

## Variation in niche and distribution model performance: The need for a priori assessment of key causal factors

E.E. Saupe, V. Barve, C.E. Myers, J. Soberón, N. Barve, C.M. Hensz, A.T. Peterson\*, H.L. Owens, A. Lira-Noriega

Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA

### ARTICLE INFO

#### Article history:

Received 14 December 2011

Received in revised form 29 March 2012

Accepted 2 April 2012

#### Keywords:

Ecological niche

Dispersal

BAM

Ecological niche model

Species distribution model

### ABSTRACT

Ecological niche models and species distribution models are becoming important elements in the toolkit of biogeographers and ecologists. Although burgeoning in use, much variation exists in implementation of these techniques, leading to considerable diversity of methodology and discussion of what is the 'best' approach. In this analysis, we explore implications of different configurations of major factors that constrain species' distributions—abiotic factors and dispersal limitation—for the success or failure of these models. We analyze variation in performance among modeling approaches as a function of the relative configuration of these two factors and the spatial extent of training region, with the result that a clear understanding of the abiotic–dispersal configuration is a prerequisite to effective model implementations; the effects of spatial extent of the training region are less consistent and clear. Model development will be powerful only when set in an appropriate and explicit biogeographic and population ecological context.

© 2012 Published by Elsevier B.V.

### 1. Introduction

In recent years, increasing effort has been invested in estimating ecological requirements of species and using those estimates to identify distributional areas. These methods are known as species distribution modeling (SDM) when emphasis is on estimating distributions of species, or ecological niche modeling (ENM) when emphasis is on niche requirements of species (Guisan and Thuiller, 2005; Marti et al., 2005; Peterson, 2006). Since 1990, growth in numbers of papers published in these fields has been almost exponential (Lobo et al., 2010). The following is a summary of relevant conceptual points, which are developed in greater detail in Peterson et al. (2011).

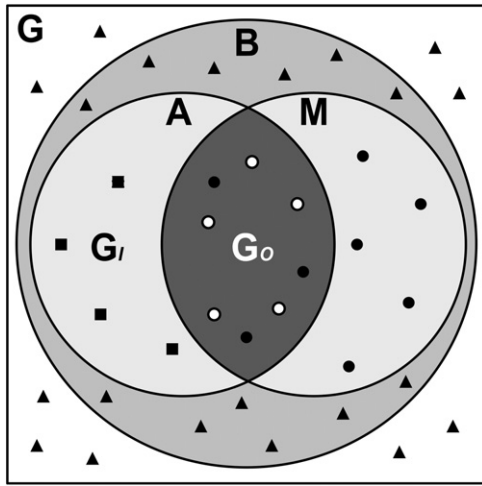
The typical correlational methods for SDM and ENM are based on finding regions in the space of environmental variables that are, in some mathematical sense, similar to conditions at sites where the species has been observed. Not surprisingly, many methods are capable of performing this task, and numerous studies have attempted to compare their performances (Guisan and Zimmermann, 2000; Segurado and Araújo, 2004; Elith et al., 2006; Pearson et al., 2006), some even trying to establish why different methods produce different answers (Elith and Graham, 2009). These questions are very relevant; to advance the discussion, however, we suggest that one must be explicit about what exactly is

being modeled, with hypotheses regarding key factors affecting the modeled object. Several kinds of distributional areas exist, each with different properties (Soberón and Peterson, 2005; Jiménez-Valverde et al., 2008; Peterson et al., 2011), so delineating the aim of modeling efforts is crucial. Known presences in relation to probable absences define the occupied area of a species, making crucial an explicit understanding of the type of absence data that are available (Lobo et al., 2010). Another often-ignored point is the extent of the region from which background data are sampled, if the algorithm being used requires such information (Hirzel and Le Lay, 2008; VanDerWal et al., 2009; Barve et al., 2011; Elith et al., 2011). Finally, one needs to understand the actual mathematical operations that different algorithms perform on the data to arrive at estimates of the object of interest (Guisan and Zimmermann, 2000; Franklin, 2009).

In this study, we compare five ENM/SDM algorithms in a novel challenge, in which we specify unequivocally: (i) the type of distributional area being modeled, (ii) the configuration of factors causing the distributions, and (iii) the way in which algorithms use "background data." Previous authors have concluded that the field of ENM/SDM is still immature, and clear guidance for selecting relevant methods cannot yet be provided (Elith and Graham, 2009). We agree with these authors that much work remains before we will understand fully the complexities of this field. However, substantial clarity can be obtained by assessing factors systematically, using virtual species as test beds. The basis of our analysis is a simple heuristic scheme, the BAM diagram (Soberón and Peterson, 2005; Soberón, 2007), which summarizes joint effects of biotic, abiotic,

\* Corresponding author. Tel.: +1 785 864 3926.

E-mail address: [town@ku.edu](mailto:town@ku.edu) (A.T. Peterson).



**Fig. 1.** A simplified BAM diagram in which the effects of biotic interactions (**B**) do not constrain distributional potential of species strongly. The circle **A** represents the parts of the world that contain the abiotic conditions required for a species' survival and growth. The circle **M** represents the region that has been accessible to the species over a relevant period of time. The intersection of these two regions is the occupied area  $G_0$ , which nonetheless includes areas of non-presence, for diverse reasons (e.g., metapopulation dynamics). The area  $G_1$  has the correct suite of environmental conditions, but has not been explored by the species. Open circles represent presence data, closed circles indicate absences due to incorrect environment, closed squares indicate absences due to lack of dispersal capacity, and triangles indicate absences owing to both incorrect conditions and limited dispersal.

and dispersal characteristics of species; we apply this framework to virtual species for which the truth is known about what factors determine each distributional area.

Fig. 1 is a simplified representation of the BAM diagram, in which only two sets of factors affect distributions of species (Grinnell, 1924; Good, 1931; Udvardy, 1969; Brown et al., 1996; Gaston, 2003): the right combination of environmental conditions (the **A** circle), and the region of geography that has been accessible to the species over a given period of time (the **M** circle). We ignore biotic interactions (**B**) for reasons discussed below. From Fig. 1, we see that three regions exist that can reasonably be regarded as the object of a modeling exercise. First, the "occupied area" (Gaston, 2003), denoted by  $G_0$ , is an area that presents the abiotic conditions that a species requires to survive and reproduce and that has been accessible to the species; by definition,  $G_0 = A \cap M$ . The second area is that which can potentially be invaded if the restrictions of **M** are relaxed (Svenning and Skov, 2004), defined as  $G_1 = A \cap M^C$ , with the **C** denoting "complement."  $G_1$  is thus the set of areas with the right environmental conditions but that is currently inaccessible to the species; this area is the focus of most modeling exercises dealing with invasive species. The third area, the union of  $G_0$  and  $G_1$ , is equivalent to **A** in the BAM diagram. The question of whether one is attempting to model  $G_0$ , **A**, or  $G_1$  is key. The relative sizes and positions of **B**, **A**, and **M** should be explicit at the outset, because, as we will see below, different configurations of the BAM diagram lead to radically different capacities for algorithms to estimate the areas of interest.

As a side note, we can use the BAM framework to distinguish between two conceptual frameworks in this emerging field. When the focus is on estimating the occupied area  $G_0$ , the study falls into the realm of SDM. Modeling  $G_0$  requires information not only about favorable conditions for the species (i.e., its fundamental ecological niche), but also about factors that restrict its spread (biotic and geographic factors constraining dispersal) or overpredictions will result (Peterson et al., 1999). When the focus is on estimating **A** or  $G_1$ , only the favorable conditions (and biotic circumstances for the case of  $G_1$ ) need to be estimated, which can be projected in

geographic space; these potential distributional areas are the subject of ENM.

Consider now the different types of absence data that the above schemes may require (Peterson, 2006; Lobo et al., 2010). In Fig. 1, several types of "absence" data are shown, but the circles are absences owing to lack of suitable environmental conditions (black circles) within space that has been accessible to the species (Barve et al., 2011) or occupancy dynamics (white circles) within completely suitable and accessible areas (Hortal et al., 2010). The squares, however, are absences under suitable conditions, but at sites where the species is not present owing to inaccessibility. Black triangles are absences resulting from both conditions acting simultaneously. It should be obvious that the biological meaning of these three classes of absences is very different, particularly if one is attempting to model  $G_1$  or **A**. Therefore, as has been emphasized previously (Barve et al., 2011), clear a priori hypotheses about **M** should be an integral part of the modeling exercise, despite the fact that no widely used SDM/ENM algorithm requests information explicitly about this region.

