Biodiversity Informatics, 15, 2020, pp. 92-102

THE ABUNDANT NICHE-CENTROID HYPOTHESIS: KEY POINTS ABOUT UNFILLED NICHES AND THE POTENTIAL USE OF SUPRASPECFIC MODELING UNITS

Carlos Yañez-Arenas[1]*, Gerardo Martín[2], Luis Osorio-Olvera[3,4],
Jazmín Escobar-Luján[1], Sandra Castaño-Quintero[1],
Xavier Chiappa-Carrara[1] and Enrique Martínez-Meyer[5]
[1]Laboratorio de Ecología Geográfica, Unidad de Biología de la Conservación,
Parque Científico y Tecnológico de Yucatán, Universidad Nacional Autónoma de México,
Sierra Papacal, Yucatán 97302, México.
[2] MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease
Epidemiology, Imperial College London, UK.
[3]Biodiversity Institute, University of Kansas, Lawrence, 66045, USA.
[4]Instituto de Ecología, Universidad Nacional Autónoma de México,
Ciudad Universitaria, Mexico City 04500, México
[5]Departmento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México,
Ciudad Universitaria, Mexico City 04510, México.

Abstract. Correlative estimates of fundamental niches are gaining momentum as an alternative to predict species' abundances, particularly via the abundant niche-centroid hypothesis (an expected inverse relationship between species' abundance variation across its range and the distance to the geometric centroid of its multidimensional ecological niche). The main goal of this review is to recapitulate what has been done, where we are now, and where should we move towards in regards to this hypothesis. Despite evidence in support of the abundance-distance to niche centroid relationship, its usefulness has been highly debated, although with little consideration of the underlying theory regarding the circumstances that might break down the relationship. We address some key points about the conditions needed to test the hypothesis in correlative studies, specifically in relation to niche characterization and configurations of the Biotic-Abiotic-Mobility (BAM) framework to illustrate the problem of unfilled niches. Using a created supraspecific modeling unit, we show that species for which only a portion of their fundamental niche is represented in their area of historical accessibility (M)—i.e., when the environmental equilibrium condition is violated—it is impossible to characterize their true niche centroid. Therefore, we strongly recommend to analyze this assumption prior to assess the abundant niche-centroid hypothesis. Finally, we discuss the potential of using modeling units above the species level for cases in which environmental conditions associated with species' occurrences may not be sufficient to fully characterize their fundamental niches.

Key words: ecological niche modeling, niche centrality, abundant niche-centroid hypothesis, abundance, population density.

Introduction

Species Distribution Models (SDMs) and Ecological Niche Models (ENMs) represent a set of tools and techniques in which georeferenced records of presence (and sometimes absence) of species are statistically related with a set of environmental predictors (e.g., temperature, precipitation, elevation) to infer their ecological requirements (i.e., their ecological niche), and project them onto the geography to estimate their potential distribution (Peterson et al. 2011). The estimated distributions can be used for different purposes: to discriminate areas with and without biological potential for the species of interest (Guisan et al. 2006), evaluate potential shifts in the geographic ranges of species as a consequence of environmental changes (Thomas et al. 2004; Peterson 2006), identify regions where invasive species could

---

* Corresponding author: lichoso@gmail.com

establish (Peterson 2003; Thuiller et al. 2005), describe biodiversity patterns and carry out macroecological studies (Guisan and Rahbek 2011; Calabrese et al. 2014), among many others.

However, for certain research goals and questions, knowing the extent of occurrence is not sufficiently informative (Gaston and Rodrigues 2003; Hurlbert and Jetz 2007; Hooker et al. 2011). For instance, one of the criteria recommended in the design and establishment of natural protected areas is to maximize the area of high-quality habitat for a target species (Pearce and Ferrier 2001). In a distribution map it is not possible to identify areas of high-quality habitat since all portions of the species' distribution have the same weight. Furthermore, distributional maps alone are useless to identify whether the protection of a fraction of the distribution (without associated information on demographic aspects) is sufficient to guarantee the viability of the species (Rodrigues et al. 2004).

Population abundance or density are frequently good indicators of habitat quality (Johnson 2007), reflecting factors such as reproductive rate, longevity, carrying capacity, and susceptibility of populations to extinction. Therefore, modeling and mapping species' abundance may be more informative for many researchers and stakeholders than just estimating geographic ranges (Hobbs and Hanley 1990). When abundance/density is available from several locations, it is preferable to model it directly as a function of key environmental predictors (Boyce et al. 2001). However, obtaining abundance data is complicated and demanding, especially for rare and cryptic species (Johnston et al. 2015). Therefore, since SDMs/ENMs became popular, researchers have been interested in evaluating their capacity to infer species' abundance (e.g., Pearce and Ferrier 2001; Pearce and Boyce 2006).

Outputs of many algorithms used in SDMs/ENMs are raster layers with continuous values that are usually interpreted as environmental suitability, that is: higher values should represent better environmental conditions for species. Yet, this interpretation assumes the existence of a positive relationship between the estimated output of the algorithm and independent measures directly related to a species' biological fitness, like abundance or population density (Gil and Lobo 2017). First attempts to evaluate this relationship, via a SDM framework, failed to consistently provide strong abundance-suitability correlations (Pearce and Ferrier 2001; Pearce and Boyce 2006; Jiménez-Valverde et al. 2009). Probably, these inconsistencies may be due to the incapacity of some modeling methods to account for environmental suitability (Jiménez-Valverde et al. 2009). Later, some studies found more promising results when using correlative ENMs (VanDerWal et al. 2009; Yañez-Arenas et al. 2012; Martínez-Meyer et al. 2013; Weber et al. 2017). Among these works, the abundant niche-centroid hypothesis stands out as a key element to understand the abundance-environmental suitability relationship.

