REVISITING NICHE FUNDAMENTALS WITH TUKEY DEPTH

J. Orestes Cerdeira, Tiago Monteiro-Henriques, M. João Martins, Pedro C. Silva, Diogo Alagador, Aldina M. A. Franco, Manuel L. Campagnolo, Pedro Arsénio, Francisca C. Aguiar, Mar Cabeza

Jorge Orestes Cerdeira
Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica, Portugal.
jo.cerdeira@fct.unl.pt

Tiago Monteiro Henriques
Centro de Estudos Florestais, CEF, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
Centro de Investigação e de Tecnologias Agroambientais e Biológicas, CITAB, Universidade de Trás-os-Montes e Alto Douro, Apartado 1013, 5001-801 Vila Real, Portugal.
tmh@isa.ulisboa.pt

M. João Martins
Centro de Estudos Florestais, CEF, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
mjmartins@isa.ulisboa.pt

Pedro C. Silva
Centro de Estudos Florestais, CEF, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
pcsilva@isa.ulisboa.pt

Diogo Alagador
CIBIO/InBio-UE: Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade de Évora, 7000-890 Évora, Portugal.
alagador@uevora.pt

Aldina M. A. Franco
School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.
a.franco@uea.ac.uk

Manuel L. Campagnolo
Centro de Estudos Florestais, CEF, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
mlc@isa.ulisboa.pt

Pedro Arsénio

Linking Landscape, Environment, Agriculture and Food, LEAF, Instituto Superior de Agronomia, Universidade de Lisboa,Tapada da Ajuda, 1349-017 Lisboa, Portugal.
arseniop@isa.ulisboa.pt

Francisca C. Aguiar
Centro de Estudos Florestais, CEF, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
fraguiar@isa.ulisboa.pt

Mar Cabeza
Centre of Excellence in Metapopulation Research, Department of Biosciences, University of Helsinki, Finland.
mar.cabeza@helsinki.fi

Corresponding author: Jorge Orestes Cerdeira

Type of article: Research article

Running headline: Revisiting niche fundamentals

ABSTRACT

1. The first attempts to describe species ecological niches were simple geometric procedures that depict the niche boundaries directly from environmental data. The convex hull was one of such procedures, popular for its simplicity, clear ecological rational and precise definition of the niche. However, it lacked the ability to differentiate areas of the niche with different probabilities of occurrence according to environmental suitability.
2. We incorporate the Tukey depth, a mathematical tool to measure the centrality of a point within a cloud of points on a multidimensional space, in the convex hull approach to (i) propose a new procedure (CH-Tukey) to estimate species' environmental suitability, and (ii) estimate niche overlap coherently. In addition to a clear ecological rational and simplicity the CH-Tukey procedure has a number of attractive features: use of presence-only data; independence from background data; invariance to scale; robustness to outliers; and the decomposition of the niche into a finite number of iso-suitability levels, permitting the computation of consistent overlap indices. We illustrate the use of CH-Tukey, using occurrence data of the main *Quercus* species and subspecies from Western Mediterranean Europe, comparing its outputs with BIOCLIM and MaxEnt.
3. Results showed distinct niche geometries among the different approaches. BIOCLIM produced rectilinear niches reflecting the assumption that ecological variables are independent in their action on the species. CH-Tukey, relaxing this assumption, adjusts niche outer boundary and the inner suitability levels to the

known occurrences. MaxEnt produced unbounded niche geometries, showing abrupt shifts in the species response to the environmental variables.

4. The niche predictions obtained with geometric approaches, BIOCLIM and CH-Tukey, are simpler but better aligned with Hutchinson's niche concept than those obtained with MaxEnt, this latter showing ecologically implausible relationships with the environmental variables. CH-Tukey and the related overlap measures provide an adequate tool to explore niche properties and species-environment relationships.

Keywords: Depth functions, environmental envelope, niche modelling, geometric procedures, Hutchinsonian niche, niche overlap, Quercus species.

INTRODUCTION

Hutchinson's niche concept (Hutchinson, 1957) inspired ecological analyses immediately after its proposal (Whittaker & Levin, 1975), with numerous applications and methods emerging after Hutchinson's seminal paper. Applications are broad, including: studies of resource partitioning in ecological communities (Schoener, 1974), methods on predictive vegetation mapping (Franklin, 1995), explorations of the relation between genome changes and niche differentiation (Rocap *et al.*, 2003), reflections on plant coexistence theories (Silvertown, 2004), or even in taxonomy (Raxworthy *et al.*, 2007). Several other associated ecological concepts followed from Hutchinson's niche concept, e.g.: the guild (Root, 1967), niche breadth and niche overlap (see Colwell & Futuyma,1971; Hurlbert, 1978), niche width (see Roughgarden, 1972), the regeneration niche (Grubb, 1977), the ontogenetic niche (Werner & Gilliam, 1984), persistence niche (Bond & Midgley, 2001), niche conservatism/dynamics (Peterson *et al.*, 1999; Pearman *et al.*, 2008). The uses of the Hutchinson's n-dimensional hypervolume concept across ecology and evolution have been recently surveyed and discussed by Blonder (2017).

In recent years, the most conspicuous application of the niche concept has been the plethora of methods developed to estimate the suitability of ecological conditions to species, or eventually, species distributions (Guisan & Zimmermann, 2000; Franklin, 2010; Ahmed *et al.*, 2015). Generalized Linear Models (GLM), Generalized Additive Models (GAM), Maximum Entropy Modelling (MaxEnt), Mahalanobis distance, point-to-point using Gower metric (DOMAIN), Ecological-Niche Factor Analysis (ENFA), among others, produce continuous predictions, raising further challenges. For instance, they depend on thresholds to delineate a precise niche boundary (see Elith *et al.*, 2006; Tsoar *et al.*, 2007; Blonder *et al.*, 2014). Some methods, such as GLM, GAM, MaxEnt, Genetic Algorithm for Rule-set Production (GARP) and artificial neural networks, even after applying user-defined suitability/probability thresholds, may produce unlimited hypervolumes (see Blonder *et al.*, 2014). Although some of those methods fit data well, they may exhibit species-environment relationships with lack of ecological plausibility (Brewer *et al.*, 2016). Addressing this critical issue and aiming at a better relationship with ecological theory, Brewer *et al.* (2016) proposed a new parametric approach, the Plateau climate envelope. A non-parametric procedure, also ecologically meaningful, is BIOCLIM, the very first geometric approach to the niche. Revisiting and improving such geometric approaches can result interesting, as will be shown hereafter.

BIOCLIM (Nix, 1986; Busby, 1991) was one of the first attempts to formalize the Hutchinsonian niche of a species and to produce maps of species potential distributions in a systematic way (Booth *et al.*, 2014). It is a simple geometric procedure, which identifies the range of suitable values for each environmental variable, independently of the other variables, thus, defining the niche as the hyperrectangle bounded by the minimum and maximum values of the environmental variables for the species known occurrences. This "range" notion within each environmental variable is implicit in Hutchinson's article and is central to the niche concept (i.e. the species is expected to survive and reproduce within some known pair of limiting values for each variable, and not outside those conditions). To distinguish suitability levels within the niche, Nix (1986) and Busby (1991) used the nested hyperrectangles obtained from percentiles of each environmental variable individually. In this way they incorporated Hutchison's perception of an optimal part of the niche versus a suboptimal one near the boundaries (Hutchinson, 1957). The most suitable regions are those corresponding to the environmental conditions lying in the innermost hyperrectangles.

The simple approach of BIOCLIM exhibits some attractive properties: i) performs geometric operations with straightforward ecological interpretation; ii) delineates precise (limited) boundaries of the niche; iii) only uses presences and does not depend on the background; iv) is invariant to scale (suitability levels do not change under linear transformations of data); v) is robust to outliers (while the outer regions are sensitive to outliers their influence vanishes for the interior); and vi) produces ordinal predictions decomposing the niche into a finite number of iso-suitability regions.