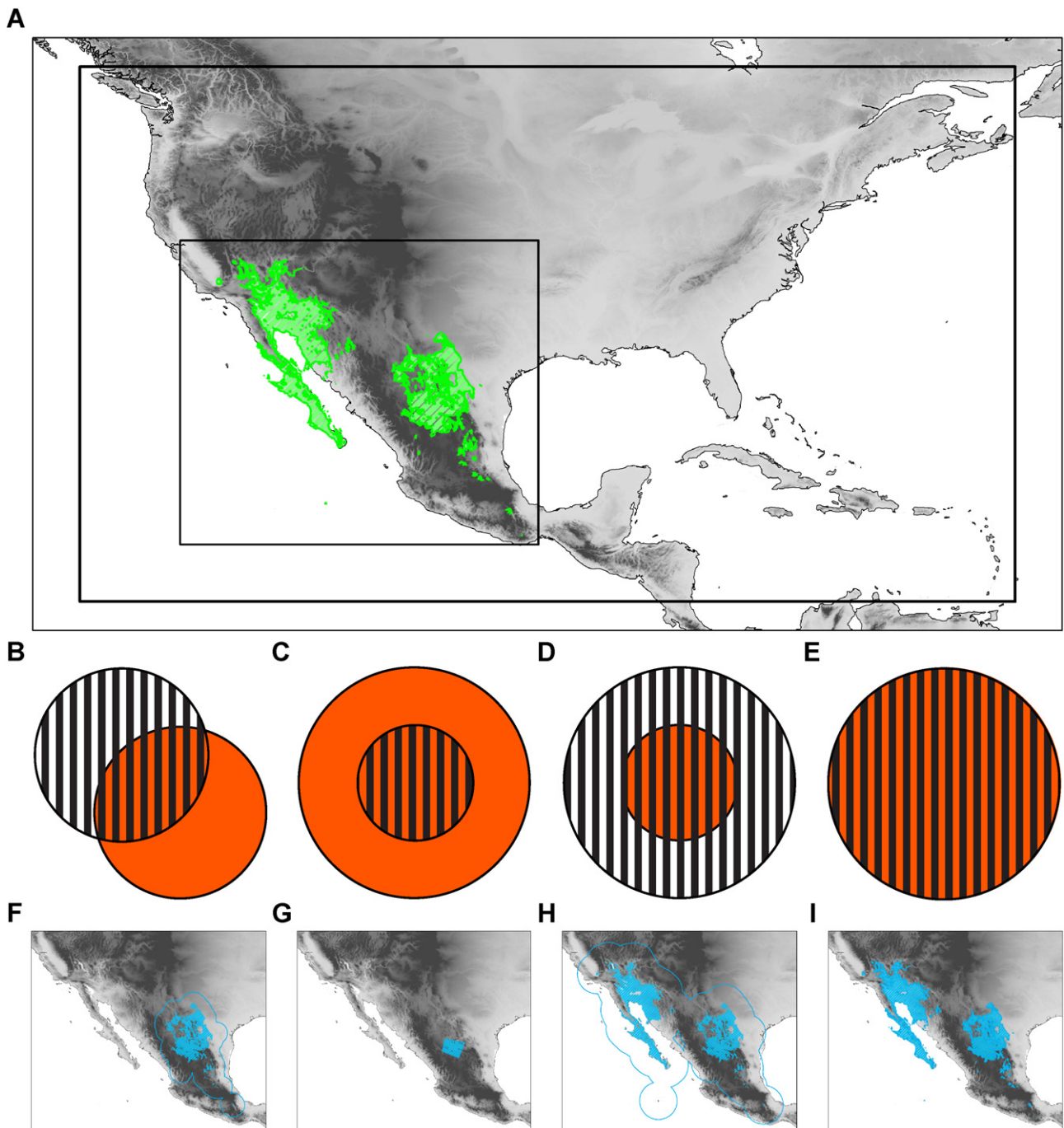
It is also important to keep in mind the implications of how different modeling algorithms operate on the data. First, one must distinguish between "background data," used to characterize the overall landscape; pseudoabsence data, which are artificial absences created for algorithms that fit functions to binary data; and true absences, which are based on reliable field evidence of non-occurrence (although one still needs to ponder the different types of absences listed above). Among commonly used algorithms in ENM/SDM, for instance, Maxent uses background data to create a null model for a probability density (Elith et al., 2011); Desktop GARP uses pseudoabsences to fit some of its component methods (Stockwell and Peters, 1999); and GAM and multivariate regression methods use either pseudoabsence or true absence data, but the interpretations of models based on one or the other are not the same (Pearce and Boyce, 2006; Ward et al., 2009). BIOCLIM, DOMAIN, and other "envelope methods" operate with only presence data (Franklin, 2009). We use the term "non-presence data" to refer to any of the above characterizations of absence; obviously, choice of a reference area implies a choice of non-presence data (Barve et al., 2011). Hence, selecting reference regions carefully and with good biogeographic considerations is a crucial point in the modeling exercise (VanDerWal et al., 2009; Godsoe, 2010; Barve et al., 2011; Elith et al., 2011). Strict presence-only methods like envelope or distance techniques (Busby et al., 1991; Hirzel et al., 2002; Farber and Kadmon, 2003) may be less affected by choice of reference region.

Jiménez-Valverde et al. (2008) argued that ENM algorithms produce outputs that fall somewhere between the occupied area  $G_0$  and the potential area **A**. In this paper, we explore this insight in greater detail, characterizing model results along this spectrum, rather than simply seeking a 'best' approach. We investigate implications of different hypotheses regarding the relative size and position of accessible and suitable areas and the size of the reference region for the ability of different algorithms to estimate  $G_0$  and **A**. The result is a picture of the strengths and limitations of a variety of ENM/SDM algorithms under certain sets of biological and biogeographic circumstances.

## 2. Methods

### 2.1. Virtual niches and virtual species

We created three virtual fundamental niches for the purpose of exploring the scenarios and ideas described above: the effects of BAM scenario, training region, and modeling method on the efficacy of models purporting to estimate ecological niches and



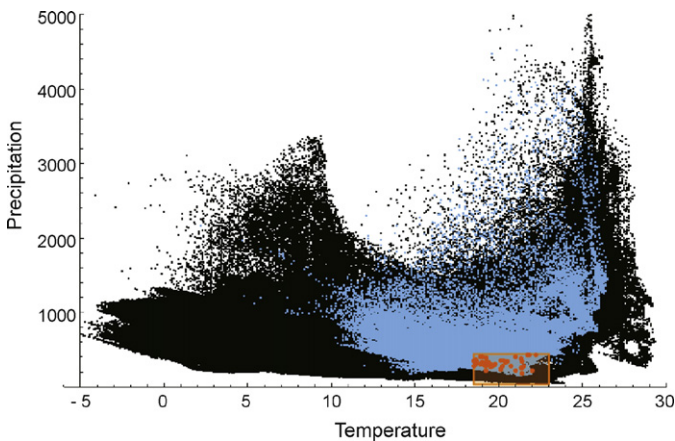
**Fig. 2.** Distributional areas and associated BAM scenarios for the Desert species. (A) Map showing  $G_1$  (small box),  $G_2$  (larger box), and the abiotic habitable area (A, in green). (B–E) Schematic diagrams summarizing the four BAM scenarios used (A is represented by the orange circle, M is represented by the hatched circle): (B) Classic BAM (CB), (C) Wallace's Dream (WD), (D) Hutchinson's Dream (HD), and (E) Full Overlap (FO). (F–I) Maps of distributional areas associated with each of the BAM scenarios (M is represented by the solid blue outline, hatched blue area indicates the inhabited area  $G_0$ ): (F) CB, (G) WD, (H) HD, and (I) FO. (The color version of this figure is available in the web version of this article.)

distributional areas. Simple fundamental niches were postulated by selecting rectilinear, non-interacting sets of conditions in a two-dimensional environmental space (annual mean temperature and annual precipitation), and identifying the geographic areas to which they correspond. We emphasize that these virtual fundamental niches were chosen on purpose to be simple, as our goal was only to create a fertile environment for testing ideas about factors affecting niche reconstructions, although this simplicity (i.e., binary, non-interactive responses to a two-dimensional environmental space) must be taken into account in pondering our results.

Note that projection of niche space into geography often produces areas that are not necessarily contiguous; in two cases, we disregarded small and peripheral disjunct areas, retaining only major regions of niche suitability.

The virtual niches represented in these three regions of North America focus on very different environmental circumstances, as follows. "Desert" was defined as the conjunction of annual mean temperature of 18.5–23.0 °C with annual precipitation of 36–445 mm, which corresponds to the Chihuahuan and Sonoran deserts of the southwestern United States and northern Mexico





**Fig. 3.** Example distribution in environmental space associated with the Classic BAM (CB) scenario for the Desert species. The black points represent all the environmental combinations available within the region  $G_2$ . Blue points represent the suite of environments represented within the  $M$  region (Supplemental Fig. 1F). The orange box indicates the fundamental niche, and the orange points inside the orange box are the environments associated with the randomly selected occurrence localities. It can be seen that the fundamental niche excludes numerous environmental combinations (in black) that have never been visited by the species (i.e., they are outside of  $M$ ). The projection of the orange box to geographic space represents the region  $A$ , while the blue points inside the orange rectangle are the geographic-space manifestation of  $G_0$ . (The color version of this figure is available in the web version of this article.)