The Abundant Niche-Centroid Hypothesis

Niche theory suggests that reproduction and survival of individuals should be higher in localities placed at optimal conditions of their ecological fundamental niche ($\mathbf{N}_F$); conceptualized as an *n*-dimensional hypervolume where each dimension represents a relevant variable that acts on the organism's fitness (Hutchinson 1973; Peterson et al. 2011). This idea was initially suggested by Hutchinson (1957), but Maguire (1973) was the first one to explicitly propose that different regions of the $\mathbf{N}_F$ space should correspond to different values of the species intrinsic population growth rate (*r*). If this is true and population abundance is an expression of fitness, then, species' abundance should be explained by their position with respect to the centroid of their $\mathbf{N}_F$'s (i.e., greater abundance should be found in populations closer to the centroid and decreases towards the edges showing a negative correlation; Fig. 1; Martínez-Meyer et al. 2013).

Initially the hypothesis accumulated some empirical support: 1) Yañez-Arenas et al. (2012) and Martínez-Meyer et al. (2013) observed negative correlations between population abundance/density and the distance to the niche centroid (estimated with correlative methods) of some vertebrates; 2) Manthey et al. (2015), Ureña-Aranda et al. (2015) and Martínez-Gutiérrez et al. (2018) noted that populations tend to have positive growth rates closer to the niche center and negative at the margins. However, more recent comprehensive analyses have obtained contradicting results: Dallas et al. (2017) and Santini et al. (2019) found weak support to the niche-centrality hypothesis when tested for different taxa; in contrast Osorio-Olvera et al. (2020) observed that correlations between abundances and the distance to the niche centroid of North American birds were mostly

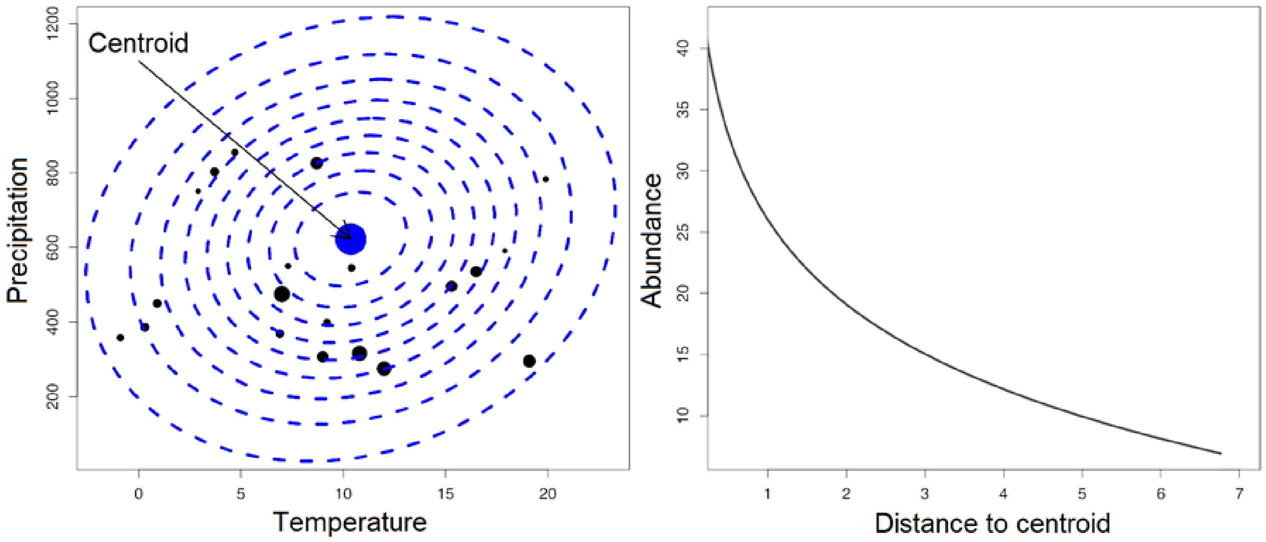

Figure 1. The abundant niche-centroid hypothesis. Left panel: bi-dimensional environmental space in which a hypothetical niche, its centroid and internal structure are shown. Black circles represent localities with different population abundance (defined by the size of the circle). Right panel: expected relationship between abundance and the distance to the niche centroid.

negative. Differences between findings could be explained by methodological artifacts and the quality of the abundance data used (Knouft 2018; Soberón et al. 2018). In any case, these studies tested the hypothesis in all possible species for which abundance data were available. However, we consider that it is necessary, in the first place, to assess some theoretical assumptions and circumstances under which the abundant niche-centroid hypothesis may be able to explain the geographic patterns of species' abundance.

On the Problem of Unfilled Fundamental Niches

Testing the abundant niche-centroid hypothesis requires an unbiased characterization of a species' $\mathbf{N}_F$, which would allow the estimated centroid to truly represent its environmental optimum (Osorio-Olvera et al. 2019). However, estimating the $\mathbf{N}_F$ via correlative techniques is not an easy task (Peterson et al. 2011). Under many circumstances a species' niche may be geographically unfilled (Strubbe et al. 2013; Ashby et al. 2017). The **BAM** diagram is a simple heuristic tool specifically designed to summarize the potential effects of biotic interactions (**B**), abiotic suitable conditions (**A**) and mobility characteristics (**M**) on a species' distribution, and can be used to assess *ex-ante* whether it is possible to characterize the $\mathbf{N}_F$ correlatively or not (Soberón and Peterson 2005). Biological features of species and scales of analysis define relative sizes and positions of the **BAM** components that lead to different configurations, some of which allow a better characterization of the $\mathbf{N}_F$ (Saupe et al. 2012). We use the **BAM** components and configurations to exemplify situations in which distances to the niche centroid estimated from distributional data may and may not explain species abundances.