BIOCLIM presents however a major drawback: it relies on the assumption that ecological variables are independent in their action on the species. With the aim of relaxing this independence assumption, the HABITAT procedure (Walker & Cocks, 1991) replaced the bounding box envelop used in BIOCLIM by the convex hull of the occurrence points. The convex hull of a set of points defines the minimum convex hypervolume containing all points. Hereafter the construction of the convex hull of occurrences in the environmental space will be referred as the CH-approach. The CH-approach is a comprehensible multidimensional geometric way to circumscribe the niche, but no geometric procedure was used to discriminate suitability levels within it. However, this is achievable in a similar way to the percentile procedure implemented in BIOCLIM by using depth functions, which are procedures to measure the centrality of a point within a data cloud on a multi-dimensional space (Zuo & Serfling, 2000). Cerdeira *et al.* (2014) suggested the use of depth functions to estimate suitability levels for species in the CH-approach. Here we describe and discuss a new method that adds normalized Tukey depth to the CH-approach, and give an application of this proposal to *Quercus* taxa occurring in the Iberian Peninsula, illustrating it further by assessing the overlap of the produced climatic niches. Finally we compare the geometry of the environmental niches produced by CH-approach using Tukey depth with those of BIOCLIM and MaxEnt.

METHODOLOGIES

Tukey depth

The Tukey depth of a point (also called location depth or halfspace depth) is a measure, introduced by Tukey (1975), to generalize the order of quantiles for the multivariate case.

Consider a set *P* of *k* points of an *n*-dimensional space ($\mathbb{R}$), with $>$, and an arbitrary point *q*. The Tukey depth of point *q* with respect to *P* is defined as $d_q(P)=m/k$, where *m* is the minimum number of points to remove from *P* so that *q* is not in the convex hull of the remaining points of *P*. We call *m* the (Tukey) interiority of point *q* in *P*.

(Figure 1)

Figure 1 depicts a set *P* of 20 points of $\mathbb{R}$ represented by “ ”, and three points *x*, *y* and *z*. The lines passing through each of the points *x*, *y* or *z*, show that removing 1, 3 and 9 points from *P* (those that are on one of the sides of the line thought the point) will leave *x*, *y* and *z* outside the convex hull of remaining points of *P* (those that are on the other side of the line). It can be easily verified that removing less points would leave *x*, *y* or *z*, respectively, inside the convex hull of the remaining points of *P*. We can therefore conclude that the interiorities of points *x*, *y* and *z* are 1, 3 and 9, respectively, and $d_x(P)$ = 1/20, $d_y(P)$ = 3/20, and $d_z(P)$ = 9/20.

From the definition, it follows that the maximum interiority cannot exceed

$$=\frac{- \quad +1}{2} \tag{1}$$

A point *q* with maximum interiority is a Tukey median point. Point *z* has interiority 9 and, since *M* = 9, it is a Tukey median of *P*.

Depth functions define a partition of $\mathbb{R}$ into iso-depth regions, i.e., regions of points having equal depth values. Figure 2A shows the iso-depth regions corresponding to the set of 10 points of $\mathbb{R}$ represented by “ ”. Note that the regions of points with Tukey depth greater than or equal to *m*/*k*, , are nested, i.e., $\mathbb{R} = \quad \supset \quad \supset \cdots \supset \quad$, with increasing depth toward the innermost region. The region (*m* $\geq$1) could alternatively be defined as the intersection of all convex hulls of the subsets of *P* with *k*-(*m*-1) points.

Depth values of the points of a set *P* can give indication on how *P* is arranged, distinguishing between configurations of points that are concentrated in the "interior", from those which occur mainly on the "margins". Silva et al. (2014) have recently established that, if the points in *P* are uniformly distributed, which is a borderline between the two above configurations, the average $d_q(P)$, for *q* in *P*, approximates $1/2^{n+1}$. This value, which only depends on the dimension (*n*) of the representation space, provides a sound threshold to distinguish configurations where the concentration is in the "interior" (average $d_q(P) > 1/2^{n+1}$), from configurations where points occur mainly on the "margins" (average $d_q(P) < 1/2^{n+1}$).

Assessing environmental suitability

Let *P* be the set of *k* points in $\mathbb{R}$, each point indicating the *n* environmental conditions of a location where a species occurs. To every point *q* of *P* we calculate the depth $d_q(P)$ and divide it by *M*/*k*. With this rescaling, points with interiority *M* (if any) score 1. We propose to interpret these normalized depth values as a measure of environmental suitability for the species. The value 1 corresponds to environmental conditions of sites that belong to the most inner region of the estimated niche. Note that such environmental conditions might not exist.

We call this the CH-Tukey procedure. Under the equilibrium assumption, areas of greater $d_q(P)$ values correspond to the optimal parts of the niche (sensu Hutchinson 1957), i.e. areas where species performance is expected to be higher (e.g. growth rate or species fitness). The innermost region, where $d_q(P)$ reach its maximum value, is the set of Tukey median points.

It should be clear that CH-Tukey procedure assumes unimodal species responses, in relation to the considered environmental variables. More specifically, CH-Tukey procedure assumes that the suitability of every point of the line connecting two points q_1 and q_2 of the environmental space, will be at least as large as the minimum between the suitability of q_1 and q_2.

Our proposal can be viewed as a refinement of the procedure of BIOCLIM to define different levels of suitability, relaxing the independence assumption. Indeed, the way BIOCLIM differentiates suitability levels in the bounding box of P, using percentiles of each environmental variable individually, is actually the output of a depth function that can be read as follows. The environmental suitability of a location corresponding to point q on the environmental space, is twice the value of $d'_q(P) = m'/k$, where m' is the minimum number of points to remove from P so that q is not in the minimum hyperrectangle containing the remaining points of P. The multiplication by two is a rescaling so that the environmental suitability of points in the most inner hyperrectangle equals one. Figure 2B depicts the regions of points with equal d' values corresponding to the set of 10 points of $\mathbb{R}$ represented by “ ” (the same points as in Figure 2A). Note that the regions of points for which d' is greater than or equal to m/k, $'$, are also nested, i.e., $\mathbb{R} = ' \supset ' \supset \cdots \supset '$, where $' =$ $-$. The region $'$ ($m \geq 1$) is the intersection of all bounding boxes of the subsets of P with k-(m-1) points. Therefore, we have: i) $M'=M$ if $n=1$ and $M'>M$ for ≥ 2; ii) $' \supset$, for $m=1,\ldots, M$, since the bounding box of a set of points always contains the convex hull of the same set; or, equivalently, for any q, $d'_q(P)= m'/k \geq m/k = d_q(P)$.

(Figure 2)

Both procedures differentiate the outer polygon in nested regions of increasing suitability for the interior, CH-Tukey producing tighter regions than BIOCLIM. Points outside the outer polygon are considered unsuitable; points in the margins have small suitability, while the largest suitability is assumed to occur in the inner regions.

Assessing the overlap of niches

Several indices have been proposed to assess the niche overlap between two species, by comparing the respective suitability (or presence/absence) on a finite number of points, which correspond to cells of the geographical area (see Warren *et al.*, 2008, for a survey). However, to assess the overlap of the environmental niches of two species all points of the environmental space should be considered and not only the projection into the cells of a map. This is an issue that is particularly important in climate change scenarios. Villéger *et al.* (2011) proposes to evaluate the functional dissimilarity between two (or more) species assemblages, considering the volumes of the union and of the intersection of the convex hulls of the points representing each species assemblage in the space of their functional traits. This approach can be easily transposed to the evaluation of the overlap between the environmental niches of two species, giving an overlap index defined as the ratio between the volume of the intersection and the volume of the union of the convex hulls of the occurrences of both species. This measure can be interpreted as the Jaccard index for the volume and is the natural way to estimate niche overlap when using the CH-approach to describe the niche. It has the disadvantage of not distinguishing between intersections occurring in regions of high suitability and those occurring in low suitability areas of the niches.

As Tukey depth divides the niche into a finite number of iso-suitability regions, it allows to generalize Pianka and Jaccard indices incorporating volumes and suitability.

Consider two species A and B, and the partition of the environmental space into and regions of (non zero) iso-depth for each species. Denote by 0 = ! "#$ < ! "#$ < ··· < ! , "#$ and 0 = ! "($ < ! "($ < ··· < !) "($ all possible values of suitability for species A and B, respectively, and let *+,, with (*i*,*j*) ≠ (0,0), be the volume of the (possibly disconnected) region composed by all the environments having suitability !+"#$ for species A and !,"($ for species B, and * arbitrary. The volumes can be arranged in a (+ 1) by (+1) matrix - = .*+,/. This matrix contains all the information on intersections of the environmental niches, taking into account distinct levels of suitability for both species. Note that, if !+"#$ and !,"($ only take 0-1 values, i.e. resulting from binary prediction models, such as the CH-approach, - = .*+,/ is a 2 by 2 matrix, where * is the volume of the intersection of the niches, and * (*) is the volume of the niche of species A (B) not shared by species B (A).