(see Figs. 2 and 3). “Southeast” was defined as the conjunction of annual mean temperature of 15.1–16.9 °C and annual precipitation of 1181–1469 mm, which corresponds to humid temperate areas in the southeastern United States, extending from southeastern Oklahoma east to the Carolinas. Finally, “Northwest” was defined as the conjunction of annual mean temperature of –3.0 to 12.0 °C and annual precipitation of 1500–3500 mm, which corresponds to the area along the Pacific Coast west of the Cascade Mountains in Washington, Oregon, northern California, and southwestern Canada, a region that is subdivided by the Willamette River Valley corridor.

## 2.2. BAM scenarios and distributional areas

The niches described above are hypothesized fundamental ecological niches, corresponding to the abiotic habitable area  $A$  in the terminology of Peterson et al. (2011). Biotic interactions ( $B$  in the BAM framework) are disregarded because we want to maintain the argument as simple as possible: modeling  $B$  is extremely complicated; actual data on biotic effects are seldom available; and such factors may manifest on much-finer spatial scales than  $A$  (Soberón, 2007, 2010), although other authors would disagree (Bullock et al., 2000; Leathwick and Austin, 2001). In the BAM framework, this assertion translates into the idea that  $B$  is large with respect to  $A$  and  $M$  (Fig. 1).

Under this scenario, we explored the four possible relationships between  $A$  and  $M$ , as follows. “Hutchinson’s Dream” (HD) is the situation in which  $A \subset M$  (Fig. 2D), and the principal constraint to occupation is lack of favorable environments; an example of this scenario would be a species distributed in a continuous landscape without major barriers, such as *Tyto alba* (Barn owl) when viewed globally, wherein many range limits are set by climatic (i.e.,  $A$ -related) factors (Marti et al., 2005). “Wallace’s Dream” (WD) is the situation in which  $M \subset A$  (Fig. 2C), and the chief constraint on occupation of areas is dispersal capacity; here, viewing the distribution of the introduced *T. alba* populations that are restricted to the Hawaiian Islands by dispersal constraints would be an example. “Classic BAM” (CB) is the situation in which  $M$  and  $A$  overlap only partially (Fig. 2B); the same *T. alba*, but viewed across mainland

North America (such that additional habitable areas exist elsewhere), would be an example. Finally, “Full Overlap” (FO) is the situation in which  $M \approx A$ , such that the accessible and habitable areas coincide (Fig. 2E), which would be the case in any study limited to a single small island where a species like *T. alba* has full ability to use all environments present. Note that the same species can be taken as exemplifying any of the four “AM” scenarios, depending on the spatial extent of reference.

Because of how the scenarios are constructed, HD and FO have potential and occupied areas that are equivalent (i.e.,  $G_0 = A$ ), whereas CB and WD have potential areas larger than their occupied areas. For each virtual niche (projected to identify a corresponding area  $A$ ), we created distributional situations that fit with each of the four BAM scenarios described above (see Fig. 2, Supplemental Figs. 1 and 2). Creating these scenarios involved selection of areas representing  $M$ , or the area accessible to the species (Soberón and Peterson, 2005; Barve et al., 2011). In each case,  $G_0$ , the occupied area, is  $G_0 = A \cap M$ . This selection of areas was accomplished arbitrarily, with an eye to creating clear and representative examples.

We assumed that models would be best calibrated within  $M$ , as has been argued elsewhere (Anderson and Raza, 2010; Barve et al., 2011; Elith et al., 2011). However, since our  $M$ ’s were drawn arbitrarily so as to mimic the effects of a hypothetical and unspecified barrier, for the purposes of comparison, we also outlined two broader rectangular areas within which models could be calibrated, here referred to as  $G_1$  (smaller) and  $G_2$  (larger; Fig. 2, Supplemental Figs. 1 and 2). Each distributional area for each species was transformed to a grid with 0’s representing unsuitable areas and 1’s indicating suitable areas. Therefore, for each species and scenario, we had known occupied ( $G_0$ ) and potential ( $A$ ) areas of distribution against which to compare the predictions of the different algorithms.

To test algorithms and their ability to estimate niches and distributional areas, we randomly sampled 50 presence points from within  $G_0$  for each niche  $\times$  BAM scenario combination using the *Generate Random Points* function in Hawth’s Tools, version 3.27, an extension to ArcGIS 9.3. One algorithm that was explored in this study required data on true absences of the species in question (GAM models, see below). For this case, we also generated 500 absence points for each species within  $M$ , within  $G_1$ , and within  $G_2$ , but in each case, constrained to be outside of  $G_0$ , such that we assumed no error in sampling of either presences or absences. It is important to note that the WD and FO scenarios lack random points to represent true absences in  $M$  (but not in  $G_1$  and  $G_2$ ), as  $M$  and  $A$  are synonymous for these scenarios. This point also means that, as the reference region is broadened to  $G_1$  and  $G_2$ , the BAM configuration may change; for instance, a WD scenario cast on a broader extent may become a HD or CB scenario.

## 2.3. Niche modeling approaches

We explored five algorithms for niche modeling. Two require presence data only (BIOCLIM and DOMAIN); a maximum entropy algorithm (Maxent; Phillips et al., 2006) uses points from the background to minimize relative entropy; GARP, a genetic algorithm (Stockwell and Peters, 1999), uses random points to model pseudoabsences for fitting a logistic regression, and, finally, a Generalized Additive Model (Yee and Mitchell, 1991) uses true absences. Note that because BIOCLIM and DOMAIN depend solely on presence data, only the  $G_2$  sampling scheme was employed for these two algorithms (training across  $M$ ,  $G_1$ , or  $G_2$  would yield identical results).

Specifications for the five algorithms were as follows. For GARP, we ran 100 replicate runs at 0.05 convergence with 1000 maximum iterations; the consensus of replicate models was achieved via a 20% relative omission threshold, retaining the central 50% of the

distribution of proportional areas predicted as suitable (Anderson et al., 2003). For Maxent, we used default settings and the logistic output, except that a 50% random test percentage was specified. Diva-GIS 7.4 (Hijmans et al., 2001) was used for generating models of BIOCLIM and DOMAIN, with all parameters at default settings. Finally, GAMs were run in R (<http://www.r-project.org/>) using the package mgcv with a spline-based smoother. We built two GAM models: one with independent variables and the other incorporating interactions between variables; the AIC value of each model was compared, and the model with the smaller AIC was used as the final model (Johnson and Omland, 2004). Models were trained across the region of interest in each treatment (see below), but all results were projected to the  $G_2$  region. The contrasting non-presence points were taken randomly from the regions  $M$ ,  $G_1$ , and  $G_2$  for Maxent and GARP, and from the region excluding  $G_0$  but within  $M$ ,  $G_1$ , or  $G_2$  for GAM.

#### 2.4. Summary statistics

The goal of this study was not to establish which algorithm or modeling setup is 'best,' but rather to assess the relative position of model outputs along a spectrum between actual and potential distributional areas, or between  $G_0$  and  $A$ , respectively. Since no error factors were included in any of the data, we could assume that models would respond closely to the known truth. While we are completely cognizant of the problems involved with Cohen's kappa, we see it as appropriate in this particular application since we know the truth; as such, we measured the correspondence between each of our models and the corresponding distributional areas  $A$  and  $G_0$  using Cohen's kappa, via the following sequence of steps.

First, we thresholded each model using the Least Training Presence Thresholding method, in which the threshold for predictions of presence (as opposed to absence) was set at the lowest suitability value assigned to any of the occurrence points used in calibrating the model (Pearson et al., 2007). This step achieved two goals: (1) all models were made into binary predictions, with 0's indicating unsuitable conditions and 1's indicating suitable conditions; and (2) the thresholds of different models were calibrated to one another based on the yardstick of the training data. For algorithms for which raw outputs are continuous in nature (e.g., Maxent, GAM), we converted raw outputs to integer values prior to thresholding, retaining four significant digits.

We then calculated Cohen's kappa statistic for correspondence to each  $G_0$  and  $A$  for each virtual niche  $\times$  BAM scenario via the following steps. First, the thresholded models for each species were combined with a binary grid summarizing  $A$  and  $G_0$  for that virtual niche  $\times$  BAM scenario using the "Combine" function in ArcGIS 9.3 Spatial Analyst Tools. Then, we calculated omission and commission error rates as the percentage of pixels predicted as suitable outside of  $A$  or  $G_0$ , and the percentage of pixels omitted from  $A$  or  $G_0$ , respectively. From these quantities, we calculated Cohen's kappa (Fielding and Bell, 1997). The kappa statistic measures departure from a random prediction, so a value of zero means a classification indistinguishable from a random one and a value of 1 indicates very close correspondence, above and beyond random expectations.