The geographic expression of a species' $\mathbf{N}_F$ is termed the existing fundamental niche ($\mathbf{N}_F$*), which is equivalent to **A** in the **BAM** framework (Soberón and Peterson 2005; Soberón 2010). Relationships between these terms can vary among species and geographical scales. If a species' $\mathbf{N}_F = \mathbf{N}_F^* = \mathbf{A} \subset \mathbf{M}$, and **B** has no significant effects on constraining the occupied geography (**G***o*, or the intersection of B, A and M in the traditional configuration), then niche centrality determined correlatively is likely to explain abundance patterns. Such **BAM** configuration is commonly called 'the Hutchinson's dream' (Saupe et al. 2012) and it is the ideal scenario to test the abundant niche-centroid hypothesis (Fig. 2; upper right panel). However, if $\mathbf{N}_F^* \subset \mathbf{N}_F$, then there will be portions of the $\mathbf{N}_F$ that do not exist in geography (**G**) and therefore cannot be characterized from species' occurrences. Here, an adequate estimation of the niche centroid will depend on the magnitude of the difference between $\mathbf{N}_F$ and $\mathbf{N}_F^*$. For instance, thermal tolerance limits of some reptiles and amphibians is roughly 10 °C above the optimum, which could constitute a general guideline for assessing the difference between $\mathbf{N}_F$ and $\mathbf{N}_F^*$ (Soberón and Ar-

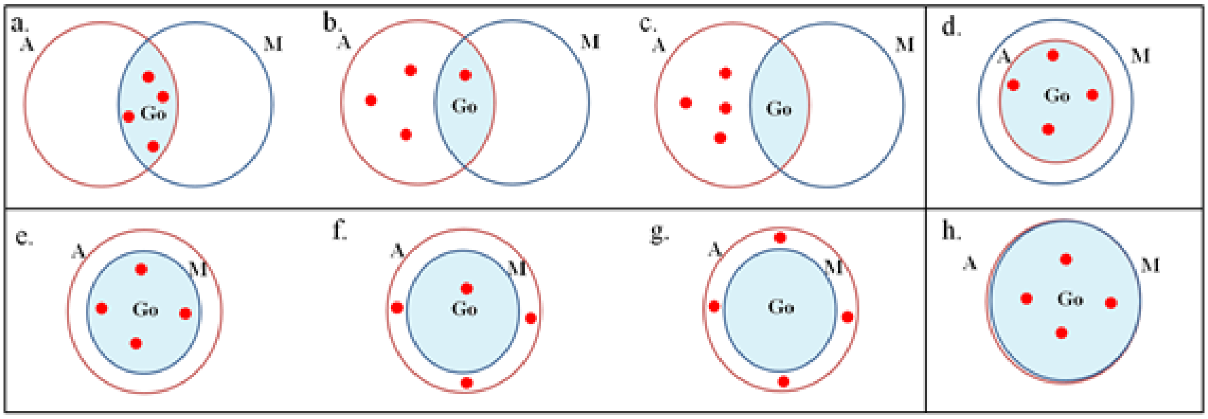

Figure 2. BAM configurations in which we show, in each letter (a – h), the potential scenarios regarding the distribution of the niche centroid conditions (red dots) in geographic space. In all cases we have omitted B to match the components accounted for in the analyses. Abiotic suitable conditions = red circles (A component of BAM); historically accessible area = blue circles (M component of BAM); Go = occupied geography. Upper left panel (a - c) = Classic BAM; lower left panel (e – g) = Wallace's dream; upper right panel (d) = Hutchinson's dream; lower right panel (h) = Full overlap.

royo-Peña 2017).

Alternatively, if only a portion of **A** exhibits biotic features suitable for the survival of a species, then **G***o* = **A** ∩ **B**, and a reduction of the $\mathbf{N}_F$ is also expected. Hutchinson (1957) was the first one to explicitly describe the above relationships in environmental space when the concept of realized niche ($\mathbf{N}_R$) was introduced. Recent empirical data reveals that key interspecific interactions may have significant effects on species' geographic ranges (e.g., Gaston 2003; Louthan et al. 2006; Gotelli et al. 2010; Ashby et al. 2017; Anderson 2017). Such cases would result in a major challenge to study abundance-centrality relationships, because it is almost axiomatic that estimated $\mathbf{N}_F$ will be reduced or biased. Therefore, research on niche centrality is highly dependent on the assumption of the so-called 'Eltonian noise' that is, when inter-specific interactions do not limit species' ranges at coarse geographical scales (Soberón and Nakamura 2009; Soberón 2010).

As described above, an important assumption to estimate a species' $\mathbf{N}_F$ via correlative modeling is that individuals can move freely across **G** and occur in all abiotic suitable locations, while being absent from all unsuitable ones, i.e., the environmental equilibrium assumption (Araújo and Pearson 2005). However, under many real-life situations, species are limited by geographical barriers and dispersal abilities (Soberón and Peterson 2005; Soberón 2010); hence, unfilled $\mathbf{N}_F$ are expected under some **BAM** configurations in which **M** is included, such as 'classic BAM' (**G***o* = **B** ∩ **A** ∩ **M**; **G***o* = **A** ∩ **M)** (Fig. 2; upper left panel) and 'Wallace's dream' (**G***o* = **M** ⊂ **B** ∩ **A**; **G***o* = **M** ⊂ **A**) (Fig. 2; lower left panel). Also, considering that **M** determines the set of regions and environments occupied by the species, the known set of presence records (**G**+) occur in geographic space (**G***o*), so the associated environments η(**G**+) are already reduced by **M**. As a consequence, niche models that have been calibrated in an **M** region will frequently under-characterize the $\mathbf{N}_F$ (Peterson and Soberón 2012), and the estimation of their centroid will be biased. An exception may occur under the 'Full overlap' configuration since **G***o* = **M** ≈ **A** and a complete characterization of the species $\mathbf{N}_F$ is possible (Fig. 2; lower right panel).

A HYPOTHETICAL EXAMPLE OF M EFFECTS

Using two virtual entities we show a simple example of how the correlation between abundance and distance to the niche centroid may be profoundly affected by an incomplete characterization of the $\mathbf{N}_F$ under two different **M** configurations.

Generating virtual entities

We created two virtual entities that are based on the biology and real data of two mussels; *Mytilus edulis* and *M. galloprovincialis*. We chose these species because they are distributed in two geographical regions separated by mainland Europe which have different environmental conditions. Both species are closely related, therefore for some niche axes they may have similar environmental tolerances (Lee et al. 2019). The former is native to northern Europe, and the latter occur in the Mediterranean Sea (Wonham 2004). Based on the assumption of phylogenetic niche conservatism (Harvey and Pagel 1991), we hypothesized that both mussel species could have largely similar $\mathbf{N}_F$. Therefore the **BAM** configuration of our example is **G***o* = **A** ∩ **M** (the classic BAM without **B**).

