DOI: 10.1111/jbi.12234

Journal of Biogeography (J. Biogeogr.) (2014) **41**, 513–523

ORIGINAL ARTICLE

The importance of biotic interactions in species distribution models: a test of the Eltonian noise hypothesis using parrots

Carlos B. de Araújo[1,2]*, Luiz Octavio Marcondes-Machado[3]
and Gabriel C. Costa[4]

[1]*Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, João Pessoa, PB 58059-900, Brazil,* [2]*Programa de Pós-graduação em Ecologia, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brazil,* [3]*Departamento de Biologia Animal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brazil,* [4]*Centro de Biociências, Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal, Brazil*

*Correspondence: Carlos B. de Araújo, Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, João Pessoa, PB 58059-900, Brazil.
E-mail: cabarau@gmail.com

ABSTRACT

Aim To test the Eltonian noise hypothesis (ENH), that biotic interactions do not affect species distributions at large geographical scales.

Location The Brazilian cerrado, a central South American savanna and biodiversity hotspot.

Methods We modelled the distributions of 11 species of cerrado parrots using the software Maxent at four different spatial resolutions. We built models using abiotic variables, biotic variables (distribution of diet resources) and models combining abiotic and biotic variables. We compared model performance using the area under the curve of the receiver operating characteristic (AUC), retrieved from test data. We partitioned the variance between sets of predictors using a generalized linear model (GLM). Finally, we evaluated whether improvement in model performance (higher AUC values) in models with both abiotic and biotic variables, was related to the species' dietary niche breadth and/or spatial resolution of the models.

Results We found that model performance was improved in most cases by the addition of biotic variables. Our variance-partitioning approach revealed that abiotic and biotic variables contribute independently to the final model. We found no relationship between model improvement and spatial resolution. We also found no relationship between dietary niche breadth and model improvement, indicating that dietary generalist and specialist species were not differently affected by the inclusion of biotic variables in the models.

Main conclusions Our results did not support the ENH. In this study, we explicitly incorporated a biotic variable (diet resource distribution) into species distribution models (SDMs), and we showed that these variables generally improve models and have independent contributions. These results agree with previous studies that incorporated biotic variables into SDMs. Ultimately, our results indicate that SDMs performed with abiotic variables only may depict only a partial representation of the geographical distribution of a species.

Keywords

AUC, Brazilian cerrado, diet, Eltonian noise hypothesis, Maxent, niche modelling, niche theory, Psittacidae, spatial resolution, variance partitioning.

INTRODUCTION

A primary goal in biogeography is to understand what factors shape the distributions of species (Gaston, 2009). Species distributions are complex biological phenomena and many factors interact to determine a species' geographical range (Sexton *et al.*, 2009). Understanding the ecological and evolutionary dynamics of species ranges can provide valuable insights into a wide range of biological phenomena, such as biological invasions (Peterson, 2003), organismal responses to large-scale environmental fluctuations in the past (Graham *et al.*, 1996) and predicted future responses to climate change (Parmesan, 2006). The overarching factors that determine the geographical distributions of species are: how

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12234

organisms relate to their environment (i.e. niche requirements); interspecific interactions, such as competition, predation and parasitism (Chave *et al.*, 2002); and historical factors, such as a lack of dispersal opportunities (Brown *et al.*, 1996). The interplay between these factors will determine species distributions [see the BAM (biotic, abiotic and migration) diagram in Soberón (2007)], and teasing apart the relative contribution of these factors for different groups of organisms remains a major challenge (Endler, 1982; Costa *et al.*, 2008).

