continued increases in greenhouse gases following recent trends (but does not include any specific climate mitigation target), reaching a total radiative forcing of 8.5 W m^{-2} by 2100 (Riahi et al., 2011) and 1,350 ppm CO_2 and 2.6–4.8°C increase by 2100 (IPCC, 2013; Harris et al., 2014). For the current and future scenarios, we used FRAGSTATS 4.2 (McGarigal et al., 2016) to quantify the area of habitat and degree of habitat fragmentation based on the binary model. We use three indicators to quantify suitable habitat surface: *i*) total area (TA), *ii*) mean patch area (MPA) and *iii*) largest patch index (LPI; the percentage of the landscape encompassed by the largest patch). The fragmentation was assessed with the aggregation index (AI), which equals 0 when suitable habitat is maximally disaggregated into single grid cell patches disconnected from all other patches and increases to 1 as suitable habitat is increasingly aggregated into a single, compact patch. We also quantified the degree of change for each future scenario relative to the current situation, classifying habitat as either gained, maintained or lost. In Figure 2, we graphically summarize the main methodological steps of the approach used in the present study.

3. Results

3.1. Site index model

The GADA formulation of the Hossfeld equation (Cieszewski, 2002) (Eq. 1) yielded the best compromise between graphical and statistical considerations (R^2 = 0.98 and RMSE = 0.74). All parameter estimates were significant at the 5% level, and the plot of residuals against estimated values showed a random pattern of residuals around zero, with homogeneous variance and no detectable significant trends, after modelling the error structure of the site quality equations following Diéguez-Aranda et al. (2006) (Fig. S1).

$$H_2 = \frac{23.8753 + X_0}{1 + 20526.03 / X_0 \cdot t_2^{-1.51}}$$
$$X_0 = \tfrac{1}{2}\left(H_1 - 23.8753 + \sqrt{\left(H_1 - 23.8753\right)^2 + 4 \cdot 20526.03 \cdot H_1 \cdot t_1^{-1.51}} \right) \qquad \text{(Eq. 1)}$$

Where, H_1 represents the predicted dominant height (m) at age t_1 (years) and H_2 represents the predicted dominant height (m) at age t_2 (years).

We used the method proposed by Diéguez-Aranda et al. (2005) to select the reference age. This consists of using different reference ages and their corresponding observed heights to estimate heights at other ages (both forward and backward) for each tree, and of comparing the results with the values obtained from stem analysis by using the relative error in predictions. Following this procedure, a reference age of 80 years was selected for the SI (Fig. S2) used to classify the stands according to their productivity.

The predicted dominant height curves overlaid on observed data from stem analysis (Fig. 3) showed that the model satisfactorily described the real trajectories. SI, which in this case was defined as the dominant height at a reference age of 80 years, can be easy obtained for a particular plot by substituting H_2 for SI and 80 years for t_2 and including the dominant height (H_1) and the age (t_1) of the plot in Eq.1.

3.2. Species Distribution Model

Of the 3,121 sites surveyed in the provinces of Asturias and León, beech was present at 539 sites and absent from 2,582. Beech was present at elevations ranging from 69 m to 1,797 m (mean elevation = 837 m), and the latitudinal distribution among the sampled sites ranged from 42.82 to 43.51 degrees north (mean latitude = 43.1 degrees north).

The performance of the beech distribution model was excellent (Table 3). As result of the feature selection process, a total of 23 out of 48 variables were retained as the optimal subset size for the RF method, indicating that the distribution of the species is driven by many interrelationated variables (Table 4). According to the normalized importance score, climate variables contribute most to the model (51.2%) with the thermal and pluviometry variables making the same contribution, although the relative importance of the thermal variables was higher. Nine soil variables were retained and contributed 34.2% to the model. With the exception of soil pH, all of these variables were related to physical properties. However, soil pH was the relatively most important variable among this type. Three terrain variables contributed 10.1% to the model but all were variables with low relative importance.

The functional form of the marginal response plots for the five most important variables was clearly a unimodal relationship with the peak or maximum probability of presence at intermediate levels (Fig. 4). Isothermality (BIO_03) was the most important variable, with a response peak in probability at 40%. This variable can be interpreted as the stability of temperature over the course of a year, or quantification of the day-to-night temperature oscillation relative to the summer-to-winter oscillation. A value of 100% would represent a site where the diurnal temperature range is equal to the annual temperature range. For the annual mean diurnal range (BIO_02), the peak response occurred at 9.3 ºC. For temperature seasonality (BIO_04), a measure of temperature change over the course of the year, the peak response was 500%. The annual temperature range (BIO_8) provides mean temperatures during the consecutives three wettest months of the year, and the peak response occurred at 6 °C. The fifth most important variable was the precipitation of the wettest month (BIO_13), which reached a peak of 115 mm.

3.3 Productivity Model

As productivity models enable SI to be predicted as a function of environmental variables (including several climatic variables), they are capable of predicting changes in SI under non-constant climate. As result of the feature selection process, only 3 of 48 variables were retained as the optimal subset size for RF method. Based on variable importance scores, soil silt percentage and the monthly mean diurnal range contributed most to the model (78% of the importance score). The plan curvature (PLC) of the terrain contributed the remaining 22 % of the importance score (Table 5).

The functional form of the marginal response plots of the first two variables is similar, with a continuous increase in site index as silt percentage in soil increased from 31% to 40-42 %, at which a peak or maximum site index of 17.5 m is reached (Fig. 5). The monthly diurnal mean ranged from 7.5 to 9.5 and the SI increased gradually from 13 to 15 m; after reaching a value of 10, the SI increased rapidly from 12.5 to a peak of 20 m at a 10.5. The relationship between the functional response of the PLC and the peak SI around a PLC zero value was very flat and unimodal (for linear surfaces, neither convex nor concave).

Model performance was good; no trends were observed in residuals (Fig. 6) and about 54.09 %. of the variance was explained. Taking into account the metrics of average residuals, the root mean square error was 3.4936 m and the mean average error, 1.0245 m, representing respectively 20.82% and 16.50% of the mean site index value (16.78 m).

3.4. Predictable effects of climate change on beech forest habitat suitability and productivity

The predictions regarding the impact of climate change on the potential distribution of beech in the Cantabrian Range suggest that there will be a drastic reduction in the area of suitable habitat for the species (Fig. 7). SDM projections under the two different emissions scenarios reveal very important shifts in suitable beech habitat towards more favourable environmental conditions, the magnitude of which mainly depends on the scenario considered. Under the moderate scenario (RCP 4.5), the mean latitude of the suitable habitat will shift 0.01352 degrees north and the elevation will increase by almost 100 m. Considering the area occupied and the degree of habitat fragmentation of beech forests, the total surface area will decrease by 40% and 45% by respectively 2050 and 2070, the mean path area will decrease by 40%, the large path area index will decrease by 59% and aggregation index by 6%. Under the pessimistic scenario (RCP 8.5), the suitable optimal conditions for beech forest will almost disappear from the Cantabrian Mountains: suitable habitat will shift 0.02409 degrees northwards and 300 m higher in elevation, and the total surface will decrease dramatically, by around 90% and 95% by respectively 2050 and 2070; the mean path area will decrease by 93%, the large path area index by 99% and the aggregation index by 36% (Fig. S3).

Figure S4 shows projections for 2050 and 2070 of the distribution of the five climatic variables of relative importance greater than 60% under the two future climatic scenarios. The future projections reveal that main climatic variables will shift under both climatic scenarios, but the greatest change will occur under the more pessimistic scenario (RCP 8.5) with the time horizon being less important. Isothermality (BIO_03) will clearly shift towards lower values under the pessimistic scenario. This score is result of dividing the diurnal range between annual ranges. As the mean diurnal range (BIO_02) decreased only slightly under this scenario, the large annual variation in temperature (BIO_04) confirmed the changes in BIO_03. Mean temperatures of the wettest quarter (BIO_08) will shift toward warmer days, whereas precipitation of wettest month (BIO_13) will decrease slightly in the moderate scenario and increase slightly in the pessimistic scenario. It appears that change in temperature ranges rather than changes in precipitation will have a greater impact on the suitable beech habitat.

Finally, Figure 8 shows the predicted SI for the future suitable habitat under the two future climatic scenarios for 2050 and 2070 time horizons. The future projections reveal that under the lower emission scenario (RCP 4.5) the mean SI for the suitable habitat will undergo a large increase from 15.19 m to 18.18 m with very little influence of the time horizon. However, under the higher emission scenario (RCP 8.5), SI will increase only slightly from 15.19 to 15.53 and 16.42 m by 2050 and 2070 respectively.

4. Discussion

4.1. Site index model

Site index (SI) is a key variable in modelling forest productivity. In this study, the generalized algebraic difference approach (Cieszewski and Bailey, 2000) was used to generate polymorphic curves with data from stem analysis. The selected equations provided good fits for both dominant height (Dieguez-Aranda et al., 2006) and basal area (Barrio-Anta et al., 2008), and they fulfilled most of the desirable properties that a site quality equation should possess (Diéguez-Aranda et al., 2006).

Beech is a widespread forest tree in Europe. However, very few site quality curves have been developed for the species relative to others (e.g. conifers). As a result, site quality curves are often applied outside the area for which they were explicitly constructed. In Spain, beech site quality curves have been developed for the regions of La Rioja (Iberian Range) (Ibáñez, 1989) and Navarra (Pyrenees) (Madrigal et al., 1992). Nevertheless, the curves elaborated by Madrigal et al. (1992) are often applied in Spain, even for different biogeo-climatic zones (Elena, 1997) such as Catalonia (Elena et al., 2001), La Rioja (Blanco et al., 2003) and Castilla y León (Sánchez et al., 2003).

Our SI model fitted well to the observed values of the stem analysis and distinguished 4 site qualities defined by heights of 5, 12, 19 and 26 m at a reference age of 80 years. The results were based on trees of ages between 43 and 199 years. The curves can therefore be used over the entire rotation of the species in Spain, between 100 and 150 years (Madrigal et al., 2008).

Visual comparison of our curves with the Navarra curves constructed by Madrigal et al. (1992) shows different growth and range of site qualities (Fig. S5). Previous studies in the Iberian Peninsula have pointed out that the Cantabrian Range encompasses a huge variety of site qualities (e.g. Gandullo et al., 2004), with the best occurring in the region of Navarra (Ruiz de la Torre, 2006). We confirmed that this is generally true as five of the site qualities reported for Navarra are between our site qualities 1 and 3, and none were as low as our site quality 4. This may be because beech forest grows in steeper sites in the Cantabrian Range and also because of the existence of damaged stands, remains of ancient forests, maintained on poor sites. By contrast, the forests in Navarra occupy comparatively flatter and more undulating land (Gandullo et al., 2004). As an exception, our best site quality corresponded to a stand selected for seed production for its above average quality (Agúndez-Leal et al., 1995), better than any of those in Navarra.

4.2. Species Distribution Model

SDMs correlate species occurrence and environmental parameters and have arisen as a widely used modelling technique to map the current species distributions (Gray and Hamann, 2013). Our study revealed the main environmental factors driving the distribution of beech in the Cantabrian Range, a region characterized by a temperate oceanic climate and a unique elevational gradient in the context of the Atlantic biogeographic region in Europe. According to our results, climate makes the greatest contribution to the beech distribution in the Cantabrian Range, as indicated by e.g. Fang and Lechowicz (2006). Temperature-related variables had the strongest effects, with isothermality, annual mean diurnal range, temperature seasonality and mean temperature of wettest quarter (Dec-Ja-Feb, BIO_08) identified as the four most important variables. Precipitation in the wettest month (December) is the fifth most important variable according to the relative importance score and the first variable related to pluviometry. Considering the Atlantic influence in the study area (Roces-Díaz et al., 2015; Rozas et al., 2015), it is not surprising that precipitation has a weaker influence than temperature on the distribution of beech, unlike other studies carried out in the Mediterranean region (e.g. Catalonia) where beech was mainly restricted to areas with > 950 mm of annual rainfall and within this, in areas with < 1050 mm, the distribution was related to winter and summer precipitation (Thuiller et al., 2003).