#### 2.5. Design of experiments

As described above, we assessed four BAM scenarios, three training regions, and five niche modeling algorithms, each for three virtual niches. Taking into account the fact that some degeneracy existed (e.g., WD and FO are identical for the BIOCLIM and DOMAIN algorithms), this scheme of testing resulted in a total of 128 models, each of which was evaluated for correspondence with  $A$  and  $G_0$ .

The goal of this *in silico* experiment was to determine how the different modeling algorithms perform under different configurations of the BAM diagram and training region, so as to encourage careful and appropriate consideration of these conditions in niche modeling applications. For our kappa-based evaluations of virtual species, two sorts of information can be derived: (1) proximity to or distance from the origin for both  $A$  and  $G_0$  as an indication of overall model quality, and (2) the position of the model along a spectrum between  $A$  and  $G_0$  as an indication of tendency to approximate the potential versus occupied distributional areas. We reiterate that the kappa statistic is appropriate for this particular application, even though we see few applications of it to real-world questions of niches and distributions, in which the truth is not known.

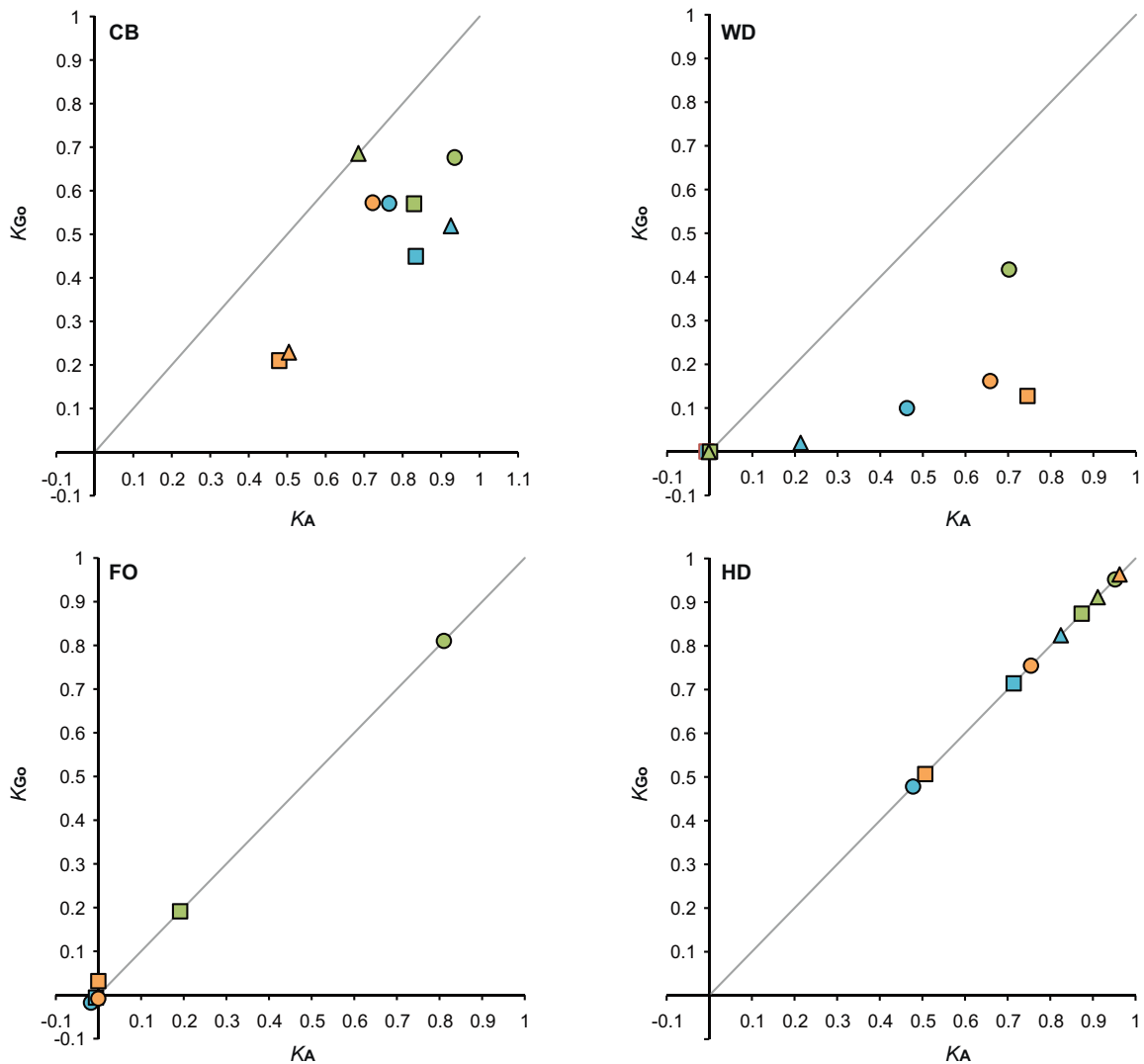
#### 2.6. Contrasts of environmental similarity

Elith et al. (2010, 2011) proposed the use of a measure of environmental similarity which they termed a Multivariate Environmental Similarity Surface (MESS) to identify areas with environments outside the range present among the points used to train models. They developed and implemented this measure as a warning for extrapolation *sensu stricto*, as opposed to transferring predictions to other regions or times, but within the range of variables used for training. We apply the MESS idea to illustrate how different two regions are, regardless of the range of values of the variables, and thus modified the details of Elith et al.'s (2010, 2011) proposal (see Supplemental Information). MESS maps were created based on the average Euclidean distance to the closest decile of the reference region for each point in the broader landscape of interest. For the questions explored herein, the most appropriate reference region is  $M$  (see below); these questions regarding MESS and transfer/extrapolation of models among regions are under detailed examination in a next study from our research group (in prep.).

In the present case, we are interested in how environments are similar or different across the broader landscape of interest in relation to those environments that have been sampled by the species (i.e., areas within  $M$ ), regardless of their suitability. Using the actual occurrence points of the species as reference (Elith et al., 2010, 2011) mixes effects of niche with those of environmental difference; therefore, to maintain focus on identifying areas of extrapolation (i.e., areas presenting environmental conditions very different from those of the training region), our MESS analyses use as a reference the  $M$  regions for each scenario. For each reference region, scenario, and species, we sampled from a set of 50,000 points to obtain the distribution of MESS values with respect to  $M$ . Finally, to identify concretely those areas with conditions that fall outside of environmental ranges for any of the variables included in the model, we identified ranges of values represented within  $M$  for each environmental variable independently, and then used those ranges to flag areas across the broader  $G_2$  that fall outside of the training-region range for any variable, and thereby represent extrapolation.

### 3. Results

For each combination of algorithm, species, and BAM scenario, we evaluated the closeness of algorithm output to the corresponding actual ( $G_0$ ) and potential ( $A$ ) distributional areas (Fig. 4). We noted consistent differences among BAM scenarios with respect to these measures. In the CB and HD scenarios, algorithms were capable of yielding relatively good predictions of both  $A$  and  $G_0$  (recalling that in HD  $A = G_0$ ). The other two scenarios posed challenges that the algorithms we tested consistently failed to meet. In the FO scenario, every prediction except one had kappa values below 0.2, which are very close to random. In the WD scenario, no



