

# Species distribution modelling to support forest management. A literature review

Matteo Pecchi<sup>a</sup>, Maurizio Marchi<sup>b,\*</sup>, Vanessa Burton<sup>c</sup>, Francesca Giannetti<sup>a</sup>, Marco Moriondo<sup>d</sup>,  
Iacopo Bernetti<sup>a</sup>, Marco Bindi<sup>a</sup>, Gherardo Chirici<sup>a</sup>

<sup>a</sup> Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Via San Bonaventura 13, I-50145, Florence, Italy

<sup>b</sup> CREA – Research Centre for Forestry and Wood, Viale Santa Margherita 80, I-52100, Arezzo, Italy

<sup>c</sup> Forest Research, Northern Research Station, Roslin, Midlothian, United Kingdom

<sup>d</sup> CNR-IBE, Via G. Caproni 8, I-50145, Firenze, Italy

## ARTICLE INFO

### Keywords:

Forest modeling  
Ecological mathematics  
Climate change scenarios  
Spatial analyses  
Ecology  
Ecosystem services from forests

## ABSTRACT

Species Distribution Modelling (SDM) techniques were originally developed in the mid-1980s. In this century they are gaining increasing attention in the literature and in practical use as a powerful tool to support forest management strategies especially under climate change. In this review paper we consider species occurrence datasets, climatic and soil predictor variables, modelling algorithms, evaluation methods and widely used software for SDM studies. We describe several important and freely available sources for species occurrence and interpolated climatic data. We outline the use of both presence-only and presence/absence modelling algorithms including distance-based algorithms, machine learning algorithms and regression-based models. We conclude that SDM techniques provide a valuable asset for forest managers. However, it is essential to consider uncertainties behind the use of future climate change scenarios.

## 1. Introduction

The future dynamics and spatial distribution of forest ecosystems is a key issue for biodiversity conservation under the many uncertainties generated by climate change (Rehfeldt et al., 2014; Walentowski et al., 2017). Forest ecosystems deliver a wide range of benefits to human beings and achieving the sustainable use of natural resources is central to research in many disciplines. Knowledge concerning the current spatial occurrence of forest species, the influence of ecological drivers (e.g. climate, soil) and the possible erosion or expansion of their envelopes of suitability is required by decision makers in order to detect both threatened areas and possible refuges (Pecchi et al., 2019; Williams and Dumroese, 2013). The use of better adapted forest tree species (genotypes) and provenance selection (genotyping) has the potential to improve the resilience of forest systems and allow assisted migration strategies (Hanewinkel et al., 2014; Marchi and Ducci, 2018), thus assisting the adaptive processes of forest ecosystems (Ferrarini et al., 2016).

Since the emergence of modelling techniques, spatial data including aerial images, cartographic layers, and national forest inventories have been fundamental resources for statistical mapping (Di Biase et al., 2018; Fleischer et al., 2017; Mura et al., 2016; Spittlehouse and

Stewart, 2004). Reliable datasets and statistical models quickly became integral to supporting decisions which aim to support sustainable use of forest resources under a changing climate. Numerous datasets of forest attributes and land suitability surfaces have been developed for many forest tree species in many areas of the world, and these are integral to developing spatial decision support systems (Johnson et al., 2014; Masek et al., 2013). Statistical modelling techniques can be divided into one of three types: 1) empirical, 2) correlative, or 3) mechanistic. Correlative Species Distribution Models (SDM) involve the collation of species occurrence data, relating these occurrences to environmental variables, and generating maps which predict past, present or future species distributions. Their ease of use makes them a popular method, and they represent the vast majority (around 90%) of SDM publications. An alternative approach is to use mechanistic SDMs, which simulate biological processes according to ecological drivers. These mechanistic models rely on huge datasets with long time-series and high resolution data, which are often not available at national or continental scales. Given that statistical SDMs have emerged as methods by which these limitations can be overcome, mechanistic models are not considered in detail by this review.

Before the mid-1980s SDM attempts were limited by the lack of reliable interpolated climatic data on large spatial scales, i.e. to estimate

\* Corresponding author.

E-mail address: [maurizio.marchi85@gmail.com](mailto:maurizio.marchi85@gmail.com) (M. Marchi).

<https://doi.org/10.1016/j.ecolmodel.2019.108817>

Received 22 May 2019; Received in revised form 19 August 2019; Accepted 1 September 2019

Available online 16 September 2019

0304-3800/ © 2019 Elsevier B.V. All rights reserved.

conditions at species occurrence sites that are often distant from meteorological stations. Modern SDM took off in the late 1990s and early 2000s with the release of global climatic surfaces such as the WorldClim database (Hijmans et al., 2005), which has provided the data used in many SDM studies to date. In 1996 BIOCLIM provided a set of 19 bioclimatic variables which are also still widely used in many SDM studies. The SDM approach was initially based on the ecological niche concept provided by Hutchinson around 1950s and then refined by Booth et al. (1988). This envisaged an 'n-dimensional hypervolume' (which included simple ranges for environmental factors such as precipitation or temperature) describing where the species grows naturally (*i.e.* its realized niche) or where it can grow and reproduce in the absence of competitors (*i.e.* its fundamental niche). Appreciating that many tree species can grow under conditions somewhat different from those within their natural distributions is crucial for understanding how they may respond to climate change. It is reasonable to assume that a long-lived tree species already well-established at particular sites may well be able to display some of the climatic adaptability it has shown at trials outside its natural distribution (Booth, 2017). Most SDM studies of forests under climate change ignore this adaptability, thus determining species climatic requirements from their natural distributions only and applying climate change scenarios. This distinction between fundamental (or potential, *i.e.* Grinnellian) and realized (*i.e.* Hutchinsonian) niche (Pearson and Dawson, 2003; Pulliam, 2000; Vetaas, 2002) has been often discussed. While the fundamental niche represents the entire habitat suitable for a considered species, the realized niche is defined as a smaller part of the expressed fundamental niche as the result of the inter and intra-specific competition for available resources in a specific environment *i.e.* geographic zone (Booth, 2017). The recognition of this distinction is critical for deciding how an SDM should function (Pearson and Dawson, 2003). The majority of SDM to date make use of 'realized niches', often deriving these from the current spatial distributions of forest species. This has limitations, as SDM are not able to consider the relationships between species and other biotic components *e.g.* pests and diseases (Austin, 2007; Morin and Thuiller, 2009).

SDM is recognized as a powerful method to forecast the most likely impact of a changing climate on the geographic distribution of a target species by means of environmental data and future Global Climate Model (GCM) outputs (Booth, 2018; Guisan et al., 2013; Thuiller et al., 2015). The technique can go by many different names including: habitat model, niche-based model, habitat suitability model, climate envelope, environmental niche model (ENM) and ecological niche model (Elith and Leathwick, 2009; Guisan et al., 2017, 2013; Hamann and Wang, 2006; Jeschke and Strayer, 2008). The use of ENM as synonymous with SDM is contentious given the confusion about the significance of the term "niche" (Mcinerney and Etienne, 2013, 2012; Peterson and Soberón, 2012; Warren, 2012). For some authors, the inability of models based on realized niches to consider biotic interactions means that the 'true' niche of the species cannot be modelled. Consequently, the current spatial distribution of forest species is viewed as inadequate to properly characterize its ecological requirements.

This paper reviews available species occurrence datasets, environmental data, modelling algorithms, evaluation processes and spatial projections, discussing the implications of the findings for forest science, silviculture and forestry. The aim is to describe how ecological modelling of forest tree species has evolved within the framework of spatial ecology to support forest management. A bibliographic search was conducted by analyzing Scopus, Google Scholar and ISI-WoS databases, for the period 2000–2019. The aim of this review is to give a general overview of the techniques and adjustments implemented by researchers in order to improve future applications of SDM in forestry research. The unresolved issues highlighted previously, including the discussion around the definition of real or potential SDM and awareness of the theoretical differences between SDM and ENM, are not discussed further. The term SDM will be used through the text to include all

methods intended to link the spatial distribution of target tree species with environmental variables.

## 2. Species distribution datasets

The spatial distribution of a target species (species occurrence) as the result of past history, current (long-term) climatic conditions and, above all, forest management strategies, represents the primary basis for any SDM (Falk and Mellert, 2011; Godsoe et al., 2017). There are many possible sources for species distribution records. Selecting the most appropriate dataset is challenging and can influence model performance (Duputié et al., 2014). While many studies have focused their models on the native range only (Gastón et al., 2014; Isaac-Renton et al., 2014), other authors have included the distribution of artificial stands, under the assumption that "if it survives, there it is suitable and worth to be considered" (Duveneck and Scheller, 2015; Marchi et al., 2016). In fact, several studies have demonstrated the ability of many forest tree species to grow well outside their native range and often better than in their origin area (Boiffin et al., 2017; Booth, 2017; Castaldi et al., 2017). Regardless of database source, there are common issues that affect the reliability of datasets. These include: uncertainty in species identification, low or unknown accuracy of sample locations, lack of design sampling, and incomplete spatial coverage of the true distribution of species (Guisan et al., 2017).

### 2.1. Vector format (shapefile) and national forest inventories

Species occurrence data for SDM can be obtained from field surveys (*e.g.* National Forest Inventories - NFI), compiled on the basis of existing literature (*e.g.* EUFORGEN maps), or derived from statistical modeling procedures (*e.g.* EFI maps, Brus et al., 2012). While presence data are easy to obtain, including absences (or pseudo-absences) is a major issue in SDM. The spatial distribution is rarely in a "true" equilibrium with climate/soil due to human pressure on the environment. For this reason, many uncertainties lie behind both presence and absence data. Many additional modelling tools have been proposed to properly simulate absences (Barbet-Massin et al., 2012; Peterson et al., 2011). However only NFI datasets have the advantage of being based on a statistical sampling scheme with additional information on absences (Marchi and Ducci, 2018).

Most of the databases used by authors are open-source and freely available on the web. The EUFORGEN maps (<http://www.euforgen.org/species/>) are the first example of distribution maps. This database was created by "Forest Genetic Resource Program" and consists of a series of pan-European distribution maps for 45 different species, which are updated continuously. The latest version is available in Caudullo et al. (2017) and was realized in the framework of the European Atlas of Forest species (San-Miguel-Ayán et al., 2016). The main shortcoming of this database relates to its polygon format, which can affect the quality and reliability of the data given that no information is available within each polygon, considering all locations as potentially suitable at the same level. Such data has been generally used to validate SDM outputs or to constrain the analysis within a native range. For instance, Bede-Fazekas and Levente Horváth (2014) used it to investigate the potential distribution of 4 different species of Mediterranean pine. Another study proposed by Falk and Hempelmann (2013) was explored the distribution and shift of beech and spruce in Europe.

ICP-Forest is another freely available European database but, in contrast to EUFORGEN which can be freely downloaded from the website, ICP-forest download requires a formal request. ICP is the result of the "International Cooperative Program on Assessment and Monitoring of Air Pollution Effects on Forests" project. The data are split into two different monitoring intensities and spatial distributions: "Level I" and "Level II". In the first case, almost 6000 forest monitoring plots are available, regularly distributed on a 16 km grid (Hanewinkel et al., 2012). Hanewinkel et al. (2012) use ICP forest to examine the

distribution of important forest species in Europe, while Casalegno et al. (2010) used ICP to realize a map of vulnerability of *Pinus cembra* to present and future climate conditions. Level II is an intensive monitoring network and provides keys insights into factors affecting the condition of forest ecosystems and relative effects of different stress factors. In this case only around 800 plots have been established within the major forest types of Europe. While ICP-Forests might be used as presence-absence dataset, a lack of a spatial sampling scheme a major shortcoming of this dataset. For this reason, an adjustment was proposed by Brus et al. (2012) where a mixture of compositional kriging in areas with NFI plot data and a multinomial multiple logistic regression model between ICP-Forests plots was performed.