In the study area, only Roces-Díaz et al. (2015) have analyzed the distribution of beech forests with SDMs. These researchers found that the presence of beech was positively related to soil fertility (suggesting that the species prefers basic substrates) and negatively related to mean temperature of daily minimum during January (reflecting the ability to resist frost damage) and to the accumulated solar radiation during one year. The mean daily minimum temperature during January may be comparable to the BIO_08 parameter, which acted similarly in our model. However, lithostratigraphic permeability (comparable to soil fertility) appeared in the 21st position (low importance), and solar radiation was not significant in our study.

4.3. Productivity Model

Models that use environmental parameters as SI predictors (indirect or geocentric models) may help forest managers to classify stand productivity when the stand age is not known or when tree measurements are not available (Clutter et al., 1983). In the past few decades, the geocentric approach has been widely used, incorporating all types of environmental parameters. Our findings showed that soil physical, climatic and terrain parameters are driving factors in determining SI for beech in the Cantabrian Range. The model used to predict SI included only three predictor variables: soil silt percentage, monthly mean diurnal range and plan curvature of the terrain. As beech is very sensitive to water deficit, it is surprising that the model did not include any rainfall parameter.

Soil physical parameters are often included in SI geocentric models as predictor variables (e.g. Bravo-Oviedo and Montero, 2005; Brandl et al., 2014). Silt is related to texture and therefore to soil drainage conditions, which are important for estimating forest productivity (e.g. Jokela et al., 1988), and it is a key factor in the Mediterranean area (Bravo and Montero, 2001). According to Brandl et al. (2014), silt content can be interpreted as a proxy for soils with more favourable physical properties regarding water and air balance than soils with high sand or clay content, because silt is associated with better soil aeration and water retention capacity. Beech can grow on any type of soil as long as the soil is sufficiently well drained (Leuschner et al., 2006). Is therefore very sensitive to excess water (it does not tolerate flooding) as well as to a lack of water in the soil (its shallow rooting makes it susceptible to drought), so clay-rich and sand-rich soils are not favourable for beech (Le Tacon, 1981). In fact, soil water availability, mainly in early summer, has been identified as the main driver of beech growth (e.g. Scharnweber et al., 2011) and it is linearly related to the percentage of silt in soil texture (Gandullo et al. 2004). A silt percentage of around 42% (our optimal) (Fig. 5) mainly corresponded to loamy soils, which are characterized by being well-aerated, fertile and fresh soils with a high water retention capacity (Costa et al., 1997).

Climatic data are commonly used parameters in geocentric models (e.g. Albert and Schmidt, 2010; Bosela et al., 2016; Brandl et al., 2018). The diurnal temperature range and its monthly mean represent the thermal amplitude that beech can endure throughout the year in the environment where it grows.

Thus, more continental climates have a greater thermal amplitude than other more temperate climates. In the same way, thermal amplitude decreases as elevation increases, as the air is colder at higher elevations (e.g. Rubio-Cuadrado et al., 2018). Our SI predictions indicated that SI increases as the monthly mean diurnal range increases up to 10 degrees, then increases rapidly and reaches its optimum at around 10.5 degrees (around 5 m for half a degree), beyond which it begins to decrease rapidly (Fig. 5). Similar findings were reported by Albert and Schmidt (2010) and Brandl et al. (2018). These researchers observed that SI increases with temperature during growing season and slows down at high temperatures, but in a wider thermal range than in the present study. In other words, for beech, SI is higher in sites where the average daily temperature is around 10.5 ^{o}C throughout the year, which apparently indicates temperate environmental conditions without frost or drought.

Several studies have shown the effects of temperature on beech growth, both in height and in diameter. For instance, both very low and high temperatures in January cause a reduction in height growth (e.g. Seynave et al., 2008; Brandl et al., 2018). Cold temperatures that induce late frosts in spring, at the beginning of the vegetative period, also have a negative effect on beech growth (e.g. Seynave et al., 2008; Rozas et al., 2015). However, Rubio-Cuadrado et al. (2018) observed that cool conditions between February and April enhance beech growth in the Cantabrian Range. Similar findings were observed in the north-eastern Italian pre-Alps (Piutti and Cescatti, 1997), Western Carpathians (Bosela et al., 2016) and Northwest Germany (Mausolf et al., 2018). These studies convert these cool conditions into water availability during the growing season. Their results indicate that early-season water shortage (February-March to May) and not summer water shortage (June to August) is the main driver of declining radial growth rates in beech. However, other studies have shown that high summer temperatures favour water deficit and stomatal closure, resulting in a reduction in height and radial growth both in the year of the summer drought (e.g. Seynave et al., 2008; Scharnweber et al., 2011) and in the following year (e.g. Hacket-Pain et al., 2016; Farahat and Linderholm, 2018). The previous findings suggest that beech develops correctly within a certain temperature range, so that when the temperatures are either very low or very high, and are consequently outside of that range, growth of the trees is negatively affected.

Many studies have also shown the effect of topographic variables on site productivity (e.g. Bergès et al., 2005; Bergès and Balandier, 2010). Topographic position, exposure and slope usually have significant effects on SI. Our findings show the optimum beech SI around a plan curvature equal to zero (Fig. 5), i.e. neither convex nor concave (what Bergès and Balandier, 2010 refer to as "neutral"). The aforementioned authors converted the topographic position into soil water availability, so that more concave surfaces correspond to higher water content and more convex surfaces correspond to a lower water availability. In the Cantabrian Range, beech forests grow in mountain areas where rainfall is high enough, but where surface run-off is also high because of the slope (Rozas et al., 2015). As previously mentioned, beech is very sensitive to drought and does not tolerate flooding. A convex curvature reduces percolation of water into the soil and is also related to higher nutrient loss and soil erosion (Bueis et al., 2017). On the other hand, a concave curvature favours flooding, which causes oxygen deprivation in plants and entails a reduction of energy demanding processes such as growth (Kreuzwieser et al., 2009).

As mentioned above, rainfall was part of the model, although indirectly. The three selected parameters had an indirect effect on the water available for beech forests. However, the non-inclusion of explicit rainfall parameters in the model can be explained. According to Rozas et al. (2015), cloud immersion and foggy conditions mitigate the drought-sensitivity of beech in the Cantabrian Range. Both directly affect the forest water budget (via the capture of cloud water by the canopy), increase air humidity and reduce leaf transpiration. This is particularly true in unmanaged forests, as in the present study, where stem density and canopy closure were higher, resulting in a higher air humidity (Latif and Blackburn, 2010).

From the point of view of model performance, the selected model explained 54.09% of the total variance, which is an usual intermediate value in this type of studies (McKenney and Pedlar, 2003). Some authors reported better performance than in the present study for SI prediction with other species, mainly conifers (e.g. Bravo and Montero (2001) for Scots pine, Bravo-Oviedo and Montero (2005) for *Pinus pinea* L., Brandl et al. (2014) for Norway spruce, Bueis et al. (2017) for *Pinus halepensis* Mill., etc.). However, several studies have reported similar results (e.g. Fontes et al. (2003) for Douglas fir, Bergès et al. (2005) for Sessile oak, etc.) and others slightly poorer results than in the present study. However, many of those models are of little practical application as they include some soil variables that require the application of complicated, expensive analytical techniques.

As far as we are aware, only three other studies have examined the relationship between abiotic site characteristics and SI in beech forests: Seynave et al. (2008) in France, Albert and Schmidt (2010) in Germany, and Brandl et al. (2018) in Germany and France. The model proposed by Seynave et al. (2008) explained 59% of the variance in SI and included four climatic variables (mean temperatures for May, July and January, and December precipitation) and three soil variables (pH, Carbon/Nitrogen ratio and soil depth). The model developed by Albert and Schmidt (2010) correctly classified 34% of the cases with water balance and mean temperature in growing season, centered mean annual nitrogen deposition, longitude and latitude as environmental variables. The model produced by Brandl et al. (2018) explained 40.13% of the variation in SI, and the selected model variables were mean temperature of the warmest quarter, total precipitation during the growing season (May to September) and elevation.

Although our results are not identical to those obtained in these three studies, they do share a common basis, as shown above. Of the parameters found to be significant in these studies, our database did not include monthly temperatures and rainfall, or the carbon/nitrogen ratio, or the centred mean annual nitrogen deposition. However, they do share the effects of soil fertility, and excessively high or excessively low temperatures throughout the year and their relation to water availability. Other parameters such as pH and soil depth were taken into account in our study, but they were not significant.

Finally, studies that use geocentric models to predict SI apply various methodologies, both procedural and statistical. The two-step methodology used in this study has already been applied in the previously mentioned studies focused on beech forests, as well as in other studies involving different species (e.g. Fontes et al., 2003; Bergès et al., 2005; Bravo-Oviedo and Montero, 2005; amongst others). Brandl et al. (2018) reflect on this procedure, indicating two reasons why they prefer it to modelling SI dependence on age and environmental variables in one step. First, determination of SI as a measure of age is independent of the uncertainty of environmental variables. Second, modelling the SI dependence on environmental variables in a separate step has the advantage that the effect and explanatory power of environmental variables on SI is immediately clear and separated from the effect of age. For instance, Brandl et al. (2014) explained 65.2% of the variance in SI for Norway spruce in Bavaria, whereas age alone explained 56.9% of the variance.

Several statistical approaches have been used to develop these geocentric methods, ranging from multiple linear regression to artificial neural networks (Aertsen et al., 2010). Nonparametric techniques such as regression tree-based methods are among the most flexible and robust for this purpose (Jiang et al., 2014). Random Forest is a nonparametric ensemble classification and regression tool, which constructs hundreds of decision trees using randomized subsets of predicted and predictor variables (Breiman, 2001). Despite its virtues (see Jiang et al., 2014), it is less widely used than other techniques such as multiple regression and General Additive Models. Very few studies have used RF to predict SI from environmental parameters (e.g. Weiskittel et al., 2011; Jiang et al., 2014), although they have obtained good results. However, several studies have shown that RF tends to overestimate lower values and underestimate higher values (e.g. Nunes and Görgens, 2016), unlike in the present study.

4.4. Predictable effects of climate change on suitable habitat for beech and the associated productivity

Climate change is a global phenomenon that has already clearly contributed to changes in forest productivity and in the distribution and abundance of plant species (e.g. Monzón et al., 2011; Hacket-Pain et al., 2016). Plants are particularly vulnerable to the alterations produced by climate change, among which beech can be highlighted for its sensitivity to water deficit, which has led to an increase in studies aiming to predict the response to this type of environmental change.

There is a broad consensus that rising temperatures and a decline in the amount of precipitation during the growing season (mainly in spring and summer) will trigger an increase in frequency of drought periods in the upcoming decades in Southern Europe (e.g. Rubio-Cuadrado et al., 2018), which will cause a latitudinal shift towards the north and an upwards elevational shift in habitats that are suitable for beech forests (e.g. Kramer et al., 2010; Falk and Hempelmann, 2013). Our findings also indicate this geographical shift (Fig. S3) and thus suggest a drastic reduction in the area of habitat suitable for beech forests in the Cantabrian Range, reducing the area by around half under the moderate scenario and almost total disappearance of suitable habitat under the pessimistic scenario (see Figures 7 and S3). These results are consistent with those reported in another studies. Several researchers have predicted a significant reduction in the surface area of the Cantabrian beech forests (e.g. Kramer et al., 2010; Falk and Hempelmann, 2013; or Dyderski et al., 2017 for RCP 4.5 scenario), while others indicate the almost total disappearance of this type of forest (Meier et al., 2011; or Dyderski et al., 2017 for RCP 8.5 scenario).