Environmental conditions (e.g. species' abiotic niche requirements) are considered to be the main factor limiting species distributions at large geographical scales (Grinnellian niche *sensu* Soberón, 2007) (Pearson & Dawson, 2003; Soberón, 2007; Soberón & Nakamura, 2009). Quantifying the abiotic niche does not, however, explain the entire distribution for every species. Species can occur in habitats outside their abiotic niche, because recurrent dispersal sustains 'sink' populations (Pulliam, 2000), or a species may be absent from habitat within its abiotic niche because of dispersal barriers or biotic interactions (e.g. presence of a competitor or lack of crucial resources; Eltonian niche *sensu* Soberón, 2007). Despite empirical and theoretical evidence for multifactor control of species distributions, the assumption that abiotic factors govern species distributions at large geographical scales is the foundation of recent methodologies for species distribution modelling (SDM; Elith & Leathwick, 2009). Research on SDM has grown impressively in the last decade, with studies applying various methods to many different questions in biogeography, macroecology, evolutionary biology and conservation (e.g. Peterson, 2003; Wiens *et al.*, 2006; Costa *et al.*, 2007, 2008; Costa & Schlupp, 2010). Despite usually not incorporating biotic interactions (but see Araújo & Luoto, 2007; Heikkinen *et al.*, 2007; Meier *et al.*, 2010), SDM has demonstrated considerable predictive value (e.g. Feria *et al.*, 2002; Raxworthy *et al.*, 2003; Elith *et al.*, 2006; Costa *et al.*, 2010).

How then can models relying only on abiotic variables depict such an accurate projection of species distributions? To explain this apparent paradox, two hypotheses have been formulated (Soberón & Nakamura, 2009). First, biotic factors may correlate closely with abiotic variables that capture an important part of the biotic signature (Brewer & Gaston, 2003; Soberón & Nakamura, 2009). On the other hand, biotic interactions may not affect distributions at the large extents and low resolutions that characterize most geographical distribution maps (Prinzing *et al.*, 2002). The latter argument constitutes what has been termed the Eltonian noise hypothesis (ENH; Soberón & Nakamura, 2009). Under this scenario, biotic interactions may be a major driver of abundance at smaller spatial resolutions but, because species distribution models (SDMs) are typically produced at coarse resolutions, the effects of biotic interactions may be averaged out (Soberón, 2007; Soberón & Nakamura, 2009). In fact, the role of biotic interactions in determining species distributions and its effects on SDM accuracy have been underscored as a major challenge for SDM research (Araújo & Guisan, 2006; Elith & Leathwick, 2009; Zimmermann *et al.*, 2010), with recent studies proposing different methods to account for biotic interactions in SDM (Boulangeat *et al.*, 2012; Kissling *et al.*, 2012). There is a growing interest in studies that explicitly tackle these questions (e.g. Araújo & Luoto, 2007; Heikkinen *et al.*, 2007; Real *et al.*, 2009; Costa & Schlupp, 2010; Meier *et al.*, 2010). Moreover, taken together, these studies have only modelled a few species and, in some cases, the selection of species was not random; species were selected because of prior expectations that specific biotic interactions would be important to explain their distribution at local and landscape levels. The general importance of biotic interactions in explaining species distributions therefore remains to be determined.

Here, we examine the ENH using the geographical distributions of dietary resources for different species of cerrado parrots as a metric for biotic interaction. Specifically, we test three predictions deriving from the ENH. First, that models using biotic only and/or abiotic + biotic variables do not perform better than models using abiotic variables only. Second, the importance of biotic interactions should be higher at finer resolutions and decrease at coarser resolutions (Soberón & Nakamura, 2009). Third, we expect that if there is improvement in models containing biotic variables, it should be more evident in species exhibiting stronger biotic interactions (i.e. food specialists). Foraging strategy can vary substantially among parrot species and some species show great dependence on specific plants whereas others may use many different plant species (Roth, 1984; de Araújo & Marcondes-Machado, 2011). This makes Neotropical parrots a good system to test the ENH, because we have different strengths of biotic interaction in a group of closely related organisms.

MATERIALS AND METHODS

Study system

The cerrado is a South American savanna and the second largest Neotropical biome. It contains an impressive biodiversity and is one of the world's most threatened regions (Myers *et al.*, 2000). Da Silva (1995) registered 837 bird species in the cerrado, with over 90% of those breeding in the region. However, the cerrado shows little bird endemism, as only 3.8% of breeding species are actually endemic (da Silva, 1995). There are 33 parrot species present in the cerrado, representing almost 40% of all Brazilian parrot species (CBRO, 2010), but again with little endemism as only two of those species are endemic (da Silva, 1995).