Although often initially conducted for different purposes e.g. to record management actions or inform economic research, National Forest Inventory (NFI) data are probably the most important and detailed source of biological data for SDM. The main superiority of NFI for SDM relies on the sampling method used which respects statistical rules and can be used for inference. NFI are an unbiased sampling of the forest area in a specific country and can be used to derive estimators of forest attributes. Their repetition at (almost) regular intervals of time enables monitoring and modelling of temporal changes (Teuscher et al., 2013). Hanewinkel et al. (2010) modeled the potential economic consequences of a shift from Norway spruce to European beech in a forest in the south of Germany. Similarly, Rivera and López-Quilez (2017) worked in Spain to compare several statistical techniques for predicting species distribution of forest species. Iverson et al. (2008a, 2008b) evaluated a potential response of forest species to climate change following two different emission scenarios in eastern of USA using inventory data for 134 different species of forest tree. Recently NFIs have been used to study the possibility for detecting and conserving marginal and peripheral forest populations with several SDM techniques in order to forecast possible adaptation strategies for two Mediterranean species (*Abies alba* and *Fagus sylvatica*) sharing a common environment (Marchi and Ducci, 2018). Based on the classic SDM evaluation method (True Skill Statistic – TSS, see below), a higher accuracy of predictions was obtained modeling only a small part of the whole distribution, referred as “provenance” and mainly due to the reduction in the “background noise”. Consequently, authors concluded that the “Provenance Species Distribution Modeling” may represent a valuable step forward in spatial analysis, particularly for the detection of marginal peripheral populations.

## 2.2. Global datasets and raster layers

The “Global Biodiversity Information Facility” (GBIF) database is an important source of auxiliary information for SDM. The database gives information on occurrence data outside the native range and therefore an indication of the ability of species to grow under different climate conditions (Booth, 2014; Dyderski et al., 2018). As with ICP-Forests, its main problem is the lack of sample design (Guisan et al., 2017); and thus it is rarely use as a principal source of data. However, there are several examples of it being utilized. Hernández-Quiroz et al. (2018) related *Quercus* occurrence data to its current distribution using GBIF data. In Booth (2014) the GBIF is related to the Atlas of Living Australia (ALA) to develop a methodology which describes the climatic requirements of *Eucalyptus nitens*. Dyderski et al. (2018) also worked with GBIF as additional information to improve the predicted performance of their SDM. In another study, Zhang et al. (2017) integrated data from the FIA (United States Forest Inventory and Analysis) and PSP (Canadian Permanent Sampling Plots) with GBIF to develop an ensemble SDM which evaluated potential habitat suitability for forest species under different conditions of climate, land use and dispersal constraints.

Local, regional or national datasets such as forest category maps represent valuable strata for regional studies which can enrich national, continental or global datasets. For example, forest ecotypes link plant

species to certain zones. Marchi et al. (2016) used ecotypes to model future scenarios for a marginal forest population of Black pine (*Pinus nigra* spp. *nigra* var. *italica*) in the Mediterranean area in order to forecast potential mitigation strategies and propose an assisted migration protocol. Similarly, in Iturbide et al. (2015), 11 different ecotypes of *Quercus* spp. were used to analyze the effect of different methods for pseudo-absence data generation and generate optimal results. National forest maps have also sometimes been utilized. In Garzón et al. (2006) a distribution map at a resolution of 1 km was used to assess the potential distribution of *Pinus sylvestris* in Spain. Similarly Wang et al. (2016) studied likely climate change effects on the distribution of some common Chinese tree species.

Statistical maps of forest tree species represent an alternative and interesting source for SDM. Brus et al. (2012) recently generated raster maps at 1 km resolution for 20 forest tree species in Europe. Input data was sourced from ICP-Forests records and NFI inventory statistics for 18 European countries. For areas covered by National Inventory plots, the proportional area of each of the 20 species was calculated using a kriging interpolation method. A multinomial regression model was then applied to predict species composition for the rest of Europe. The results of the model were then scaled using independent data. These datasets have been widely used. Van der Maaten et al. (2017) investigated the hypothesis that future climate projections are linked to temporal or spatial variation in forest growth (ring width). These datasets were also used by maps were also used by Noce et al. (2017) to study the potential effect of climate change on hot-spot distribution in southern Europe with regards to common group of forest species.

## 3. Climate, soil, land cover and variable choice

Many factors need to be taken into account when choosing an appropriate predictor variable: the purpose of the study, the availability of data, and the redundancies between variables. An additional issue in variable selection is ‘collinearity’ which can occur between predictors. This phenomenon occurs where two or more predictors are related to one another, linearly or not. This can affect the proportion of variance explained by each independent variable. While no impact has been found on final prediction, this characteristic can make it difficult to establish the relative importance of predictors in affecting the distribution of species (Dormann et al., 2013). Where there is high collinearity among predictors, the easiest method is to remove some of the highly correlated variables from the computational steps (Schröder, 2008). A pre-determined threshold which generally ranges between 0.8 and 1.0 (Dormann et al., 2013) can be applied to filter them. To avoid subjective selection, the use of a pre-selective technique such as Principal Component Analysis (PCA) can be used. Variables are then chosen according to the proportion of variance explained by each component (Cruz-Cárdenas et al., 2014; Metzger et al., 2013). Alternatively, PCA components can be used as predictors (Marchi and Ducci, 2018). However, even with a mathematically perfect model (the components are orthogonal for construction) the importance of each ecological predictor is hard to estimate. Another option is the Cluster independent method with two sub-variants: i) to select variables where the correlation values are under 0.7 or ii) residual regression. Finally, latent variable models can be used to infer ‘hidden’ variables from observed and collinear variables (Dormann et al., 2013).

### 3.1. Worldclim and Worldclim-based raster surfaces

The environmental variables (or predictors) in SDM are used to derive the ecological niche which can then be used to model species distributions according to their drivers (Pearson, 2010). The choice of predictor variables represents a critical step which must be based on the ecological tolerance and habitat requirements of the species in question (Jamevich et al., 2015). Despite its important repercussions on model performance, the “predictors issue” has only recently begun to acquire

greater attention (Barbet-Massin and Jetz, 2014). Climate variables (e.g. temperature and precipitation but also derived climatic indexes) are generally the most used variables in SDM (Thuiller, 2013). At current time, many different sources of climate data are freely available on the web (Barbet-Massin and Jetz, 2014) Worldclim (Fick and Hijmans, 2017) being the most common. Worldclim collects monthly climate data at several resolutions ranging from 10 min (about 300 square km) to 30 arc-seconds (about 1 square km). After initial release (1960–1990, Version 1.4) where only temperature (maximum, minimum and average), precipitation and 19 bioclimatic indices were available, new indices including solar radiation, vapor pressure and wind speed have been added, and all other indices extended for the period 1970–2000 (Version 2.0) (Fick and Hijmans, 2017). Worldclim implementation in SDM is very common in the scientific literature. For example, Casalegno et al. (2010) build an SDM describing the vulnerability of *Pinus cembra* to climate change. Märkel and Dolos (2017) used the 19 bioclimatic variables from Worldclim to calculate starting from climate data derived from German climate service. The final aim is to build a methodology to combine them in unique technique to better evaluate the climate change impact in Germany. However, two main shortcomings arise: i) the global extension makes the dataset often unsuitable for local application and ii) the lack of an adequate coverage in some regions of the globe make some climatic variables unreliable, especially precipitation data (Bedia et al., 2013; Marchi et al., 2019).

Given these limitations, an alternative or complementary sources to Worldclim is the recently released ENVIREM dataset (Environmental Raster for Ecological Modeling). It combines a set of biological and topographic variables calculated from WorldClim rasters and solar radiation (Title and Bemmels, 2018).

### 3.2. Standalone software for climatic custom queries

The use of standalone software for local downscaling of climatic data is gaining attention in the scientific literature. Among these, ClimateEU, ClimateNA, and ClimateSA software packages are valuable tools (<https://sites.ualberta.ca/~ahamann/data.html>) with which to generate customized raster maps for Europe, North America and South America respectively, for both historical time slices or future scenarios. These packages provide a method to downscale PRISM data using a combination of bilinear interpolation and “dynamic lapse rate” adjustment. The PRISM (Parameter-elevation Regression on Independent Slopes Model) climate database, available for United States only, provides mean monthly precipitation and minimum and maximum temperature values for the period 1971–2000. These data were interpolated to produce climate variables for the entire United States using a DEM available at 30 arcsec of spatial resolution (Daly et al., 2008). The software calculates monthly, seasonal and annual climate variables for a specified location on the base of latitude, longitude and elevation (Wang et al., 2012). Hamann et al. (2013) and Van der Maaten (2017) used this database to build a SDM for distribution of European tree forest species including Norway spruce, Scots pine, European beech and Pedunculate oak. Isaac-Renton et al. (2014) also used ClimateNA and ClimateEU to build a SDM for Douglas fir and evaluate species transferability between continents in view of climate change. Concerning Europe, ClimateEU was used by Marchi and Ducci (2018) to generate raster surfaces at 250 m of spatial resolution in combination with NFI data.

Local climate datasets are undoubtedly the most accurate source of information for SDM, but they are also the least accessible. Local data is used widely in forest management and especially in forest monitoring due to its higher precision and quality (Ferrara et al., 2017). Bedia et al. (2013) compared the use of national climate data to data from Worldclim in order to verify the differing sensitivity in SDMs using distribution data for *Fagus sylvatica* in smaller region. Local climate databases were also used in Crimmins et al. (2013). Four predictor variables (minimum and maximum temperature, mean annual actual

evapotranspiration and mean annual climatic water deficit) hypothesized to have a direct influence on species distribution were tested in order to evaluate the success of a consensus approach for predicting the distribution of plant species.

### 3.3. Soil and land cover datasets

In addition to climate data, soil and land cover information can be included in SDM approaches. The principal characteristics of interest with regards to soil are PH, texture, fertility and soil moisture. Rocés-Diàz et al. (2014) used NFI data for Spain together with 9 different predictors (among them an index of fertility of soils) to realize a distribution model which investigates the relationship between climate and the distribution of 6 different tree species. Coudun et al. (2006) used many different predictor variables to realize a SDM to investigate the importance of climate and other edaphic variables to distribution of *Acer campestre*. pH was a key variable together with other 6 other variables linked to aspects of soil. Despite soil being a critical predictor variable, its use in SDM is limited as the lack of available datasets in either digital format or suitable resolution makes it hard to find data of comparable quality to climate data (Mod et al., 2016; Thuiller, 2013). The most relevant and up-to-date information with global coverage is SoilGrids250 m (Hengl et al., 2017). This dataset provides global predictions for standard numeric soil properties (organic carbon, bulk density, Cation Exchange Capacity (CEC), pH, soil texture fractions and coarse fragments) at seven standard depths (0, 5, 15, 30, 60, 100 and 200 cm), in addition to predictions of depth to bedrock and distribution of soil classes based on the World Reference Base (WRB) and USDA classification systems. The interpolated surfaces were built using more than 150,000 soil profiles and 158 remote sensing-based soil covariates (primarily derived from MODIS land products, SRTM DEM derivatives) used as predictors in a combination random forest model and multinomial logistic regression algorithm. In Manchego et al. (2017) this database was used to evaluate the potential effects of deforestation and climate change on the distribution of 17 characteristic forest tree species of dry forest in Ecuador.