Inputs for generating the niche models

We obtained presence records of *M. edulis* (hereafter "Northern entity") and *M. galloprovincialis* ("Southern entity") from GBIF (https://www.gbif.org). All data points which were evidently miss-georeferenced (outside its known range or placed on land) were eliminated. Data were filtered for duplicate points with the "clean_dup" function of the "ntbox" R package (Osorio-Olvera et al. 2020), using a distance threshold equal to the resolution of the environmental data (below). Finally, we randomly selected 150 occurrence records of each species to remove from the dataset used to characterize $\mathbf{N}_F$. Hence, we obtained for each modeling entity two spatially-independent datasets of presence points called "d_pres" (full database without the 150 removed points) and "d_abun" (the 150 records removed from the original filtered database). We also obtained 12 marine variables from the Bio-Oracle database v2.0 (Assis et al. 2018; http://www.bio-oracle.org), at a spatial resolution of 10' (~20 km). These surfaces describe annual trends of sea surface temperature and salinity for the period 2000–2014: average, range, average maximum, average minimum, maximum, and minimum. We used a principal components analysis (PCA) to reduce multicollinearity applying the 'PCARaster' function of the 'ENMGadgets' package (Barve and Barve 2013) in R (R Core Team 2018), and retained the first three components that explained 93% of the overall variance for further analyses.

Niche models

First, we merged the "d_pres" datasets of occurrence records of both virtual entities to create the "Coupled unit". Then, using a minimum volume ellipsoid (MVE, Van Aelst and Rosseeuw 2009) we generated a hypothetical $\mathbf{N}_F$ that included 97.5% of the occurrence records (leaving out the presence records from atypical environmental conditions that might represent sink populations or undetected errors in the original data cleaning process; Fig. 3; green ellipsoid). The assumption is that $\mathbf{N}_F$ are convex, as suggested by abundant observational and physiological experimental data (Maguire 1973; Hooper et al. 2008; Angilletta 2009; Soberón and Nakamura 2009), and theoretical arguments (Drake 2015). The MVE was generated with the function "cov.rob" of the MASS package in R (R Core Team 2019), and graphed with "rgl" (Adler et al. 2019). Using the "cov_center" function of package "ntbox" (Osorio-Olvera et al. 2020), we calculated the volume and its centroid. Then we projected the Mahalanobis distance to the MVE centroid with the "mahalnobis" function of R to project the niche in geographic space (**G**) and obtained a continuous map of environmental suitability. The inputs to the "mahalanobis" function were the covariance matrix of the environmental variables and the vector of the centroid coordinates. Finally, we built the MVEs using the same described methods with the sets of presences for the Northern and Southern entities separately (Fig. 3; blue and red ellipsoids, respectively).

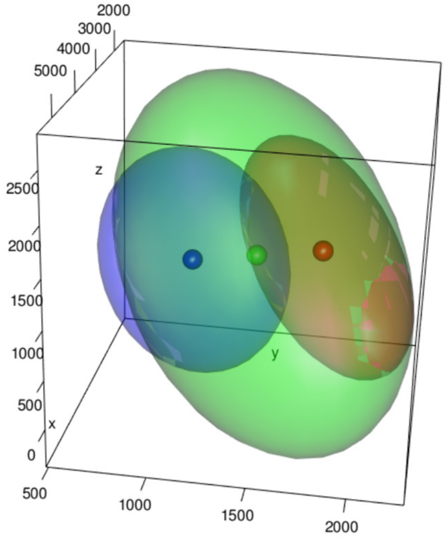

Figure 3. Minimum volume ellipsoids defined by three marine variables (Sea Surface Mean Temperature, Sea Surface Temperature Range, Sea Surface Salinity). Green ellipsoid: hypothetical fundamental niche built with presence records of both species ("Northern" and "Southern"). Blue ellipsoid: niche characterized only from occurrences of the Northern species. Red ellipsoid: niche characterized only from occurrences of the Southern species. The estimated centroid is also presented for each entity.

Abundance data

As abundance data we used the maximum number of individuals reported for either species in an area equal to the size of the grid cells. Then, we matched abundance data with the distance to the centroid of the MVE of the Coupled unit. The result was that the maximum observed abundance (12,500 virtual individuals) was perfectly, negatively correlated with the distance to the niche centroid (Fig. 4). Using this approach of virtual entities we eliminated the effects of biotic interactions, dispersion, human impact, and variability of physiological tolerances among populations.

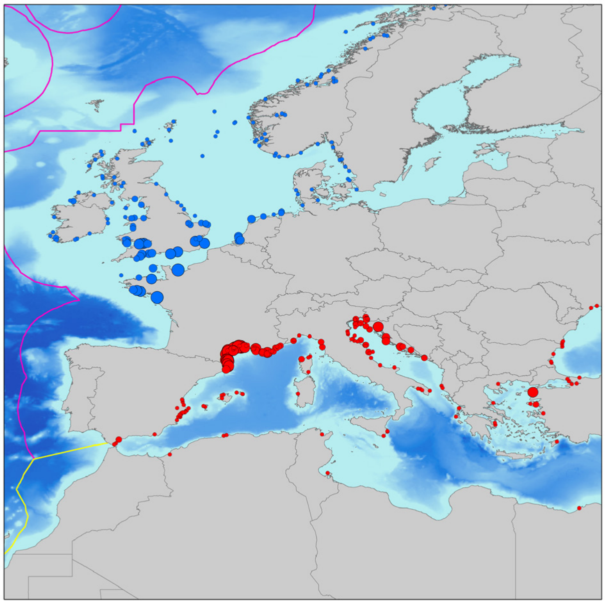

Figure 4. Distribution of the 300 locations with a hypothetical abundance value. Blue circles represent abundance data of the Northern species within its **M** (pink polygon) and red circles represent abundance data of the Southern species within its **M** (yellow polygon).