Cerrado parrots exhibit substantial variation in foraging strategy. Most species in the group are food generalists, such as *Brotogeris chiriri*, which feeds on 45 different plant species in four localities in the state of São Paulo (Paranhos *et al.*, 2007), consuming at least 116 different plant species over its entire range (de Araújo, 2011). Other species are highly specialized, including the red-bellied macaw (*Orthopsittaca manilata*), which moves great distances to find fruits of Mauritia palm (*Mauritia flexuosa*) (Roth, 1984; da Silva, 2009), or *Anodorhynchus hyacinthinus*, which feeds mainly on two plant species in the Brazilian pantanal (Antas *et al.*, 2010). This diversity of biotic requirements makes parrots a useful model to test the ENH and to explore the influence of biotic interactions on species distributions.

Although 33 parrot species have been recorded in the cerrado (da Silva, 1995), some of them were found only in single locations and might not be typical cerrado species. We used only species that occur in several localities over a wide geographical range within the cerrado, including species with different habitat restrictions and different geographical range sizes and positions. We thus avoided using species that are typical of different ecosystems and have only been registered in a few cerrado localities. Hence, we have chosen species that are present in a wide range of cerrado habitats and have available diet data (see Appendix S1 in the Supporting Information). Based on these two criteria, we selected the following 11 cerrado parrot species: *Amazona aestiva*, *Alipiopsitta xanthops*, *Pionus maximiliani*, *Forpus xanthopterygius*, *Aratinga aurea*, *Ara ararauna*, *Ara chloropterus*, *Anodorhynchus hyacinthinus*, *Orthopsittaca manilata*, *Diopsittaca nobilis* and *Brotogeris chiriri*. These 11 species are typical of cerrado, have diet information, and exhibit considerable variation in dietary breadth. We followed the Brazilian Committee for Ornithological Records (CBRO, 2010) for parrot species taxonomy.

Species distribution modelling

Species distribution modelling (SDM) requires two distinct data sets: occurrence information for the species of interest (georeferenced localities), and GIS layers of abiotic and/or biotic variables. We used Maxent 3.3.3k, a method that has been demonstrated to perform well in a diverse set of modelling scenarios and is widely used in a great number of studies in ecology, biogeography and conservation (Elith & Leathwick, 2009). The algorithm in Maxent works by fitting a probability distribution for species occurrence to the set of pixels across the region of interest. For a detailed explanation on how the maximum entropy principle applies to SDM, see Elith *et al.* (2011). We performed 100 replications, using a cross-validation procedure where we divided our dataset using 75% of data for model calibration and retaining 25% of the data to evaluate models. We report the mean and standard deviation of area under the curve of the receiver operating characteristic (AUC) test values for the 100 runs. We used Maxent default parameters (Phillips & Dudík, 2008).

We used different methods to obtain locality data for parrots and their diet plants. In order to acquire parrot localities, we searched through the Ornis (http://www.ornisnet.org/) and Species Link (http://www.splink.cria.org.br/) databases, and requested data directly from museum curators and researchers. To obtain data for the plant species, we used the Species Link (http://splink.cria.org.br/), Tropicos (http://www.tropicos.org/) and the New York Botanical Garden (http://www.nybg.org/) databases. Together, these databases gave us access to a great number of plant collections. A complete data list is available from the authors upon request.

Abiotic variables

We used the bioclimatic environmental variables from the WorldClim project version 1.4 (Hijmans *et al.*, 2005; available online at: http://www.worldclim.org/). These variables were downloaded from the WorldClim project at four different spatial resolutions: 10, 5 and 2.5 arc-minutes, and 30 arc-seconds (approximately 20, 10, 5 and 1-km cell sizes, respectively). The bioclimatic layers were cropped to span latitudes from 12° 47′ N to 34° 46′ S and longitudes from 78° 31′ W to 35° W; this represents a larger spatial range than the cerrado region and also includes tropical and subtropical zones with bioclimatic conditions compatible with the occurrence of cerrado species. To avoid overfitting the models and using redundant climatic variables, we identified highly correlated variables ($r > 0.9$) and excluded one of them from the model based on their biological relevance. We then built another correlation matrix and repeated the procedure until all variables kept in the model presented correlations lower than 0.9; a similar procedure was described in Rissler & Apodaca (2007). We used a total of nine abiotic environmental variables: BIO3, isothermality; BIO4, temperature seasonality; BIO7, annual temperature range; BIO10, mean temperature of the warmest quarter; BIO11, mean temperature of the coldest quarter; BIO14, precipitation of the driest month; BIO15, precipitation seasonality; BIO16, precipitation of the wettest quarter; and BIO17, precipitation of the driest quarter.