In terms of land cover, land cover change (for example agricultural intensification) can have important effect on the distribution of different organisms. By including land cover data in SDM the explanatory power of the model is often increased, while the predictive performance remain unchanged (Thuiller et al., 2004). Despite this, use of land cover data has not been common in plant distribution studies, except when predicting species abundance (Bradley et al., 2012; Mod et al., 2016). For example Hill et al. (2017) predicted the future abundance of typical forest trees in UK by means of a land cover change map. In addition to land use and land cover maps, Digital Elevation Models (DEM) or Digital Terrain Models (DTM) and mathematically derived maps (i.e. slope, aspect, topographic position index, etc.) are sometimes included. When choosing an appropriate DEM, spatial resolution and any uncertainties associated with interpolation are principal factors to consider (Franklin, 2010). Garzón et al., 2006 used aspect and slope as predictors, while Duan et al. (2014) included altitude among several other predictive variables. However, generally the use of terrain models has often been neglected, given that the information they provide is already included in climatic maps and no climate change effects can be added to such predictors.

## 4. Modelling algorithms

Many different algorithms are currently implemented in SDM processes (Guisan et al., 2017) and 12 modeling methods have been described and selected according to their use in the analyzed literature (Fig. 1). There is consensus that a single and perfect technique for all possible SDM cases is impossible, and thus the selection of the appropriate modeling algorithm is fundamental. As demonstrated by many studies (Beale and Lennon, 2012; Buisson et al., 2010; Duan et al.,



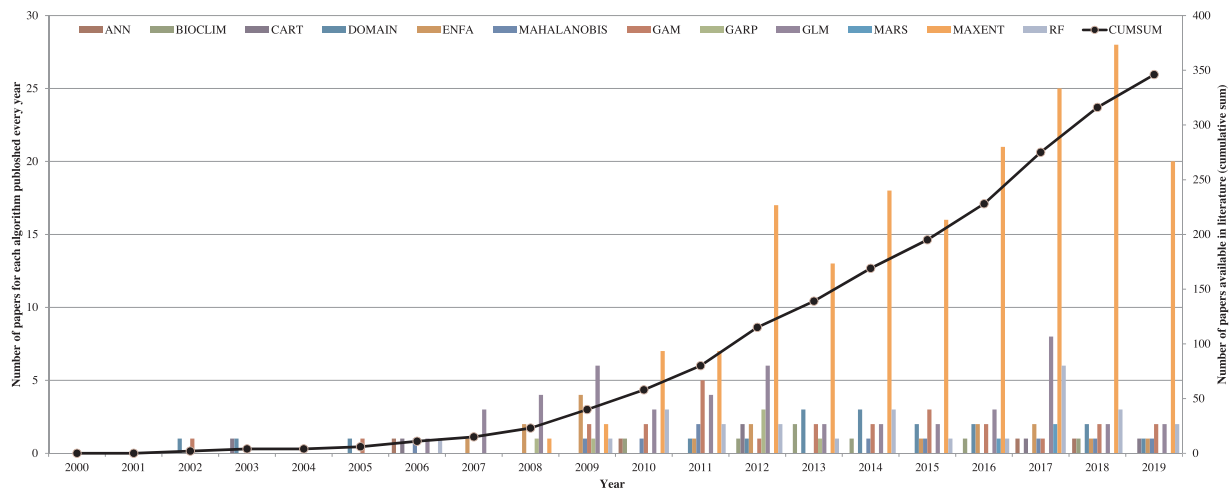


Fig. 1. Number of Scopus papers published between 2000 and the present time dealing with the use of SDM in forestry and grouped according to the used modelling algorithm(s).

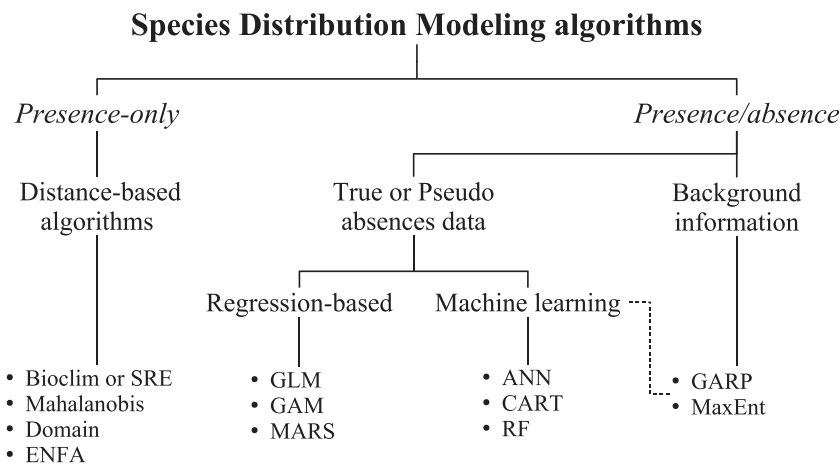


Fig. 2. A hierarchical structure of the 12 SDM algorithms used by the literature analysed by this review.

2014; Jarnevich et al., 2015; Koo et al., 2017) choosing the correct algorithm can reduce the uncertainty within the model. The algorithms involved in SDM computational steps can be divided into different groups and according to various characteristics and grouping criteria. In this review paper the analyzed algorithms were firstly divided according to input data: i) presence only and ii) presence/absence algorithms. Then subgroups were made based on intrinsic characteristics e.g. distance based, linear (or regression model), classification and decision trees, machine learning. A graphical scheme of the proposed structure is shown in Fig. 2.

#### 4.1. Presence-only algorithms

All techniques included within this group are characterized by their ability to model the spatial distribution of a target species simply on the basis of species occurrence (or presence). Very simple and computationally light, these are the oldest technique used in ecological modeling but also acknowledged as being less powerful and often unsuited for predicting the effects of climate change (Guisan et al., 2017; Hijmans and Graham, 2006; Miller, 2010). Algorithms included in this group and reported here are: Bioclim or Surface Range Envelope (SRE), the Mahalanobis Distance, the Domain algorithm and the Environmental Niche Factor Analysis (ENFA).

The Bioclim or Surface Range Envelope (SRE) algorithm has been extensively used for SDM and represents the classic 'climate-envelope-model' (Booth, 2014; Hijmans and Elith, 2011). Although it generally

does not perform as well as some other modelling methods (Elith et al., 2006) it is still used as it is easy to understand and thus useful in teaching SDM. This method computes the similarity of a location by comparing the values of environmental variables at any location to a percentile distribution of the values at known locations of occurrence. The closer to the 50th percentile (the median), the more suitable the location is. A key shortcoming is that the tails of the distribution are not distinguished and the 10<sup>th</sup> percentile is treated as equivalent to the 90<sup>th</sup> percentile. Bioclim has been used less by recent literature and where it does occur, this is often in comparative papers. For example, Duan et al. (2014) evaluated the predictive capacity and solidity of different techniques for estimating species distribution for many forest tree species (*Pinus massoniana*, *Betula platyphylla*, *Quercus wutaishanica*, *Quercus mongolica* and *Quercus variabilis*). An interesting use of such algorithm is as ancillary model to generate pseudo-absences in biomod2 package (Thuiller et al., 2009). In this package, Bioclim (called SRE) can be used to generate a user-defined number of pseudo absences in a target spatial extent, laying outside the ecological distribution of the species to be modeled and described by the occurrences.

We found the Mahalanobis distance to be the most popular algorithm among the distance-based methods. This method calculates the suitability area as a multivariate and environmental distance between the study area and a vector of optimum climate condition, generally calculated as a mean of all values which occur in a presence dataset (Farber and Kadmon, 2003; Peterson et al., 2011). The predictive power of Mahalanobis is higher than Bioclim (Farber and Kadmon, 2003) but

some disadvantages still occur. Franklin (2010) highlighted the inability to weight the relative influence of different predictor variables.

Similarly to Mahalanobis, the Domain algorithm (Carpenter et al., 1993) computes environmental distance. In this case the ‘Gower’ distance is used, which calculates the distance between environmental variables at any location and those at any of the known locations of occurrence (training sites). For each variable the minimum distance between a site and any of the training points is taken. To integrate findings across environmental variables, the maximum distance to any one of the variables is used and this distance is subtracted from one. Environmental Niche Factor Analysis (ENFA) is the last algorithm in the family of presence only and distance-based methods. This algorithm is able to estimate the ecological niche through a comparison of presence data and environmental values for the entire area (Guisan et al., 2017; Hirzel et al., 2002). Rupprecht et al. (2011) used this method to evaluate the prediction capacity of different type of algorithms for a distribution of *Juniperus oxycedrus* species.

#### 4.2. Presence-absence algorithms

This second group is populated by more complex and time-consuming but also more complete algorithms. This is due to both the higher amount of information they can handle and the inclusion of absences in the modeling steps. Indeed, absence data are often numerically more common than presence data (sometimes even ten times more). For this reason, such models need to handle this problem properly, weighting the sum of presences and absences equally. This statistical method is generally called “prevalence” (Barbet-Massin et al., 2012; Manel et al., 2001; Marchi and Ducci, 2018). According to the statistical family, two different sub-groups can be defined: regression based and machine learning. The first group includes parametric models such as Generalized Linear Models (GLM), Generalized Additive Models (GAM) and Multivariate Adaptive Regression Splines (MARS). Among the immense literature on machine learning algorithms, Artificial Neural Network (ANN), Classification Trees (CART), Maximum Entropy (MaxEnt), Genetic Algorithm (GARP), and Random Forest (RF) are herewith discussed as the most used nonparametric algorithms.

GLM represent one of the principal algorithms in SDM, as a flexible and relatively simple tool derived from linear model (Guisan et al., 2002). The main characteristic of GLM which distinguishes it from a general linear model is the possibility to include a response variable with a different distribution family from the Gaussian, for example Binomial or Poisson (Guisan et al., 2017). This algorithm is particularly useful for non-normal distribution data (Bolker et al., 2009). Its use in the case of SDM is allowed by means of the specification of a binomial family and a logistic link function. The GLM algorithm is extensively used in the literature. Higa et al. (2013) assessed the importance of non-climatic factors on the provision of the potential habitat for typical Japanese tree species: *Fagus crenata*, *Betula grossa*, *Carpinus laxiflora*, *Carpinus tschonoskii*, *Celtis sinensis*, *Ulmus laciniata* and *Zelkova serrata*. Thuiller et al. (2006) evaluated the potential change in distribution of 112 tree species following climate change in Europe. In Rocas-Díaz et al. (2014) GLM clearly revealed both the difference in habitat suitability among different tree species in Spain and the importance of predictor variables, in particular minimum temperature and soil fertility. Finally Thuiller et al. (2009b) and Eith et al. (2006) compared the predictive accuracy of GLM with GAM and CTA using three independent datasets of tree species at different scales and resolution. Results showed that the predictive performance of GLM were superior to other algorithms, especially at finer scales.

An “extension” of GLM is represented by the mixed-effect Generalized Linear Mixed Model (GLMM). GLMM is currently not commonly used in the forestry sector and has only recently been applied to build a novel SDM approach (Benito Garzón et al., 2019) where forest tree species were modelled according to the performances obtained in common garden experiments. The main novelty of GLMM rely

on the use of the common garden as random effect predictor which allows to handle the differences within sites in a model, “cleaning” the prediction from artifacts and unexplained differences. As an example from another field, GLMM has been used to quantify the effect of imperfect detection on the estimation of niche overlap between two forest dormice (*Muscardinus avellanarius* and *Glis glis*) by Panizza et al. (2018)

GAM represents a natural expansion of the GLM algorithm (Guisan et al., 2002) with its principal feature being high flexibility. This aspect allows use of this algorithm to represent situations where there are non-linear combinations between variables (Elith et al., 2006). GAM adopts a particular smoothing function to fit a non-linear relation between predictive variables and species occurrence. The GAM algorithm is used by Keenan et al. (2011) to compare the future distribution of different forest species with an output from mechanistic processes. Walentoski et al. (2017) used GAM to evaluate the suitability of a Franconian Plateau in the south of Germany for three different species in the context of climate change. In Rivera and López-Quílez (2017) compared algorithms including GAM, CART, MARS and MaxEnt to predict the potential distribution of 17 species of forest tree from NFI data. No significant differences were found between these techniques, although GAM showed a slightly higher predictive capacity.