Statistical analyses

We performed Spearman correlation tests with the "cor_test" R function between abundance and distance to the centroid of the Northern and Southern entities. The two real species (and their virtual entities) have geographically disjoint distributions, separated by mainland Europe, which represents the main barrier to dispersal (**M**) and impedes their sympatry. Simultaneously, Europe interrupts the spatial continuity of the environmental conditions to which these species have historically developed physiological tolerance. Therefore, these tests are equivalent to an assessment of the effect of a niche characterization limited by **M**.

Results

The resulting environmental suitability from each MVE had a distinct geographic pattern. In the Coupled unit model, the highest environmental suitability occurred throughout most of the Cantabric Sea, a portion of the Northern Sea and a small region in the southern coast of France, in the Mediterranean Sea. In the Northern entity model, the highest environmental suitability almost exclusively occurred in the Northern Sea, while in the Southern entity there are high suitability values (small distances to the niche centroid) across the three southern European seas: Adriatic, Aegean and Tyrrhenian. Given that hypothetical abundance was derived from environmental suitability of the Coupled unit model, the correlation between this and abundance was $\rho = 0.99$. The correlation between abundance and the environmental suitability estimated for the other models were $\rho = 0.27$ and $\rho = 0.16$, for the Northern and Southern entities, respectively (Fig. 5).

Conclusion

As expected, the correlation between abundance and environmental suitability characterized from occupied niches reduced and biased by **M** was significantly lower. This example shows that accessibility has a crucial effect on the characterization of the $\mathbf{N}_F$ and the resulting environmental centroids. Under **BAM** configurations in which there is a previous suspicion that there is a strong effect of **M** over **G***o*, the abundant niche-centroid hypothesis has low expectations to explain the geographic patterns of abundance. Then, the following questions arise from these problems: 1) should species with unfilled fundamental niches be completely avoided when measuring the centrality-abundance relationship? 2) What options exist to reconstruct the $\mathbf{N}_F$ of species in which this problem exists?

THE POTENTIAL USE OF SUPRASPECIFIC MODELING UNITS

Mechanistic modeling is a good alternative to estimate and map species' $\mathbf{N}_F$ truncated by geography. These techniques are based on data from controlled experiments for measuring the physiological effects of different manipulated variables to estimate their tolerance limits (Kearney and Porter 2004; 2009). Building these models, however, requires complex and long experimental processes and equipment (Gallien et al. 2010). For these reasons, mechanistic models are usually built only with one or two environmental variables, and are unlikely to capture all the biologically relevant environmental factors for a species (Aragón et al. 2010). In addition their application is limited to species in which physiological experiments are feasible (Larson et al. 2014).

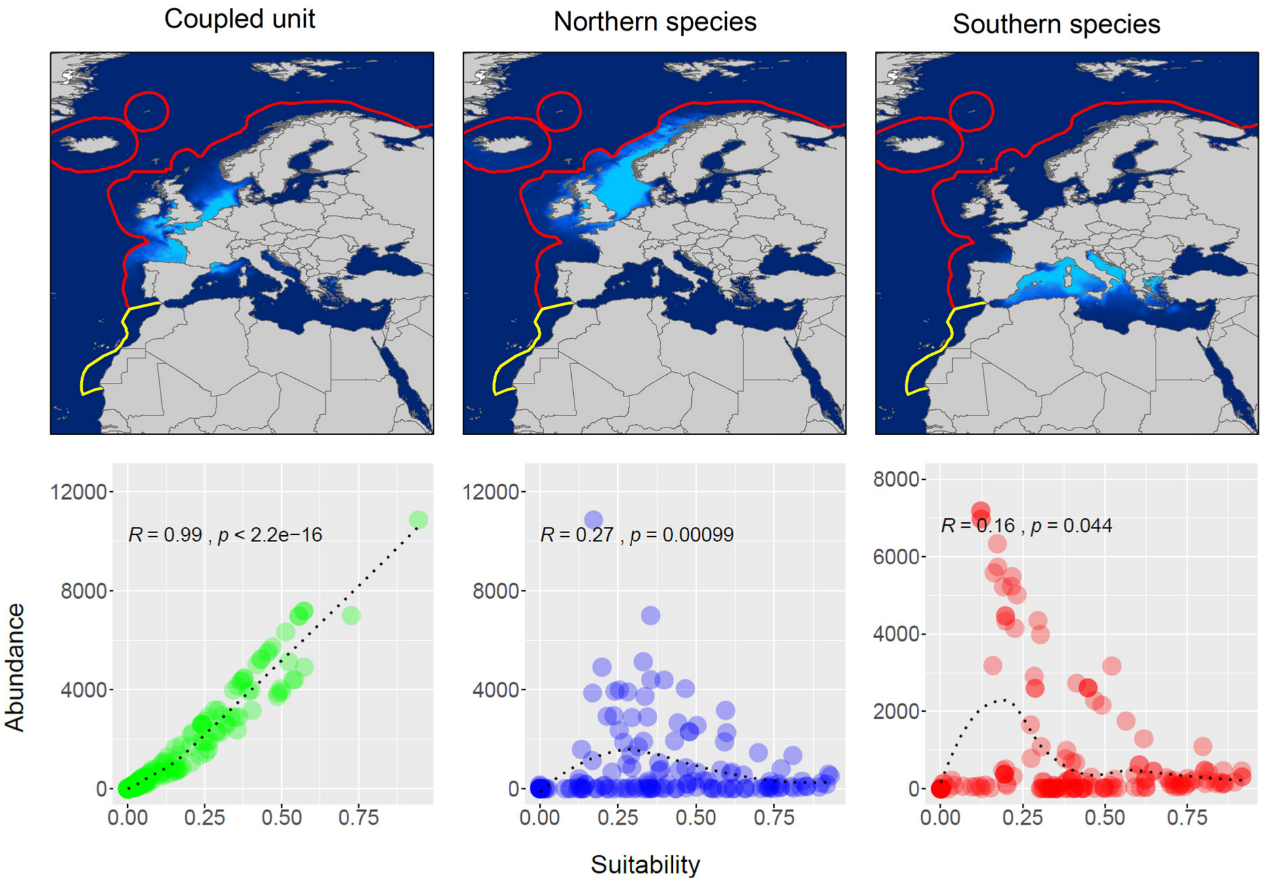

Figure 5. Top panels: environmental suitability estimated from the distances to the niche characterized by different MVEs (the gradient goes from highest to lowest suitability = from light blue to dark blue). Panels below: correlations between hypothetical abundance and environmental suitability estimated by each modeling entity.