Biotic variables

Coarse-grained layers of abundances for specific plants would be the ideal dataset by which to estimate diet resource availability, but such data are not easily available, especially for larger geographical extents. In order to capture geographical variation in the availability of parrot food resources, we built SDMs for key plant species that were included in the diet of each parrot species (Appendix S1). The results of these models are raster grids in which values can range from 0 to 1 (Maxent, logistic output), where higher values indicate higher probability of occurrence. Previous studies have shown that the values of these grids also correlate with species abundance, where localities with higher values have higher abundance values for the species being studied (VanDerWal *et al.*, 2009; Kulhanek *et al.*, 2011; Oliver *et al.*, 2012; Tôrres *et al.*, 2012). Here, we use these grids directly in SDMs as representations of biotic variables. In this case, the grids are a proxy for plant species abundance or a measure of dietary resource availability. The procedure we adopted to select which plants represent the relevant biotic variables for each parrot species is described in Appendix S1. In addition, because the biotic variables were suitability scores already obtained from the abiotic variables, we transformed the biotic continuous scores into binary (presence/absence) maps, and then used those maps as the biotic variable. This approach has been used before in the literature to incorporate biotic variables in SDMs (Barbet-Massin & Jiguet, 2011). The use of binary versus continuous plant models did not qualitatively change the results; we therefore present only the results of the analysis with the continuous models. Details and results of the analysis using binary maps can be seen in Appendix S2.

We modelled parrot distributions using (1) abiotic variables alone, (2) biotic variables alone, and (3) both abiotic and biotic variables. All models for plant species were constructed using the methods and abiotic variables described in the previous section. The ENH can be rejected if the models using biotic only and/or abiotic + biotic variables perform better than models using abiotic variables only. We also expect that the importance of biotic interactions should be higher at finer resolutions and decrease at coarser resolutions (Soberón & Nakamura, 2009). Moreover, we expect that if there is improvement, it should be more evident in species exhibiting stronger biotic interactions (i.e. food specialists).

Variance partitioning

In order to determine the amount of independent and shared information contained in the abiotic and biotic variables, we used a variance-partitioning approach. We fitted generalized linear models (GLM) with presence and pseudo-absence as the response variables and using three different combinations of predictors for each parrot species: just the abiotic variables (abiotic), just the biotic variables (biotic), and the abiotic and biotic variables combined (full). Model fit was evaluated using the adjusted coefficient of determination (adj. R^2). There are many different ways to calculate coefficient of determination for logistic regression models; we use McFadden's (McFadden, 1974). This metric is preferable to other R^2 analogues because it has the most intuitive interpretation and is the most generally applicable and consistently useful (Menard, 2000). To partition the variance explained by the abiotic and biotic variables, we calculated the three partial models as discussed above, estimating the adjusted R^2 for each model type per species. Next, the pure effect of abiotic variables was calculated as $R^2_{\text{pure.abiotic}} = 1 - R^2_{\text{biotic}}$. The pure effect of biotic variables was calculated as $R^2_{\text{pure.biotic}} = 1 - R^2_{\text{abiotic}}$, and the effect shared by abiotic and biotic variables was calculated as $R^2_{\text{shared.abiotic+biotic}} = 1 - R^2_{\text{pure.abiotic}} - R^2_{\text{pure.biotic}}$. Similar variance-partitioning procedures have been used in other SDM contexts have been performed in recent work (Muñoz & Real, 2006; Real *et al.*, 2013). In this procedure, the output values are the proportional contribution to the total variation explained by the model. The amount of unexplained variation is not known (Real *et al.*, 2013). We used R 2.15 (R Development Core Team, 2012) in all analyses and the *varPart* function of the package modEvA (Barbosa *et al.*, 2013) for variance partitioning. For the GLM, we selected 10,000 random pseudo-absences points from throughout the sampling region, as suggested by Barbet-Massin *et al.* (2012).