From M_V we define the matrix M_A (M_B) of the asymmetric *overlap of species* B (A) on A (B), multiplying the row *i* of M_V by !+"#$ (the column *j* of M_V by !,"($). Each entry (*i*,*j*) of matrix M_A (M_B) is the volume *+, of the region with iso-suitability !+"#$, for species A and !,"($ for species B, weighted by !+"#$ (!,"($).

We propose to use the matrices above to define two overlap indices. The first is

$$wJ = \frac{\sum_{i=1}^{k_A}\sum_{j=1}^{k_B} M_A(i,j)\, M_B(i,j)}{2\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} M_A(i,j)\, M_B(i,j) - \sum_{i=1}^{k_A}\sum_{j=1}^{k_B} M_A(i,j)\, M_B(i,j)} = \frac{\sum_{i=1}^{k_A}\sum_{j=1}^{k_B} s_i(A)\, s_j(B)\, M_V(i,j)}{2\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} s_i(A)\, s_j(B)\, M_V(i,j) - \sum_{i=1}^{k_A}\sum_{j=1}^{k_B} s_i(A)\, s_j(B)\, M_V(i,j)}.$$

Index *wJ* incorporates the different levels of suitability on the ratio of the volumes of the intersection and of the union of the niches. If !+"#$ and !,"($ only assumes values 0 or 1,

$$wJ = \frac{M_V(1,1)}{M_V(1,0) + M_V(0,1) + M_V(1,1)},$$

which is the Jaccard index for the volume. For the particular case when suitability is given by CH-approach, the niches are estimated by the convex hulls of the occurrences of both species, and *wJ* becomes the index of Villéger *et al.* (2011).

When suitability is obtained with Tukey depth, *wJ* distinguishes between intersections occurring in regions of high suitability and those occurring in low suitability areas of the niches.

Figure 3 illustrates iso-suitability regions for each of three hypothetical species A, B and B’ (top) in the same environmental space, and how they are arranged together (bottom). The volumes of the intersections/unions of the convex hulls for A and B, and for A and B’ are exactly the same. Thus, Jaccard index returns the same value (0.2179) for both pairs. Index *wJ* gives 0.3579 for pair (A, B), where intersection occurs mainly in the regions of high suitability, and 0.1604 for pair (A, B’) where the regions of high suitability almost do not overlap.

(Figure 3)

The second overlap index we propose is a *cosine* index, which is an adaptation of Pianka index used to measure resource overlap between two species (Pianka, 1973). Consider matrices M_{A} and M_{B} as vectors of length (+1)×(+1). The *cos* index is defined as the cosine of the angle between these vectors.

$$cos = \frac{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} M_{\mathrm{A}}(i,j)\, M_{\mathrm{B}}(i,j)}{\sqrt{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} M_{\mathrm{A}}(i,j)^2}\sqrt{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} M_{\mathrm{B}}(i,j)^2}}$$

$$= \frac{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} s_i(A)s_j(B)\,(M_{\mathrm{V}}(i,j))^2}{\sqrt{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} s_i(A)M_{\mathrm{V}}(i,j)^2}\sqrt{\sum_{i=0}^{k_A}\sum_{j=0}^{k_B} s_j(B)M_{\mathrm{V}}(i,j)^2}}$$

When the niches do not overlap, vectors are orthogonal and *cos*=0. When high and low suitability occur in the same regions for both species, *cos*≈1.

ILLUSTRATIVE EXAMPLE

The climatic suitability of *Quercus* taxa in the Iberian Peninsula

As an example of application of the above-described CH-Tukey procedure, we computed climatic suitability for *Quercus* taxa occurring in the Iberian Peninsula.

Quercus spp. are widespread in the Northern Hemisphere (Manos *et al.*, 1999). In the Iberian Peninsula *Quercus* taxa occur in most of the territory being an important economic and conservation target (see Bugalho *et al.*, 2011; Correia *et al.*, 2017). For example, several of the local economies are dependent on cork extraction, cork represents 2.5% of Portugal exports, and acorns from *Quercus rotundifolia* and *Q. suber* are particularly important in many animals' diet (Rodríguez-Estévez *et al.*, 2009; Bugalho *et al.*, 2011). *Q. robur* is an important forestry species. *Q. faginea* and *Q. pyrenaica* are valuable to local communities for the production of firewood (Gonçalves *et al.*, 2012; Cerqueira *et al.*, 2013).

In this paper we focus our analyses on six species/subspecies: *Q. rotundifolia* (= *Q. ilex* subsp. *ballota*), *Q. suber*, *Q. pyrenaica*, *Q. faginea* subsp. *faginea*, *Q. faginea* subsp. *broteroi* and *Q. robur*. Occurrences were obtained from the Proyecto Anthos (Anthos, 2011) and from the Herbarium "João de Carvalho e Vasconcellos" (Lisbon, LISI) and include a mix of point and grid-based data of several natures and types. Positional accuracy of the species observations data was analysed and improved when appropriate.

We used four climatic variables known to affect and limit species distributions, downloaded from the Worldclim website (Hijmans *et al.*, 2005;

http://www.worldclim.org) at the 1km × 1km grid cell resolution. Two variables involve temperature (isothermality and the minimum temperature of coldest month) and two variables involve precipitation (annual precipitation and the precipitation of driest quarter).

For each selected taxon, we applied the CH-Tukey and the BIOCLIM procedures in order to produce geographical maps distinguishing suitability levels. The Tukey depth of each site within the Iberian Peninsula (ca. 1 km × 1 km resolution), with respect to the occurrences of each taxon, was determined using the function *depth* from the R package *depth* version 2.0-0 (Genest *et al.*, 2012). The depth values were then rescaled dividing by $0'' - 3\$/23$, where *k* is the number of occurrences of the taxon, thus obtaining a normalized Tukey depth. The BIOCLIM suitability associated to each site was obtained using the functions *bc* and *predict* from the R package *dismo* version 1.1-1 (Hijmans *et al.*, 2016).

The maps clearly show that the inland non-white areas for CH-Tukey, (Figure 4) are contained in the inland non-white areas for BIOCLIM, $'$ (Figure 5), denoting the fact that the convex hull is a tighter envelope than the hyperrectangle.

A robust validation of these models imply controlled experiments using a small set of factors considered limitative of species fitness. Given that these experiments are typically time-consuming, expensive and therefore are seldom performed, expert knowledge is a valuable source of information to discuss niche predictions.

The obtained spatial extents are quite consistent with experts' knowledge (see e.g. Rivas-Martínez & Sáenz Laín, 1991), specially for the CH-Tukey approach (e.g. *Q. robur* in the NW of the Peninsula, *Q. faginea* subsp. *broteroi* in the SW, *Q. suber* mainly in the W, while the other three species occur mostly in the interior parts of the Peninsula). For the BIOCLIM procedure such spatial coincidence is not so evident and can only be approximated if the lower levels of suitability are discarded.

For some taxa (*Q. suber*, *Q. rotundifolia* and both *Q. faginea* subspecies) the maps show a low density of occurrences in regions of predicted high suitability. This may result, among other possible reasons, from: i) missing relevant environmental variables (e.g. soil characteristics or other climatic variables) in the model; ii) incomplete sampling of occurrences in the corresponding geographic areas; iii) some ecological factors that lead to the exclusion of the species from that area (e.g. competition, parasitism, disease); or even iv) local extinction due to human overuse of *Quercus* forests, conversion to agricultural fields, managed forests, pastures or urban fabric. In the case of *Q. faginea* subsp. *faginea* and *Q. suber* it is quite plausible that a soil-related environmental variable is missing in the model, as the former is known to prefer base-rich or clay-rich soils, while the latter is calcifugous (Franco, 1990). In the case of *Q. faginea* subsp. *broteroi*, the biogeographic disjunction that the occurrence points show (i.e. one cloud of occurrences concentrated in the west of Portugal and a second cloud in the central part of Spain) may lead to an unoccupied niche core. The findings of Silva et al. (2014), mentioned above, can be used to detect those less frequent distribution patterns within the niche, previously to modelling.

In general, with the CH-Tukey approach, the maximum suitability obtained for each of the studied *Quercus* taxa is relatively low. This is not problematic *a priori* and might give further structural information on the species niche and on the study area, as: i) the maximum suitability level of 1, might not be achieved for some configurations of occurrence points in the environmental space; ii) when achieved, the volume of the region with suitability 1 can be very small; iii) even when the volume of the Tukey median region is conspicuous, it may correspond to combinations of the *n* variables that do not come into existence within the study area.