However, SDM projections regarding climate change should not be literally interpreted as predicted species demographics (Gray and Hamann, 2013). Although the predicted loss of suitable habitat does not necessarily entail removal of current populations (Hampe, 2004), expected environmental effects of climate change on beech forest will lead to deterioration of the conditions necessary for future growth. These poorer conditions may lead to rather lower levels of regeneration (Silva et al., 2012), because a decline in frequency of favourable years for reproduction is expected, but also in the reduction in local density of populations (e.g. Geßler et al., 2007; Falk and Hempelmann, 2013) or in a higher risk of mortality of trees (e.g. Allen et al., 2010). In addition, other studies provide some hope in the face of such a pessimistic outlook for beech forests in the Cantabrian Range. For example, Jump et al. (2006) and Hacket-Pain et al. (2016) have shown that climatic and site constraints at the species distribution margins can also lead to adaptive responses that may enhance the tolerance of populations to drier environmental conditions. Similarly, Psidova et al. (2018) have shown that beech forests at higher elevations are less sensitive to drought and heat stress. Finally, several studies claim that the response of beech to climate warming can be mitigated by producing more diversified stands in terms of tree height (Bosela et al., 2016), by leaving stands unmanaged (Mausolf et al., 2018) or by establishing mixed stands (Geßler et al., 2007).

In terms of productivity, the study findings reveal that the mean SI of the suitable beech habitat increased under both emission scenarios considered (Fig. 8). Only three studies have indicated how climate change may affect SI in beech forests, with increases and decreases and considerable regional variation, as with other species (e.g. Jiang et al., 2014). Albert and Schmidt (2010) predicted a decrease in SI at elevations below 300 m and an increase at higher elevations in Lower Saxony.

Nothdurft et al. (2012) predicted an increase in SI in larger areas with lower elevation in the Schwarzwald and Swabian Alps, and in the south-eastern part of the Alpine foothills region (Baden-Württemberg), whereas Brandl et al. (2018) predicted a decrease in SI for Southern Germany. In other words, the response depends on location, and no clear pattern has been observed for the species (e.g. Weiskittel et al., 2011). Seynave et al. (2008) did not make any SI projections in relation to climate change.

Although information on the responses of forest ecosystems to climate change has increased in recent years, uncertainties due to temporal and geographical scale remain. As pointed out by Bosela et al. (2016), individual studies differ greatly in the type of data used and statistical methods applied, making comparison and generalization difficult.

5. Conclusions

We used powerful machine learning techniques and currently available spatially-continuous environmental variables to develop two raster-based models of 250 m resolution and thus generate suitable habitat and SI estimates for beech trees in the Cantabrian Range in Northwestern Spain. Both models incorporate climatic variables and enable prediction of future values under different climate change scenarios. Climate change is expected to cause a large reduction in the area of habitat suitable for beech by 2070 (loss of around 40 and 90% for the moderate and pessimistic scenarios respectively). By contrast, an average increase in SI of 3 m is expected for the moderate scenario and no change for the pessimistic scenario because of the almost total disappearance of suitable habitat. Predicted loss of suitable habitat may lead to less favourable conditions for seedling establishment, a reduction in local density of populations and/or in a higher risk of mortality of adults, but does not entail current population removal, because it does not consider adaptive responses of the species or ecosystem management. In this respect, the models developed may be useful tools for helping decision-makers to develop plans for protecting biodiversity, forest management plans and species re-habitation plans to prevent or mitigate the impact of climate change on beech forests. Further research aimed at obtaining a better understanding of the complex relationships between environmental variables and species occurrence and productivity is needed to enhance these climate-sensitive predictive models.

6. Acknowledgements

While undertaking the present study, the first author was in receipt of a Severo Ochoa Fellowship from the Asturias Government-FICYT (code 09/111). The authors are grateful to the Asturias Government Forest Service for valuable help and also for financing the felling of the trees used in this study. Unfortunately, the authors were unable to obtain permission from the Environmental Service to measure (more) experimental plots or felling trees for scientific purposes. Nonetheless, they hope that the study findings will be of use for the management of these valuable ecosystems. The authors also thank José Manuel Álvarez González and Dr María Castaño Díaz for help with field and laboratory tasks. Special thanks to Dr. Christine Francis for revising the English of the manuscript. Finally, the authors thank two anonymous referees for their valuable, highly constructive comments and suggestions.

7. References

Aertsen, W., Kint, V., van Orshoven, J., Ozkan, K., Muys, B., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecol. Model. 221, 1119-1130. https://doi.org/10.1016/j.ecolmodel.2010.01.007.

Agúndez-Leal, D., Martin, S.; De Miguel, J.; Galera, R.; Jiménez, M. y Díaz-Fernández, P., 1995. Las regiones de procedencia de *Fagus sylvatica* en España. ICONA, Madrid.

Albert, M., Schmidt, M., 2010. Climate-sensitive modelling of site-productivity relationships for Norway spruce (*Picea abies* (L.) Karst.) and common beech (*Fagus sylvatica* L.). For. Ecol. Manage. 259, 739-749. https://doi.org/10.1016/j.foreco.2009.04.039.

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H., González, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.H., Allard, G., Running, S.W., r, Semerci, A., Cobb, N., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manage. 259 (4), 660-684. https://doi.org/10.1016/j.foreco.2009.09.001.

Bailey, R.L., Clutter, J.L., 1974. Base-age invariant polymorphic site curves. For. Sci. 20 (2), 155-159. https://doi.org/10.1093/forestscience/20.2.155

Barrio-Anta, M., Sixto-Blanco, H., Cañellas-Rey De Viñas, I., Castedo-Dorado, F., 2008. Dynamic growth model for I-214 poplar plantations in the northern and central plateaux in Spain. For. Ecol. Manage. 255, 1167-1178. https://doi.org/10.1016/j.foreco.2007.10.022

Bergès, L., Balandier, P., 2010. Revisiting the use of soil water budget assessment to predict site productivity of sessile oak (*Quercus petraea* Liebl.) in the perspective of climate change. Eur. J. For. Res. 129, 199-208. https://doi.org/10.1007/s10342-009-0315-1.

Bergès, L., Chevalier, R., Dumas, Y., Franc, A., Gilbert, J.M., 2005. Sessile oak (*Quercus petraea* Liebl.) site index variations in relation to climate, topography and soil in even-aged high-forest stands in northern France. Ann. For. Sci. 62, 391-402. https://doi.org/10.1051/forest:2005035.

Blanco, A., Graña, D., Elena, R., Sánchez, O., Rubio, A., Gómez, V., 2003. Autoecología de los hayedos de La Rioja. Invest. Agrar.: Sist. Recur. For. 12, 21-38.

Bontemps, J.D., Bouriaud, O., 2014. Predictive approaches to forest site productivity: recent trends, challenges and future perspectives. Forestry 87 (1), 109-128. https://doi.org/10.1093/forestry/cpt034.

Bosela, M., Sˇtefancˇík, I., Petrás, R., Vacek, S. 2016. The effect of climate warming on the growth of European beech forests depend critically on thinning strategy and site productivity. Agr. Forest Meteorol. 222, 21-31. http://dx.doi.org/10.1016/j.agrformet.2016.03.005.

Brandl, S., Falk, W., Klemmt, H.J., Stricker, G., Bender, A., Rötzer, T., Pretzsch, H., 2014. Possibilities and Limitations of Spatially Explicit Site Index Modelling for Spruce Based on National Forest Inventory Data and Digital Maps of Soil and Climate in Bavaria (SE Germany). Forests 5 (11), 2626-2646. https://doi.org/10.3390/f5112626.

Brandl, S., Mette, T., Falk, W., Vallet, P., Rötzer, T., Pretzsch, H., 2018. Static site indices from different national forest inventories: harmonization and prediction from site conditions. Ann. For. Sci. 75, 56. https://doi.org/10.1007/s13595-018-0737-3.

Bravo, F., Montero, G., 2001. Site index estimation in Scots pine (*Pinus sylvestris* L.) stands in the high Ebro basin (northern Spain) using soil attributes. Forestry 74, 395-406. https://doi.org/10.1093/forestry/74.4.395.

Bravo-Oviedo, A., Montero, G., 2005. Site index in relation to edaphic variables in stone pine (*Pinus pinea* L.) stands in south west Spain. Ann. For. Sci. 62, 61-72. https://doi.org/10.1051/forest:2004086.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5-32. https://doi.org/10.1023/A:1010933404324.

Bueis, T., Bravo, F., Pando, V., Turrión, M.B., 2017. Site factors as predictors for *Pinus halepensis* Mill. productivity in Spanish plantations. Ann. For. Sci. 74, 6. https://doi.org/10.1007/s13595-016-0609-7.

Cieszewski, C.J., 2002. Comparing Fixed- and Variable-Base-Age Site Equations having single versus multiple asymptotes. For. Sci. 48 (1), 7-23. https://doi.org/10.1093/forestscience/48.1.7.

Cieszewski, C.J., Bailey, R.L., 2000. Generalized algebraic difference approach: theory based derivation of dynamic equations with polymorphism and variable asymptotes. For. Sci. 46, 116-126. https://doi.org/10.1093/forestscience/46.1.116.

Cieszewski, C.J., Harrison, M., Martin, S.W., 2000. Practical methods for estimating non-biased parameters in selfreferencing growth and yield models. University of Georgia, Athens, Ga. PMRC-TR 2000-7.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., Bailey, R.L., 1983. Timber Management: A Quantitative Approach. John Wiley and Sons, New York.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Whichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8, 1991-2007. https://doi.org/10.5194/gmd-8-1991-2015.

Costa, M., Morla, C., Sáinz, H., 1997. Los bosques ibéricos. Planeta, Barcelona.

Deschamps, B., McNairn, H., Shang, J., Jiao, X., 2012. Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier. Can. J. Remote Sens. 38, 60-68. https://doi.org/10.5589/m12-012.

DGCN, 2006. III Inventario Forestal Nacional (1997–2006). Principado de Asturias. Dirección General de Conservación de la Naturaleza, Secretaría General de Medio Ambiente, Ministerio de Medio Ambiente. Madrid.

Díaz, T.E., Fernández, J.A., 1987. Asturias y Cantabria, in: Peinado Lorca, M., Rivas-Martínez, S. (Eds.), La vegetación de España. Servicio de publicaciones de la Universidad de Alcalá de Henares, Madrid, pp. 77-116.

Diéguez-Aranda, U., Álvarez-González, J.G., Barrio-Anta, M., Rojo Alboreca, A., 2005. Site quality equations for *Pinus sylvestris* L. plantations in Galicia (northwestern Spain). Ann. For. Sci. 62, 143-152. https://doi.org/10.1051/forest:2005006.

Diéguez-Aranda, U., Burkhart, H.E., Amateis, R.L., 2006. Dynamic site model for Loblolly pine (*Pinus taeda* L.) plantations in the United States. For. Sci. 52 (3), 262-272. https://doi.org/10.1093/forestscience/52.3.262.

Dyderski, M.K., Paz, S., Frelich, L.E., Jagodzinski, A.M., 2017. How much does climate change threaten European forest tree species distributions? Glob. Change Biol. 24, 1150-1163. https://doi.org/10.1111/gcb.13925.

EEA, 2017. Climate change, impacts and vulnerability in Europe 2016. An indicator-based report. European Environment Agency Report Nº 1/2017.

Elena, R., 1997. Clasificación Biogeoclimática de España Peninsular y Balear. Ministerio de Agricultura, Pesca y Alimentación, Madrid.

Elena, R., Sánchez, F., Rubio, A., Gómez, V., Aunós, A., Blanco, A., Sánchez, O., 2001. Autoecología de los hayedos catalanes. Invest. Agrar.: Sist. Recur. For. 10, 21-42.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129-151. https://doi.org/10.1111/j.2006.0906-7590.04596.x.

Falk, W., Hempelmann, N., 2013. Species favourability shift in Europe due to climate change: A case study for *Fagus sylvatica* L. and *Picea abies* (L.) Karst. based on an ensemble of climate models. J. Climatol. 2013, 1-18. https://doi.org/10.1155/2013/787250.

Fang, J., Lechowicz, M.J., 2006. Climatic limits for the present distribution of beech (*Fagus* L.) species in the world. J. Biogeogr. 33, 1804-1819. https://doi.org/10.1111/j.1365-2699.2006.01533.x

Farahat, E., Linderholm, H.W., 2018. Growth–climate relationship of European beech at its northern distribution limit. Eur. J. For. Res. 137, 619-629. https://doi.org/10.1007/s10342-018-1129-9.

Fielding, A., Bell, J., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38-49. https://doi.org/10.1017/S0376892997000088.

