Forest Ecology and Management 304 (2013) 150–161

Contents lists available at SciVerse ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Predicting potential distribution of *Quercus suber* in Italy based on ecological niche models: Conservation insights and reforestation involvements

Federico Vessella *, Bartolomeo Schirone

Dipartimento Ambiente, Foreste, Natura ed Energia (DAFNE), Università degli Studi della Tuscia, 01100 Viterbo, Italy

ARTICLE INFO

Article history:
Received 7 December 2012
Received in revised form 30 April 2013
Accepted 4 May 2013

Keywords:
Quercus suber
Distribution patterns
GARP
MaxEnt
Conservation and reforestation areas

ABSTRACT

Different statistical techniques have been used to model species potential distribution related to environmental variables. This paper provides a comprehensive assessments of GARP and MaxEnt methods, and investigates for the first time the probability of occurrence of cork oak (*Quercus suber* L.) in Italy based on ecological niche modelling approaches. A detailed distribution of the species was achieved during a 3-year National Project (SuberItalia) and 17 environmental layers were employed to obtain the potential distribution of cork oak. The performance of the models were measured using the receiver operating characteristic (ROC) approach and Cohen's Kappa statistic. Results achieved by GARP and MaxEnt showed as the drought and the cold stresses are the main factors affecting cork oak occurrence in Italy. Moreover, the accuracy of the obtained prediction maps were compared to a specifically calibrated geo-statistical method at regional scale, pointing out a preliminary geographical assessment of the suitable surfaces to set apart for cork oak forest expansion in Italy, thereby useful to address reforestation and conservation concerns to face the ongoing area reduction of these forests.

1. Introduction

Cork oak (*Quercus suber* L., Fagaceae) is a sclerophyllous tree occurring in the western Mediterranean Basin where an "oceanic Mediterranean climate" rules (Daget, 1977), rather continuously along the Atlantic coasts of North Africa and Iberian Peninsula, where it is also found inland. In Italy cork oak is present along the Tyrrhenian coast, from Liguria to Calabria, in the major islands (Sicily and Sardinia), and eastward in small enclaves scattered along the Apulian coast (Adriatic Sea, SE Italy), whereas its occurrence in the Balkans is historically contradictory and generally unconsidered (Simeone et al., 2009). *Quercus suber* is characteristic of a mosaic of habitats ranked among the most valuable in Europe and listed in the EC Habitats Directive (2180 "Wooded dunes of the Atlantic, Continental and Boreal region", 6310 "Dehesas with evergreen *Quercus* spp.", 9330 "*Q.suber* forests", 9540 "Mediterranean pine forests with endemic Mesogean pines"). Moreover, cork oak landscapes sustain rich biodiversity and traditional livehoods, represent an important source of income derived from cork production, and play a key role in ecological processes, such as water retention, soil conservation or carbon storage (WWF, 2006; Gil and Varela, 2008). In recent times, forest fires and diseases, overgrazing, land abandonment and exploitation are causing a serious cork oak forest area reduction all over the Mediterranean Basin with subsequent biodiversity loss and degradation of ecosystem services (Costa et al., 2010; Bugalho et al., 2011; Sedda et al., 2011). These factors lead to a severe decrease of cork production, in particular of the high-quality one, especially in Italy where the processing industry of cork is advanced and of great value worldwide (Sedda et al., 2011). A reasonable solution could concern the extension of the areas devoted to cork oak stands along the Tyrrhenian coast, but it requires an investigation about the ecologically suitable surfaces for such a species.

In this view, we investigated the species occurrence focusing at the eastern marginal areas of its range (Tyrrhenian coast, Sardinia, Sicily and Apulia), where disturbance, habitat fragmentation and climate change are ongoing drivers that hardly affect species persistence (Jump et al., 2009; Lepetz et al., 2009), so predictions of suitable expansion areas should be considered of high importance for reforestation, but also for conservation and management purposes (Channell and Lomolino, 2000; Petit et al., 2005; Parmesan, 2006, 2007).

Predictive distribution mapping refers to the creation of models for graphically representing and studying species' ranges in a geographical information systems (GIS) environment, especially in the context of biogeography and evolutionary ecology, conservation and speciation, and issues dealing with responses to global change

* Corresponding author. Tel.: +39 0761357391; fax: +39 0761357250.
E-mail address: vessella@unitus.it (F. Vessella).

0378-1127/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.foreco.2013.05.006

and climatic factors (Franklin, 1995; Anderson et al., 2002; Ko et al., 2009; Li et al., 2012; Tarkesh and Jetschke, 2012). In this respect, a novel prosing approach is the ecological niche modelling (ENM), which has been extensively applied in several studies about marine and terrestrial fauna, but rarely employed in plant studies (some examples are found in Thuiller et al., 2003; Irfan-Ullah et al., 2007; Scarnati et al., 2009; Giannini et al., 2010; Qi et al., 2010; Navarro-Cerrillo et al., 2011; Babar et al., 2012; Li et al., 2012; Rojas-Soto et al., 2012). The key input for ENM application is geo-referenced distributional records for species, combined with raster geospatial data sets representing environmental parameters (biotic and abiotic) that commonly affect species' macro-distributions (Anderson et al., 2003; Elith et al., 2006; Peterson, 2006; Bodbyl-Roels et al., 2011). Several methods are actually available to statistically predict species distribution relying on relationships between species occurrence data points and environmental variables. Among them, the presence-only ENM approaches have been often employed to predict geographic distribution of many taxa, including endangered species (Irfan-Ullah et al., 2007; Benito et al., 2009; Babar et al., 2012; Torres et al., 2012), invasive alien species (Chen, 2008, 2009; Herborg et al., 2009; Zhu et al., 2012), endemisms (Escalante et al., 2007, 2009), and disease organisms (Gao et al., 2010; Joyner et al., 2010; Bodbyl-Roels et al., 2011; Khatchikian et al., 2011).