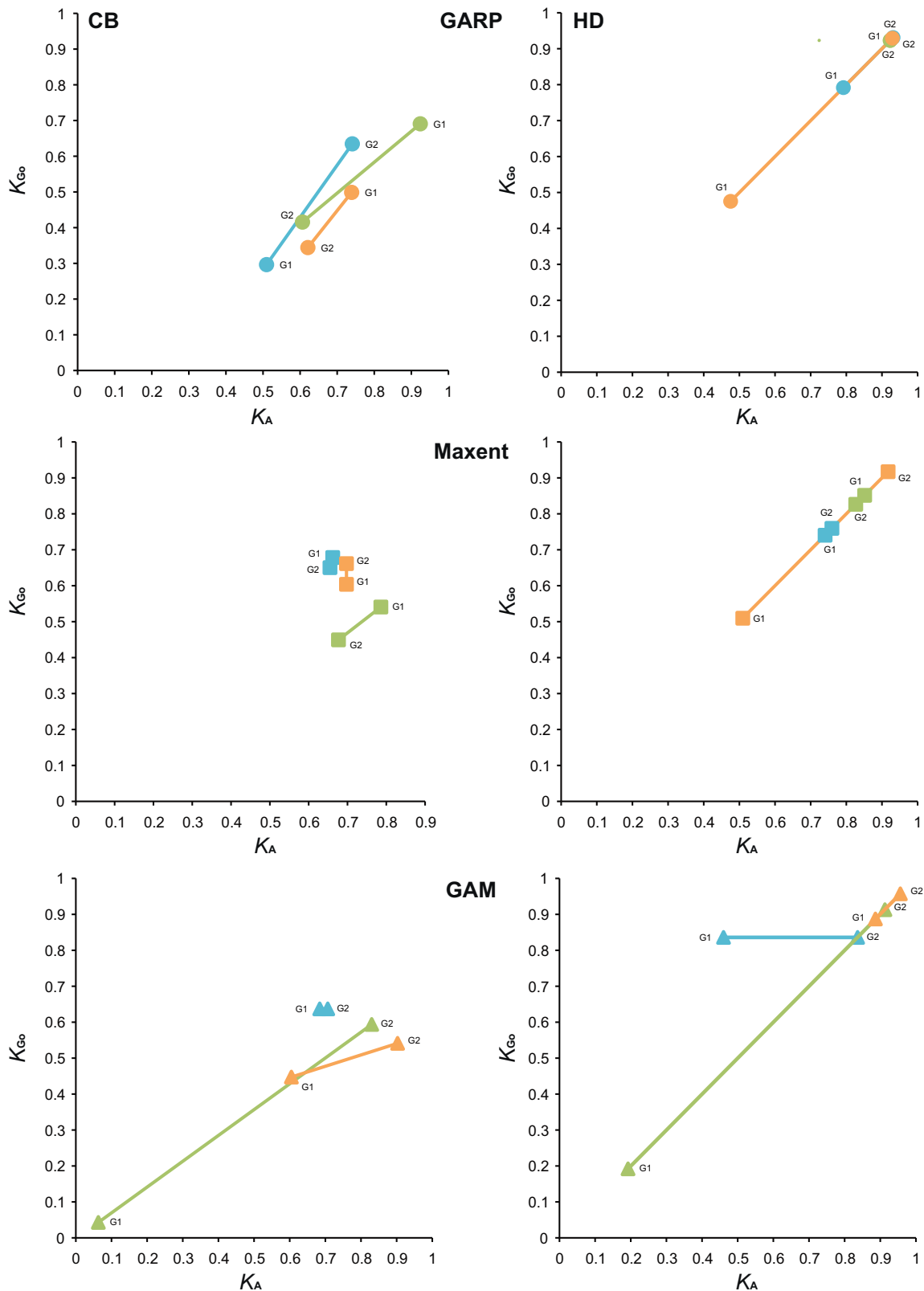
**Fig. 4.** Predictive ability of models as measured by kappa values comparing the model output to the potential ( $K_A$ ) and occupied ( $K_{G_0}$ ) areas when trained in **M**, for each of the four BAM scenarios. Blue = Desert species; green = Southeast species; orange = Northwest species. Circles = GARP; squares = Maxent; triangles = GAM. (The color version of this figure is available in the web version of this article.)

$K_{G_0}$  value exceeded 0.4, and most  $K_A$  values were below 0.5. That is, FO and WD showed predictions that were near random. It is not surprising that WD models did a better job of recovering **A** than  $G_0$ , as the latter is determined mostly by the dispersal factors of **M**. In the FO scenario, most algorithms captured, at best, only a subset of the niche.

We also compared the performance of algorithms as regards different training areas. We trained models based on the arbitrary area **M** that curtails **A** to produce  $G_0$ , but also based on two larger areas, here termed  $G_1$  and  $G_2$ . In light of our previous argument that model calibration should be restricted to areas of relevance, i.e., areas that have been accessible to the species, and thus are likely unsuitable if they are not inhabited (Barve et al., 2011), our initial expectation was that models trained in **M** would perform better than models trained across broader areas. That idea, however, was not confirmed (see [Supplementary Table 1](#)), likely owing to the fact that our hypothesized **M**'s do not directly inform the algorithms. As a more controlled comparison of small versus large model training areas, we compared performance of models trained in  $G_1$  and  $G_2$ , but found little consistency in performance differences (Fig. 5). Indeed, the only pattern that we noted was that GAM

models performed consistently better when trained across broader areas ( $G_2$ ).

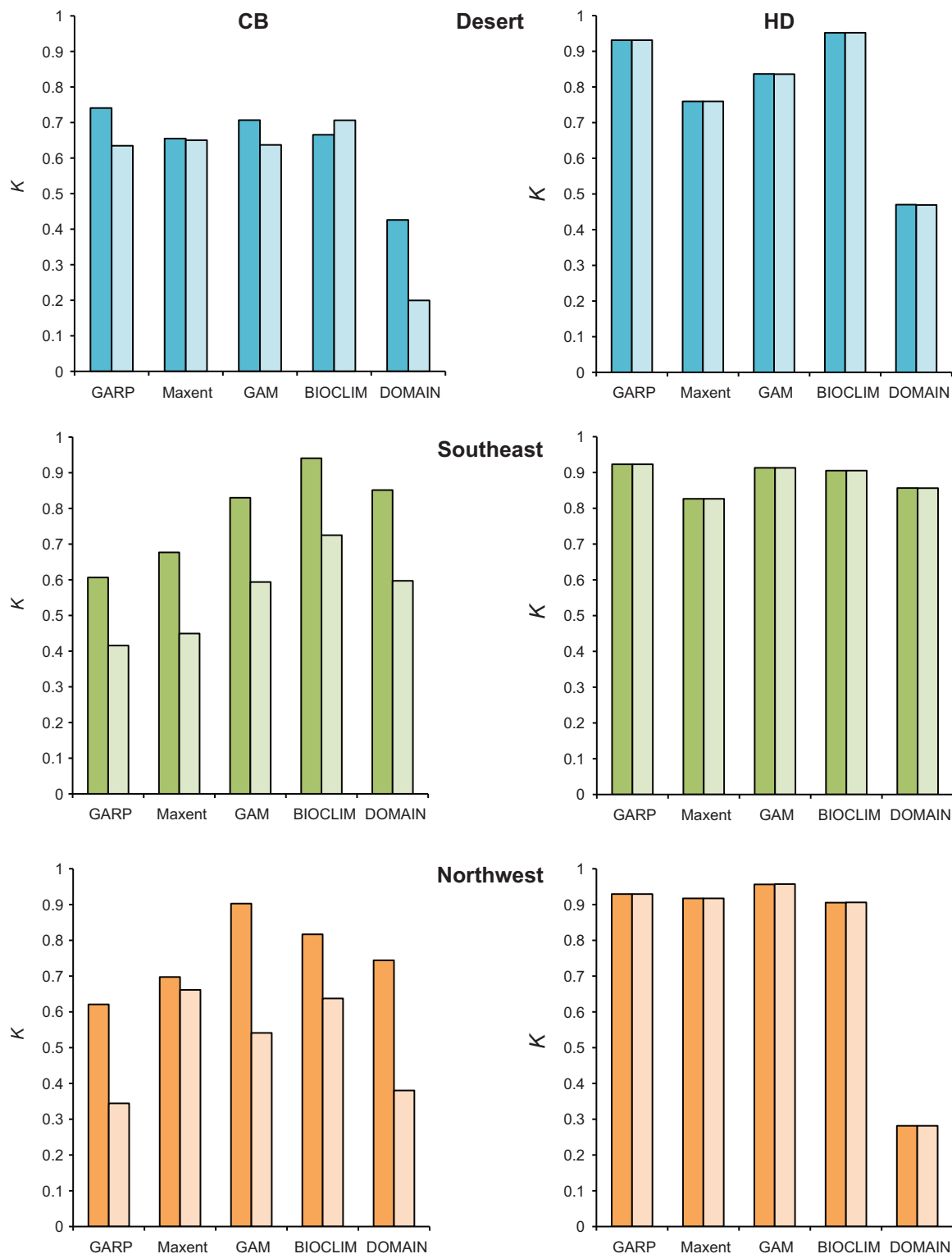
We compared the five algorithms for each of the three species and for the two BAM scenarios for which models could be developed successfully (see above). Models generally performed better under HD than CB (Fig. 6), likely reflecting the fact that HD scenarios allow the species to access the entirety of **A**, whereas CB scenarios inevitably include some suitable areas that are not occupied owing to accessibility constraints (Figs. 2 and 3). For example, for the Desert species under the CB scenario (Fig. 3), 296,686 cells were present across  $G_2$ ; **M** included 11.7% of these points, and the fundamental niche contained 2601 cells, or 0.87% of available combinations. In this scenario (CB), **M** includes only a part of the fundamental niche termed the existing fundamental ecological niche (i.e., the part of the fundamental ecological niche that is actually manifested within the region of interest; Peterson et al., 2011), creating a region fitting the conditions of the existing fundamental niche that is outside of the area sampled by the species. In such situations, modeling **A** may be difficult, because parts of the fundamental niche remain “hidden” from analyses based on data drawn from the known occurrences of the species.



**Fig. 5.** Predictive ability of models as measured by kappa values comparing the model output to the potential ( $K_A$ ) and occupied ( $K_{G0}$ ) areas for two different calibration areas ( $G_1$  and  $G_2$ ). Symbols as in Fig. 4. (The color version of this figure is available in the web version of this article.)