Similarly to GAM, the MARS algorithm is a further development of GLM. It is computationally faster than other algorithms of the regression family and is particularly suitable when a wide range predictive variables are available (Choe et al., 2016; Miller, 2010). This algorithm represents an important alternative to fitting non-linear responses using a piecewise linear fit instead of a smooth function. It has been used by Bedia et al. (2013) in comparison with GLM to evaluate the sensitivity of these algorithms to different climate databases. In Marchi et al. (2016) the algorithm was used to evaluate the potential effect of climate change on the spatial distribution of a marginal and peripheral forest population of European black pine (*Pinus nigra* spp. *nigra* var. *italica*) in comparison with GLM and RF and to generate a consensus map. In Périé and De Blois (2016) MARS is used together with other seven algorithms to evaluate the potential decline in habitat suitability for 5 forest tree species in Canada: *Picea mariana* (Mill.) Britton, Sterns & Poggenb., *Abies balsamea* (L.) Mill., *Betula papyrifera* Marshall, *Acer saccharum* Marsh, and *Betula alleghaniensis* Britton.

CART is the first nonparametric algorithm within the machine learning group. The technique is based on a recursive partitioning process, where the dataset is broken into small homogeneous groups. Noce et al. (2017) used this algorithm within a suite of different models to investigate the likelihood of future provision of suitability distributions for important forest tree species in Europe. The predictive performance of CART is compared with several other algorithms used by Aerts et al. (2010) to model three different forest species (*Pinus brutia*, *Pinus nigra*, *Cedrus libani*) in Turkey. CART algorithms were found to be one of the most user-friendly models. McKenney and Pedlar (2003) used a CART algorithm to predict site productivity on the base of climatic and soil characteristics for two forest species in Canada.

ANN (or sometimes simply Neural Networks, NN) is a complex technique inspired by working principles of the brain. A basic ANN procedure consists of a network of simple elements (artificial neurons) representing the brain (Li and Wang, 2013). The use of a set of adaptive weights allow the tuning of the algorithms with a learning process. A nonlinear relationship between a response variable and an explanatory variable is allowed and the possibility of use data in every statistical distribution except with Gaussian data are the most relevant features of the algorithm (Li and Wang, 2013; Pearson et al., 2002). ANN algorithm is among the 4 different techniques used by Thuiller (2004) to evaluate the potential distribution of different species of plant under various hypothetical climate change scenarios. Bedia et al. (2011) also use ANN algorithm as one of 6 different algorithms in assessing and comparing the predictive performance for distributions of herbaceous plant species in a northern region of Spain. In most cases ANN turned out to be the best algorithm for predictive performance.

**Table 1**

List of software currently available for SDM and related characteristics.

Software	Reference	Operating System	Implemented Algorithms
Biomod2 (R package)	(Thuiller et al., 2009)	Linux distributions, Mac OS, Windows	SRE, ANN, CART, GAM, GARP, GLM, MARS, MaxEnt, RF
dismo (R package)	(Laio et al., 2011)	Linux distributions, Mac OS, Windows	Bioclim, Mahalanobis
SDM (R package)	(Naimi and Araújo, 2016)	Linux distributions, Mac OS, Windows	ANN, GLM, GAM, MARS, CART, RF, ENFA, MaxEnt, Domain, Mahalanobis distance
SDM toolbox (ARCGIS)	(Brown, 2014)	Windows only	MaxEnt
ENiRG (GRASS + R)	(Cánovas et al., 2016)	Linux distributions, Mac OS, Windows	ENFA
Species	(Pearson et al., 2002)	Windows only	ANN
MOPA (R package)	(Iturbide et al., 2015)	Linux distributions, Mac OS, Windows	GLM, SVM, MaxEnt, MARS, RF, CART

Often improperly reported as presence-only algorithm (Elith et al., 2011; Merow et al., 2013; Phillips et al., 2006), we include the MaxEnt algorithm in the presence/absence group. This is due to the fact that this algorithm is deeply different from SRE, Domain, Mahalanobis and ENFA. MaxEnt requires additional information about the external environment where the species is located: *i.e.* the background. This information is generally obtained automatically during computation by means of a spatial random sampling procedure (Guillera-Arroita, 2017). As recognizable from the name, this technique estimates the suitability of an area through a maximum entropy principle. The algorithm calculates the maximum entropy probability of the distribution species and compares it with a maximum entropy probability of the entire object region (Guillera-Arroita et al., 2015). The MaxEnt algorithm has been extensively used in the literature. Lahssini et al. (2015) studied the distribution of *Ceratonia siliqua* L. across Morocco. Results showed a good predictive performance. Antón et al. (2018) used the algorithm to predict potential distribution for 13 tree species in three different time periods: the most recent glaciation, the present and the future period using the A2 scenario from the Third Assessment Report of IPCC. del Río et al. (2018) implemented MaxEnt to evaluate the principal driving factors shaping the distribution of Spanish beech in current and future climate conditions. Cruz- Cardenas et al. (2014) developed a methodology to reduce or resolve the problem of spatial autocorrelation for predictor variables with MaxEnt, offering a PCA for predictor variables and randomness selection for presence records. Clark et al. (2014) examined the current and future potential distribution of an important invasive species (*Ailanthus altissima*) in the Appalachian region of the United States. Finally Dyderski et al. (2018) modelled the potential distribution of 12 forest species for current and future climate conditions.

GARP (Genetic Algorithm for Rule-set Prediction) works similarly to MaxEnt in that it requires a presence/background method (Barbet-Massin et al., 2012). The ‘genetic algorithm’ works on the base of a set of mathematical rules, which are randomly selected and interpreted as a different and limited environmental condition or particular relationship between the environment and a species. Each rule is defined as a “gene”, and each combination of genes generates a different algorithm (Franklin and Miller, 2010; Li and Wang, 2013). Elith et al. (2006) use GARP among 16 different algorithms to predict the potential distribution of 226 different species (both animals and plants) in 6 different regions of the globe. GARP was found to be the most suitable and best-performing algorithm. Vessella and Schirone (2013) used both MaxEnt and GARP to investigate the potential distribution of *Quercus suber* on the basis of current climate conditions. The GARP algorithm outperformed MaxEnt, with drought and cold stress found to be the main factors influencing the distribution of the species.

RF is one of the most important, most used and high-performing algorithms. It consists of a series of decisional trees, which are randomly generated and used to build a virtual forest. Each single tree is

constituted by a random bootstrap sample (Wang et al., 2016). The most important feature of RF is that it is nonparametric and not vulnerable to collinearity. Moreover, it is a robust algorithm and performs well with large datasets (Li and Wang, 2013). Shortcomings of the algorithm include over-fitting which can occur in some cases and the black-box structure which doesn’t allow the user to fully understand the calculation process as well as the weights applied to predictors. Morin and Thuiller (2009) used different techniques to assess the potential range shift of 15 different north-eastern American tree species in the context of climate change. This technique (RF) was compared to PHE-NOFIT which is a process-based model and in addition to all the correlative algorithms contained within the biomod2 package (Thuiller et al., 2009). The main aim was to compare the final output of a niche-based method with other mechanistic models. A high degree of uncertainty was demonstrated by the results which was similar for both models. Attorre et al. (2011) compared RF, GAM and CART to evaluate the potential effects of climate change on the abundance of 27 species on the Italian peninsula. In Garzón et al., 2006 FR, ANN and CART are used to study the potential distribution area of *Pinus sylvestris* RF demonstrated the best predictive performance. RF is also used in Koo et al. (2017) where 6 other algorithms were combined to model the geographical distribution of *Machilus thunbergii* Siebold & Zucc. a typical evergreen broadleaved tree in Korea Peninsula. Finally, RF was used in Benito Garzón et al. (2008) together with CART and ANN to study the future tree distributions in the Iberian Peninsula. The predictive performance of RF is consistently found to be slightly higher than other models.

All the above-described algorithms are briefly summarized in Table1 where the main software involved in SDM are reported. For each of them the included algorithms are reported. Concerning R language, the basic packages such as, for instance, stats for GLM, mgcv for GAM, randomForest for RF etc. were dropped.

## 5. Model evaluation

As stressed in Section 4, the choice of SDM algorithm can give different predictions for habitat suitability. The causes of these observed difference can be due to the small sample sizes and measurement errors, as well as possible omission of an important predictor variable, and the choice of GCM and climate scenario used (Buisson et al., 2010; Marmion et al., 2009). In the scientific literature many comparisons between differential algorithms have been published in order to quantify and evaluate this variability and uncertainty. This issue can raise further problems, in particular: i) how to compare between different statistical algorithms? ii) which parameters should be used for evaluation? and iii) how to evaluate the compatibility between the statistical model and ecological model? In such a framework two principal solutions to resolve these problems are reported in literature: i) comparison between different models can be made by indicators of

goodness of “fit”, or ii) a consensus model approach can be used to balance projections obtained from different algorithms (Austin, 2007; Cheaib et al., 2012; Marmion et al., 2009; Thuiller, 2004). Both approaches aim to assess the relative accuracy of different modelling algorithms. Within these solutions, methods can be grouped into one of two categories: threshold dependent, or threshold independent (Laio et al., 2011; Watling et al., 2013). A threshold independent method evaluates the performance of algorithm only based on comparison of the resulting probabilities. In contrast, a threshold dependent method requires the conversion of raw probabilities produced by the algorithm into two classes based on a defined cut-off value or threshold (Watling et al., 2013). The choice of threshold is a key source of uncertainty. Three different approaches can be followed: a) fixed threshold, b) data-driven (*i.e.* linked to species data or predicted probability); c) accuracy based *i.e.* the value is selected to produce the best compromise between original and the evaluated data (Hanewinkel et al., 2014a).

In ecology, three main methods are consistently applied: Area Under Curve (AUC) or Receiver Operating Characteristic (ROC), Kappa or Cohen Kappa Statistic and the True Skill Statistic (TSS) (Leroy et al., 2018). While the first technique is generated by means of an analysis of the AUC curve, the last two methods are derived from the ‘confusion matrix’ classification system, which facilitates visualization of the performance of an algorithm (Márcia Barbosa et al., 2013). The AUC “Area Under Curve” or ROC values represents an independent threshold technique. AUC produces a bi-dimensional analysis with true positive error on the y axis and false positive error on the x axis (Fig. 3).