The resulting output, based on the species' fundamental niche's (*sensu* Hutchinson, 1957), is projected onto a map reflecting the potential distribution of a species, but could be over-predicted due to several limitations in climatic data (Bodbyl-Roels et al., 2011) and moreover, because the model can include areas occupied by competitors, land uses that prevent species' natural occurrence (e.g. urban areas) and suitable sites to which the species failed to disperse or it has gone extinct (Anderson et al., 2003). However, ENM allows a comparison between theoretical and realized species distributions, thus involving or suggesting concerns about biogeography, evolutionary ecology, conservation and restoration measurements.

In this study we explore two of the most popular used data algorithms, GARP (Genetic Algorithm for Rule-set Prediction, Stockwell and Peters, 1999), and MaxEnt (Maximum Entropy, Phillips et al., 2004) applied to provide information on cork oak in Italy.

GARP is a heuristic approach that belongs to a class of machine-learning applications inspired by models of genetic and evolution (Holland, 1975; Anderson et al., 2003); it includes a set of non-deterministic iterative procedures that incorporate various model distribution methods (e.g. logistic regression, bioclimatic-envelope rules) producing predicted presence/absence binary maps for each run. GARP reduces error in predicted distributions by maximizing both significance and predictive accuracy (Anderson et al., 2002, 2003; Peterson et al., 2007). In contrast, MaxEnt fits a probability distribution for species occurrences by finding the distribution of maximum entropy (i.e. the distribution that is most spread-out, or closest to uniform), an axiom of the Bayesian probability theory, in agreement with given presence data points and subject to restrictions defined by the environmental features being analyzed (Phillips et al., 2004, 2006; Elith et al., 2011). Both GARP and MaxEnt use known occurrences and pseudo-absence data re-sampled from study area where the target species is not known to occur (Peterson et al., 2007).

This investigation represents the first approach in the Mediterranean Basin using these two machine learning techniques to a common, but economically and ecologically strategic tree species, relying also on a considerable and very accurate present distribution dataset. The aims were: (i) to predict the potential geographical distribution of cork oak using the ENM approach toward GARP and MaxEnt algorithms, (ii) to identify the environmental variables mostly influencing its geographical distribution, and (iii) to evaluate proper areas for conservation and reforestation actions, taking into account the current land cover. Additionally, a deeper investigation at regional scale (Latium) is reported to compare the predictions or to point out refinements useful to better achieve the aim (iii).

2. Materials and methods

2.1. Occurrence data collection

The occurrence locations of *Q. suber* in Italy were collected during a 3-year field survey across the country, as the main goal of a National project (SuberItalia) funded by the Ministry of Agriculture and Forests (MIPAF). Natural cork oak stands and isolated individuals were recorded, geo-referenced, and a detailed distribution map was finally produced at high scale level (1:2000) using ArcGIS 9.3 (http://www.esri.com/software/arcgis), with 13,099 entries (polygons and points) representing the actual distribution of the species (Fig. 1a). According to the requirements forced by the selected ENM approaches, we excluded those entries referred to single trees or populations in areas not covered by the environmental predictors, producing a subset shapefile of polygons that was converted into a database of centroids using the conversion function implemented in X-Tools Pro 4.0 extension of ArcGis 9.3. Since the accuracy of the spatial distribution modelling is constrained by the quality of occurrence data and mainly by the resolution of environmental data sets, our final database included 5647 spatially unique points under a resolution of 0.0083 decimal degrees (equal to a cell size of 0.93 km) to be used as input in our study (Fig. 1b). However, the restrictions imposed by GARP and MaxEnt methods did not affect significantly the information stored into the original database (data not shown): in fact, the presence data points used as final input, under the mentioned resolution, is representative of the 99% ca. of the cork oak distribution showed in Fig. 1a.

2.2. Environmental layers

Predictive data, selected as main range determinants for cork oak, consisted of 17 variables retrieved from different databases (Table 1). Nine climatic raster data were obtained from WorldClim 1.4 at 30 arc-second resolution (0.93×0.93 km) for the period 1950–2000 (http://www.worldclim.org; Hijmans et al., 2005). Topographic variables was achieved from ASTER Global Digital Elevation Model (GeoTIFF format with geographic lat/long coordinates and a 1 arc-second (30 m) grid of elevation postings; http://gdem.ersdac.jspacesystems.or.jp/) and re-scaled to 30 arc-second resolution in order to match with WorldClim sources; aspect and slope maps were retrieved with 30 arc-second resolution as well from Digital Elevation Model (DEM) using Raster Surface extension in 3D-Analyst toolbox implemented in ArcGis 9.3. Lithologic and eco-pedologic maps were also employed after conversion to raster format at 30 arc-second resolution, to provide information about soil properties and the associated ecological vegetation units (source: Ministero dell'Ambiente e della Tutela del Territorio e del Mare – MATTM; http://www.pcn.minambiente.it). Additional variables included two phytoclimatic indexes, namely Winter Cold Stress (WCS) and Summer Drought Stress (SDS) of Mitrakos (Mitrakos, 1980) derived from WorldClim data and spatially computed for Italy, and Corine Land Cover (CLC, 4th level, 1:100,000) from European Environment Agency (EEA; http://www.eea.europa.eu/) used also to evaluate suitable areas for reforestation after computing the potential distribution maps.

Concerning the investigation performed in Latium, DEM, slope and aspect without re-scaling (30 m) were used, while the same climatic variables obtained from WorldClim were calculated using