On the other hand, characterizing the $\mathbf{N}_F$ from presence data and using correlative methods is limited under many **BAM** scenarios, as we have shown in the previous sections. Nevertheless, modeling with supraspecific entities (i.e., modeling units above species level), may represent an alternative to overcome the problem of "unfilled niches" (Qiao et al. 2017; Smith et al. 2018; Castaño et al. in press). This approach is valid under the assumption of phylogenetic niche conservatism, which states that evolutionary patterns of species with common ancestors share a substantial portion of their biological and physiological characteristics that determine their $\mathbf{N}_F$. In other words, ancestral adaptations to a set of environmental conditions tend to be preserved by descendant species (Harvey and Pagel 1991). Niche conservatism tends to break down over time, although there is vast evidence showing that at timescales comparable with speciation events, the most common pattern is that of little ecological divergence of $\mathbf{N}_F$ (climatic preferences of species tend to be phylogenetically preserved; Peterson 2011; Pavoine and Bonsall 2011). Therefore, it is possible to assume that sister species have similar $\mathbf{N}_F$ despite having natural geographical distributions with different climatic characteristics. Under this scenario, each species would inhabit different portions and combinations of the entire environmental space, but complementary regarding their $\mathbf{N}_F$. This strategy of combining presence records of sister species is known as "lumping" (Smith et al. 2018). Recent analyses of invasive species have demonstrated that lumping allows better characterization of $\mathbf{N}_F$ when it has been reduced in **G***o* by **M** or **B** (Castaño et al. in press).

Lumping, therefore, appears to be a good strategy to improve predictability of biological invasions or estimating $\mathbf{N}_F$ for evaluating the abundant niche-centroid hypothesis. The assumption is that when **BAM** configurations produce unfilled niches, lumping would allow a better approximation to estimate the $\mathbf{N}_F$ and its centroid, hence improving the chances that abundance is explained by the abundant niche-centroid hypothesis.

A third possibility is that of modeling subspecific units (e.g., subspecies, populations, environmental units) when local adaptation plays an important role in determining local abundance. Under this scenario, the single centroid would not represent the species' optimum, instead, several subspecific units would have individual optima. This may occur in species with very extensive geographic distributions and wide ecological niches (Yañez-Arenas et al. 2012).

Concluding Remarks

The abundant niche-centroid hypothesis is currently a hot topic in biogeography, macroecology and distributional ecology. To-date, at least three studies have evaluated the generality of the idea, one of them finding strong support for the hypothesis (Osorio-Olvera et al. 2020), and the other two showing contrasting results (Dallas et al. 2017; Santini et al. 2019). Some of the factors that may affect the expected negative correlation between species' abundance and the distance to their niche centroid are related to methodological issues, such as the effect of sample size and bias in occurrences used to build niche models (Yañez-Arenas et al. 2014), the quality of abundance/density data (Soberón et al. 2018, Knouft 2018), and decisions regarding the metrics used to compute environmental distances (Soberón et al. 2018). Others are inherent to the system of study. For instance, biotic interactions (e.g., high levels of competition, predation, herbivory, or parasitism), metapopulation dynamics (e.g., stochastic processes of individuals dispersal among populations, Allee effects) and human impact (e.g., direct exploitation of species, habitat destruction/modification) can decrease species' population abundances in localities that are environmentally near the centroid of the niche (Osorio-Olvera et al. 2016; Weber et al. 2017; Osorio-Olvera et al. 2019). Also, adaptation of populations to local environments would result in higher abundance in those localities (Leimu and Fischer 2008), some of which may be distributed at the edges of the niche where selection pressures are stronger (Aguirre-Liguori et al. 2017). In such cases, splitting lineages into subespecific units (e.g., environmentally similar units) and estimate niche models for each unit may be a good strategy to describe environmental relationships (Yañez-Arenas et al. 2012; Smith et al. 2018).

Some of the mentioned factors are impossible to identify prior to test the abundant niche-centroid hypothesis. However, here we demonstrate that thinking about assumptions related to the problem of unfilled niches and analyzing the potential **BAM** configuration could be used to decide *ex-ante* in which cases it is worth testing the hypothesis. We showed that **M** is important in limiting the $\mathbf{N}_F$ of species and the explanatory power of the niche structure towards abundance within its distribution. Thus, species in which environmental conditions within their **M** do not represent a subset of their $\mathbf{N}_F$ are the best candidates to study abundance-niche relationships. Finally, we introduce the idea that supraspecific units can help to overcome some of the limitations inherent to testing the abundant niche-centroid hypothesis.

Acknowledgments

We thank Town Peterson for his useful comments on the supra-specific modeling idea. We thank the editors of *Biodiversity Informatics*, Ángela Cuervo and Luis Escobar, for their effort and patience to organize this special supplement, and anonymous reviewers whose comments and suggestions improved the quality of this manuscript.

References

Adler, D., D. Murdoch, O. Nenadic, S. Urbanek, M. Chen, A. Gebhardt, B. Bolker, G. Csardi, A. Strzelecki, and A. Senger. 2019. rgl: 3D visualization using OpenGL. R package version 0.96.16.

Aguirre-Liguori, J. A., M. I. Tenaillon, A. Vázquez-Lobo, B. S. Gaut, J. P. Jaramillo-Correa, S. Montes-Hernández, and L. E. Eguiarte, L. E. 2017. Connecting genomic patterns of local adaptation and niche suitability in teosintes. Molecular Ecology 26: 4226-4240.

Anderson, R. P. 2017. When and how should biotic interactions be considered in models of species niches and distributions? Journal of Biogeography 44: 8-17.

Angilletta Jr, M. J., and M. J. Angilletta, M. J. 2009. Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press. Oxford.

Aragón, P., A. Baselga, and J. M. Lobo. 2010. Global estimation of invasion risk zones for the western corn rootworm *Diabrotica virgifera virgifera*: integrating distribution models and physiological thresholds to assess climatic favourability. Journal of Applied Ecology 47:1026-1035.