A second insight from these comparisons is that four of the five approaches consistently performed quite well (Fig. 6 and Supplemental Table 1)—only DOMAIN failed in 3 of the 6 comparisons. We note that the good performance of BIOCLIM is likely a

function of the rectilinear definition of our virtual niches, which matches the mechanism of this algorithm, and likely gives it an unfair advantage that would not hold true in more realistic applications. Overall, then, the remaining three algorithms (GARP, Maxent,



**Fig. 6.** Kappa values for the different modeling methods under two different scenarios (Classic BAM CB, and Hutchinson's Dream HD) trained within the  $G_2$  region. Dark bars are  $K_A$  and lighter bars are  $K_{Co}$ . Colors as in Fig. 4. (The color version of this figure is available in the web version of this article.)

GAM) performed similarly and consistently well, and will be the focus of our discussions.

Finally, we note an interesting feature of Fig. 4: the comparisons between prediction of  $G_0$  versus  $A$ . In the HD and FO scenarios,  $A$  and  $G_0$  are equivalent. In the CB and WD scenarios, however, all predictions fell closer to  $A$ . This imbalance would be highly unexpected if no underlying difference existed (sign test,  $P < 0.05$ ).

#### 4. Discussion

This study departs from previous comparative analyses of ENM or SDM (e.g., Brotons et al., 2004; Segurado and Araújo, 2004; Elith et al., 2006; Elith and Graham, 2009) in that: (i) the relative importance of environmental suitability vs. dispersal (i.e., configurations of the BAM diagram) is considered, (ii) the measure of performance



of the models is their correspondence to two known distributional areas (the occupied and potential distributions), and (iii) we distinguish between different types of non-occurrence data. We also explore the consequences of using different reference regions for obtaining calibration data (VanDerWal et al., 2009; Barve et al., 2011; Elith et al., 2011).

#### 4.1. Comparisons among BAM scenarios

Our four BAM scenarios neglected the effects of **B**, owing to the conceptual and computational difficulties of including effects of biotic interactions in such explorations (Araújo and Guisan, 2006; Soberón, 2010). As a result, our comparisons explore different configurations of an “AM” framework only. Nonetheless, we observed dramatic contrasts in model success rates among scenarios, in which certain configurations cause algorithms to have very low success in reconstructing either potential or actual distributional areas. This result suggests that the factors included in BAM diagrams should always be considered carefully in ENM/SDM projects. In particular, models under the WD and FO scenarios frequently failed to perform better than random expectations. These scenarios are ones in which all areas that are accessible to the species are also suitable for the species, effectively situations in which dispersal limitations govern distributional patterns of species. In these situations, little or no contrast is available to algorithms on which to base inferences. These scenarios represent situations that are frequently manifested in the real world—for example, species on small islands may see their distributional limits set entirely or predominantly by **M**, and not by factors related to **A**. These species are effectively WD species, and our results indicate that ENM/SDM will not provide good estimates of either **A** or **G<sub>0</sub>**. Niche estimates obtained in such scenarios may still be employed as first approximations to small portions of the realized niche that can be useful in further documentation of the distribution (Siqueira et al., 2009; Menon et al., 2010). Some improvement may be made to these models if the spatial resolution can be increased or the **M**-based training region broadened, the effect being to shift WD situations towards HD or CB scenarios.

Lack of appreciation of the differences in BAM configurations may have confused previous efforts to understand differences between ENM/SDM methods. For instance, Elith et al. (2006) presented comparisons across taxa sampled from six regions. One extreme in terms of BAM configuration was birds in Ontario, in which all species had broader ranges than the study area. If we use “**S**” to denote the area actually sampled to produce the data analyzed, then we could describe this part of the study as a case of the FO scenario, in which  $S \subset A \approx M$ , which appears to present a serious challenge to SDM/ENM. At the other extreme, another Elith et al. (2006) dataset examined distributions of liana species across South America, where the likely BAM configuration was  $S \supset A \supset M$ , which is close to a WD situation, and again is difficult to model. As a consequence, different BAM configurations, rather than intrinsic differences among algorithms in inferential ability, probably account for some of the variation observed and analyzed in that study.

In another example of failing to assess implications of particular BAM configurations, Beale et al. (2008) concluded that European bird species' distributions were rarely governed by climatic factors. However, the authors neglected to consider (1) the barriers to dispersal that surround much of Western Europe (Mediterranean Sea to the south, Atlantic Ocean to the west, North Sea to the north), and (2) the abrupt cut-off of the sampling available to them to the east. The result was a situation in which the study region is artificially constrained, imposing distorted and unrealistic limits on what can be observed of **M** and **A**. The limits of **A** will only rarely be discernable in such cases, and one ends up with a WD situation. A follow-up

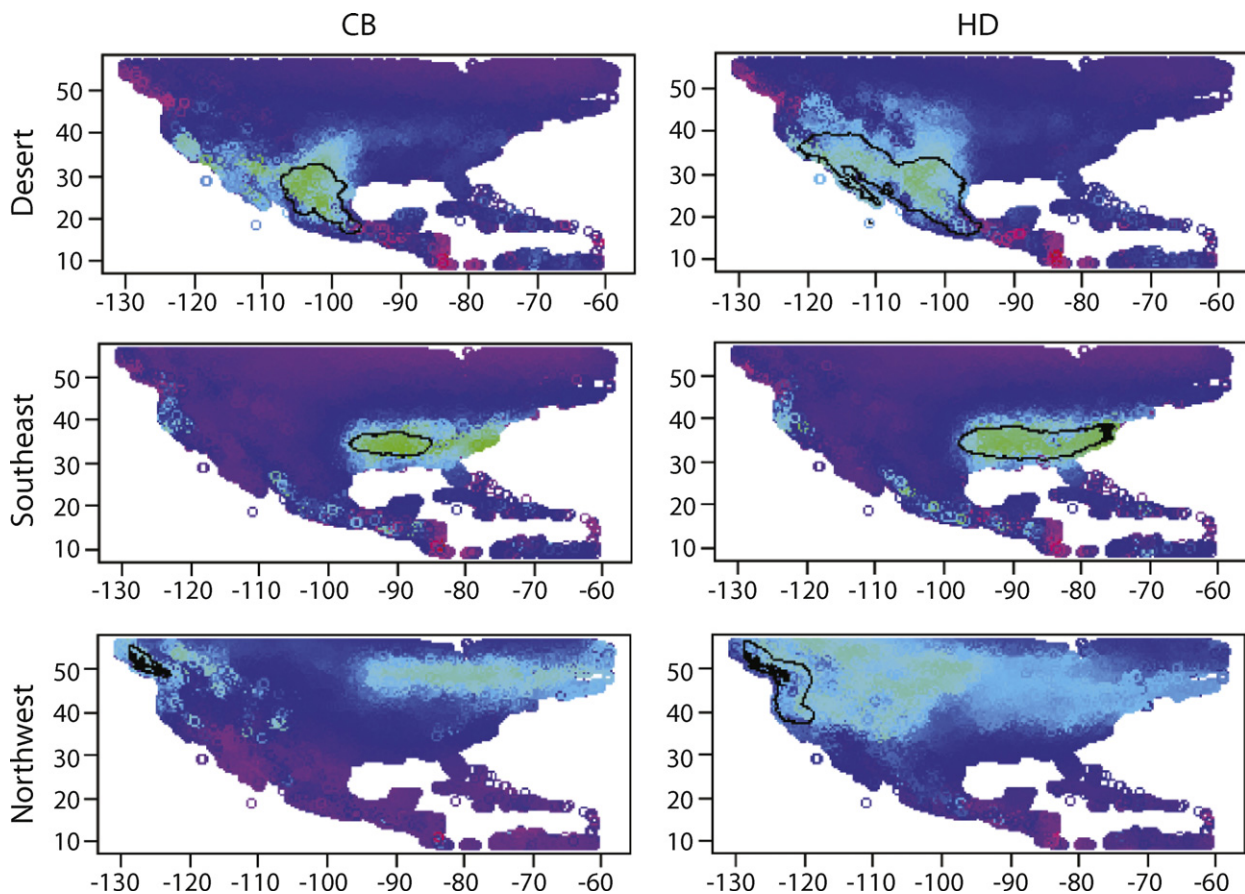
study replicating the Beale et al. methodology for North America, where much more avian endemism is present, such that frequently  $A \subset M$  (an HD case), showed much greater climatic determination of bird species' ranges, effectively corroborating the idea that the BAM structure has critical implications both for the niche modeling process and for interpretation of results (Jiménez-Valverde et al., 2011).