(Figure 4)

(Figure 5)

Quantifying the environmental niche overlaps of Iberian *Quercus* taxa

In the space defined by the four environmental variables (isothermality, minimum temperature of coldest month, annual precipitation and precipitation of driest quarter), we uniformly generated N=45x10^6 points within the bounding box of the occurrences of all taxa, and computed the Tukey suitability of each of these points with respect to the occurrences of each taxon. To assess niche overlap for each pair of taxa, A and B, we estimated the volumes $*_{+,}$, for $"4, 6\$ \neq "0,0\$$, as the number of points with Tukey suitability $!_{+}"\#\$$, for taxon A, and Tukey suitability $!_{,}"(\$$, for taxon B, divided by *N*. Results for the two proposed overlap indices (*wJ* and *cos*) are presented in Table 1.

(Table 1)

Both *wJ* and *cos* showed that *Q. robur* and *Q. pyrenaica* presented the highest overlap, while *Q. robur* and *Q. faginea* subsp. *broteroi* presented the lower overlap.

In order to visualize graphically the relative positions of occurrences and iso-depth regions of the pairs of taxa that, in the 4-dimensional space, were more separated (*Q. robur* vs. *Q. faginea* subsp. *broteroi*), and less separated (*Q. robur* vs. *Q. pyrenaica*)

(Table 1), we projected the occurrences of each of these taxa in the 2-dimensional environmental space of variables minimum temperature of the coldest month and precipitation of the driest quarter (Figure 6). Computing the overlap indices in this 2-dimensional space, we obtained *wJ*= 0.9177 and *cos*= 0.3450 for the first pair (*Q. robur*, *Q. pyrenaica*) and *wJ*= 0.3274 and *cos*= 0.0283 for the second pair (*Q. robur*, *Q. faginea* subsp. *broteroi*). These values agree with the graphs of Figure 6, where it is shown a clear overlap of the niches of *Q. robur* and *Q. pyrenaica* while the niche of *Q. faginea* subsp. *broteroi* is mainly concentrated in the peripheral areas of the *Q. robur* niche.

In general, Table 1 shows values that are coherent with the known regional co-occurrence patterns of the studied *Quercus* (see e.g. Rivas-Martínez & Sáenz Laín, 1991; Capelo *et al*., 2007). For example, *Q. robur* co-occurs frequently with *Q. pyrenaica* in the highlands of the northwest of the Iberian Peninsula, and with *Q. suber* in the lowlands. *Q. pyrenaica* co-occurs with *Q. rotundifolia* in the continental plateaux of the Iberian Meseta and with *Q. suber* in the warmer areas of the centre-south. *Q. rotundifolia* and *Q. suber* co-occur in the SW of the Peninsula. The referred pairs correspond both to the highest values of the indices shown in Table 1 and to intermingled vegetation series mapped as a mosaic in the work of Capelo *et al.* (2007).

(Figure 6)

Comparing niche geometries

In order to illustrate and compare niche geometries produced by different approaches, we compared outputs from CH-Tukey to BIOCLIM, and also MaxEnt (Phillips *et al.*, 2006). MaxEnt is one of the most used modelling algorithms, to obtain species environmental suitability maps (Ahmed *et al.*, 2015). We ran these three models using the two variables referred above (minimum temperature of the coldest month and precipitation of the driest quarter) and depicted their outputs both in the environmental and in the geographic space (Figure 7) for *Q. robur, Q. pyrenaica* and *Q. faginea* subsp. *broteroi*. Default MaxEnt parameterization was used.

Inspecting Figure 7 (B, C, D) we can visualize how BIOCLIM relies on the assumption that ecological variables are independent in their action on the species, producing rectilinear niche geometries, while CH-Tukey relaxes it, adjusting its outer boundary and the inner suitability levels to the known occurrences. Both BIOCLIM and CH-Tukey are built upon the "range" notion, implicit in Hutchinson's seminal work, i.e., if you select two points from the species niche representing environmental conditions known to be suitable for that species, you should not expect a decrease in suitability between these two points, at least while only those variables are considered. MaxEnt neglects this, producing not only unbounded geometries, but also allowing (visible for the three tested taxa) a decrease in suitability between two points of the niche. Although relying on unimodal geometries, BIOCLIM and CH-Tukey approximate better Hutchinson's notion of optimal versus suboptimal niche regions than the geometry produced by MaxEnt, which proposes high suitability values for unbounded or extrapolative areas (see “clamping” in Merow *et al.*, 2013; check *Q. faginea* subsp. *broteroi* output in Figure 7D, values of min Temperature greater than 10.0 ºC or smaller than -10.0 ºC and Precipitation around 50 mm). This is in agreement with the known limitations of correlative models concerning extrapolation (Merow *et al*., 2013). According to the same authors, MaxEnt can produce intricate model outputs of hard ecological interpretation, especially when regularization is used. This is apparent in the sudden shifts in suitability observed in the MaxEnt output (within the species niche). Even if some shifts could relate to interesting or interpretable facts (e.g. coming from the background data), we believe that they deviate this modelling technique from the seminal ecological theory.

(Figure 7)

A thorough and careful setting of the input parameters in MaxEnt can improve the ecological interpretation of the output (Merow *et al*., 2013), but such parameterization is not always achievable using presence-only data coming from usual databases. Even so, MaxEnt produces geographical maps that are closely linked to the known occurrences of points (see Figure 7A for the occurrence points of each taxa), which is expected from a correlative approach. This tends to hinder fundamental discussions, such as: i) niche boundaries and niche optimality; ii) interpolation vs. extrapolation of the observations; and iii) possible discrepancies between niche optimality and the geographic distribution (e.g. niche is highly suitable according to the studied environmental variables, but there are no occurrences in the corresponding geographic regions, possibly indicating that other important variables lack in the model).

Incorporating putative niche voids directly in the modelling approaches has been considered useful, especially while addressing niches using real-world examples (Walker & Cocks, 1991; Blonder et al., 2014). Nevertheless, even when dealing with real-world examples, general niche premises and properties should hold (e.g. the "range" notion, niche boundaries and optimality) and these can be consistently addressed and explored using appropriate methodological approaches such as BIOCLIM and CH-Tukey.

CONCLUDING REMARKS

The CH-Tukey procedure allows the construction of environmental suitability maps, as shown for the selected *Quercus* taxa, differentiating iso-suitability levels within the CH-approach. Such levels are drawn from occurrence data only, in the selected environmental hyperspace, using normalized Tukey depth, and can, subsequently, be used to derive overlap measures coherently.

Carpenter *et al.* (1993) listed four disadvantages linked with the CH-approach, two related to computational difficulties which are no longer a problem for current computers. The other two were related to the tightly constrained envelope produced by the CH-approach and the exclusion of many similar sites (close to, but outside, the convex hull boundary). But this may be arguable since the convex hull is a natural boundary between interpolation and extrapolation (see Hirzel & Le Lay, 2008). We stress that the CH-Tukey procedure is appropriate to interpolate from the input data, as the model simply considers unsuitable all environmental combinations outside the convex hull (i.e. it does not extrapolate). Nevertheless, similar sites in the immediate outside neighbourhood of the convex hull, which are considered unsuitable in the CH-Tukey procedure, could be modelled separately with other techniques, in a coupled approach, assuming that extrapolation is acceptable.

As shown in the maps obtained with CH-Tukey methodology, it is possible that higher suitability regions have low density of occurrences (see e.g. *Q. faginea* subspecies). This results from the combinatorial-geometrical nature of the approach that does not rely so strongly on point density to predict suitability. As the procedure is geometrically interpretable, hypotheses can be sketched from the obtained patterns, given the used environmental variables, and tested subsequently. This interpretability ensures a seamless relationship with Hutchinsonian niche concept, a feature frequently asked by ecologists and epistemologists (Austin, 2002; Elith & Leathwick, 2009; Brewer *et al.*, 2016).

In conclusion, the CH-Tukey procedure presents several relevant features that can be considered advantages in some ecological niche modelling contexts, such as:

i) The modelling process is easily interpretable geometrically, with precise definition of the estimated niche boundaries.

ii) It uses only presence data and does not depend on the background data.

iii) The suitability levels are invariable with respect to the ecological variables units.

iv) The regions of the multivariate space corresponding to high Tukey depth values are robust to outliers, in accordance to what occurs with the univariate median of a set of values.

v) It produces ordinal predictions, decomposing the niche into a finite number of iso-suitability levels, permitting the computation of consistent overlap indices.

vi) A tight envelope is adjusted to the points cloud and, consequently, the independence of the ecological variables action on the species is not assumed (i.e., allowing for interactions between variables); moreover, the output of the procedure results uniquely from interpolations of the input data.