Testing the Eltonian niche hypothesis

We tested the ENH (Soberón & Nakamura, 2009) by comparing the performance of models produced with only abiotic variables to models produced using only biotic variables and both abiotic and biotic variables. SDMs are traditionally evaluated by using the AUC statistic. The curve is obtained by plotting all sensitivity values (true positive fraction) on the *y*-axis against their equivalent (1 − specificity) values (false positive fraction) for all available thresholds on the *x*-axis. The AUC provides a threshold-independent measure of model performance as compared with that of null expectations (Fielding & Bell, 1997). Many studies have pointed out problems with the AUC approach in SDM, but these problems mostly concern the comparison of different methods and/or species, which is not the case here (Raes & ter Steege, 2007; Lobo *et al.*, 2008; Peterson *et al.*, 2008). To assess whether models generated with different sets of predictors (abiotic; biotic; abiotic + biotic) yielded different mean AUC values, we performed analyses of variance (ANOVAs) with Tukey's post-hoc test to explore which pairwise comparisons were significant. We used 100 replicates for each species and set of predictors in cross-validation runs, holding 25% of occurrence points to validate and calculate AUC statistics. In addition, to test whether the number of comparisons that resulted in increases of AUC with the addition of biotic variables would be higher than expected by chance, we used Wilcoxon's signed-rank test, using all comparisons between abiotic and abiotic + biotic models. Finally, we used GLMs to test for differences in omission and commission error rates in models built with the different set of predictors. Should there be differences in AUC values (e.g. higher AUC with the addition of biotic variables), model improvement can be achieved by decreasing omission and/or commission error rates. In the case of omission errors, the addition of biotic variables would improve the ability of the models to correctly classify a presence even when abiotic conditions are marginally suitable. In the case of commission errors, the addition of biotic variables could remove from models areas that have suitable abiotic conditions but not suitable biotic conditions. We compared average omission and commission test rates for our 100 runs. We performed the analysis using three different criteria to select a threshold to transform continuous output into binary predictions (minimum training presence, 10% training presence, and equal sensitivity and specificity).

There is also the possibility that the ENH is only refuted when stronger levels of biotic interactions are present. In this case, we expect that species with narrow dietary niche breadth (specialists) would present a greater model improvement (higher differences between AUCs from abiotic versus abiotic + biotic models), as the dependence on the food resource would be the highest. We therefore expect a negative relationship between model improvement (for models with biotic variables) and dietary niche breadth. To test for a relationship between dietary niche breadth and model improvement, we used a general linear model. As a measure of model improvement we calculated the proportional measure of AUC improvement $AUC_{imp} = (AUC_{abiotic+biotic} - AUC_{abiotic})/AUC_{abiotic}$. As a measure of parrot dietary niche breadth, we used their diet richness (see Appendix S1).

The ENH also predicts that biotic interactions will be more important at higher resolutions. To test for this effect, we calculated AUC_{imp} within resolutions between models with the abiotic-only set of predictors and models with both abiotic and biotic predictors, for each species. Next, we performed a nonparametric analysis of variance (Kruskal–Wallis) to assess differences between median proportional increase in AUC among resolutions. ENH would predict that proportional improvement of AUC would be higher at 30-arc-seconds resolution (our highest resolution). All statistical analyses were carried out in R 2.15 (R Development Core Team, 2012).

RESULTS

On average, we collected 92 unique localities per grid cell for the 11 chosen cerrado parrot species ($n = 1014$; minimum of 68 for *Orthopsittaca manilata* and a maximum of 132 for *Brotogeris chiriri*; Appendix S2). Using our plant importance index (I), we determined the most important food resource for each parrot species (Appendix S1). We obtained on average 98 unique localities for each key plant species ($n = 4715$; minimum of 13 for *Sterculia apetala* and maximum of 528 for *Qualea parviflora*).