The value of AUC can vary between -1 and 1 (Noce et al., 2017; Rivera and López-Quílez, 2017). Much of the literature reports the method to be biased and of limited use. Lobo et al. (2008) report five key disadvantages: i) it ignores the predicted probability values and goodness-of-fit; ii) performances are calculated over regions of the ROC space in which one would rarely operate; iii) omission and commission errors are weighted equally; iv) lack of information concerning spatial distribution of model errors; v) model extent strongly influences the rate of well-predicted absences and AUC scores. Despite these disadvantages AUC continued to be used, even by recent papers. Alternatively threshold dependent methods such as the Kappa or Cohen Kappa rates algorithm performance between 0 and 1, with 1 representing good agreement between predicted and observed presence data. The indicator is calculated according to the following equation:

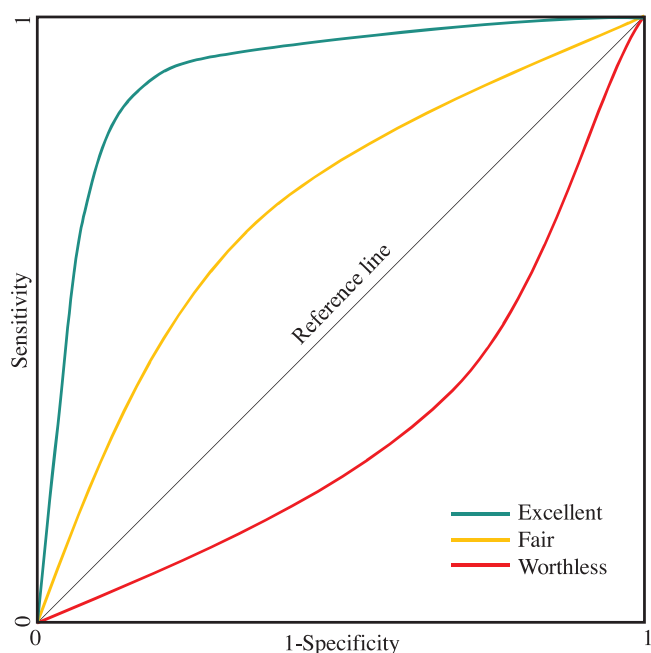


Fig. 3. Possible AUC (or ROC) curves in SDM evaluation.

$$Kappa = \frac{(TP + TN) - \frac{[(TP + FN) \cdot (TP + FP) + (FP + TN) \cdot (FN + TN)]}{N}}{N - \frac{[(TP + FN) \cdot (TP + FP) + (FP + TN) \cdot (FN + TN)]}{N}}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives detected by the algorithm on the total number of testing samples (N). The method is limited by the fact that it is strongly reliant on linked to species prevalence, as well as uncertainty relating to the application of a threshold (Miller, 2010; Watling et al., 2013, 2012).

TSS “Total Sum of Squares” is also threshold dependent, but it has the advantage of being independent from species prevalence. Values between -1 and 1 correspond to the sum of the value of sensitivity and the value of specificity, which are calculated as a proportion of presence areas and absence areas respectively (Barbet-Massin et al., 2012). TSS is calculated using the following equation:

$$TSS = \text{Sensitivity} + \text{Specificity} - 1$$

where:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{FP + TN}$$

Both AUC and TSS have been applied by Noce et al. (2017) to evaluate the capacity of different SDM algorithms to predict potential future suitability of hot-spots for many important forest tree species in southern of Europe. Morin and Thuiller (2009) used only AUC to evaluate predictive performance of alternative niche modelling techniques (process based vs. correlative). Similarly only the Kappa statistic was calculated by Freeman and Moisen (2008) to evaluate predictive performance of SDM for 13 forest tree species in USA. A combination of methods was used by Falk and Hempelmann (2013) to evaluate predictive performance. Zhang et al. (2015a,b) compared all three different methods to evaluate the accuracy of alternative predictive models to study plant distribution in China. AUC and TSS were found to outperform Kappa. Thurm et al. (2018) also adopted TSS to evaluate algorithms used to estimate the present and future potential distribution of 12 forest tree species in Europe.

All the statistics discussed above are among the most used (but not the only) in literature and are often employed to generate consensus or ensemble maps. The interest in ensemble modeling is growing rapidly, not only in ecology but also in other fields such as economy and medicine. Individual algorithms are combined using different techniques: for example a selective algorithm (PCA) or a mathematical or statistical functions such as taking the median, mean or weighted average. Taking the mean remains the most commonly used option in SDM (Kindt, 2018; Marmion et al., 2009). In many papers studying the forestry sector, TSS values are the most common weight applied to each algorithm. Keenan et al. (2011) evaluated the predictive performance of models using the weighted mean of TSS. Engler et al. (2013) used the same technique to present an interesting method to map the distribution of each individual tree belonging to a principal forest species in Switzerland. Zhang et al. (2015a,b) calculated Kappa, AUC and TSS values for each different SDM algorithm, employing an ensemble technique which considered three different functions: the median, the frequency and the simple mean.

## 6. Future scenarios: dealing with the uncertainty behind modeling steps

The process of selecting adequate software, presence/absence datasets, environmental predictors, modeling algorithms and weighting procedures is often carried out with the aim of exploring future scenarios in order to derive insight on how climate change might impact forest tree species distributions. The future provision of ecosystem



services will be highly influenced by climate change (Albert et al., 2017; Ray et al., 2019) and SDM techniques can support decision makers in developing forest management strategies. The use of better adapted forest tree species (genotypes) and provenance selection (genotyping) will improve the resilience of forest systems and allow assisted migration strategies (Hanewinkel et al., 2014; Marchi and Ducci, 2018) thus enforcing the adaptive processes of forest ecosystems (Ferrari et al., 2016). Dealing with the uncertainties generated by climate change is a challenging matter. By sampling along latitudinal (north-south or east-west) or altitudinal gradients, research strategies often aim to search for regions where local adaptation is taking place (Becerra, 2016; Boisvert-Marsh et al., 2014; Kozyr, 2014). The selective pressure exerted upon genotypes at higher elevations (colder temperatures) or southern latitudes (warmer temperatures) forces forest species to adapt to local conditions. This kind of adaptive process, if recognized as genetic difference, will be a valuable resource in forest management strategies (Williams and Dumroese, 2013).

Future projections are generally based on a specific emission scenario, which represents a hypothetical image of the possible trend in greenhouse gas emissions. Alternative pathways for environmental, socio-economic, technological and demographic development are also included. Currently, the most commonly used projections have been generated by the IPCC 5th Assessment Report (AR5). Emission scenarios are represented by “Representative Concentration Pathways” (RCPs) with four possibilities trajectories of increasing severity (RCP 2.6, RCP 4.5, RCP 6.0 and RCP 8.5). These replaced the previous Special Report on Emissions Scenarios (SRES): A1, A2, B1, B2 (Goberville et al., 2015). In addition to the AR5 RCPs, many Global Circulation Models (GCM) have been developed by research groups around the world, often targeted at specific geographic regions. These regional GCMs rely heavily on statistical probability, which introduces deep degree of uncertainty in SDM efforts using those datasets. For this reason, researchers have tried to tackle the issue by combining multiple GCM and RCP inputs and analyzing the resulting variance. The most common method is to average different outputs (Goberville et al., 2015). For example, Iverson et al. (2008a,b) used three different climate models and data from the SRES emission scenarios A1 and B1. Walentowski et al. (2017) did repeated this method using new AR5 scenarios, thereby providing a comparison. Hanewinkel et al. (2010) used data from emission scenarios A2 and B1 to predict the possible economic consequences of a shift from *Picea abies* to *Fagus sylvatica* in Southern Germany. Finally, Benito Garzón et al. (2008) used data from a range of emission scenarios (A1, A2, B1, B2) to simulate the impact on the distribution of the Iberian tree species in three different time slices: 2020s, 2050s, 2080s. The use of Regional Climate Models (RCM) allow higher precision than GCM due to their greater detail at small scale (Franklin et al., 2013; Koca et al., 2006). RCM are derived from GCM models through a process of statistical downscaling. For instance, an interesting comparison was made by Liu et al. (2014) where the use of GCM or RCM was assessed when dealing with climate change impacts and the future scenarios for invasive plants in USA.

A “static” SDM (i.e. no migration included) is based on the assumption that species distribution is in equilibrium with climate and will react locally to a changing climate. The inclusion of the migration capacity of different tree species is a rapidly increasing theme in recent literature and could be used to improve estimations of the ability of a target species to colonize new sites. Two important methods are currently available to evaluate the migration capacity: the MigClim (Engler et al., 2012) and KissMig algorithms (Subba et al., 2018). Both are currently available as R packages and while the first enables the implementation of species-specific dispersal constraints into projections of SDM, KissMig offers a simple, raster-based and stochastic migration model (Nobis and Normand, 2014).

## 7. Conclusions

This review has found an evident increment in scientific contribution relating to SDM from 2000 to 2019, demonstrating a growing interest for the technique. The primary aim of SDM in forest research and management is to derive insights relating to the future potential distribution of tree species in order to implement effective adaptation strategies (Janowiak et al., 2017). The literature highlights that the correct interpretation and use of presence/absence datasets is central to deriving a correct estimation of the ecological niche for the species in question. Another key finding is the need to reduce the uncertainties associated with modelling steps (e.g. reliability of species distribution data, climate surfaces, GCMs) and further work in this area is highly important to improve estimations of future forest distribution and resulting ecosystem services.

A key downfall of SDM as it stands is the lack of inclusion of biotic interaction into the modeling procedure. Inclusion of these processes would see a shift from SDM to true Ecological Niche Models. However, such models, even if theorized, are yet to be achieved. A promising area of further research are multi-species and multi-level SDM, where many target trees are modeled at the same time. Currently only single tree models are in use, leaving the final merging procedure to a simple overlay process in a GIS environment. Host-disease models offer an similar promising advancement, where both host and disease are modelled according to climate change scenarios to simulate the potential future impact of biotic stresses (e.g. insects, fungi, bacteria) on forest tree species. In this case a co-evolution in modelling techniques is expected and, consequently, a combined SDM offers a promising method to model such interactions.

## Acknowledgements

This study was partially supported by the PhD grant provided by the University of Florence to Matteo Pecchi in the framework of the project “*Effetti di eventiestremi a seguito di cambiamenti climaticisu ambientaturali e semi-naturali. Impatto, mitigazione e resilienza*”.

Maurizio Marchi and Vanessa Burton were partially funded by the EU in the framework of the Horizon 2020 B4EST project “*Adaptive BREEDING for productive, sustainable and resilient FORESTS under climate change*”, UE Grant Agreement 773383.

## References

- Aertsens, W., Kint, V., van Orshoven, J., Özkan, K., Muys, B., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol. Modell.* 221, 1119–1130. <https://doi.org/10.1016/j.ecolmodel.2010.01.007>.
- Bede-Fazekas, Ákos, Levente Horváth, M.K., 2014. Impact of climate change on the potential distribution of Mediterranean pines. *Quarterl J. Hungarian Meteorol. Serv.* 118, 41–52.
- Albert, M., Nagel, R.V., Nuske, R.S., Suttmöller, J., Spellmann, H., 2017. Tree species selection in the face of drought risk-uncertainty in forest planning. *Forests* 8. <https://doi.org/10.3390/f8100363>.
- Antúnez, P., Suárez-Mota, M.E., Valenzuela-Encinas, C., Ruiz-Aquino, F., 2018. The potential distribution of tree species in three periods of time under a climate change scenario. *Forests* 9, 1–12. <https://doi.org/10.3390/f9100628>.
- Attorre, F., Alfó, M., De Sanctis, M., Francesconi, F., Valenti, R., Vitale, M., Bruno, F., 2011. Evaluating the effects of climate change on tree species abundance and distribution in the Italian peninsula. *Appl. Veg. Sci.* 14, 242–255. <https://doi.org/10.1111/j.1654-109X.2010.01114.x>.
- Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol. Modell.* 200, 1–19. <https://doi.org/10.1016/j.ecolmodel.2006.07.005>.
- Barbet-Massin, M., Jetz, W., 2014. A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Divers. Distrib.* 20, 1285–1295. <https://doi.org/10.1111/ddi.12229>.
- Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol. Evol.* 3, 327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>.
- Beale, C.M., Lennon, J.J., 2012. Incorporating uncertainty in predictive species distribution modelling. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 247–258. <https://doi.org/10.1098/rstb.2011.0178>.