Fontes, L., Tomé, M., Thompson, F., Yeomans, A., Sales, L.J., Savill, P., 2003. Modelling the Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) site index from site factors in Portugal. Forestry 76, 491-507. https://doi.org/10.1093/forestry/76.5.491.

Gandullo, J.M., Blanco, A., Sánchez, O., Rubio, A., Elena, R., Gómez, V., 2004. Las estaciones ecológicas de los hayedos españoles. Monografías INIA: Serie Forestal nº 8. Ministerio de Educación y Ciencia, Madrid.

Geβler, A., Keitel, C., Kreuzwieser, J., Matyssek, R., Seiler, W., Rennenberg, H., 2007. Potential risks for European beech (*Fagus sylvatica* L.) in a changing climate. Trees 21, 1-11. https://doi.org/10.1029/2005GL025044.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random Forests for land cover classification. Pattern Recogn. Lett. 27 (4), 294-300. https://doi.org/10.1016/j.patrec.2005.08.011.

Gray, L.K., Hamann, A., 2013. Tracking suitable habitat for tree populations under climate change in western North America. Climatic Change 117, 289-303. https://doi. org/10.1007/s10584-012-0548-8.

Guisan A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Modell. 135, 147-186. https://doi. org/10.1016/S0304-3800(00)00354-9.

Hacket-Pain, A.J., Cavin, L., Friend, A.D., Jump, A.S., 2016. Consistent limitation of growth by high temperature and low precipitation from range core to southern edge of European beech indicates widespread vulnerability to changing climate. Eur. J. For. Res. 135, 897-909. https://doi.org/10.1007/s10342-016-0982-7.

Hall, M., 1999. Correlation-based feature selection for Machine Learning. PhD Thesis. University of Waikato. Hamilton, New Zealand.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA data mining software: An update. SIGKDD Explorations 11 (1), 10-18. https://doi.org/10.1145/1656274.1656278.

Hampe, A., 2004. Bioclimate envelope models: what they detect and what they hide. Glob. Ecol. Biogeogr. 12 (5), 469-471. https://doi.org/10.1111/j.1466-822X.2004.00090.x.

Harris, R.M.B., Grose, M.R., Lee, G., Bindoff, N.L., Porfirio, L.L., Fox-Hughes, P., 2014. Climate projections for ecologists. Wiley Interdisciplinary Reviews: Climate Change 5, 621-637. https://doi.org/10.1002/wcc.291.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PloS ONE 12 (2), e0169748. https://doi.org/10.1371/journal.pone.0169748.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965-1978. https://dx.doi.org/10.1002/joc.1276.

Ibáñez, J.I., 1989. El haya (*Fagus sylvatica* L.) en La Rioja. Selvicultura y Ordenación. PhD Thesis. Universidad Politécnica de Madrid.

IGME, 2015a. Mapa Geológico de España a escala 1:200.000. Instituto Geológico y Minero de España, Ministerio de Ciencia, Innovación y Universidades. Madrid.

IGME, 2015b. Mapa Geológico de la Península Ibérica, Baleares y Canarias a escala 1:1.000.000. Instituto Geológico y Minero de España, Ministerio de Ciencia, Innovación y Universidades. Madrid.

IPCC, 2013. Climate Change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel of Climate Change. Cambridge University Press. Cambridge, New York.

Jiang, H., Radtke, P.J., Weiskittel, A.R., Coulston, J.W., Guertin, P.J., 2014. Climate- and soil-based models of site productivity in eastern US tree species. Can. J. For. Res. 45, 325-342. https://doi.org/10.1139/cjfr-2014-0054.

Johnston, A., Fink, D., Reynolds, M.D., Hochachka, W.M., Sullivan, B.L., Bruns, N.E., Hallstein, E., Merrifield, M.S., Matsumoto, S., Kelling, S., 2015. Abundance models improve spatial and temporal prioritization of conservation resources. Ecol. Appl. 25, 1749-1756. https://doi.org/10.1890/14-1826.1.

Jokela, E.J., White, E.H., Berglund, J.V., 1988. Predicting Norway spruce growth from soil and topographic properties in New York. Soil Sci. Soc. Am. J. 52 (3), 809-815. https://doi.org/10.2136/sssaj1988.03615995005200030038x.

Jump, A.S., Hunt, J.M., Penuelas, J., 2006. Rapid climate change-related growth decline at the southern range edge of *Fagus sylvatica*. Glob. Change Biol. 12, 2163-2174. https://doi.org/10.1111/j.1365-2486.2006.01250.x

Kramer, K., Degen, B., Buschbom, J., Hickler, T., Thuiller, W., Sykes, M.T., de Winter, W., 2010. Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change—Range, abundance, genetic diversity and adaptive response. For. Ecol. Manage. 259, 2213-2222. https://doi.org/10.1016/j.foreco.2009.12.023.

Kreuzwieser, J., Hauberg, J., Howell, K.A., Carroll, A., Rennenberg, H., Millar, A.H., Whelan, J., 2009. Differential response of gray poplar leaves and roots underpins stress adaptation during hypoxia. Plant Physiol. 149, 461-473. https://doi.org/10.1104/pp.108.125989.

Larson, A.J., Lutz, J.A., Gersonde, R.F., Franklin, J.F., Hietpas, F.F., 2008. Potential site productivity influences the rate of forest structural development. Ecol. Appl. 18, 899-910. https://doi.org/10.1890/07-1191.1.

Latif, Z.A., Blackburn, G.A., 2010. The effects of gap size on some microclimate variables during late summer and autumn in a temperate broadleaved deciduous forest. Int. J. Biometeorol. 54, 119-129. https://doi.org/10.1007/s00484-009-0260-1.

Le Tacon, F., 1981. Caractérisation édaphique, in: Teissier du Cros, E. (Ed.), Le Hêtre. INRA, París, pp. 77-84.

Leuschner, C., Meier, I.C., Hertel, D., 2006. On the niche breadth of *Fagus sylvatica*: soil nutrient status in 50 Central European beech stands on a broad range of bedrock types. Ann. For. Sci. 63, 355-368. https://doi.org/10.1051/forest:2006016.

Li, G., Du, S., Wen, Z., 2016. Mapping the climatic suitable habitat of oriental arbovitae (*Platycladus orientalis*) for introduction and cultivation at a global scale. Sci. Rep. 6, 30009. https://doi.org/10.1038/srep30009.

Madrigal, A., Puertas, F., Martínez-Millán, F.J., 1992. Tablas de producción para *Fagus sylvatica* L. en Navarra. Serie Agraria nº 3. Departamento de Agricultura, Ganadería y Montes. Gobierno de Navarra, Pamplona.

Madrigal, A., Calama, R., Madrigal, G., Aunós, A., Reque, J.A., 2008. Selvicultura de *Fagus sylvatica* L., in: Serrada, R., Montero, G., Reque, J.A. (Eds.), Compendio de Selvicultura Aplicada en España. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria-Ministerio de Educación y Ciencia-Fundación Conde del Valle Salazar, Madrid, pp. 155-185.

Mausolf, K., Wilm, P., Härdtle, W., Jansen, K., Schuldt, B., Sturm, K., von Oheimb, G., Hertel, D., Leuschner, C., Fichtner, A., 2018. Higher drought sensitivity of radial growth of European beech in managed than in unmanaged forests. Sci. Total Environ. 642, 1201-1208. https://doi.org/10.1016/j.scitotenv.2018.06.065

McGarigal, K., Wan, H.Y., Zeller, K.A., Timm, B.C., Cushman, S.A., 2016. Multi-scale habitat selection modeling: a review and outlook. Landsc. Ecol. 31, 1161-1175. https://doi.org/10.1007/s10980-016-0374-x.

McKenney, D.W., Pedlar, J.H., 2003. Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada. For. Ecol. Manage. 175, 497-507. https://doi.org/10.1016/S0378-1127(02)00186-X.

Meier, E.S., Edwards, T.C. Jr., Kienast, F., Dobbertin, M., Zimmermann, N.E., 2011. Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. J. Biogeogr. 38, 371-382. https://doi.org/10.1111/j.1365-2699.2010.02405.x.

Monzón, J., Moyer-Horner, L., Palamar, M.B., 2011. Climate Change and Species Range Dynamics in Protected Areas. BioScience 61, 752-761. https://doi.org/10.1525/bio.2011.61.10.5.

Newberry, J.D., 1991. A note on Carmean's estimate of height from stem analysis data. For. Sci. 37, 368-369. https://doi.org/10.1093/forestscience/37.1.368.

Noce, S., Collati, A., Santini, M., 2017. Likelihood of changes in forest species suitability, distribution and diversity under future climate: The case of Southern Europe. Ecol. Evol. 7, 9358-9375. https://doi.org/10.1002/ece3.3427.

Nothdurft, A., Wolf, T., Ringeler, A., Bohner, J., Saborowski, J., 2012. Spatio-temporal prediction of site index based on forest inventories and climate change scenarios. For. Ecol. Manage. 279, 97-111. https://doi.org/10.1016/j.foreco.2012.05.018.

Nunes, M.H., Görgens, E.B., 2016. Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. PloS One 11(5), e0154738. https://doi.org/10.1371/journal.pone.0154738.

Parresol, B.R., Scott, D.A., Zarnoch, S.J., Edwards, L.A., Blake, J.I., 2017. Modelling forest site productivity using mapped geospatial attributes within a South Carolina Landscape, USA. For. Ecol. Manage. 406, 196-207. https://doi.org/10.1016/j.foreco.2017.10.006.

Piutti, E., Cescatti, A., 1997. A quantitative analysis of the interactions between climatic response and intraspecific competition in European beech. Can. J. For. Res. 27, 277-284. https://doi.org/10.1139/x96-176.

Potter, K.M., Woodall, C.W., 2014. Does biodiversity make a difference? Relationships between species richness, evolutionary diversity, and aboveground live tree biomass across U.S. forests. For. Ecol. Manage. 321, 117-129. https://doi.org/10.1016/j.foreco.2013.06.026.

Prasad, A., Iverson, L., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181-199. https://doi.org/10.1007/s10021-005-0054-1.

Pretzsch, H., 2009. Forest dynamics, growth and yield. Springer-Verlag Berlin, Heidelberg.

Psidova, E., Zivcak, M., Stojnic, S., Orlovic, S., Gomory, D., Kucerova, J., Ditmarova, L., Strelcova, K., Brestic, M., Kalaji, H.M., 2018. Altitude of origin influences the responses of PSII photochemistry to heat waves in European beech (*Fagus sylvatica* L.). Environ. Exp. Bot. 152, 97-106. https://doi.org/10.1016/j.envexpbot.2017.12.001.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj, P., 2011. RCP 8.5-A scenario of comparatively high greenhouse gas emissions. Climatic Change 109, 33-57. https://doi.org/10.1007/s10584-011-0149-y.

Roces-Díaz, J.V., Jiménez-Alfaro, B., Álvarez-Álvarez, P., Álvarez-García, M.A., 2015. Environmental niche and distribution of six deciduous tree species in the Spanish Atlantic region. iForest 8, 224-231. https://doi.org/10.3832/ifor1183-008.

Rozas, V., Camarero, J.J., Sangüesa-Barreda, G., Souto, M., García-González, I., 2015. Summer drought and ENSO-related cloudiness distinctly drive *Fagus sylvatica* growth near the species rear-edge in northern Spain. Agr. Forest Meteorol. 201, 153-164.
http://dx.doi.org/10.1016/j.agrformet.2014.11.012.

Rubio-Cuadrado, A., Camarero, J.J., del Río, M., Sánchez-González, M., Ruiz-Peinado, R., Bravo-Oviedo, A., Gil, L., Montes, F. 2018. Long-term impacts of drought on growth and forest dynamics in a temperate beech-oak-birch forest. Agr. Forest Meteorol. 259, 48-59. https://doi.org/10.1016/j.agrformet.2018.04.015.

Ruiz de la Torre, J., 2006. Flora mayor. Organismo Autónomo Parques Nacionales, Madrid.

Ryan, T.P., 1997. Modern regression methods. John Wiley & Sons, New York.

Sánchez, O., Rubio, A., Blanco, A., Elena, R., Gómez, V., 2003. Autoecología paramétrica de los hayedos de Castilla y León. Invest. Agrar.: Sist. Recur. For. 12, 87-110.