The main disadvantage of the CH-Tukey procedure is that depth computation is time demanding, and heuristic approximations may be required depending on the number of occurrence points used as input data and on the spatial resolution of the predictions. Heuristics are already implemented in package *depth* (Genest *et al.*, 2012).

ACKNOWLEDGMENTS

This work was supported by the Portuguese Foundation for Science and Technology (FCT) through the projects UID/MAT/00297/2013, CMA (Centro de Matemática Aplicada); UID/AGR/00239/2013, CEF (Centro de Estudos Florestais) under FEDER/POCI; and PTDC/AAC-AMB/ 113394/2009. T. Monteiro-Henriques was partially funded by a post-doctoral grant within Project SUSTAINSYS: Environmental Sustainable Agro-Forestry Systems - NORTE-07-0124-FEDER-0000044 (ON.2-QREN-FEDER and PIDDAC-FCT-MEC). D. Alagador was funded through the project PTDC/AAG-GLO/3979/2014 via FEDER through the COMPETE program and national funding. D. Alagador and F. C. Aguiar are also supported with FCT post-doctoral fellowships (SFRH/BPD/104077/2014 and SFRH/BPD/112417/2015, respectively).

AUTHORS' CONTRIBUTIONS

JOC, TMH, MJM and PCS conceived the ideas and designed methodology; MJM, TMH and MLC dealt with computational issues; TMH, PA, MJM and FCA collected, analysed and interpreted the data; JOC, TMH, MJM and PCS led the writing of the manuscript. DA, AF and MC revised it and contributed to the development of ideas. All authors contributed critically to the manuscript.

REFERENCES

Ahmed, S.E., McInerny, G., O'Hara, K., Harper, R., Salido, L., Emmott, S. & Joppa, L.N. (2015) Scientists and software – surveying the species distribution modelling community. Diversity and Distributions, **21**, 258–267.

Anthos (2011) Sistema de información de las plantas de España. Real Jardín Botánico, CSIC- Fundación Biodiversidad. Downloaded in November 16, 2011.

Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101–118.

Blonder, B. (2017) Hypervolume concepts in niche- and trait-based ecology. *Ecography*, **40**, 001–013.

Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hypervolume. *Global Ecology and Biogeography*, **23**, 595–609.

Bond, W.J. & Midgley, J.J. (2001) Ecology of sprouting in woody plants: the persistence niche. *Trends in Ecology & Evolution*, **16**, 45–51.

Booth, T.H., Nix, H.A, Busby, J.R. & Hutchinson, M.F. (2014) BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and Distributions*, **20**, 1–9.

Brewer, M.J., O’Hara, R.B., Anderson, B.J., & Ohlemüller, R. Plateau: A New Method for Ecologically Plausible Climate Envelopes for Species Distribution Modelling. *Methods in Ecology and Evolution*, **7**, 1489–1502.

Bugalho, M.N., Caldeira, M.C., Pereira, J.S., Aronson, J. & Pausas, J.G. (2011) Mediterranean cork oak savannas require human use to sustain biodiversity and ecosystem services. *Frontiers in Ecology and the Environment*, **9**, 278–286.

Busby, J.R. (1991) *BIOCLIM - A Bioclimate Analysis and Prediction System*. *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (ed. by C.R. Margules and M.P. Austin), pp. 64–68. CSIRO.

Capelo, J., Mesquita, S., Costa, J.C., Ribeiro, S., Arsénio, P., Neto, C., Monteiro-Henriques, T., Aguiar, C., Honrado, J., Espírito-Santo, D. & Lousã, M.F. (2007) A methodological approach to potential vegetation modeling using GIS techniques and phytosociological expert-knowledge: application to mainland Portugal. *Phytocoenologia*, **37**, 399–415.

Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, **2**, 667–680.

Cerdeira, J.O., Monteiro-Henriques, T., Martins, M.J., Silva, P.C., Alagador, D. & Franco, A.M.A. (2014) Mathematical contributions to link biota with environment. *Journal of Vegetation Science*, **25**, 1148–1153.

Cerqueira, M., Gomes, L., Tarelho, L. & Pio, C. (2013) Formaldehyde and acetaldehyde emissions from residential wood combustion in Portugal. *Atmospheric Environment*, **72**, 171–176.

Colwell, R. & Futuyma, D. (1971) Measurement of Niche Breadth and Overlap. *Ecology*, **52**, 567–576.

Correia, R.A., Bugalho M.N., Franco A.M.A. & Palmeirim J.M. Contribution of Spatially Explicit Models to Climate Change Adaptation and Mitigation Plans for a Priority Forest Habitat. *Mitigation and Adaptation Strategies for Global Change*, *in press* 1–16.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Jin Li, Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. Overton, J., Peterson, A.T. & Phillips, S.J. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Franco, J. do A. (1990) *3. Quercus L.*. Flora iberica: Plantas vasculares de la Península Ibérica e Islas Baleares* (ed. by S. Castroviejo, M. Laínz, G. López González, P. Montserrat, F. Muñoz Garmendia, J. Paiva, and L. Villar), pp. 15–36. Real Jardín Botánico - Consejo Superior de Investigaciones Científicas, Madrid.

Franklin, J. (1995) Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, **19**, 474–499.

Franklin, J. (2010) *Mapping Species Distributions: Spatial Inference and Prediction*, Cambridge University Press, Cambridge.

Genest, M., Masse, J.-C. & Plante, J.-F. (2012) *depth: Depth functions tools for multivariate analysis (version 2.0-0)*, http://CRAN.R-project.org/package=depth.

Gonçalves, C., Alves, C. & Pio, C. (2012) Inventory of fine particulate organic compound emissions from residential wood combustion in Portugal. *Atmospheric Environment*, **50**, 297–306.

Grubb, P.J. (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews*, **52**, 107–145.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2016) *dismo: Species distribution modeling*, http://CRAN.R-project.org/package=dismo.

Hirzel, A.H. & Le Lay, G. (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, **45**, 1372–1381.

Hurlbert, S. (1978) Measurement of Niche Overlap and Some Relatives. *Ecology*, **59**, 67–77.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

Manos, P.S., Doyle, J.J. & Nixon, K.C. (1999) Phylogeny, biogeography, and processes of molecular differentiation in Quercus subgenus Quercus (Fagaceae). *Molecular Phylogenetics and Evolution*, **12**, 333–349.

Merow, C., Smith, M.J., Silander Jr, J.A. 2013 A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**, 1058-1069.

Nix, H.A. (1986) *A biogeographic analysis of Australian elapid snakes*. *Atlas of elapid snakes of Australia* (ed. by R. Longmore), pp. 4–15. Australian Government Publishing Service, Canberra.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution*, **23**, 149–158.

Peterson, A.T., Soberon, J. & Sanchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.

Phillips, S.J., Anderson, R.P., Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Pianka, E.R. (1973) The Structure of Lizard Communities. *Annual Review of Ecology and Systematics*, **4**, 53–74.

Raxworthy, C.J., Ingram, C.M., Rabibisoa, N. & Pearson, R.G. (2007) Applications of Ecological Niche Modeling for Species Delimitation: A Review and Empirical Evaluation Using Day Geckos (Phelsuma) from Madagascar. *Systematic Biology*, **56**, 907.

Rivas-Martínez, S. & Sáenz Laín, C. (1991) Enumeración de los Quercus de la Península Ibérica. *Rivasgodaya*, **6**, 101–110.

Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., Arellano, A., Coleman, M., Hauser, L., Hess, W.R., Johnson, Z.I., Land, M., Lindell, D., Post, A.F., Regala, W., Shah, M., Shaw, S.L., Steglich, C., Sullivan, M.B., Ting, C.S., Tolonen, A., Webb, E.A., Zinser, E.R. & Chisholm, S.W. (2003) Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. *Nature*, **424**, 1042–1047.

Rodríguez-Estévez, V., García, A. & Gómez, A.G. (2009) Characteristics of the acorns selected by free range Iberian pigs during the montanera season. *Livestock Science*, **122**, 169–176.

Root, R.B. (1967) The Niche Exploitation Pattern of Blue-Gray Gnatcatcher. *Ecological Monographs*, **37**, 317–350.

Roughgarden, J. (1972) Evolution of Niche Width. *The American Naturalist*, **106**, 683–718.

Schoener, T. (1974) Resource Partitioning in Ecological Communities. *Science*, **185**, 27–39.