Araújo, M. B., and R. G. Pearson. 2005. Equilibrium of species' distributions with climate. Ecography 28:693-695.

Ashby, B., E. Watkins, J. Lourenço, S. Gupta, and K. R. Foster. 2017. Competing species leave many potential niches unfilled. Nature: Ecology and Evolution 1:1495.

Assis, J., L. Tyberghein, S. Bosch, H. Verbruggen, E. A. Serrão, and & O. De Clerck. 2018. Bio-ORACLE v2. 0: Extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography 27: 277-284.

Boyce, M. S., D. I. MacKenzie, B. F. Manly, M. A. Haroldson, and D. Moody. 2001. Negative binomial models for abundance estimation of multiple closed populations. Journal of Wildlife Management 65:498-509.

Calabrese, J. M., G. Certain, C. Kraan, and C. F. Dormann. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. Global Ecology and Biogeography 23: 99-112.

Castaño-Quintero, S., J. Escobar-Luján, L. Osorio-Olvera, A. T. Peterson, X. Chiappa-Carrara, E. Martínez-Meyer, and C. Yañez-Arenas. In press. Supraspecific units in correlative niche modelling improves the prediction of biological invasions. PeerJ.

Dallas, T., R. R. Decker, and A. Hastings. 2017. Species are not most abundant in the centre of their geographic range or climatic niche. Ecology Letters, 20:1526-1533.

Drake, J. M. 2015. Range bagging: a new method for ecological niche modelling from presence-only data. Journal of the Royal Society Interface 12(107): 20150086.

Gallien, L., T. Münkemüller, C. H. Albert, I. Boulangeat, and W. Thuiller. 2010. Predicting potential distributions of invasive species: where to go from here? Diversity and Distributions 16:331-342.

Gaston, K. J. 2003. The structure and dynamics of geographic ranges. Oxford University Press. Oxford.

Gaston, K. J., and A. S. Rodrigues. 2003. Reserve selection in regions with poor biological data. Conservation Biology 17:188-195.

Gil, G., and J. Lobo. 2017. ¿Son útiles los modelos de distribución realizados a gran escala para determinar la variación en la abundancia local de especies emblemáticas? El caso del Parque Nacional Iguazú. Revista Biodiversidad Neotropical 7:269-283.

Gotelli, N. J., G. R. Graves, and C. Rahbek. 2010. Macroecological signals of species interactions in the Danish avifauna. Proceedings of the National Academy of Sciences USA 107: 5030-5035.

Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. E. Zimmermann. 2006. Using niche-based models to improve the sampling of rare species. Conservation Biology 20: 501-511.

Guisan, A., and C. Rahbek. 2011. SESAM–a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. Journal of Biogeography 38:1433-1444.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology (Vol. 239). Oxford: Oxford University Press. Oxford.

Hobbs, N. T., and T. A. Hanley. 1990. Habitat evaluation: do use/availability data reflect carrying capacity? The Journal of Wildlife Management 515-522.

Hooker, S. K., A. Cañadas, K. D. Hyrenbach, C. Corrigan, J. J. Polovina, and R. R. Reeves. 2011. Making protected area networks effective for marine top predators. Endangered Species Research 13: 203-218.

Hooper, H. L., R. Connon, A. Callaghan, G. Fryer, S. Yarwood-Buchanan, J. Bigg, and R. M. Sibly. 2008. The ecological niche of *Daphnia magna* characterized using population growth rate. Ecology 89:1015-1022.

Hurlbert, A. H., and W. Jetz. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proceedings of the National Academy of Sciences USA 104:13384-13389.

Hutchinson, G. 1957. Population studies-animal ecology and demography concluding remarks. In: Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring Harbor Lab Press. Bungtown, Plainview, NY 11724:415-427.

Jiménez-Valverde, A., F. Diniz, E. B. de Azevedo, and P. A. Borges. 2009. Species distribution models do not account for abundance: the case of arthropods on Terceira Island. Annales Zoologici Fennici 46: 451-465.

Johnson, M. D. 2007. Measuring habitat quality: a review. Condor 109: 489-504.

Johnston, A., D. Fink, M. D. Reynolds, W. M. Hochachka, B. L. Sullivan, N.E. Bruns, E. Hallstein, M. S. Merrifield, S. Matsumoto, and S. Kelling. 2015. Abundance models improve spatial and temporal prioritization of conservation resources. Ecological Applications 25:1749-1756.

Kearney, M., and W. Porter. 2004. Mapping the fundamental niche: physiology, climate, and the distribution of a nocturnal lizard. Ecology 85: 3119-3131.

Kearney, M., and W. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology Letters 12:334-350.

Knouft, J. H. 2018. Appropriate application of information from biodiversity databases is critical when investigating species distributions and diversity: a comment on Dallas et al. (2018). Ecology Letters 21:1119-1120.

Larson, E. R., R. V. Gallagher, L. J. Beaumont, and J. D. Olden. 2014. Generalized "avatar" niche shifts improve distribution models for invasive species. Diversity and Distributions 20:1296-1306.

Lee, Y., H. Kwak, J. Shin, K. Seung-Chul, T. Kim, and J. K. Park. 2019. A mitochondrial genome phylogeny of Mytilidae (Bivalvia: Mytilida). Molecular Phylogenetics and Evolution 139: 106533.

Leimu, R., and M. Fischer. 2008. A meta-analysis of local adaptation in plants. PLoS ONE 3, e4010.

Louthan, A.M., D. F. Doak, and A. L. Angert. 2006. Where and when do species interactions set range limits? Trends in Ecology and Evolution 30:780–792.

Maguire, Jr, B. 1973. Niche response structure and the analytical potentials of its relationship to the habitat. American Naturalist 107:213-246.

Manthey, J. D., L. P. Campbell, E. E. Saupe, J. Soberón, C. M. Hensz, C. E. Myers, H. L. Owens, K. Ingenloff, A. T. Peterson, N. Barve, A. Lira-Noriega, and V. Barve. 2015. A test of niche centrality as a determinant of population trends and conservation status in threatened and endangered North American birds. Endangered Species Research 26: 201-208.