The parrot SDMs using only biotic variables performed worst. Abiotic-only models generally provided a good discrimination capacity, with most AUC values ranging between 0.792 and 0.931 (Fig. 1). Despite the initial good performance of the models, the inclusion of biotic variables increased AUC for all SDMs generated with one of the resolutions modelled (Fig. 1, Appendix S2). For six species (*Amazona aestiva*, *Ara ararauna*, *Ara chloropterus*, *Alipiopsitta xanthops*, *Brotogeris chiriri* and *Forpus xanthopterygius*), models at all four resolutions had higher mean AUC values when biotic variables were included in the model. For *Anodorhynchus hyacinthinus* and *Pionus maximiliani*, the addition of biotic variables resulted in higher mean AUC values at three different resolutions. *Aratinga aurea*, *Diopsittaca nobilis* and *Orthopsittaca manilata* produced mixed results, with different sets of models achieving better performance at different resolutions. The ANOVA comparison of mean AUC values from models using different sets of predictors showed that biotic-only models had significantly lower values than abiotic-only and abiotic + biotic models (Fig. 1). No models using abiotic-only variables had significantly higher AUC values. Higher AUC values were found in 36% of comparisons when abiotic + biotic models were used. The remaining comparisons were not statistically significant (Fig. 1, Appendix S2). Also, the Wilcoxon signed-rank test showed that the number of comparisons that resulted in increased AUC with the addition of biotic variables is higher than expected by chance ($z = 4.33$; $P < 0.001$).

The results of our comparison of omission and commission errors between models fitted only with abiotic variables versus models fitted with abiotic and biotic variables, depended on species, resolution and choice of threshold. In general, however, most significant comparisons showed that models with abiotic and biotic variables have lower commission errors and higher omission errors (Appendix S2). Overall, there was no difference in proportional improvement of AUC for models using only abiotic variables versus models with abiotic and biotic variables at different resolutions (Kruskal–Wallis $\chi^2 = 1.75$, d.f. = 3, $P = 0.62$). The variance-partitioning approach showed that, for the majority of species at all different modelled resolutions, the pure abiotic contribution was higher than the pure biotic contribution. The magnitude of the shared effect represents the proportion of the model variation that is attributable to either kind of variable. In our case, the shared component had high negative values, indicating strong independent contributions of the two sets of predictors (Table 1).

Rarefaction of diet data shows that the three macaws (*Anodorhynchus hyacinthinus*, *Ara chloropterus* and *Orthopsittaca manilata*) have the most specialized diets; *Amazona aestiva*, *Pionus maximiliani*, *Brotogeris chiriri* and *Aratinga aurea* have generalist diets (Appendix S2). *Orthopsittaca manilata* specializes on Mauritia palm (*Maurita flexuosa*), so only a single diet component was considered for this species. There was no relationship between diet specialization and proportional increase in AUC (GLM, $z = -0.13$, $P = 0.89$).

DISCUSSION

The Eltonian noise hypothesis states that biotic interactions will not be relevant to predict species distributions at coarse-grained resolutions and large extents. Based on this hypothesis, we predicted that the addition of biotic variables in relatively coarse models would not lead to significant improvements in model predictions. Despite some exceptions, our overall results did not support this hypothesis, as most species yielded higher AUC when biotic variables were added to the set of predictors. We also found that the increase in AUC is likely to be caused by a reduction of commission errors, suggesting that the addition of biotic variables narrows model predictions to areas that have suitable abiotic and biotic conditions. A visual inspection of the models also confirms that pattern (Appendix S3). Our results support the findings of previous studies that investigated the influence of biotic interactions in distribution models, which found a substantial influence of biotic variables in explaining