Silva, P.C., Cerdeira, J.O., Martins, M.J. & Monteiro-Henriques, T. (2014) Data depth for the uniform distribution. *Environmental and Ecological Statistics*, **21**, 27–39.

Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology & Evolution*, **19**, 605–611.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007) A comparative evaluation of presence-only methods for modelling species distribution. *Diversity & Distributions*, **13**, 397–405.

Tukey, J.W. (1975) Mathematics and the picturing of data. *Proceedings of the International Congress of Mathematicians*, **2**, 523–531.

Villéger, S., Novack-Gottshall, P.M. & Mouillot, D. (2011) The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. *Ecology Letters*, **14**, 561–568.

Walker, P.A. & Cocks, K.D. (1991) HABITAT: A Procedure for Modelling a Disjoint Environmental Envelope for a Plant or Animal Species. *Global Ecology and Biogeography Letters*, **1**, 108–118.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution*, **62**, 2868–2883.

Werner, E. & Gilliam, J. (1984) The Ontogenetic Niche and Species Interactions in Size Structured Populations. *Annual Review of Ecology and Systematics*, **15**, 393–425.

Whittaker, R.H. & Levin, S.A. eds. (1975) *Niche: Theory and Application*, Dowen, Hutchinson & Ross, Pennsylvania.

Zuo, Y. & Serfling, R. (2000) General notions of statistical depth function. *The Annals of Statistics*, **28**, 461–482.

BIOSKETCH

The authors include mathematicians, ecologists and biologists from several Universities and Research Centers, involved in a three-year multidisciplinary project (S-PerforMA: Species Performance Modelling Algorithm - http://www.isa.utl.pt/proj/S-PerforMa/) aiming at exploring issues related with the concept of niche using combinatorial and geometric tools.

TABLES

Table 1. Overlap (*wJ* and *cos* indices) of all pairs of *Quercus* taxa in the four dimensional space defined by the used environmental variables: isothermality, minimum temperature of coldest month, annual precipitation and the precipitation of driest quarter.

	wJ					
cos	***faginea broteroi***	0.3388	0.3238	0.1128	0.0460	0.2190
	0.0349	***faginea faginea***	0.5419	0.2802	0.1214	0.2876
	0.0700	0.2044	***rotundifolia***	0.8391	0.5262	0.7920
	0.0130	0.0463	0.3041	***pyrenaica***	0.8874	0.8810
	0.0043	0.0084	0.0516	0.3542	***robur***	0.7605
	0.0772	0.0426	0.1509	0.3141	0.2079	***suber***

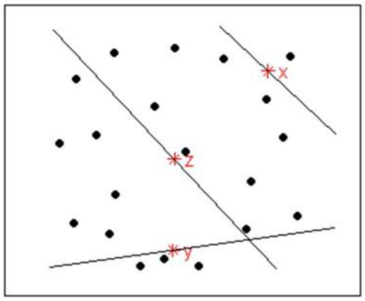


Figure 1. Illustration of the Tukey interiority of points x, y and z (represented by ∗), with respect to the set *P* of 20 points (represented by), as the number of points of *P* in the side of the line passing through ∗ which has the minimum number of points. Interiorities of x, y and z are 1, 3 and 9, respectively; $d_x(P) = 1/20$, $d_y(P) = 3/20$, and $d_z(P) = 9/20$.

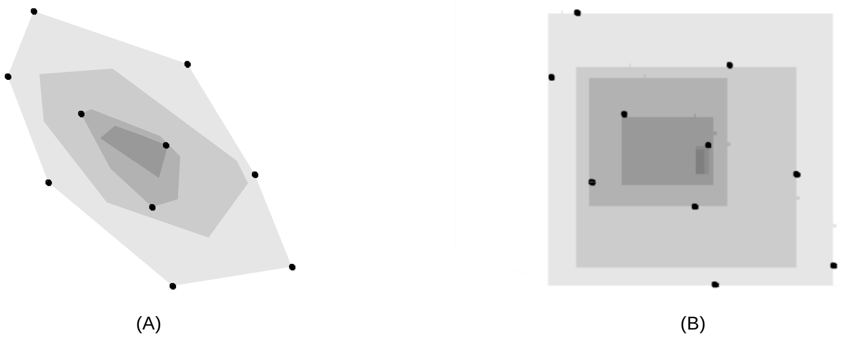


Figure 2. The 10 points represent occurrences of a species in a 2-dimensional environmental space. Shades of grey define areas of “iso-suitability”. The suitability levels increase from the white (=0) to the dark grey areas. The left panel (A) refers to CH-Tukey and the right panel (B) refers to BIOCLIM.

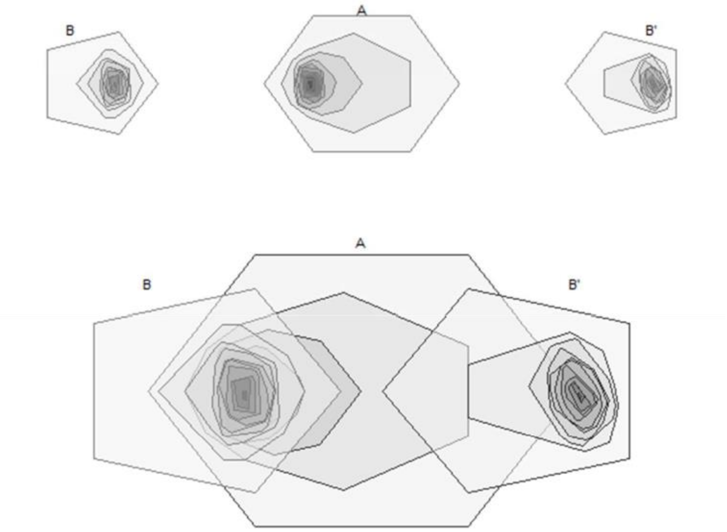
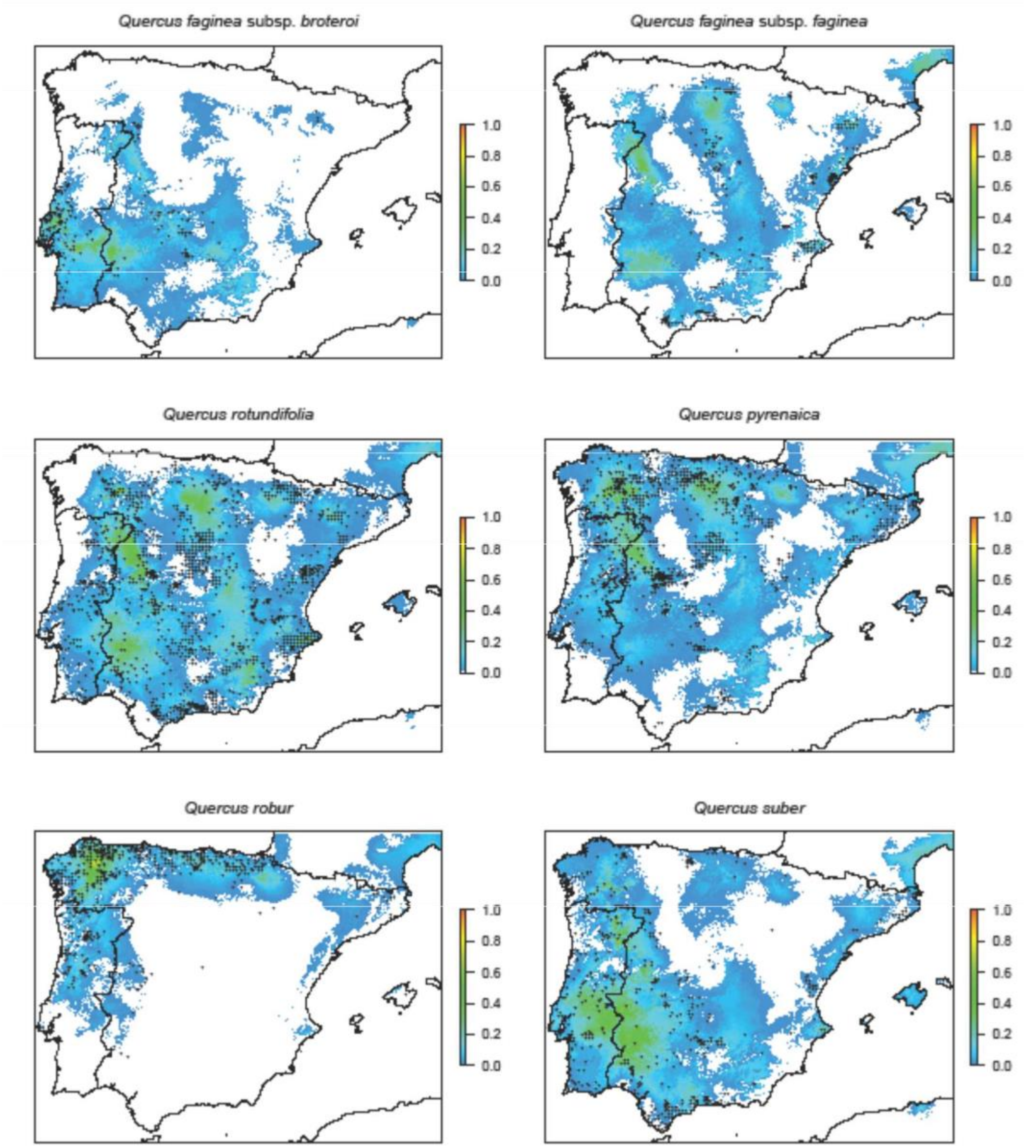


Figure 3. Niches and iso-depth regions of three species (A, B and B') in the same environmental space. Top row:

niches depicted isolated. Bottom row: overlapped niches.