SAS Institute Inc., 2004. SAS/ETS 9.1 User's Guide. SAS Institute Inc., Cary, NC, USA.

Schar, C., Vidale, P.L., Luthi, D., Frei, C., Haberli, C., Liniger, M.A., Appenzeller, C., 2004. The role of increasing temperature variability in European summer heatwaves. Nature 427, 332-336. https://doi.org/10.1038/nature02300.

Scharnweber, T., Manthey, M., Criegee, C., Bauwe, A., Schröder, C., Wilmking, M., 2011. Drought matters — declining precipitation influences growth of *Fagus sylvatica* L. and *Quercus robur* L. in north-eastern Germany. For. Ecol. Manage. 262, 947-961. https://doi.org/10.1016/j.foreco.2011.05.026.

Seynave, I., Gégout, J.C., Hervé, J.C., Dhôte, J.F., 2008. Is the spatial distribution of European beech (*Fagus sylvatica* L.) limited by its potential height growth? J. Biogeogr. 35, 1851-1862. https://doi.org/10.1111/j.1365-2699.2008.01930.x.

Shirk, A.J., Cushman, S.A., Waring, K.M., Wehenkel, C.A., Leal-Sáenz, A., Toney, C., Lopez-Sanchez, C.A., 2018. Southwestern white pine (*Pinus strobiformis*) species distribution models project a large range shift and contraction due to regional climatic changes. For. Ecol. Manage. 411, 176-186. https://doi.org/10.1016/j.foreco.2018.01.025.

Silva, D.E., Rezende Mazzella, P., Legay, M., Corcket, E., Dupouey, J.L. 2012. Does natural regeneration determine the limit of European beech distribution under climatic stress? For. Ecol. Manage. 266, 263-272. https://doi.org/10.1016/j.foreco.2011.11.031.

Skovsgaard, J.P., Vanclay, J.K., 2008. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. Forestry 81, 13-31. https://doi.org/10.1093/forestry/cpm041.

Tegel, W., Seim, A., Hakelberg, D., Hoffmann, S., Panev, M., Westphal, T., Buntgen, U., 2014. A recent growth increase of European beech (*Fagus sylvatica* L.) at its Mediterranean distribution limit contradicts drought stress. Eur. J. For. Res. 133, 61-71. https://doi.org/10.1007/s10342-013-0737-7.

Thomson, A.M., Calvin, K.V., Smith, S.J., Kyle, G.P., Volke, A., Patel, P., Delgado-Arias, S., Bond-Lamberty, B., Wise, M.A., Clarke, L.E., Edmonds, J.A., 2011. RCP 4.5: a pathway for stabilization of radiative forcing by 2100. Climatic Change 109, 77-94. https://doi.org/10.1007/s10584-011-0151-4.

Thuiller, W., Vaydera, J., Pino, J., Sabaté, S., Lavorel, S., Gracia, C., 2003. Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). Glob. Ecol. Biogeogr. 12, 313-325. https://doi.org/10.1046/j.1466-822X.2003.00033.x.

Wei, X., Kimmins, J.P., Zhou, G., 2003. Disturbances and the sustainability of long-term site productivity in lodgepole pine forests in the central interior of British Columbia—an ecosystem modeling approach. Ecol. Modell. 164, 239-256. https://doi.org/10.1016/S0304-3800(03)00062-0.

Weiskittel, A.R., Crookston, N.L., Radtke, P.J., 2011. Linking climate, gross primary productivity, and site index across forests of the western United States. Can. J. For. Res. 41(8), 1710-1721. https://doi.org/10.1139/x11-086.

Zhiwei, X., Xinghua, W., 2010. Research for information extraction based on wrapper model algorithm. 2010 Second International Conference on Computer Research and Development. Kuala Lumpur, Malaysia, pp. 652-655.

8. Tables

Table 1. Summary statistics for individual tree and stand variables used to develop the site index system.

Statistic	Tree variables (60 dominant trees)			Stand variables (30 plots)		
	Diameter at 1.3m (cm)	Height (m)	Age (years)	Number of trees (ha^{-1})	Basal area ($m^2\ ha^{-1}$)	Dominant height (m)
Minimum	12.90	9.77	43	300	22.51	10.37
Maximum	69.60	38.63	215	2445	94.95	35.90
Mean	39.10	20.21	109.17	1073.70	44.04	19.25
Standard deviation	14.20	6.50	41.88	575.40	13.46	6.02

Table 2. Environmental variables considered as possible predictors in the distribution and site quality models.

Type	Code	Description	Unit	Source
	SLP	Slope based on a digital elevation model	%	
	ASP	Aspect based on a digital elevation model	º	
	CU	Curvature		
	PLC	Plan curvature		
	PRC	Profile curvature		
Terrain	TSI	Terrain shape index		PNOA Lidar
	WI	Wetness index		
	SR_SS	Potential incoming solar radiation in summer solstice	kJ m^2 year^{-1}	
	SR_EQ	Potential incoming solar radiation in equinox	kJ m^2 year^{-1}	
	SR_WS	Potential incoming solar radiation in winter solstice	kJ m^2 year^{-1}	
	DHN	Euclidean distance to hydrographic network	meters	
	BIO_01	Annual mean temperature	mm	
	BIO_02	Mean diurnal range (Mean of monthly (max temp - min temp))	mm	
	BIO_03	Isothermality (BIO_02/ BIO_07) (*100)	ºC	
	BIO_04	Temperature seasonality (standard deviation *100)	ºC	
	BIO_05	Max temperature of warmest month	ºC	
	BIO_06	Min temperature of coldest month	ºC	
	BIO_07	Temperature annual range (BIO_05- BIO_06)	ºC	
	BIO_08	Mean temperature of wettest quarter	ºC	
Climate	BIO_09	Mean temperature of driest quarter	ºC	WorldClim
	BIO_10	Mean temperature of warmest quarter	ºC	
	BIO_11	Mean temperature of coldest quarter	ºC	
	BIO_12	Annual precipitation	mm	
	BIO_13	Precipitation of wettest month	mm	
	BIO_14	Precipitation of driest month (mm)	mm	
	BIO_15	Precipitation seasonality (Coef. of variation)	%	
	BIO_16	Precipitation of wettest quarter	mm	
	BIO_17	Precipitation of driest quarter	mm	
	BIO_18	Precipitation of warmest quarter	mm	
	BIO_19	Precipitation of coldest quarter	mm	
	SC	Soil organic carbon content	mG/ha	
	Ph_H$_2$O	Soil Ph in H_2O solution		
	Ph_KCl	Soil Ph in KCl solution		
	BD	Bulk density of fine earth fraction (< 2mm)	kg m^{-3}	
	CLAY	Percentage of clay in soil	Weight %	
	SAND	Percentage of sand in soil	Weight %	
	SILT	Percentage of silt in soil	Weight %	SoilGrids250m
	CF	Coarse fragments	Volumetric %	
	CEC	Cation-exchange capacity	cmol+ kg^{-1}	
Soil	DB	Absolute deep to bed rock	cm	
	DB200	Depth to bedrock (R horizon) up to 200 cm	cm	
	R	Probability occurrence of R horizon	%	
	Geo_units	Geological units		SGM
	Geo_lit_units	Lithological units		
	LIT_dco	Lithostratigraphy		SSM
	LIT_per	Lithostratigraphy permeability		
	WRB-FULL	Full soil code of the Soil typological units from the World Reference Base (WRB) for Soil Resources		ESDB
	WRB-LEV1	Soil reference group of the Soil typological units from the World Reference Base (WRB) for Soil Resources		

Table 3. Model fit metrics for species distribution (SDM).

Model	Data set	AUC	OA	TSS	Kappa	Sensitivity	Specificity	PoP
SDM	Test	0.9630	0.7240	0.7620	0.7371	0.8190	0.9430	0.2500
	Train	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.2500

AUC = area under the receiver operator curve; OA = overall accuracy; TSS = true skill statistic; Kappa = Cohen's kappa.; PoP = probability of presence. Model fit was assessed on the training data used to fit the model as well as the withheld test data used for model evaluation. All values represent the mean 10-fold cross-validation.

Table 4. Variables included in the beech SDM, including type and relative importance.

Type	Variable	Normalized Importance	Relative Importance	Summarized values in the suitable habitat			
				mean	max	min	sd
Climate	BIO_03	0.064	100	40.15	44	38	0.93
Climate	BIO_02	0.061	92	9.90	11.0	7.3	0.48
Climate	BIO_04	0.055	72	4975	5472	3849	274
Climate	BIO_08	0.055	72	55.68	121	-28	21.83
Climate	BIO_13	0.054	68	116.52	158	90	10.06
Climate	BIO_14	0.046	44	46.18	73	32	4.95
Terrain	SR_ws	0.046	44	897	2206	165	401
Climate	BIO_19	0.045	40	261.44	415	189	38.61
Terrain	DHN	0.045	40	1740	5161	200	939
Climate	BIO_18	0.044	36	167.65	244	126	14.36
Soil	Ph_H_2O	0.044	36	51.56	64	43	3.39
Climate	BIO_15	0.042	28	25.77	34	21	2.18
Soil	CLAY	0.040	24	23.10	36	13	3.16
Soil	SC	0.040	24	47.11	69	21	6.54
Soil	R	0.039	20	28.04	62	10	6.45
Soil	DB	0.038	16	1513	3259	830	286
Soil	DB200	0.037	12	191.84	200	117	10.83
Soil	SAND	0.035	8	39.42	56	27	2.78
Soil	LIT_dco	0.034	4	3	9	1	-
Soil	SILT	0.034	4	37.76	44	29	1.61
Terrain	SLP	0.034	4	19.49	60.03	0.00	8.49
Soil	LIT_per	0.033	0	4	9	1	-
Terrain	TSI	0.033	0	0.00	0.47	-0.73	0.092

To ensure values of variable importance were expressed on comparable scales for each of the response variable, the scores of all the predictors selected were normalized so that they summed to a unit value (normalized importance) or were expressed as relative values: Relative importance = $(VIM-VIM_{min})/(VIM_{max}-VIM_{min})$.

Table 5. Variables included in productivity model, including their type and relative importance.

Type	Variable	Normalized Importance	Relative Importance	Summarized values in the suitable habitat			
				mean	max	min	sd
Soil	SILT	0.436	100	37.76	44	29	1.61
Climate	BIO_02	0.344	58	9.90	11.0	7.3	0.48
Terrain	PLC	0.220	0	0.006	1.62	-1.07	0.197

To ensure values of variable importance were expressed on comparable scales for each of the response variable, the scores of all the predictors selected were normalized so that they summed to a unit value (normalized importance) or were expressed as relative values: Relative importance = (VIM-VIMmin)/(VIMmax-VIMmin).

9. Figure Captions

Figure 1. Location of the study area.

Figure 2. Workflow adopted for modelling and mapping the current and future distribution and site productivity for beech forests under climate change in this study.

Figure 3. Dominant height growth curves for site indices of 5, 12, 19 and 26 m at a base age of 80 years overlaid on the trajectories of the observed values over time.

Figure 4. Marginal response curves for the five most important variables included in *Fagus sylvatica* species distribution model. The variables are ordered by their contribution to the model (importance score). BIO_03 = isothermality, BIO_02 = Mean diurnal range (Mean of monthly (max temp - min temp)), BIO_04 = Temperature seasonality, BIO_08 = Mean temperature of wettest quarter, BIO_13 = Precipitation of wettest month. The mean (black line) and standard deviation (grey area) of the probability presence.

Figure 5. Marginal response curves for the three variables included in *Fagus sylvatica* productivity model. Variables are ordered by their contribution to the model (importance score). SILT= percentage of silt in soil, BIO_02 = Mean diurnal range (Mean of monthly (max temp - min temp)) and PLC = Plan curvature. The mean (black line) and standard deviation (grey area) of the probability presence. The prediction value of site index is shown as a function of each variable while all other variables are held at their median values at presence locations.

Figure 6. Field measures vs. predicted values of SI for beech in training (left) and validation (right). Solid lines indicate the regression fits (n = 30, 10-fold-CV).