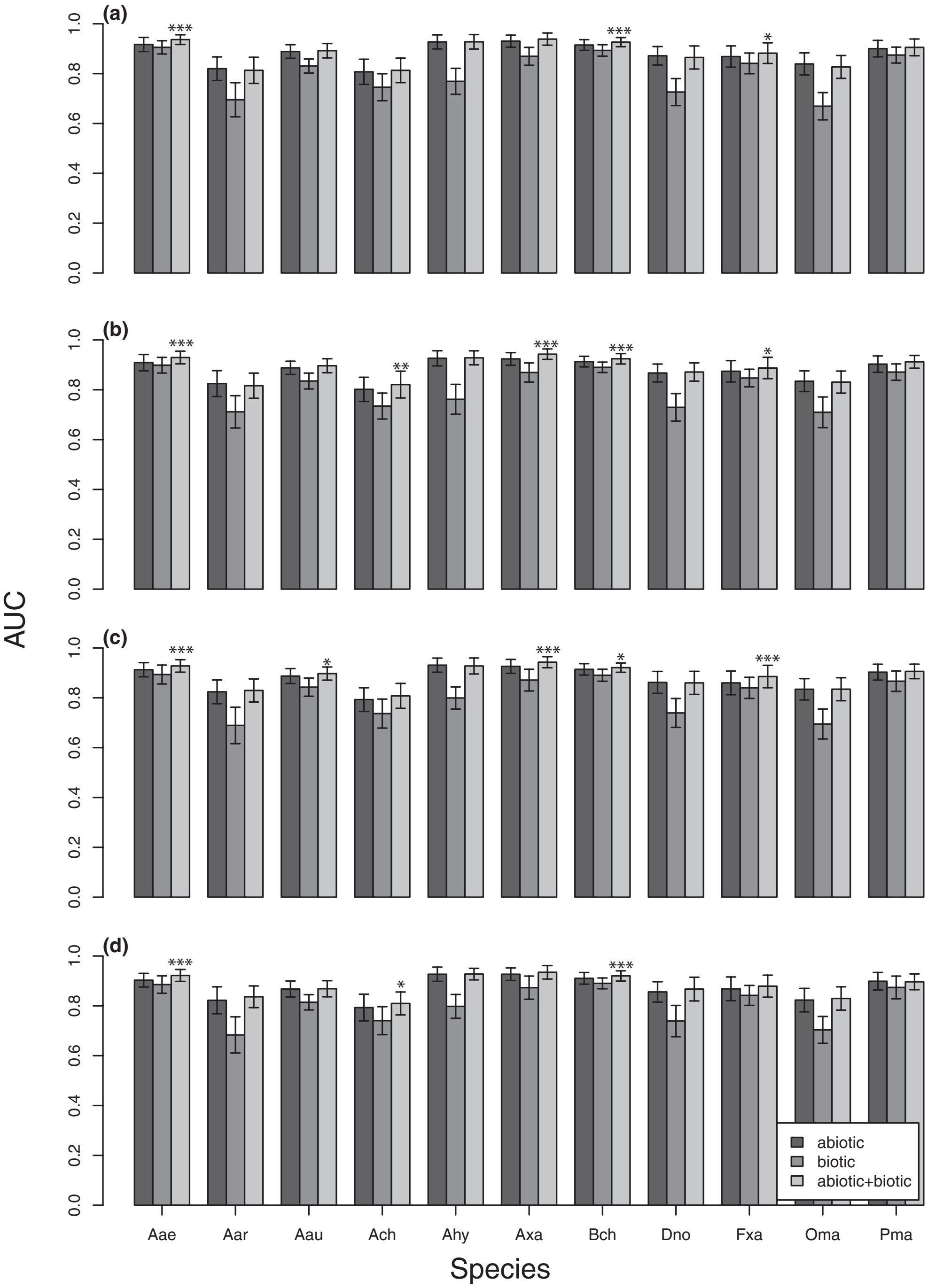

Figure 1 Statistical results of SDM for 11 species of parrots in the Brazilian cerrado. Average AUC values for test data for models at different resolutions (error bars = standard deviation): (a) 30 arc-seconds; (b) 2.5 arc-minutes; (c) 5 arc-minutes; and (d) 10 arc-minutes. We used 100 replicates for each species and set of predictors in cross-validation runs holding 25% of occurrence points to validate and calculate AUC statistics. Aae, *Amazona aestiva*; Aar, *Ara ararauna*; Aau, *Aratinga aurea*; Ach, *Ara chloropterus*; Ahy, *Anodorhynchus hyacinthinus*; Axa, *Alipiopsitta xanthops*; Bch, *Brotogeris chiriri*; Dno, *Diopsittaca nobilis*; Fxa, *Forpus xanthopterygius*; Oma, *Orthopsittaca manilata*; Pma, *Pionus maximiliani*. Asterisks (*) on top of bars indicate significant differences between abiotic-only and abiotic + biotic models. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

coarse-scale distributions (Araújo & Luoto, 2007; Heikkinen *et al.*, 2007; Meier *et al.*, 2010). Araújo & Luoto (2007) tested the role of three host plants on the distribution of the European clouded apollo butterfly (*Parnassius mnemosyne*) and found that biotic interactions significantly affected both the explanatory and predictive power of SDMs at macro-scales. Heikkinen *et al.* (2007) showed that owl distribution models using climate, land cover and biotic interactions (presence of a facilitator species – woodpeckers) generated higher AUC values than models using only climate and land

Table 1 Summary of variance-partitioning approach for models of 11 parrot species from the Brazilian cerrado. Coefficients of determination (adj. R^2) are listed for models containing only the abiotic set of predictors, only the biotic set of predictors, and both abiotic and biotic variables. In some cases, the shared component may become negative as a result of opposing effects of the two sets of predictors in the full model (Legendre & Legendre, 1998).