Quercus faginea subsp. *broteroi*

Quercus faginea subsp. *faginea*

Quercus rotundifolia

Quercus pyrenaica

Quercus robur

Quercus suber

Figure 4. Environmental suitability maps for six *Quercus* taxa in the Iberian Peninsula obtained using the CH- Tukey procedure, with four climatic variables. Dots represent occurrences. The white inland areas refer to points outside the convex hull of the occurrences of the corresponding taxon, in the considered 4-dimensional environmental space.

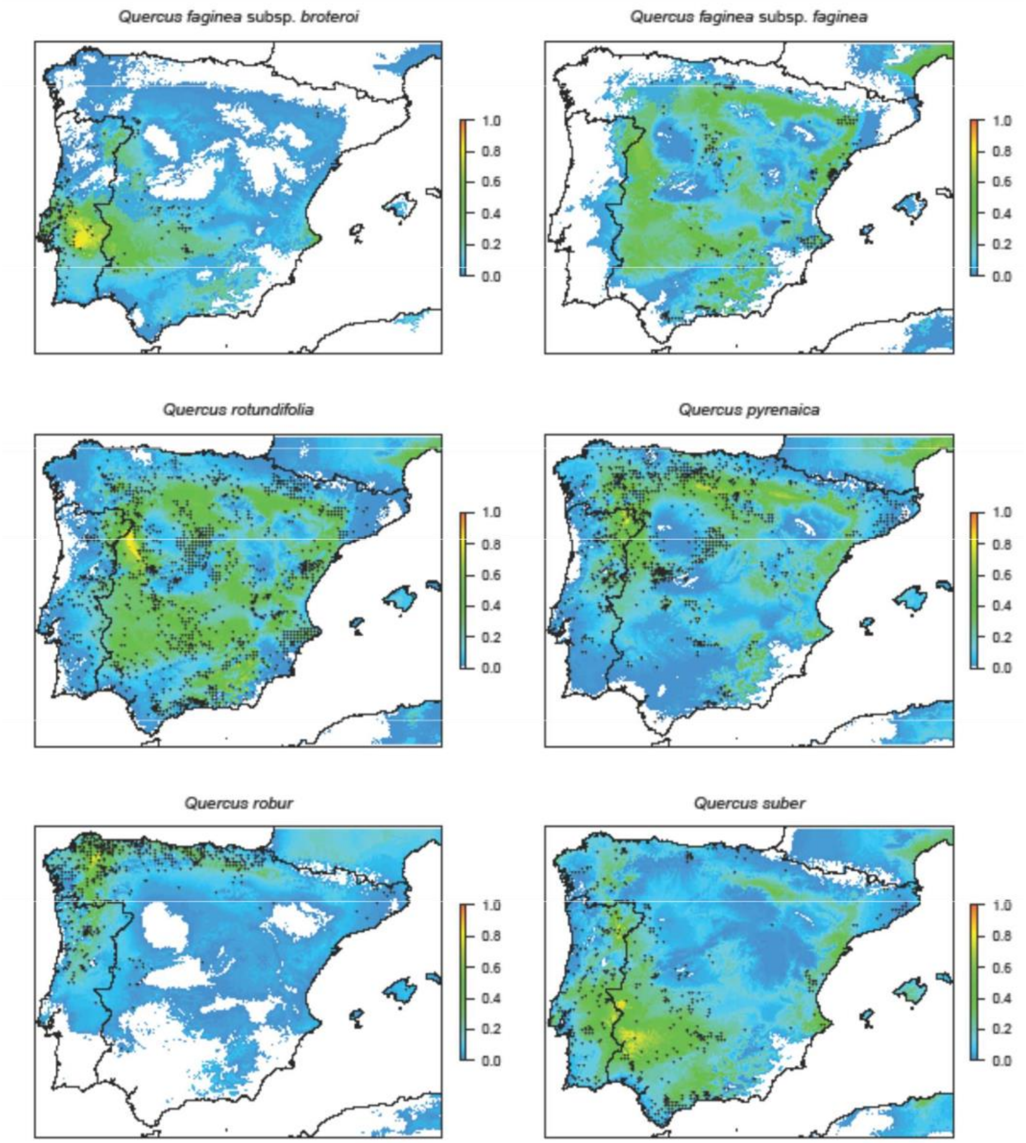


Figure 5. Environmental suitability maps for six *Quercus* taxa in the Iberian Peninsula obtained using the BIOCLIM model, with four climatic variables. Dots represent occurrences. The white inland areas refer to points outside the bounding box of the occurrences of the corresponding taxon, in the considered 4-dimensional environmental space.

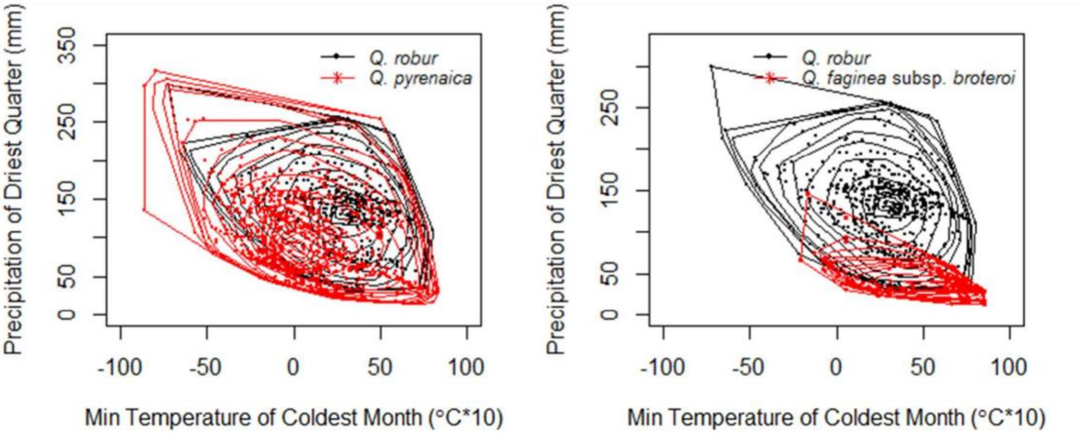


Figure 6. Occurrences and iso-depth regions of *Q. robur* and *Q. pyrenaica* in the left panel and of *Q. robur* and *Q. faginea* subsp. *broteroi* in the right panel, in the environmental space defined by the minimum temperature of the coldest month and precipitation of the driest quarter. Dots and stars correspond to occurrences, and the lines delimit iso-depth regions (regions with the same Tukey depth).