#### 4.2. Comparisons among calibration areas

This portion of our study yielded the least clear-cut results. We initially designated explicit hypotheses of **M** for each of the virtual species (i.e., combinations of three virtual niches and four BAM scenarios), and expected that models calibrated on these restricted areas would perform more reliably than those calibrated across broader regions (**G<sub>1</sub>** and **G<sub>2</sub>**). However, such was not the case (cf. Elith et al., 2011), perhaps, we realize in retrospect, owing to the fact that our **M** hypotheses only very indirectly affect the operation of the algorithms used. In other words, proposing an accessible region **M** in theory adds an ecologically sensible factor for the creation of **G<sub>0</sub>** in a virtual species, but the information about **M** enters the modeling process only indirectly, through the characteristics of occurrence and non-presence points. Moreover, the effect of different types of non-occurrence data on different algorithms is probably not the same; we found that it is difficult to disentangle the effects of different algorithms, different types of non-occurrence data, and distinct calibration areas.

First, we stress that the region **M** plays two roles in our analyses. It is used to define the occupied region **G<sub>0</sub>** by intersection with the suitable region,  $G_0 = A \cap M$ , but it is also one of the three regions from which non-presence points were extracted. The first role is internalized in the modeling process only if true-absence data are available and a regression method is used (Phillips et al., 2009), and then only implicitly. In other words, none of the algorithms we used includes explicit consideration of a region beyond which the species has never been able to explore. Very recent process-oriented methods are beginning to take into account dispersal and movements (Cabral and Schurr, 2010; Smolik et al., 2010; Barve et al., 2011), but this gap is a limitation for the set of methods most commonly used: it is simply impossible to inform them directly about **M**, which is a major factor determining **G<sub>0</sub>**. This limitation causes confusion as to whether model output should be interpreted as **G<sub>0</sub>** or as **A**, which is central in many of the arguments regarding “overprediction” (Jiménez-Valverde et al., 2008) and whether the object of modeling is a niche or a distribution.

The second role that **M** plays is as the smallest of the three regions from which non-occurrence data were extracted. Several things happen as the calibration region grows. As illustrated in Fig. 7, non-occurrence data may come from regions that are similar to or different from those assumed to be sampled by the species. Different algorithms use non-occurrence data in different ways. Maxent uses background data to estimate a distribution of environmental combinations **z** (strictly speaking, in “features” space), called  $f(\mathbf{z})$  (Elith et al., 2011). This distribution provides a null model that is used to find a Gibbs' distribution as close as possible to the null model and restricted to produce the same means of features as those of the distribution of environments in the occurrence points, called  $f_1(\mathbf{z})$ . More specifically, Maxent minimizes the Kullback–Leibler divergence  $div(f_1, f) = \int_{\mathbf{z}} f_1(\mathbf{z}) \ln[f_1(\mathbf{z})/f(\mathbf{z})] d\mathbf{z}$ , between these two distributions, subject to the means restriction (Elith et al., 2011). Therefore, the effect of adding background data from a broader area essentially informs the algorithm via the function  $f(\mathbf{z})$ . Adding more background points alters the shape of  $f(\mathbf{z})$ , in one way if the background points are different from the region near the occurrences, as is the case in **G<sub>2</sub>** for the Southeast (Supplemental Fig. 1), or in another way if



**Fig. 7.** MESS values summarizing similarity of environments across the broader study region to the environments represented across the accessible area **M** for each of the three species (Desert, Southeast, Northwest) for two BAM scenarios (Classic BAM and Hutchinson's Dream). Pale blue indicates high environmental similarity, grading to purple, which indicates highest dissimilarity. Scales are the same across all of the maps. The region **M** is outlined in black in each case. (The color version of this figure is available in the web version of this article.)

the environmental values are similar, as it is the case for **G**<sub>2</sub> for the Northwest (Supplemental Fig. 2). On the other hand, when fitting a logistic (or a similar function), as GAM or GARP does (GARP does it internally, as one of its four competing algorithms, and GAM uses a linear combination of non-linear smoothers for the log-odds), the non-presence data are used not as a null model, but together with presences to maximize the log-likelihood of a model for binary data. Broadening the reference region, then, may add non-presences with dissimilar environments, as in the Southeast case, or non-presences with similar environments, as in the Northwest case, or species, which may be more challenging, since it may lower the discriminating power of the model. It is clear that these two cases probably affect the likelihoods in contrasting ways, but it is not obvious to us that increasing non-occurrence points with different MESS distributions affects entropy-based methods in the same way that it affects regression-based methods.

As the calibration area is broadened, another effect is taking place. If, as is the case for the Desert and Northwest species, increasing the area allows areas highly similar to the training region but lacking occurrences to be included in the analysis, the BAM configuration is changing. The different BAM configurations explored here are not independent of hypotheses of calibration area. That is, a "Wallace's Dream" configuration, in which species' distributional limits are constrained by **M** exclusively, morphs into something more akin to a "Classic BAM" or a "Hutchinson's Dream" when broader calibration areas are employed. As a consequence, one should take considerable care in defining scenarios and applying them only under appropriate circumstances.

The conclusion that we extract from the preceding discussion is that increasing the extent from which non-occurrence data is obtained has complex and difficult-to-assess effects on modeling algorithms. *Biologically speaking*, models should be calibrated based on contrasts between presences and absences of the species across **M**, as has been appreciated in several recent papers (VanDerWal et al., 2009; Barve et al., 2011; Elith et al., 2011)—these contrasts are those in which the species of interest has in effect "sampled" the non-inhabited landscape and not established populations there, ostensibly for reasons related to **A** (although effects of **B** may also have to be assessed). However, we note that the effects of constraining model calibration to conditions represented within **M** are reflected in model correctness, but potentially not in model performance. The current generation of SDM/ENM algorithms essentially ignores information about **M** in their workings. A new generation should seek to incorporate such information more directly in models, perhaps via Bayesian approaches.

#### 4.3. Comparisons among algorithms

The main point of our study is that comparisons of performance of ENM/SDM algorithms are best done when the factors underlying a species' distribution are considered explicitly, and when evaluation criteria are unequivocal. Using virtual species, where the true distributions are known, permits these possibilities. As described above, disentangling the relevant factors, even in very simplified cases, may be surprisingly difficult. This point makes the uncontrolled nature of some previous comparisons (e.g., Elith et al., 2006; Ortega-Huerta and Peterson, 2008; Wisz et al., 2008) difficult to

interpret, since, quite simply, it remains unclear what one learns from comparisons based on unknown BAM configurations and situations in which the actual region being modeled (be it **A** or **G**) remains unknown and largely unknowable.

In this study, we observed rather poor performance by DOMAIN, and good, but perhaps artifactual performance by BIOCLIM. The remaining three algorithms all showed relatively good performance, albeit with some variation. Despite being provided with true absences, GAM was, surprisingly, not consistently better at estimating **G** than the other methods. No other consistent differences emerged among algorithms in these comparisons, save for a couple of GAM results showing oddly low performance values. In general, however, Maxent, GARP, and GAM showed similar performance, a result that is of particular note because it indicates no significant improvement in model performance when high-quality absence data are used in the GAM algorithm. Much theory (Keating and Cherry, 2004; Pearce and Boyce, 2006; Phillips et al., 2009; Ward et al., 2009; Li et al., 2011) and data (Brotons et al., 2004; Wisz et al., 2008) support the widespread view that true-absence data would be the only way to estimate **G**, since prevalence cannot be identified without true absences: this idea was not supported by our analyses using virtual species.

#### 4.4. Implications for modeling applications

The exercises presented in this paper are admittedly simplified, since their purpose was to illuminate the operation of complicated methods in identifying unknown biological objects: they are based on virtual niches and actual and potential distributions across real-world landscapes. However, the virtue of the simplification is that some issues are clarified substantially—while our very-simple virtual species may give an overly optimistic view of simple algorithms such as BIOCLIM, they represent a best-case situation for the important insights that we have obtained, which have clear implications for real-world applications of niche and distribution modeling techniques.

The principal lesson to be drawn from our analyses is that the configuration of the main factors determining distributions of species (which we conceptualize via the BAM diagram) must be understood clearly, or at least hypothesized, prior to any application of niche and distribution modeling. Although any set of occurrence points and environmental data layers can be stuck into some algorithm and maps obtained, their meaning remains dubious without a firm conceptual framework by which to organize the results. That is, we must understand what factors generally structure the distributional area of the species of interest: are at least some of its distributional limits related to aspects of **A**, and not to **M**? For instance, a species may be strictly endemic to an island, which suggests a Wallace's Dream-style BAM configuration, and thus the species may not be very amenable to analysis; however, if the species does not cover the entire island, and if distributional and environmental data are sufficiently detailed, some level of analysis may be feasible . . . perhaps it is distributed only up to middle elevations, but does not ascend to the highest points on the island. In this manner, we need not immediately "throw the baby out with the bath water," but instead take care to assess likely BAM configurations before beginning analyses. More in depth, we should know from which sectors of which BAM configuration the presence and absence data being used to calibrate models are drawn; in cases in which this information is not known, assumptions should be stated clearly and explicitly, such that their implications can be evaluated.