- Becerra, P.I., 2016. Relationship between climate and geographical variation of local woody species richness within the Mediterranean-type region of Chile. *Rev. Chil. Hist. Nat.* 89, 1–11. <https://doi.org/10.1186/s40693-016-0062-x>.
- Bedia, J., Busqué, J., Gutiérrez, J.M., 2011. Predicting plant species distribution across an alpine rangeland in northern Spain. A comparison of probabilistic methods. *Appl. Veg. Sci.* 14, 415–432. <https://doi.org/10.1111/j.1654-109X.2011.01128.x>.
- Bedia, J., Herrera, S., Gutiérrez, J.M., 2013. Dangers of using global bioclimatic datasets for ecological niche modeling. Limitations for future climate projections. *Glob. Planet. Change* 107, 1–12. <https://doi.org/10.1016/j.gloplacha.2013.04.005>.
- Benito Garzón, M., Robson, T.M., Hampe, A., 2019.  $\Delta$ TraitSDM: species distribution models that account for local adaptation and phenotypic plasticity. *New Phytol.* 222, 1757–1765. <https://doi.org/10.1111/nph.15716>.
- Benito Garzón, M., Sánchez De Dios, R., Sainz Ollero, H., 2008. Effects of climate change on the distribution of Iberian tree species. *Appl. Veg. Sci.* 11, 169–178. <https://doi.org/10.3170/2008-7-18348>.
- Boiffin, J., Badeau, V., Bréda, N., 2017. Species distribution models may misdirect assisted migration: insights from the introduction of Douglas-fir to Europe. *Ecol. Appl.* 27, 446–457. <https://doi.org/10.1002/eap.1448>.
- Boisvert-Marsh, L., Périé, C., De Blois, S., 2014. Shifting with climate? Evidence for recent changes in tree species distribution at high latitudes. *Ecosphere* 5, art83. <https://doi.org/10.1890/ES14-00111.1>.
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol. (Amst.)* 24, 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>.
- Booth, T.H., 2018. Species distribution modelling tools and databases to assist managing forests under climate change. *For. Ecol. Manage.* 430, 196–203. <https://doi.org/10.1016/j.foreco.2018.08.019>.
- Booth, T.H., 2017. Assessing species climatic requirements beyond the realized niche: some lessons mainly from tree species distribution modelling. *Clim. Change* 145, 259–271. <https://doi.org/10.1007/s10584-017-2107-9>.
- Booth, T.H., 2014. Using biodiversity datasets to verify and improve descriptions of tree species climatic requirements. *For. Ecol. Manage.* 315, 95–102. <https://doi.org/10.1016/j.foreco.2013.12.028>.
- Booth, T.H., Nix, H.A., Hutchinson, M.F., Jovanic, T., 1988. Niche analysis and tree species introduction. *For. Ecol. Manage.* 23 (1), 47–59. [https://doi.org/10.1016/0378-1127\(88\)90013-8](https://doi.org/10.1016/0378-1127(88)90013-8).
- Bradley, B.A., Olsson, A.D., Wang, O., Dickson, B.G., Pelech, L., Sesnie, S.E., Zachmann, L.J., 2012. Species detection vs. Habitat suitability: are we biasing habitat suitability models with remotely sensed data? *Ecol. Modell.* 244, 57–64. <https://doi.org/10.1016/j.ecolmodel.2012.06.019>.
- Brown, J.L., 2014. SDMTtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol. Evol.* 5, 694–700. <https://doi.org/10.1111/2041-210X.12200>.
- Brus, D.J., Hengeveld, G.M., Walvoort, D.J.J., Goedhart, P.W., Heidema, A.H., Nabuurs, G.J., Gunia, K., 2012. Statistical mapping of tree species over Europe. *Eur. J. For. Res.* 131, 145–157. <https://doi.org/10.1007/s10342-011-0513-5>.
- Buisson, L., Thuiller, W., Casajus, N., Lek, S., Grenouillet, G., 2010. Uncertainty in ensemble forecasting of species distribution. *Glob. Chang. Biol.* 16, 1145–1157. <https://doi.org/10.1111/j.1365-2486.2009.02000.x>.
- Cánovas, F., Magliozzi, C., Mestre, F., Palazón, J.A., González-Wangüemert, M., 2016. ENIRG: R-GRASS interface for efficiently characterizing the ecological niche of species and predicting habitat suitability. *Ecography (Cop.)* 39, 593–598. <https://doi.org/10.1111/ecog.01426>.
- Carpenter, G., Gillison, A.N., Winter, J., 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodivers. Conserv.* 2, 667–680. <https://doi.org/10.1007/BF00051966>.
- Casalegno, S., Amatulli, G., Camia, A., Nelson, A., Pekkarinen, A., 2010. Vulnerability of *Pinus cembra* L. in the Alps and the Carpathian mountains under present and future climates. *For. Ecol. Manage.* 259, 750–761. <https://doi.org/10.1016/j.foreco.2009.10.001>.
- Castaldi, C., Vacchiano, G., Marchi, M., Corona, P., 2017. Projecting nonnative douglas fir plantations in Southern Europe with the forest vegetation simulator. *For. Sci.* 63, 101–110. <https://doi.org/10.5849/forsci.16-048>.
- Caudullo, G., Welk, E., San-Miguel-Ayán, J., 2017. Chorological maps for the main European woody species. *Data Br.* 12, 662–666. <https://doi.org/10.1016/j.dib.2017.05.007>.
- Cheab, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrene, E., François, C., Gritti, E.S., Legay, M., Pagé, C., Thuiller, W., Viovy, N., Leadley, P., 2012. Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. *Ecol. Lett.* 15, 533–544. <https://doi.org/10.1111/j.1461-0248.2012.01764.x>.
- Choe, H., Thorne, J.H., Seo, C., 2016. Mapping national plant biodiversity patterns in South Korea with the mars species distribution model. *PLoS One* 11, 1–15. <https://doi.org/10.1371/journal.pone.0149511>.
- Clark, J., Wang, Y., August, P.V., 2014. Assessing current and projected suitable habitats for tree-of-heaven along the appalachian trail. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130192. <https://doi.org/10.1098/rstb.2013.0192>.
- Coudun, C., Gégout, J.C., Piedallu, C., Rameau, J.C., 2006. Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *J. Biogeogr.* 33, 1750–1763. <https://doi.org/10.1111/j.1365-2699.2005.01443.x>.
- Crimmins, S.M., Dobrowski, S.Z., Mynsberge, A.R., 2013. Evaluating ensemble forecasts of plant species distributions under climate change. *Ecol. Modell.* 266, 126–130. <https://doi.org/10.1016/j.ecolmodel.2013.07.006>.
- Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J.L., Ortiz, E., 2014. Potential species distribution modeling and the use of principal component analysis as predictor variables. *Rev. Mex. Biodivers.* 85, 189–199. <https://doi.org/10.7550/rmb.36723>.
- Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.* 28, 2031–2064. <https://doi.org/10.1002/joc.1688>.
- del Río, S., Álvarez-Esteban, R., Cano, E., Pinto-Gomes, C., Penas, Á., 2018d. Potential impacts of climate change on habitat suitability of *Fagus sylvatica* L. Forests in Spain. *Plant Biosyst.* 152, 1205–1213. <https://doi.org/10.1080/11263504.2018.1435572>.
- Di Biase, R.M., Fattorini, L., Marchi, M., 2018. Statistical inferential techniques for approaching forest mapping. A review of methods. *Ann. Silv. Res.* 42, 46–58. <https://doi.org/10.12899/asr-1738>.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop.)* 36, 027–046. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Duan, R.-Y., Kong, X.-Q., Huang, M.-Y., Fan, W.-Y., Wang, Z.-G., 2014. The predictive performance and stability of six species distribution models. *PLoS One* 9, e112764. <https://doi.org/10.1371/journal.pone.0112764>.
- Duputié, A., Zimmermann, N.E., Chuine, I., 2014. Where are the wild things? Why we need better data on species distribution. *Glob. Ecol. Biogeogr.* 23, 457–467. <https://doi.org/10.1111/geb.12118>.
- Duvenne, M.J., Scheller, R.M., 2015. Climate-suitable planting as a strategy for maintaining forest productivity and functional diversity. *Ecol. Appl.* 25, 1653–1668. <https://doi.org/10.1890/14-0738.1>.
- Dyderski, M.K., Paź, S., Frelich, L.E., Jagodziński, A.M., 2018. How much does climate change threaten European forest tree species distributions? *Glob. Change Biol. Bioenergy* 24, 1150–1163. <https://doi.org/10.1111/gcb.13925>.
- 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., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.C., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *M. Ecography (Cop.)* 29, 129–151.
- Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>.
- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>.
- Engler, R., Hordijk, W., Guisan, A., 2012. The MIGCLIM R package - seamless integration of dispersal constraints into projections of species distribution models. *Ecography (Cop.)* 35, 872–878. <https://doi.org/10.1111/j.1600-0587.2012.07608.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.* 1–18. <https://doi.org/10.1155/2013/787250>.
- Falk, W., Mellert, K.H., 2011. Species distribution models as a tool for forest management planning under climate change: risk evaluation of *Abies alba* in Bavaria. *J. Veg. Sci.* 22, 621–634. <https://doi.org/10.1111/j.1654-1103.2011.01294.x>.
- Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecol. Modell.* 160, 115–130. [https://doi.org/10.1016/S0304-3800\(02\)00327-7](https://doi.org/10.1016/S0304-3800(02)00327-7).
- Ferrara, C., Marchi, M., Fares, S., Salvati, L., 2017. Sampling strategies for high quality time-series of climatic variables in forest resource assessment. *iForest - Biogeosciences For.* 10, 739–745. <https://doi.org/10.3832/for2427-010>.
- Ferrarini, A., Selvaggi, A., Abeli, T., Alatalo, J.M., Orsenigo, S., Gentili, R., Rossi, G., 2016. Planning for assisted colonization of plants in a warming world. *Sci. Rep.* 6, 28542. <https://doi.org/10.1038/srep28542>.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Fleischer, P., Pichler, V., Flaischer, P., Holko, L., Máls, F., Gömöryová, E., Cudlín, P., Holeska, J., Michalová, Z., Homolová, Z., Skvarenina, J., Střelcová, K., Hlaváč, P., 2017. Forest ecosystem services affected by natural disturbances, climate and land-use changes in the Tatra Mountains. *Clim. Chang. Res. Lett.* 73, 57–71. <https://doi.org/10.3354/cr01461>.
- Franklin, J., 2010. *Mapping Species Distributions Spatial Inference and Prediction*. Franklin, J., Davis, F.W., Ikegami, M., Syphard, A.D., Flint, L.E., Flint, A.L., Hannah, L., 2013. Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Glob. Chang. Biol.* 19, 473–483. <https://doi.org/10.1111/gcb.12051>.
- Franklin, J., Miller, J.A., 2010. *Mapping Species Distribution*. Cambridge University Press <https://doi.org/10.1007/s13398-014-0173-7.2>.
- Freeman, E.A., Moisen, G.G., 2008. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol. Modell.* 217, 48–58. <https://doi.org/10.1016/j.ecolmodel.2008.05.015>.
- Garzón, M.B., Blazek, R., Neteler, M., Dios, R.S., de Ollero, H.S., Furlanello, C., 2006. Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula. *Ecol. Modell.* 197, 383–393. <https://doi.org/10.1016/j.ecolmodel.2006.03.015>.
- Gastón, A., García-Viñas, J.I., Bravo-Fernández, A.J., López-Leiva, C., Oliet, J.A., Roig, S., Serrada, R., 2014. Species distribution models applied to plant species selection in