Martínez-Gutiérrez, P. G., E. Martínez-Meyer, F. Palomares, and N. Fernández. 2018. Niche centrality and human influence predict rangewide variation in population abundance of a widespread mammal: The collared peccary (*Pecari tajacu*). Diversity and Distributions 24:103-115.

Martínez-Meyer, E., D. Díaz-Porras, A. T. Peterson, and C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. Biology Letters 9:20120637.

Osorio-Olvera, L., Lira-Noriega, A., Soberón, J., Peterson, A. T., Falconi, M., Contreras-Díaz, R. G., … Barve, N. (2020). ntbox: an R package with graphical user interface for modeling and evaluating multidimensional ecological niches. Methods in Ecology and Evolution, 11:1199-1206.

Osorio-Olvera, L. A., M. Falconi, and J. Soberón. 2016. Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión. Revista Mexicana de Biodiversidad 87:1080-1088.

Osorio-Olvera, L., J. Soberón, and M. Falconi. 2019. On population abundance and niche structure. Ecography 42:1415-1425.

Osorio-Olvera L., Yañez-Arenas C., Martinez-Meyer E, and A. T. Peterson. 2020. Relationships between population densities and niche-centroid distances in North American birds. Ecology Letters doi: 10.1111/ele.13453.

Pavoine, S., and M. B. Bonsall. 2011. Measuring biodiversity to explain community assembly: a unified approach. Biological Reviews 86:792-812.

Pearce, J. L., and M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43:405-412.

Pearce, J., and S. Ferrier. 2001. The practical value of modelling relative abundance of species for regional conservation planning: a case study. Biological Conservation 98:33-43.

Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. The Quarterly Review of Biology 78:419-433.

Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59-72.

Peterson, A.T. 2011. Ecological niche conservatism: A time-structured review of evidence. Journal of Biogeography 38: 817–827.

Peterson, A. T., and J. Soberón. 2012. Integrating fundamental concepts of ecology, biogeography, and sampling into effective ecological niche modeling and species distribution modeling. Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology 146:789-796.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. Nakamura, and M. B. Araújo .2011. Ecological niches and geographic distributions. Princeton University Press. Princeton.

Qiao, H., A. T. Peterson, L. Ji, and J. Hu. 2017. Using data from related species to overcome spatial sampling bias and associated limitations in ecological niche modeling. Methods in Ecology and Evolution 8:1804-1812.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rodrigues, A. S., S. J. Andelman, M. I. Bakarr, L. Boitani, T. M. Brooks, R. Cowling, L. D. Fishpool, G. A. da Fonseca, K. J. Gaston, and M. Hoffmann. 2004. Effectiveness of the global protected area network in representing species diversity. Nature 428:640.

Santini, L., S. Pironon, L. Maiorano, and W. Thuiller. 2019. Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. Ecography 42: 696-705.

Saupe, E. E., V. Barve, C. E. Myers, J. Soberón, N. Barve, C. M. Hensz, A. T. Peterson, H. L. Owens, and A. Lira-Noriega. 2012. Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. Ecological Modelling 237:11-22.

Smith, A. B., W. Godsoe., F. Rodríguez-Sánchez., H. H. Wang, and D. Warren. 2018. Niche estimation above and below the species level. Trends in Ecology and Evolution 34: 260-273

Soberón, J. M. 2010. Niche and area of distribution modeling: a population ecology perspective. Ecography 33:159-167.

Soberón, J., and B. Arroyo-Peña. 2017. Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. PLoS ONE 12, e0175138.

Soberón, J., and M. Nakamura. 2009. Niches and distributional areas: concepts, methods, and assumptions. Proceedings of the National Academy of Sciences USA 106:19644-19650.

Soberón, J., and A. T. Peterson .2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2:1-10.

Soberón, J., A. T. Peterson, and L. Osorio-Olvera. 2018. A comment on "Species are not most abundant in the center of their geographic range or climatic niche". BioRxiv 266510.

Strubbe, D., O. Broennimann, F. Chiron, and E. Matthysen. 2013. Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. Global Ecology and Biogeography 22:962-970.

Thomas, C. D., A. Cameron, R.E. Green, M. Bakkenes, L. J. Beaumont, Y.C. Collingham, B. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. van Jaarsveld, G. F. Midgley, L Miles, M. A. Ortega-Huerta, T. Peterson, O. L. Phillips, and L. Hughes. 2004. Extinction risk from climate change. Nature 427:145.

Thuiller, W.,D. M. Richardson, P. PYŠEK, G. F. Midgley, G. O. Hughes, and M. Rouget. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology 11: 2234-2250.

Ureña-Aranda, C. A., O. Rojas-Soto, E. Martínez-Meyer, C. Yáñez-Arenas, R. L. Ramirez, and A. E. de los Monteros. 2015. Using range-wide abundance modeling to identify key conservation areas for the micro-endemic Bolson tortoise (*Gopherus flavomarginatus*). PLoS ONE 10:e0131452.

Van Aelst S., and P. Rousseeuw. 2009. Minimum volume ellipsoid. Wiley Interdisciplinary Reviews: Computational Statistics 1:71-82.

VanDerWal, J., L. P. Shoo, C. N. Johnson, and S. E. William. 2009. Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. American Naturalist 174: 282-291.

Weber, M. M., R. D. Stevens, J. A. F. Diniz-Filho, and C. E. V. Grelle. 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. Ecography 40:817-828.

Wonham, M. J. 2004. Mini-review: distribution of the Mediterranean mussel, *Mytilus galloprovincialis* (Bivalvia: Mytilidae), and hybrids in the northeast Pacific. Journal of Shellfish Research 23:535-544.

Yañez-Arenas, C., R. Guevara, E. Martínez-Meyer, S. Mandujano, and J. M. Lobo. 2014. Predicting species' abundances from occurrence data: effects of sample size and bias. Ecological Modelling 294:36-41.

Yañez-Arenas, C., E. Martínez-Meyer, S. Mandujano, and O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. Oikos 121:2081-2089.