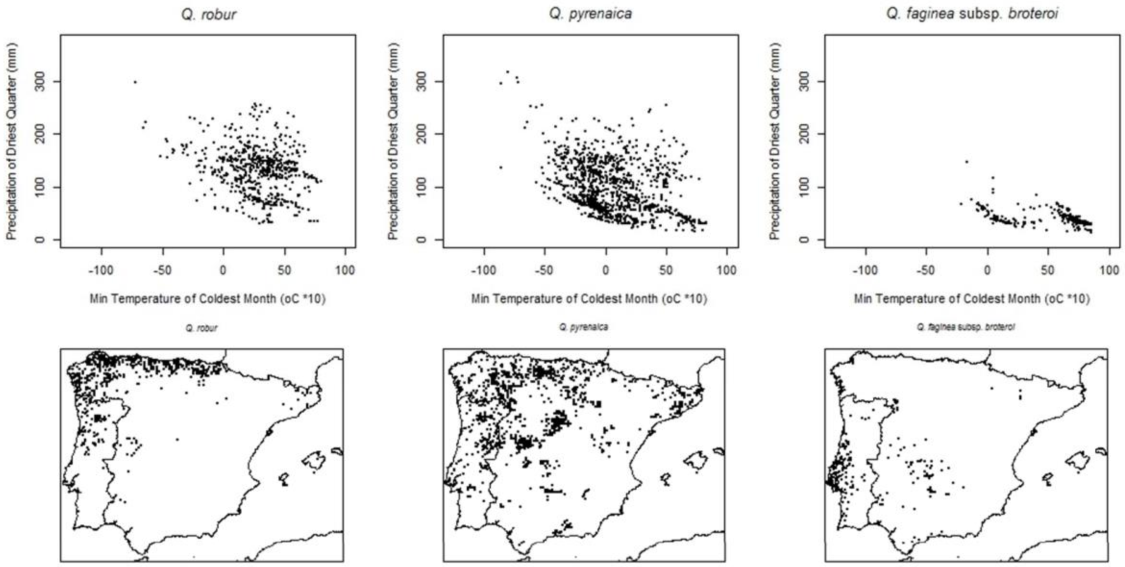


Figure 7A. Occurrences of *Q. robur*, *Q. pyrenaica* and *Q. faginea* subsp. *broteroi* in the environmental and the geographical spaces.

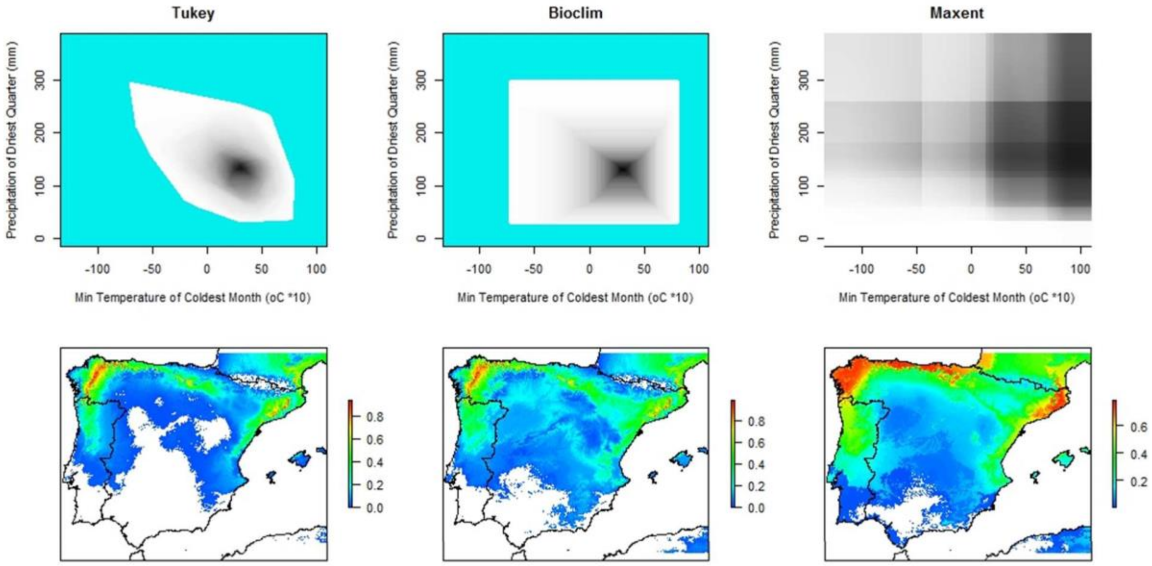


Figure 7B. Niches and suitability maps produced by CH-Tukey, BIOCLIM and MaxEnt for *Q. robur*.

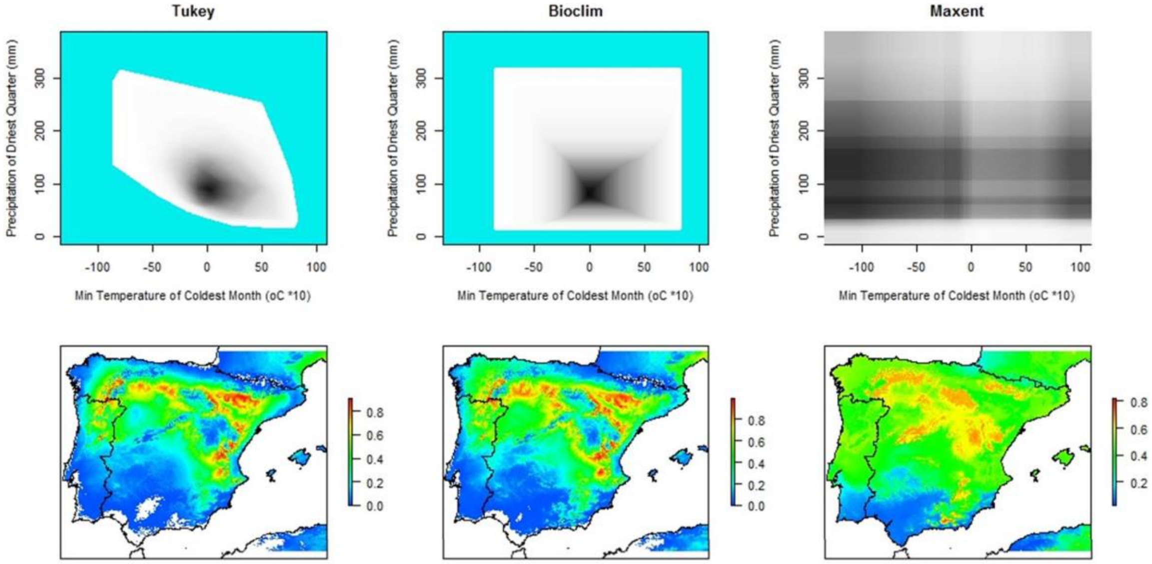


Figure 7C. Niches and suitability maps produced by CH-Tukey, BIOCLIM and MaxEnt for *Q. pyrenaica*.

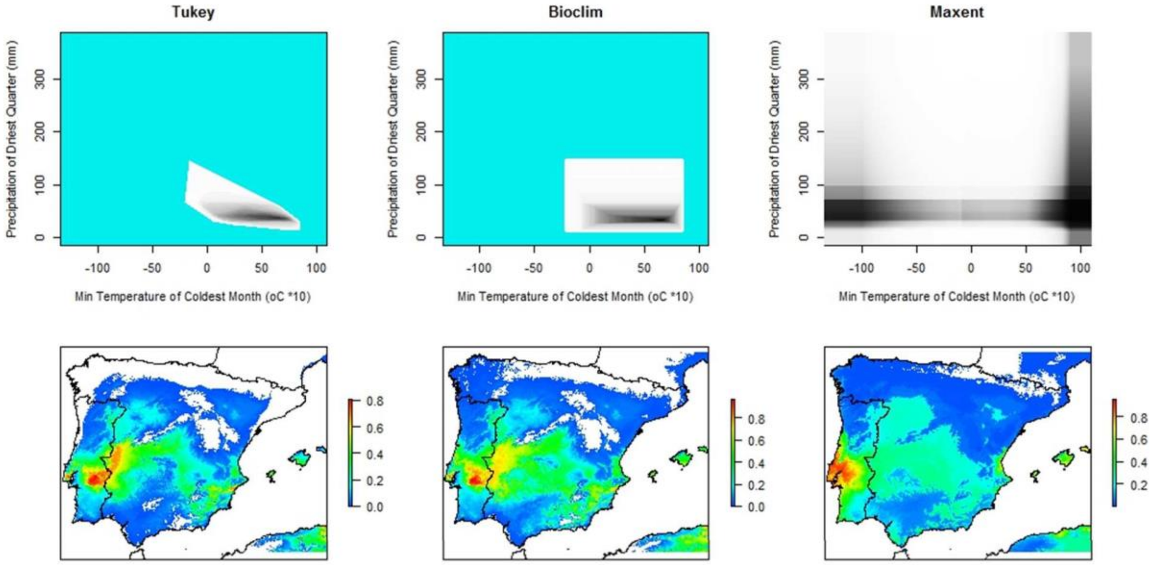


Figure 7D. Niches and suitability maps produced by CH-Tukey, BIOCLIM and MaxEnt for *Q. faginea* subsp. *broteroi*.