Figure 7. Random forest predictions for SI of *Fagus sylvatica* in the Cantabrian Range (North Spain).

Figure 8. Distribution of SI by area covered at sites where beech is presented for five different scenarios. The average SI by area covered is shown on the upper right-hand side of the graph.

10. Figures

Fig 1.

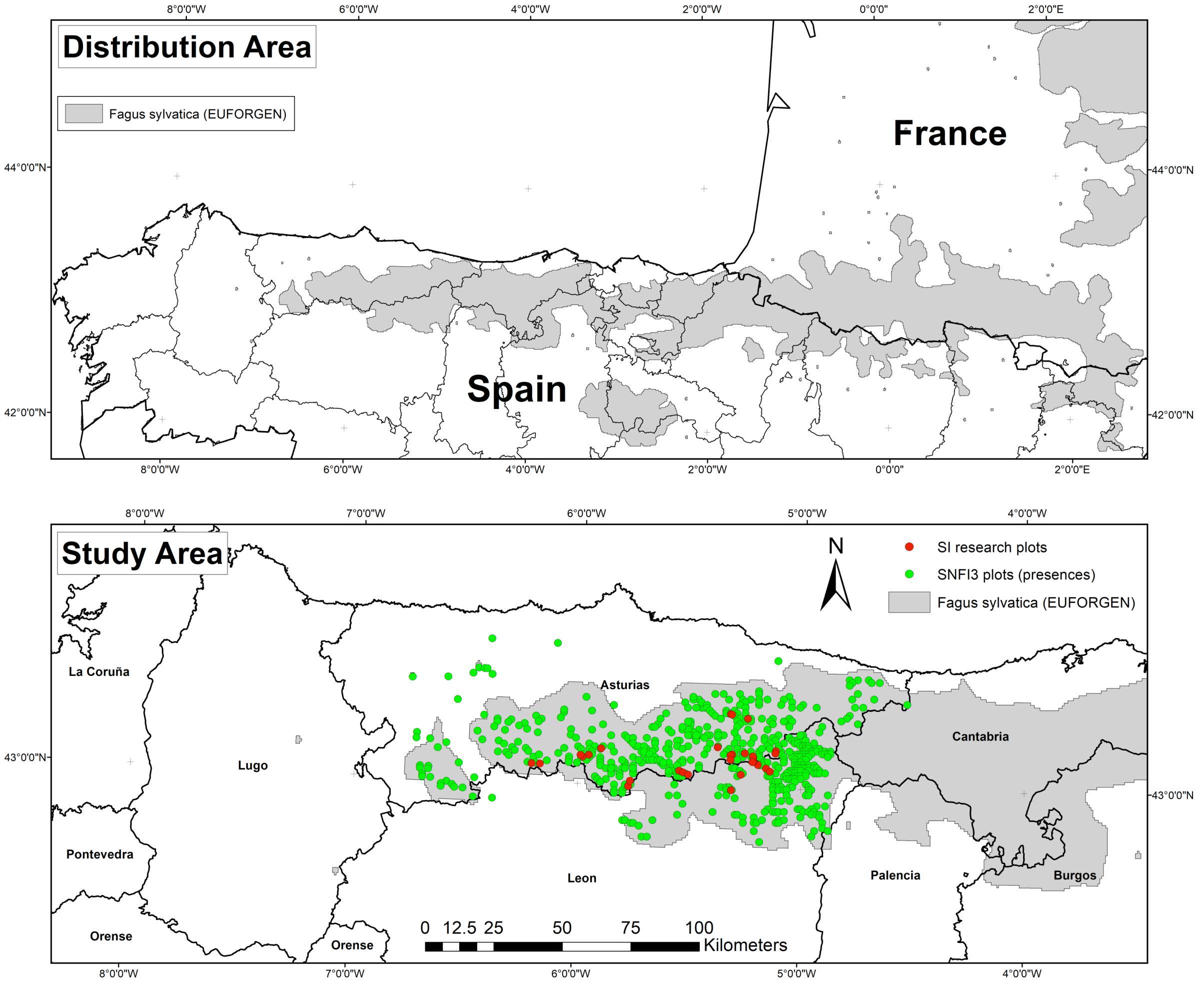


Fig 2.

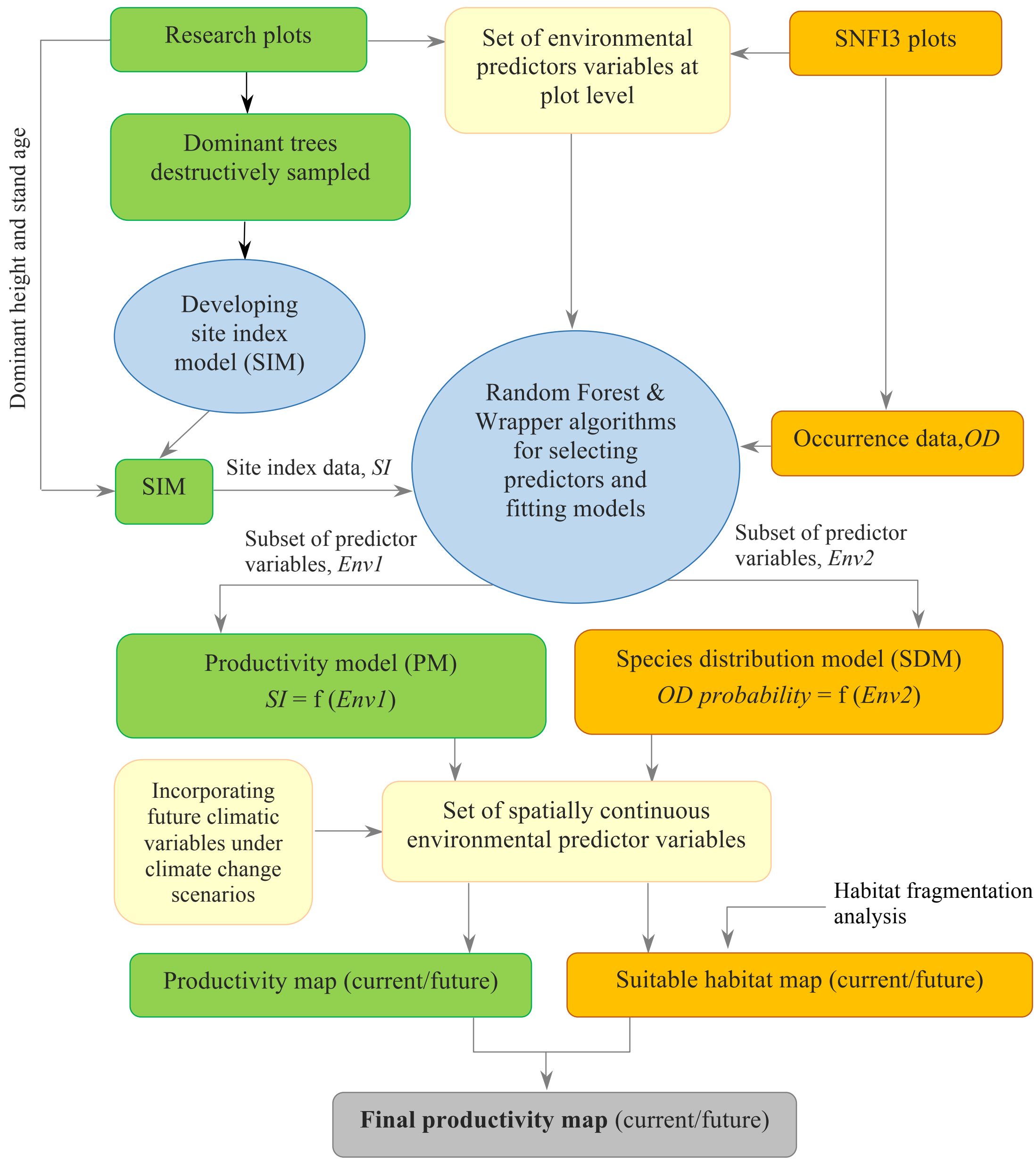

Research plots
Set of environmental predictors variables at plot level
SNFI3 plots
Dominant trees destructively sampled
Dominant height and stand age
Developing site index model (SIM)
Random Forest & Wrapper algorithms for selecting predictors and fitting models
Occurrence data, *OD*
SIM
Site index data, *SI*
Subset of predictor variables, *Env1*
Subset of predictor variables, *Env2*
Productivity model (PM)
SI = f (*Env1*)
Species distribution model (SDM)
OD probability = f (*Env2*)
Incorporating future climatic variables under climate change scenarios
Set of spatially continuous environmental predictor variables
Habitat fragmentation analysis
Productivity map (current/future)
Suitable habitat map (current/future)
Final productivity map (current/future)

Fig 3.

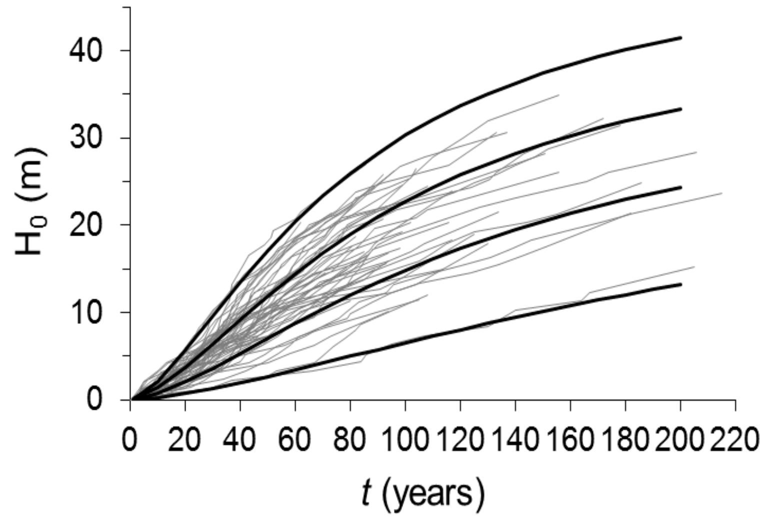


Fig 4.

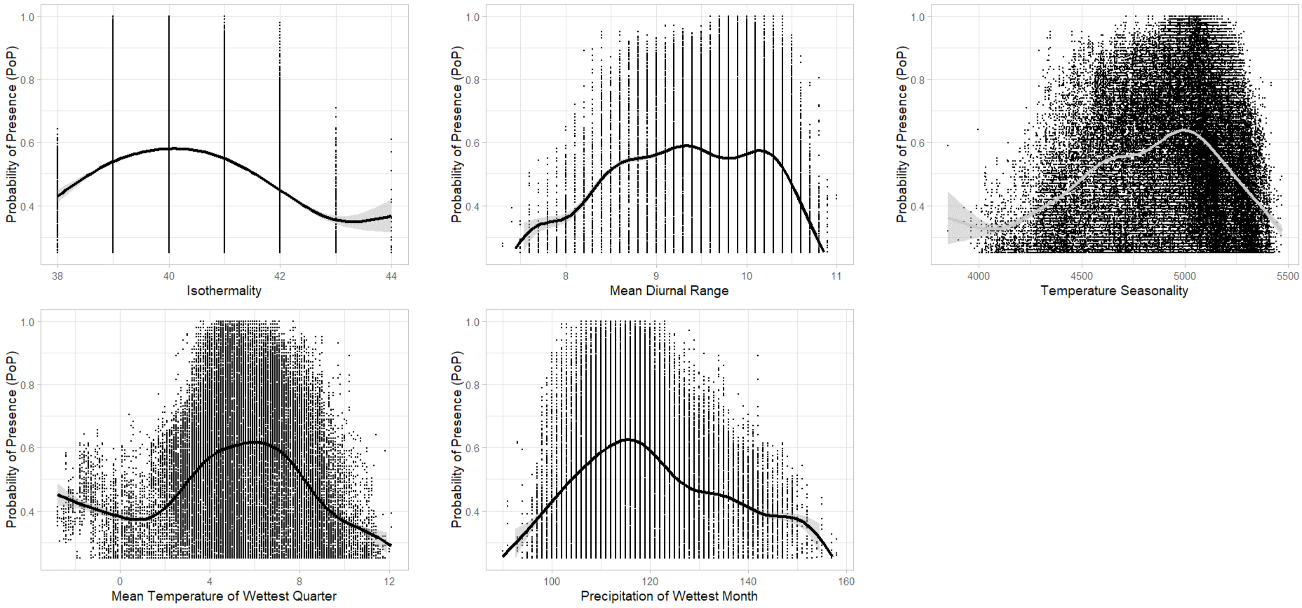

Probability of Presence (PoP)
1.0
0.8
0.6
0.4
38
40
42
44
Isothermality
8
9
10
11
Mean Diurnal Range
4000
4500
5000
5500
Temperature Seasonality
0
4
8
12
Mean Temperature of Wettest Quarter
100
120
140
160
Precipitation of Wettest Month

Fig 5.