- forest restoration: Are model predictions comparable to expert opinion? *New For.* 45, 641–653. <https://doi.org/10.1007/s11056-014-9427-7>.
- Goberville, E., Beaugrand, G., Hautek, N.C., Piquot, Y., Luczak, C., 2015. Uncertainties in the projection of species distributions related to general circulation models. *Ecol. Evol.* 5, 1100–1116. <https://doi.org/10.1002/ece3.1411>.
- Godsoe, W., Franklin, J., Blanchet, F.G., 2017. Effects of biotic interactions on modeled species' distribution can be masked by environmental gradients. *Ecol. Evol.* 7, 654–664. <https://doi.org/10.1002/ece3.2657>.
- Guillera-Arroita, G., 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography (Cop.)* 40, 281–295. <https://doi.org/10.1111/ecog.02445>.
- Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R., Wintle, B.A., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. *Glob. Ecol. Biogeogr.* 24, 276–292. <https://doi.org/10.1111/geb.12268>.
- Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Modell.* 157, 89–100. [https://doi.org/10.1016/S0304-3800\(02\)00204-1](https://doi.org/10.1016/S0304-3800(02)00204-1).
- Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. *Habitat Suitability and Distribution Models*. Cambridge Univ. Press <https://doi.org/10.1017/9781139028271>.
- Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P., Buckley, Y.M., 2013. Predicting species distributions for conservation decisions. *Ecol. Lett.* 16, 1424–1435. <https://doi.org/10.1111/ele.12189>.
- Hamann, A., Wang, T., 2006. Potential effects of climate change on ecosystem. *Ecology* 87, 2773–2786.
- Hamann, A., Wang, T., Spittlehouse, D.L., Murdock, T.Q., 2013. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. *Bull. Am. Meteorol. Soc.* 94, 1307–1309. <https://doi.org/10.1175/BAMS-D-12-00145.1>.
- Hanewinkel, M., Cullmann, D., Schelhaas, M.-J., Nabuurs, G.-J., Zimmermann, N.E., 2012. Climate change may cause severe loss in the economic value of European forest land. *Nat. Clim. Chang.* 3, 203–207. <https://doi.org/10.1038/nclimate1687>.
- Hanewinkel, M., Hummel, S., Cullmann, D.A., 2010. Modelling and economic evaluation of forest biome shifts under climate change in Southwest Germany. *For. Ecol. Manage.* 259, 710–719. <https://doi.org/10.1016/j.foreco.2009.08.021>.
- Hanewinkel, M., Kuhn, T., Bugmann, H., Lanz, A., Brang, P., 2014. Vulnerability of uneven-aged forests to storm damage. *Forestry* 87, 525–534. <https://doi.org/10.1093/forestry/cpu008>.
- Hengl, T., De Jesus, J.M., Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 12, e0169748. <https://doi.org/10.1371/journal.pone.0169748>.
- Hernández-Quiroz, N.S., Badano, E.I., Barragán-Torres, F., Flores, J., Pinedo-Álvarez, C., 2018. Habitat suitability models to make conservation decisions based on areas of high species richness and endemism. *Biodivers. Conserv.* 27, 3185–3200. <https://doi.org/10.1007/s10531-018-1596-9>.
- Higa, M., Tsuyama, I., Nakao, K., Nakazono, E., Matsui, T., Tanaka, N., 2013. Influence of nonclimatic factors on the habitat prediction of tree species and an assessment of the impact of climate change. *Landsc. Ecol. Eng.* 9, 111–120. <https://doi.org/10.1007/s11355-011-0183-y>.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. <https://doi.org/10.1002/joc.1276>.
- Hijmans, R.J., Elith, J., 2011. *Species Distribution Modeling With R*. R-Core Dev. Team 87.
- Hijmans, R.J., Graham, C.H., 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Glob. Chang. Biol.* 12, 2272–2281. <https://doi.org/10.1111/j.1365-2486.2006.01256.x>.
- Hill, L., Hector, A., Hemery, G., Smart, S., Tanadini, M., Brown, N., 2017. Abundance distributions for tree species in Great Britain: a two-stage approach to modeling abundance using species distribution modeling and random forest. *Ecol. Evol.* 7, 1043–1056. <https://doi.org/10.1002/ece3.2661>.
- Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., Perrin, A.N., 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83, 2027–2036. [https://doi.org/10.1890/0012-9658\(2002\)083\[2027:ENFAHT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2027:ENFAHT]2.0.CO;2).
- Isaac-Renton, M.G., Roberts, D.R., Hamann, A., Spiecker, H., 2014. Douglas-fir plantations in Europe: a retrospective test of assisted migration to address climate change. *Glob. Chang. Biol.* 20, 2607–2617. <https://doi.org/10.1111/gcb.12604>.
- Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutiérrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. *Ecol. Modell.* 312, 166–174. <https://doi.org/10.1016/j.ecolmodel.2015.05.018>.
- Iverson, L., Prasad, A., Matthews, S., 2008a. Modeling potential climate change impacts on the trees of the northeastern United States. *Mitig. Adapt. Strateg. Glob. Chang.* 13, 487–516. <https://doi.org/10.1007/s11027-007-9129-y>.
- Iverson, L.R., Prasad, A.M., Matthews, S.N., Peters, M., 2008b. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *For. Ecol. Manage.* 254, 390–406. <https://doi.org/10.1016/j.foreco.2007.07.023>.
- Janowiak, M.K., Iverson, L.R., Fosgitt, J., Handler, S.D., Dallman, M., Thomasma, S., Hutnik, B., Swanston, C.W., 2017. Assessing stand-level climate change risk using forest inventory data and species distribution models. *J. For.* 115, 222–229. <https://doi.org/10.5849/jof.2016-023r1>.
- Jarnevich, C.S., Stohlgren, T.J., Kumar, S., Morissette, J.T., Holcombe, T.R., 2015. Caveats for correlative species distribution modeling. *Ecol. Inform.* 29, 6–15. <https://doi.org/10.1016/j.ecoinf.2015.06.007>.
- Jeschke, J.M., Strayer, D.L., 2008. Usefulness of bioclimatic models for studying climate change and invasive species. *Ann. N. Y. Acad. Sci.* <https://doi.org/10.1196/annals.1439.002>.
- Johnson, K.D., Birdsey, R., Finley, A.O., Swantaran, A., Dubayah, R., Wayson, C., Riemann, R., 2014. Integrating forest inventory and analysis data into a LIDAR-based carbon monitoring system. *Carbon Balance Manag.* 9, 1–11. <https://doi.org/10.1186/1750-0680-9-3>.
- Keenan, T., Maria Serra, J., Lloret, F., Ninoyrola, M., Sabate, S., 2011. Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO2 matters!. *Glob. Chang. Biol.* 17, 565–579. <https://doi.org/10.1111/j.1365-2486.2010.02254.x>.
- Kindt, R., 2018. Ensemble species distribution modelling with transformed suitability values. *Environ. Model. Softw.* 100, 136–145. <https://doi.org/10.1016/j.envsoft.2017.11.009>.
- Koca, D., Smith, B., Sykes, M.T., 2006. Modelling regional climate change effects on potential natural ecosystems in Sweden. *Clim. Change* 78, 381–406. <https://doi.org/10.1007/s10584-005-9030-1>.
- Koo, K.A., Park, S.U., Kong, W.S., Hong, S., Jang, I., Seo, C., 2017. Potential climate change effects on tree distributions in the Korean Peninsula: understanding model & amp; climate uncertainties. *Ecol. Modell.* 353, 17–27. <https://doi.org/10.1016/j.ecolmodel.2016.10.007>.
- Kozyr, I.V., 2014. Forest vegetation dynamics along an altitudinal gradient in relation to the climate change in southern taibalkaia. *Russia. Achiev. Life Sci.* 8, 23–28. <https://doi.org/10.1016/j.als.2014.11.006>.
- LAHSSINI, S., HAJIB, S., Lahlaoui, H., Mharzi Alaoui, H., Khattabi, A., 2015. Modelling spatial distribution of the carob tree (*Ceratonia siliqua* L.) in Azilal Province. *Morocco. J. Geogr. Geol.* 7, 33. <https://doi.org/10.15439/jgg.v7n4p33>.
- Laio, F., Ganora, D., Claps, P., Galeati, G., 2011. Spatially smooth regional estimation of the flood frequency curve (with uncertainty). *J. Hydrol. (Amst)* 408, 67–77. <https://doi.org/10.1016/j.jhydrol.2011.07.022>.
- Leroy, B., Delsol, R., Huguency, B., Meynard, C.N., Barhoumi, C., Barbet-Massin, M., Bellard, C., 2018. Without quality presence-absence data, discrimination metrics such as TSS can be misleading measures of model performance. *J. Biogeogr.* 45, 1994–2002. <https://doi.org/10.1111/jbi.13402>.
- Li, X., Wang, Y., 2013. Applying various algorithms for species distribution modelling. *Integr. Zool.* 8, 124–135. <https://doi.org/10.1111/1749-4877.12000>.
- Liu, S., Liang, X., Gao, W., Stohlgren, T.J., 2014. Regional climate model downscaling may improve the prediction of alien plant species distributions. *Front. Earth Sci.* 8, 457–471. <https://doi.org/10.1007/s11707-014-0457-4>.
- Lobo, J.M., Jiménez-valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* 17, 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>.
- Manchego, C.E., Hildebrandt, P., Cueva, J., Espinosa, C.I., Stimm, B., Günter, S., 2017. Climate change versus deforestation: implications for tree species distribution in the dry forests of southern Ecuador. *PLoS One* 12, 1–19. <https://doi.org/10.1371/journal.pone.0190092>.
- Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence absence models in ecology: the need to count for prevalence. *J. Appl. Ecol.* 38, 921–931.
- Marchi, M., Ducci, F., 2018. Some refinements on species distribution models using tree-level national forest inventories for supporting forest management and marginal forest population detection. *iForest - Biogeosciences For.* 11, 291–299. <https://doi.org/10.3832/ifor2441-011>.
- Marchi, M., Nocentini, S., Ducci, F., 2016. Future scenarios and conservation strategies for a rear-edge marginal population of *Pinus nigra* Arnold in Italian central Apennines. *For. Syst.* 25, e072. <https://doi.org/10.5424/fs/2016253-09476>.
- Marchi, M., Sinjur, I., Bozzano, M., Westergren, M., 2019. Evaluating WorldClim version 1 (1961–1990) as the baseline for sustainable use of forest and environmental resources in a changing climate. *Sustainability* 11, 3043. <https://doi.org/10.3390/su11113043>.
- Márcia Barbosa, A., Real, R., Muñoz, A.R., Brown, J.A., 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Divers. Distrib.* 19, 1333–1338. <https://doi.org/10.1111/ddi.12100>.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* 15, 59–69. <https://doi.org/10.1111/j.1472-4642.2008.00491.x>.
- Masek, J.G., Goward, S.N., Kennedy, R.E., Cohen, W.B., Moisen, G.G., Schleeweis, K., Huang, C., 2013. United States forest disturbance trends observed using landsat time series. *Ecosystems* 16, 1087–1104. <https://doi.org/10.1007/s10021-013-9669-9>.
- McInerny, G.J., Etienne, R.S., 2013. 'Niche' or 'distribution' modelling? A response to Warren. *Trends Ecol. Evol. (Amst.)* 28, 191–192. <https://doi.org/10.1016/j.tree.2013.01.007>.
- McInerny, G.J., Etienne, R.S., 2012. Ditch the niche - is the niche a useful concept in ecology or species distribution modelling? *J. Biogeogr.* 39, 2096–2102. <https://doi.org/10.1111/jbi.12033>.
- 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](https://doi.org/10.1016/S0378-1127(02)00186-X).
- Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography (Cop.)* 36, 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>.
- Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Sayre, R., Trabucco, A., Zomer, R., 2013.