A further point is that, although in theory models should be calibrated only within a (clearly and explicitly stated) hypothesis of **M**, our results and those of Elith et al. (2011) and VanDerWal et al. (2009) do not consistently show better results when model calibration is restricted to smaller regions of reference. The use of

MESS maps (Elith et al., 2011) may clarify the reasons for these equivocal results, and should become a key tool in the design and interpretation of ENM/SDM exercises—only areas relatively similar to, and falling within the environmental breadth of, the training region should be interpretable in terms of whether they are part of the distributional potential of a species. MESS can also be used to describe whether a model is projecting (extrapolating) to conditions not manifested within **M** (see Supplemental Information for worked examples).

#### 4.5. Niche modeling versus distribution modeling

Finally, we make some reflection regarding what these diverse techniques are actually modeling. Niche modeling and distribution modeling are terms that have been equated quite carelessly in the literature, with "niche" uses treated under the distribution modeling rubric and vice versa (Peterson, 2006). Here and elsewhere (Peterson et al., 2011), we emphasize the critical need for much more careful use of terminology and concepts: a clear separation between the two ideas is that of focus on estimation of **G** (species distribution modeling) versus focus on estimation of the environments of **A** (ecological niche modeling). Jiménez-Valverde et al. (2008) pointed out that these two objects of estimation, in reality, form the extremes of a spectrum, and that most modeling efforts probably estimate some quantity in between the two extremes.

Curiously, although it is easy to erect different definitions of the two approaches, the tools and data streams employed by the two are similar or identical. The implications are serious—the same data are input and the same tools are applied, but the results may be interpreted in very different manners. Our results point to a resolution of the question of which quantity is being estimated: in 14 of 15 comparisons, our model results were more closely aligned with the potential distribution than with the actual distribution. In short, at least in the case of three simple virtual species, our models were much more niche models than distribution models. This result suggests that additional processing steps are necessary to estimate actual distributions from these models (i.e., to reduce model outputs from **A** to **G**, on the basis of suitable additional information), and that "species distribution modeling" as a universal term is misleading and incorrect.

#### Acknowledgments

We thank our colleagues in thinking about these ideas, including in particular Fangliang He, Monica Papeş, Yoshinori Nakazawa, Sean Maher, Fabricio Villalobos, and Alberto Jiménez-Valverde. This research was supported in part by a grant to JS and ATP by Microsoft Research. AL-N received support from the Consejo Nacional de Ciencia y Tecnología, México (189216).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ecolmodel.2012.04.001>.

#### References

- Anderson, R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162, 211–232.
- Anderson, R.P., Raza, A., 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography* 37, 1378–1393.
- Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33, 1677–1688.

- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* 222, 1810–1819.
- Beale, C., Lennon, J.J., Gimona, A., 2008. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences of United States of America* 105, 14908–14912.
- Brottons, L., Thuiller, W., Araújo, M., Hirzel, A., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27, 437–448.
- Brown, J.H., Stevens, G.C., Kaufman, D.M., 1996. The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics* 27, 597–623.
- Bullock, J.M., Edwards, R.J., Carey, P.D., Rose, R.J., 2000. Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? *Ecography* 23, 257–271.
- Busby, J.R., Margules, C.R., Austin, M.P., 1991. BIOCLIM – a bioclimate analysis and prediction system. In: Margules, C.R., Austin, M.P. (Eds.), *Nature Conservation. Cost-effective Biological Surveys and Data Analysis*. Canberra, Australia.
- Cabral, J.S., Schurr, F.M., 2010. Estimating demographic models for the range dynamics of plant species. *Global Ecology and Biogeography* 19, 85–97.
- Elith, J., Graham, C., 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32, 66–77.
- Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L., Loiselle, A., Manion, B., Moritz, G., Nakamura, C., Nakazawa, M.Y., Overton, M.C.M., Peterson, J., Phillips, A.T., Richardson, S., Scachetti-Pereira, K., Schapire, R., Soberón, R., Williams, J., Wisz, S., Zimmermann, M.N., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151.
- Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1, 330–342.
- Elith, J., Phillips, S., Hastie, T., Dudik, M., En Chee, Y., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Biodiversity Research* 17, 43–57.
- Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling* 160, 115–130.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38–49.
- Franklin, J., 2009. *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge.
- Gaston, K., 2003. *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, Oxford.
- Godsoe, W., 2010. I can't define the niche but I know it when I see it: a formal link between statistical theory and the ecological niche. *Oikos* 119, 53–60.
- Good, R.D., 1931. A theory of plant geography. *New Phytologist* 30, 149–171.
- Grinnell, J., 1924. Geography and evolution. *Ecology* 5, 225–229.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8, 993–1009.
- Guisan, A., Zimmermann, N., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147–186.
- Hijmans, R.J., Guarino, L., Cruz, M., Rojas, E., 2001. Computer tools for spatial analysis of plant genetic resources data: DIVA-GIS. *Plant Genetic Resources Newsletter* 127, 15–19.
- Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 83, 2027–2036.
- Hirzel, A.H., Le Lay, G., 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45, 1372–1381.
- Hortal, J., Roura-Pascual, N., Sanders, N.J., Rahbek, C., 2010. Understanding (insect) species distributions across spatial scales. *Ecography* 33, 51–53.
- Jiménez-Valverde, A., Lobo, J.M., Hortal, J., 2008. Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions* 14, 885–890.
- Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J., Aragón, P., Lobo, J.M., 2011. Use of niche models in invasive species risk assessments. *Biological Invasions* 13, 2785–2797.
- Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. *Trends in Ecology and Evolution* 19, 101–108.
- Keating, K.A., Cherry, S., 2004. Use and interpretation of logistic regression in habitat-selection studies. *Journal of Wildlife Management* 68, 774–778.
- Leathwick, J.R., Austin, M., 2001. Competitive interactions between tree species in New Zealand's old-growth indigenous forest. *Ecology* 82, 2560–2573.
- Li, W., Guo, Q., Elkan, C., 2011. Can we model the probability of presence of species without absence data? *Ecography* 34, 1096–1105.
- Lobo, J.M., Jiménez-Valverde, A., Hortal, J., 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33, 103–114.
- Marti, C.D., Poole, A.F., Bevier, L.R., 2005. Barn owl (*Tyto alba*). *Birds of North America* 1, 1–16.
- Menon, S., Choudhury, B.I., Khan, M.L., Peterson, A.T., 2010. Ecological niche modeling predicts new populations of *Gymnocladus assamicus*, a critically endangered tree species. *Endangered Species Research* 11, 175–181.
- Ortega-Huerta, M.A., Peterson, A.T., 2008. Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. *Revista Mexicana de la Biodiversidad* 79, 205–216.
- Pearce, J., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology* 43, 405–412.
- Pearson, R., Raxworthy, C., Nakamura, M., Peterson, A., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34, 102–117.
- Pearson, R.G., Thuiller, W., Araújo, M.B., Martínez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P., Lees, D.C., 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography* 33, 1704–1711.
- Peterson, A.T., 2006. Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* 3, 59–72.
- Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., Araújo, M.B., 2011. *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton.
- Peterson, A.T., Soberón, J., Sánchez-Cordero, V., 1999. Conservatism of ecological niches in evolutionary time. *Science* 285, 1265–1267.
- Phillips, S., Dudik, M., Elith, J., Graham, C., Lehman, A., Leathwick, J., Ferrier, S., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* 19, 181–197.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190, 231–259.
- Segurado, P., Araújo, M., 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography* 31, 1555–1568.
- Siqueira, M.F.d., Durigan, G., Marco Jr., P.d., Peterson, A.T., 2009. Something from nothing: using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation* 17, 25–32.
- Smolik, M.G., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L.-M., Vogl, G., 2010. Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography* 37, 411–422.
- Soberón, J., 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* 10, 1115–1123.
- Soberón, J., 2010. Niche and area of distribution modeling: a population ecology perspective. *Ecography* 33, 159–167.
- Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2, 1–10.
- Stockwell, D.R.B., Peters, D.P., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13, 143–158.
- Svenning, J.-C., Skov, F., 2004. Limited filling of the potential range in European tree species. *Ecology Letters* 7, 565–573.
- Udvardy, M., 1969. *Dynamic Zoogeography, with Special Reference to Land Animals*. van Nostrand Reinhold, New York.
- VanDerWal, J., Shoo, L.P., Graham, C., Williams, S.E., 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling* 220, 589–594.
- Ward, G., Hastie, T., Barry, S., Elith, J., Leathwick, J.R., 2009. Presence-only data and the EM algorithm. *Biometrics* 65, 554–563.
- Wisz, M.S., Hijmans, R., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14, 763–773.
- Yee, T.W., Mitchell, N.D., 1991. Generalized additive models in plant ecology. *Journal of Vegetation Science* 2, 587–602.