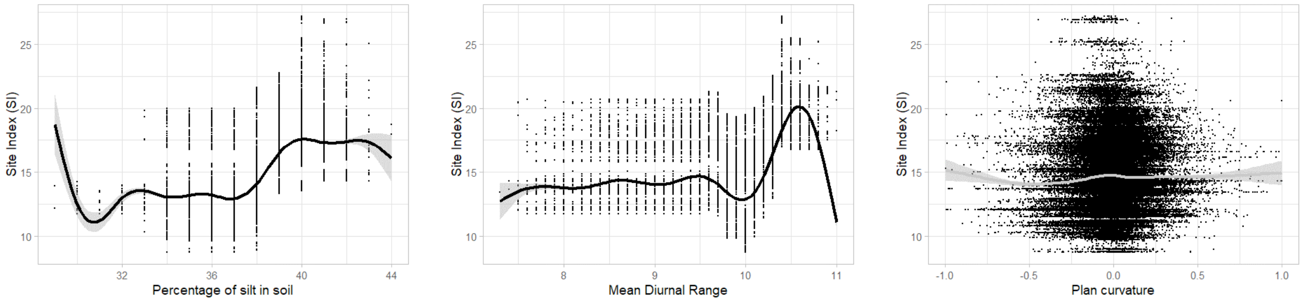

Site Index (SI)
25
20
15
10
32
36
40
44
Percentage of silt in soil
Site Index (SI)
25
20
15
10
8
9
10
11
Mean Diurnal Range
Site Index (SI)
25
20
15
10
-1.0
-0.5
0.0
0.5
1.0
Plan curvature

Fig 6.

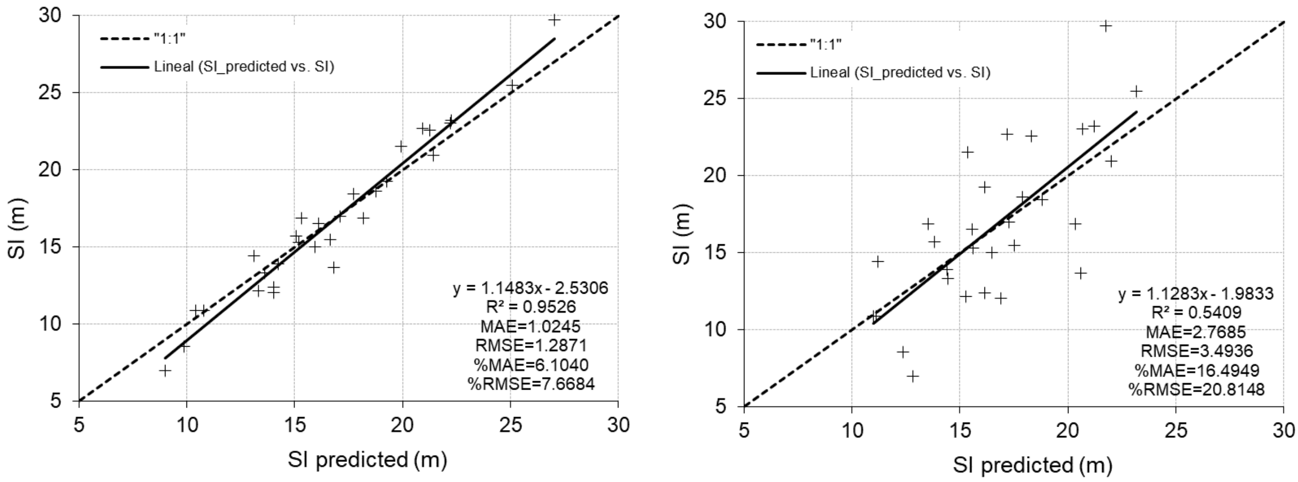
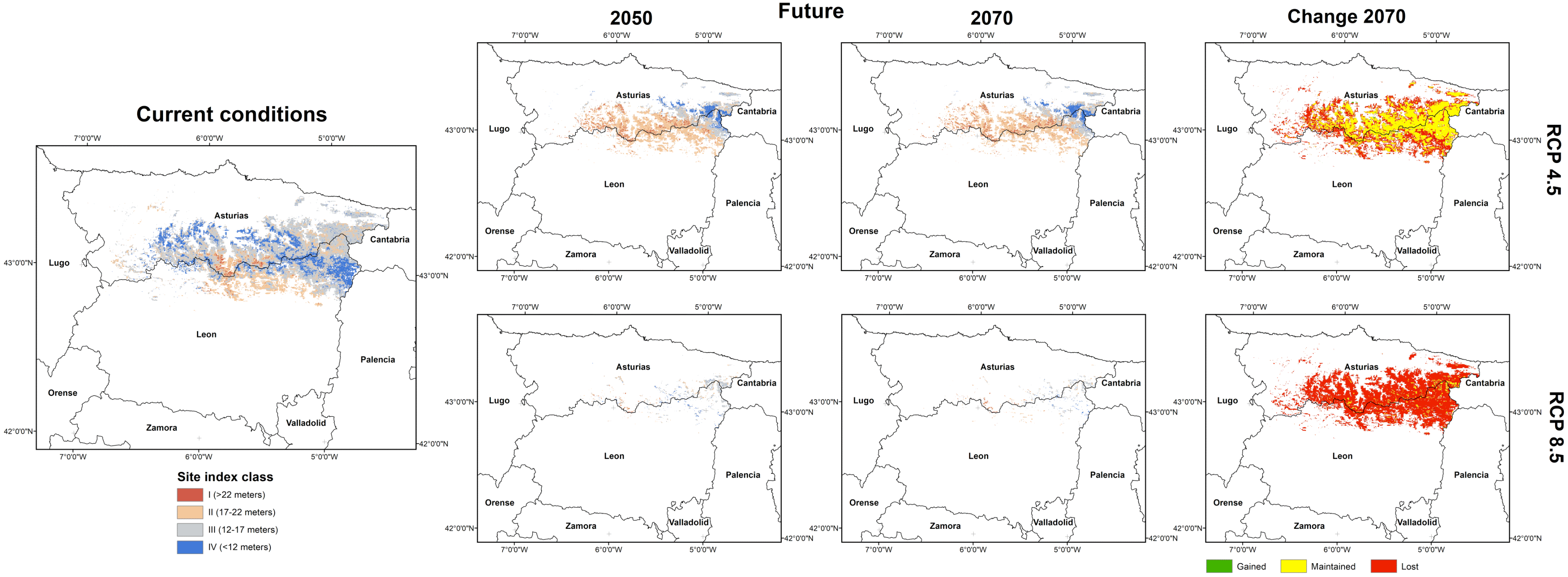
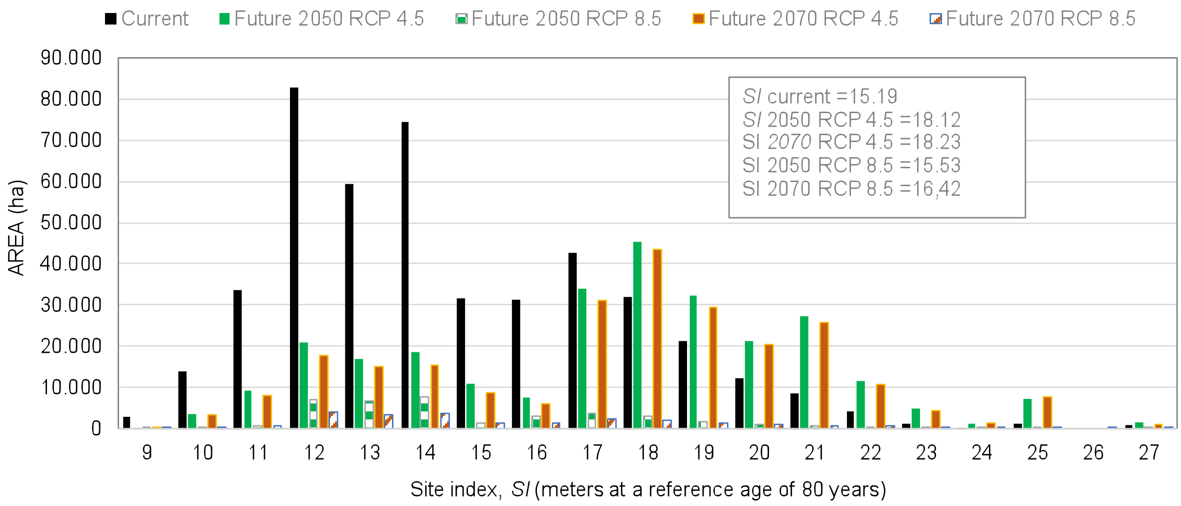

"1:1"
Lineal (SI_predicted vs. SI)
y = 1.1483x - 2.5306
R² = 0.9526
MAE=1.0245
RMSE=1.2871
%MAE=6.1040
%RMSE=7.6684
SI (m)
SI predicted (m)
"1:1"
Lineal (SI_predicted vs. SI)
y = 1.1283x - 1.9833
R² = 0.5409
MAE=2.7685
RMSE=3.4936
%MAE=16.4949
%RMSE=20.8148
SI (m)
SI predicted (m)

Fig 7.



Current conditions
Future
2050
2070
Change 2070
RCP 4.5
RCP 8.5
Asturias
Cantabria
Lugo
Leon
Palencia
Orense
Zamora
Valladolid
7°0'0"W
6°0'0"W
5°0'0"W
43°0'0"N
42°0'0"N
Site index class
I (>22 meters)
II (17-22 meters)
III (12-17 meters)
IV (<12 meters)
Gained
Maintained
Lost

Fig 8.


Current
Future 2050 RCP 4.5
Future 2050 RCP 8.5
Future 2070 RCP 4.5
Future 2070 RCP 8.5
SI current =15.19
SI 2050 RCP 4.5 =18.12
SI 2070 RCP 4.5 =18.23
SI 2050 RCP 8.5 =15.53
SI 2070 RCP 8.5 =16,42
AREA (ha)
90.000
80.000
70.000
60.000
50.000
40.000
30.000
20.000
10.000
0
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
Site index, SI (meters at a reference age of 80 years)

11. Supplementary figure captions

Figure S1. Residuals versus age-lag1-residuals (left column), age-lag2-residuals (middle column) and age-lag3-residuals (right column) for dominant height projection function without considering autocorrelation parameters (first row) and using a second-order autoregressive error structure (second row).

Figure S2. Relative error in stand height prediction (RE) related to the choice of a reference age for dominant height model.

Figure S3. Changes in the distribution (mean latitude and altitude), area (total area) and fragmentation (mean patch area; largest patch index, i.e. the percent of the study area occupied by the single largest patch; and aggregation index, a measure of fragmentation that varies from 0 to 100, with zero reflecting conditions where all suitable grid cells are maximally dispersed from each other across the landscape) of the habitat for the beech in north-western Spain, under five scenarios: (1) the current reference period; (2) 2050 under the RCP 4.5 emissions scenario; (3) 2050 under the RCP 8.5 emissions scenario; (4) 2070 under the RCP 4.5 emissions scenario; and (5) 2070 under the RCP 8.5 emissions scenario.

Figure S4. Distribution of those variables that contributed to the model algorithm for more than 75% for explaining the distribution of beech and also altitude under five scenarios: (1) the current reference period; (2) 2050 under the RCP 4.5 emissions scenario; (3) 2050 under the RCP 8.5 emissions scenario; (4) 2070 under the RCP 4.5 emissions scenario; and (5) 2070 under the RCP 8.5 emissions scenario. The variables shown are the five presenting a relative importance higher 60%.