- A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. *Glob. Ecol. Biogeogr.* 22, 630–638. <https://doi.org/10.1111/geb.12022>.
- Miller, J., 2010. Species distribution modeling. *Geogr. Compass.* 4, 490–509. <https://doi.org/10.1111/j.1749-8198.2010.00351.x>.
- Mod, H.K., Scherrer, D., Luoto, M., Guisan, A., Scheiner, S., 2016. What we use is not what we know: environmental predictors in plant distribution models. *J. Veg. Sci.* 27, 1308–1322. <https://doi.org/10.1111/jvs.12444>.
- Morin, X., Thuiller, W., 2009. Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology* 90, 1301–1313. <https://doi.org/10.1890/08-0134.1>.
- Mura, M., McRoberts, R.E., Chirici, G., Marchetti, M., 2016. Statistical inference for forest structural diversity indices using airborne laser scanning data and the k-Nearest Neighbors technique. *Remote Sens. Environ.* 186, 678–686. <https://doi.org/10.1016/j.rse.2016.09.010>.
- Naimi, B., Araújo, M.B., 2016. Sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography (Cop.)* 39, 368–375. <https://doi.org/10.1111/ecog.01881>.
- Nobis, M.P., Normand, S., 2014. KISSMig - a simple model for R to account for limited migration in analyses of species distributions. *Ecography (Cop.)* 37, 1282–1287. <https://doi.org/10.1111/ecog.00930>.
- Noce, S., Collalti, 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>.
- Paniccia, C., Di Febbraro, M., Frate, L., Sallustio, L., Santopoli, G., Altea, T., Posillico, M., Marchetti, M., Loy, A., 2018. Effect of imperfect detection on the estimation of niche overlap between two forest dormice. *iForest - Biogeosciences For.* 11, 482–490. <https://doi.org/10.3832/for2738-011>.
- Pearson, R.G., 2010. Species' distribution modeling for conservation educators and practitioners. *Lessons Conserv.* 3, 54–89.
- Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* 12, 361–371.
- Pearson, R.G., Dawson, T.P., Berry, P.M., Harrison, P.A., 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecol. Modell.* 154, 289–300. [https://doi.org/10.1016/S0304-3800\(02\)00056-X](https://doi.org/10.1016/S0304-3800(02)00056-X).
- Pecchi, M., Marchi, M., Giannetti, F., Bernetti, I., Bindi, M., Moriondo, M., Maselli, F., Fibbi, L., Corona, P., Travaglini, D., Chirici, G., 2019. Reviewing climatic traits for the main forest tree species in Italy. *iForest - Biogeosciences For.* 12, 173–180. <https://doi.org/10.3832/for2835-012>.
- Périé, C., de Blois, S., 2016. Dominant forest tree species are potentially vulnerable to climate change over large portions of their range even at high latitudes. *PeerJ* 4, e2218. <https://doi.org/10.7717/peerj.2218>.
- Peterson, A.T., Soberón, J., 2012. Species distribution modeling and ecological niche modeling: Getting the Concepts Right. *Nat. a Conserv.* 10, 102–107. <https://doi.org/10.4322/natcon.2012.019>.
- Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., Araújo, M.B., 2011. *Ecological Niches and Geographic Distributions*.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Pulliam, H.R., 2000. On the relationship between niche and distribution. *Ecol. Lett.* 3, 349–361. <https://doi.org/10.1046/j.1461-0248.2000.00143.x>.
- Ray, D., Petr, M., Mullett, M., Bathgate, S., Marchi, M., Beauchamp, K., 2019. A simulation-based approach to assess forest policy options under biotic and abiotic climate change impacts: a case study on Scotland's National Forest Estate. *For. Policy Econ.* 103, 17–27. <https://doi.org/10.1016/j.forpol.2017.10.010>.
- Rehfeldt, G.E., Worrall, J.J., Marchetti, S.B., Crookston, N.L., 2014. Adapting forest management to climate change using bioclimate models with topographic drivers. *Forestry* 88, 528–539. <https://doi.org/10.1093/forestry/cpv019>.
- Rivera, Ó., López-Quílez, A., 2017. Development and comparison of species distribution models for forest inventories. *ISPRS Int. J. Geo-Inf.* 6, 176. <https://doi.org/10.3390/ijgi6060176>.
- Roces-Díaz, J.V., Jiménez-Alfaro, B., Álvarez-Álvarez, P., Álvarez-García, M.A., 2014. Environmental niche and distribution of six deciduous tree species in the Spanish atlantic region. *iForest - Biogeosci. For.* 8, 214–221. <https://doi.org/10.3832/for1183-008>.
- Rupprecht, F., Oldeland, J., Finckh, M., 2011. Modelling potential distribution of the threatened tree species *Juniperus oxycedrus*: how to evaluate the predictions of different modelling approaches? *J. Veg. Sci.* 22, 647–659. <https://doi.org/10.1111/j.1654-1103.2011.01269.x>.
- San-Miguel-Ayán, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A., 2016. *European Atlas of Forest Tree Species*. Publication Office of the European Union, Luxembourg, Luxembourg. <https://doi.org/10.2788/038466>.
- Schröder, B., 2008. Challenges of species distribution modeling belowground. *J. Plant Nutr. Soil Sci.* (1999) 171, 325–337. <https://doi.org/10.1002/jpln.200700027>.
- Spittlehouse, D.L., Stewart, R.B., 2004. Adaptation to climate change in forest management. *BC J. Ecosyst. Manag.* 4, 2–11.
- Subba, B., Sen, S., Ravikanth, G., Nobis, M.P., 2018. Direct modelling of limited migration improves projected distributions of Himalayan amphibians under climate change. *Biol. Conserv.* 227, 352–360. <https://doi.org/10.1016/j.biocon.2018.09.035>.
- Teuscher, M., Brandl, R., Förster, B., Hothorn, T., Rösner, S., Müller, J., 2013. Forest inventories are a valuable data source for habitat modelling of forest species: an alternative to remote-sensing data. *Forestry* 86, 241–253. <https://doi.org/10.1093/forestry/cps081>.
- Thuiller, W., 2013. On the importance of edaphic variables to predict plant species distributions - limits and prospects. *J. Veg. Sci.* 24, 591–592. <https://doi.org/10.1111/jvs.12076>.
- Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. *Glob. Chang. Biol.* 10, 2020–2027. <https://doi.org/10.1111/j.1365-2486.2004.00859.x>.
- Thuiller, W., Araújo, M.B., Lavorel, S., 2004. Do we need land-cover data to model species distributions in Europe? *J. Biogeogr.* 31, 353–361. <https://doi.org/10.1046/j.0305-0270.2003.00991.x>.
- Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - A platform for ensemble forecasting of species distributions. *Ecography (Cop.)* 32, 369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>.
- Thuiller, W., Lavorel, S., Sykes, M.T., Araújo, M.B., 2006. Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Divers. Distrib.* 12, 49–60. <https://doi.org/10.1111/j.1366-9516.2006.00216.x>.
- Thuiller, W., Pollock, L.J., Gueguen, M., Münkemüller, T., 2015. From species distributions to meta-communities. *Ecol. Lett.* 18, 1321–1328. <https://doi.org/10.1111/ele.12526>.
- Thurm, E.A., Hernandez, L., Baltensweiler, A., Ayan, S., Rasztovits, E., Bielak, K., Zlatanov, T.M., Hladnik, D., Balic, B., Freudenthuss, A., Büchsenmeister, R., Falk, W., 2018. Alternative tree species under climate warming in managed European forests. *For. Ecol. Manage.* 430, 485–497. <https://doi.org/10.1016/J.FORECO.2018.08.028>.
- Title, P.O., Bemmels, J.B., 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography (Cop.)* 41, 291–307. <https://doi.org/10.1111/ecog.02880>.
- van der Maaten, E., Hamann, A., van der Maaten-Theunissen, M., Bergsma, A., Hengeveld, G., van Lammeren, R., Mohren, F., Nabuurs, G.J., Terhiürne, R., Sterck, F., 2017. Species distribution models predict temporal but not spatial variation in forest growth. *Ecol. Evol.* 7, 2585–2594. <https://doi.org/10.1002/ece3.2696>.
- Vessella, F., Schirone, B., 2013. Predicting potential distribution of *Quercus suber* in Italy based on ecological niche models: conservation insights and reforestation involvements. *For. Ecol. Manage.* 304, 150–161. <https://doi.org/10.1016/j.foreco.2013.05.006>.
- Vetaas, O.R., 2002. Realized and potential climate niches: a comparison of four Rhododendron tree species. *J. Biogeogr.* 29, 545–554. <https://doi.org/10.1046/j.1365-2699.2002.00694.x>.
- Walentowski, H., Falk, W., Mette, T., Kunz, J., Bräuning, A., Meinardus, C., Zang, C., Sutcliffe, L.M.E., Leuschner, C., 2017. Assessing future suitability of tree species under climate change by multiple methods: a case study in southern Germany. *Ann. For. Res.* 60, 101–126. <https://doi.org/10.15287/afr.2016.789>.
- Wang, T., Campbell, E.M., O'Neill, G.A., Aitken, S.N., 2012. Projecting future distributions of ecosystem climate niches: uncertainties and management applications. *For. Ecol. Manage.* 279, 128–140. <https://doi.org/10.1016/j.foreco.2012.05.034>.
- Wang, T., Wang, G., Innes, J., Nitschke, C., Kang, H., 2016. Climatic niche models and their consensus projections for future climates for four major forest tree species in the Asia-Pacific region. *For. Ecol. Manage.* 360, 357–366. <https://doi.org/10.1016/j.foreco.2015.08.004>.
- Warren, D.L., 2012. In defense of niche modeling. *Trends Ecol. Evol. (Amst.)* 27 (9), 497–500. <https://doi.org/10.1016/j.tree.2012.03.010>.
- Watling, J.I., Brandt, L.A., Mazzotti, F.J., Romanach, S.S., 2013. Use and Interpretation of Climate Envelope Models: A Practical Guide. Univ. Florida, pp. 1–43.
- Watling, J.I., Romanach, S.S., Bucklin, D.N., Speroterra, C., Brandt, L.A., Pearlstone, L.G., Mazzotti, F.J., 2012. Do bioclimate variables improve performance of climate envelope models? *Ecol. Modell.* 246, 79–85. <https://doi.org/10.1016/j.ecolmodel.2012.07.018>.
- Williams, M.I., Dumroese, R.K., 2013. Preparing for climate change: forestry and assisted migration. *J. For.* 111, 287–297. <https://doi.org/10.5849/jof.13-016>.
- Zhang, J., Nielsen, S.E., Chen, Y., Georges, D., Qin, Y., Wang, S.S., Svenning, J.C., Thuiller, W., 2017. Extinction risk of North American seed plants elevated by climate and land-use change. *J. Appl. Ecol.* 54, 303–312. <https://doi.org/10.1111/1365-2664.12701>.
- Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., Wang, L., 2015a. Consensus forecasting of species distributions: the effects of niche model performance and niche properties. *PLoS One* 10, e0120056. <https://doi.org/10.1371/journal.pone.0120056>.
- Zhang, M.T., Kang, X.G., Meng, J.H., Zhang, L.X., 2015b. Distribution patterns and associations of dominant tree species in a mixed coniferous-broadleaf forest in the Changbai Mountains. *J. Sci.* 12, 659–670. <https://doi.org/10.1007/s11629-013-2795-1>.