Figure S5. Comparison between the Cantabrian site index curves (continuous lines) and the Navarrese curves developed by Madrigal et al. (1992) (dashed lines).

12. Supplementary figures

Fig. S1.

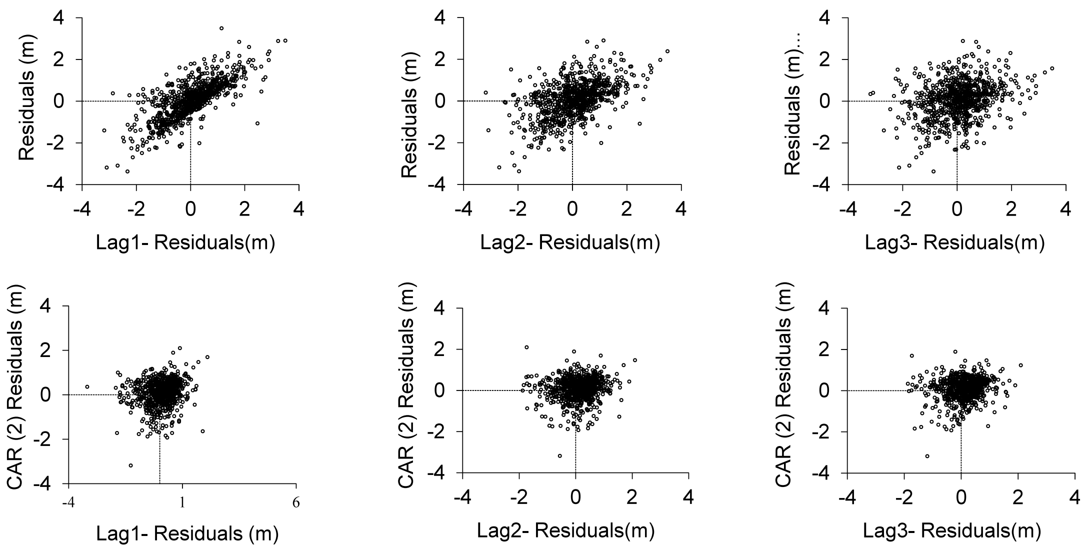
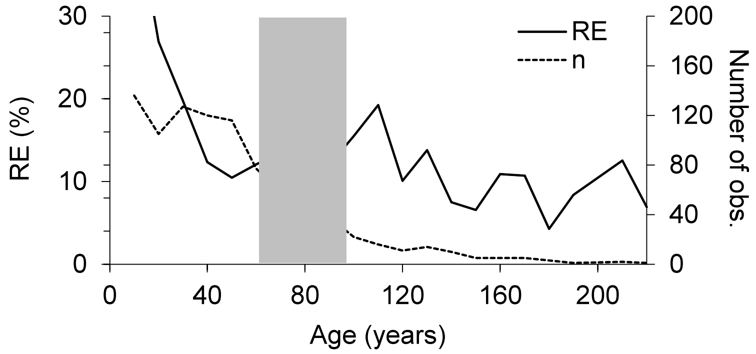


Fig. S2.



RE
n
RE (%)
Number of obs.
Age (years)
0
10
20
30
0
40
80
120
160
200
0
40
80
120
160
200

Fig. S3.

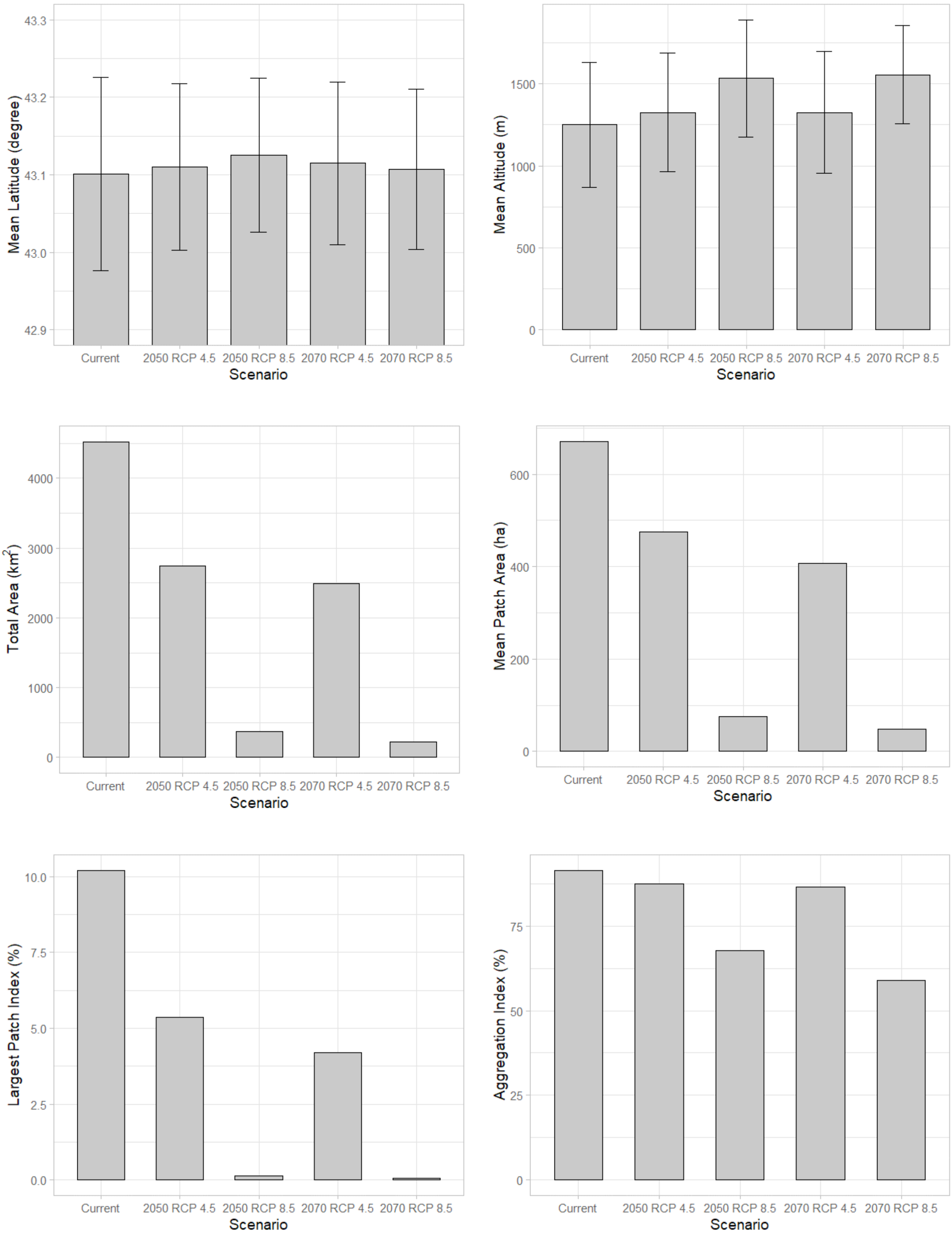

Mean Latitude (degree)
43.3
43.2
43.1
43.0
42.9
Current
2050 RCP 4.5
2050 RCP 8.5
2070 RCP 4.5
2070 RCP 8.5
Scenario
Mean Altitude (m)
1500
1000
500
0
Total Area (km^2)
4000
3000
2000
1000
Mean Patch Area (ha)
600
400
200
Largest Patch Index (%)
10.0
7.5
5.0
2.5
0.0
Aggregation Index (%)
75
50
25

Fig. S4.

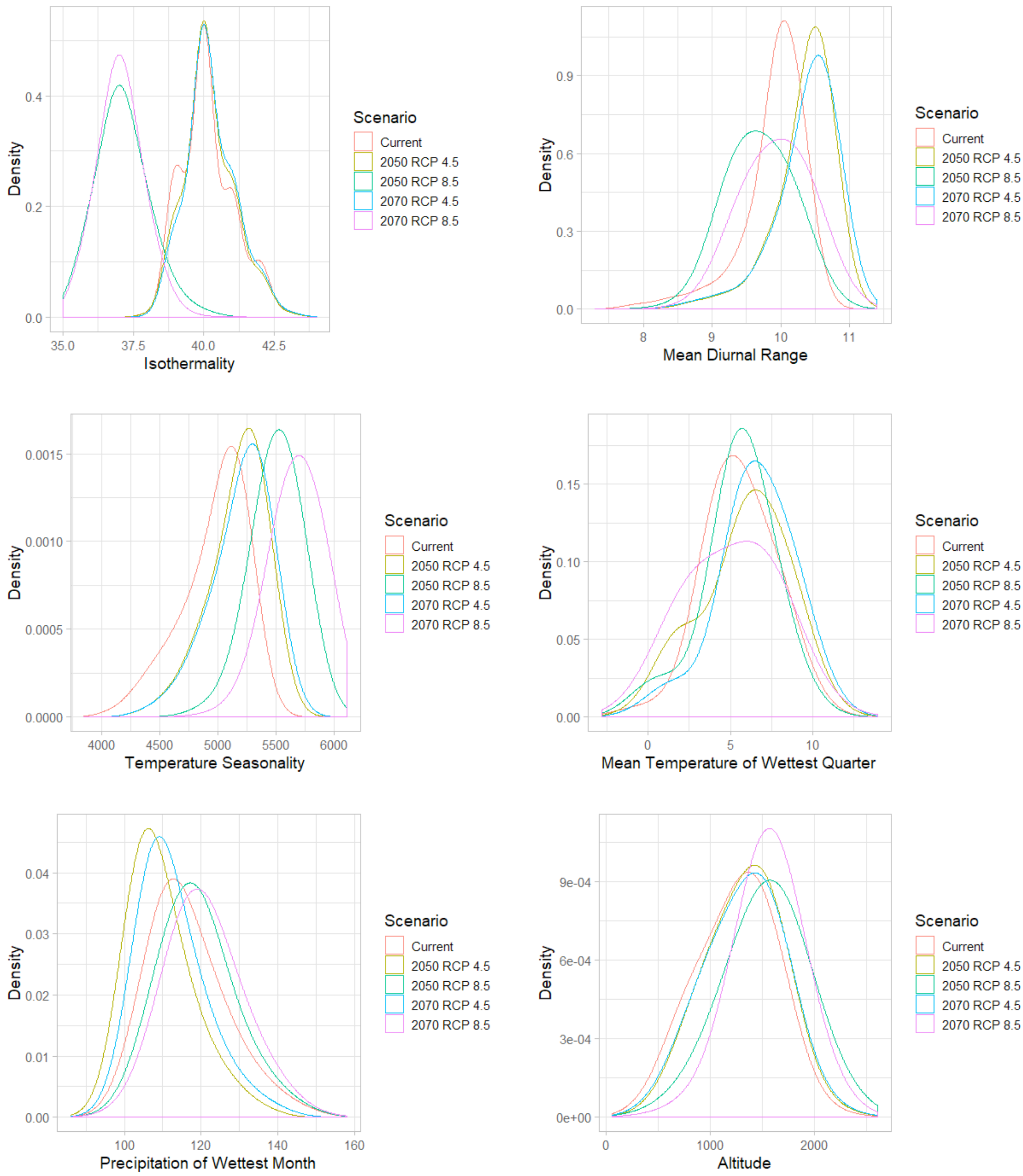

Density
Isothermality
35.0
37.5
40.0
42.5
0.0
0.2
0.4
Scenario
Current
2050 RCP 4.5
2050 RCP 8.5
2070 RCP 4.5
2070 RCP 8.5
Density
Mean Diurnal Range
8
9
10
11
0.0
0.3
0.6
0.9
Density
Temperature Seasonality
4000
4500
5000
5500
6000
0.0000
0.0005
0.0010
0.0015
Density
Mean Temperature of Wettest Quarter
0
5
10
0.00
0.05
0.10
0.15
Density
Precipitation of Wettest Month
100
120
140
160
0.00
0.01
0.02
0.03
0.04
Density
Altitude
0
1000
2000
0e+00
3e-04
6e-04
9e-04

Fig. S5.