	Abiotic only				Biotic only				Shared			
Species	30″	2.5′	5′	10′	30″	2.5′	5′	10′	30″	2.5′	5′	10′
Amazona aestiva	0.85	0.84	0.82	0.83	0.86	0.87	0.86	0.86	−0.72	−0.71	−0.68	−0.69
Ara ararauna	0.99	0.99	0.99	0.99	0.91	0.89	0.90	0.90	−0.90	−0.88	−0.89	−0.89
Aratinga aurea	0.90	0.90	0.90	0.90	0.86	0.86	0.86	0.86	−0.75	−0.76	−0.76	−0.75
Ara chloropterus	0.99	0.98	0.98	0.98	0.92	0.92	0.92	0.91	−0.90	−0.90	−0.90	−0.88
Anodorhynchus hyacinthinus	0.94	0.93	0.92	0.91	0.81	0.80	0.80	0.80	−0.76	−0.74	−0.72	−0.71
Alipiopsitta xanthops	0.91	0.91	0.91	0.91	0.83	0.83	0.84	0.83	−0.74	−0.74	−0.75	−0.74
Brotogeris chiriri	0.94	0.94	0.88	0.88	0.86	0.85	0.85	0.85	−0.79	−0.79	−0.73	−0.73
Diopsittaca nobilis	1.00	1.00	0.99	0.99	0.87	0.88	0.88	0.87	−0.87	−0.87	−0.87	−0.86
Forpus xanthopterygius	0.88	0.88	0.88	0.88	0.89	0.90	0.90	0.90	−0.78	−0.78	−0.78	−0.78
Orthopsittaca manilata	0.98	0.98	0.98	0.98	0.90	0.90	0.90	0.90	−0.89	−0.88	−0.88	−0.88
Pionus maximiliani	0.91	0.89	0.89	0.89	0.90	0.90	0.91	0.91	−0.81	−0.79	−0.80	−0.80
Average	0.94	0.93	0.92	0.92	0.87	0.87	0.87	0.87	−0.81	−0.80	−0.80	−0.79
SD	0.05	0.05	0.06	0.06	0.03	0.03	0.04	0.04	0.07	0.07	0.08	0.08

cover. Meier *et al.* (2010), studying the distribution of 11 tree species in Switzerland, found not only that the inclusion of biotic variables increased model quality, but also that the component of variation explained by these two sets of variables showed little redundancy.

Our variance-partitioning approach demonstrated a negative shared variance component. A negative value in the shared component of the variance indicates that the two groups of variables, together, explain the response variable better than the sum of the individual effects of these variables. This occurs when the set of predictors have opposing effects on the response (Legendre & Legendre, 1998). Negative values could also appear when there are strong correlations between groups of variables. We examined the correlation between abiotic and biotic variables for each species and we generally find no strong correlations. Our results suggest that biotic variables provide valuable new information within the models, which is independent of the abiotic predictors.

We also evaluated the prediction that model improvement would be related to the species' dietary niche breadth, which is a surrogate for the strength of the biotic interaction. Again, we found no support for this prediction, as there was no relationship between AUC improvement and diet richness. In fact, *Amazona aestiva* had the most generalist diet and showed model improvement with the addition of biotic variables at all resolutions. We also tested a prediction derived from the ENH, that models accounting for biotic interactions would have higher proportional AUC increase at higher resolution. We found no support for this prediction, as we found no relationship between model improvement and resolution. The ENH was, however, explicitly proposed in a context where biotic interactions would take place at a much higher resolution than the ones used in this study (Soberón & Nakamura, 2009). Our finest resolution (1-km cell size) may still be too coarse for a proper assessment of the relationship between the importance of biotic variables and species distributions, although parrots use a large portion of their habitat on a daily basis (de Araújo & Marcondes-Machado, 2011; de Araújo *et al.*, 2011). A more appropriate test of this prediction would consider much finer resolutions. Nevertheless, a previous study using coarse-grained resolutions was able to show that biotic interactions were more important at relatively higher resolution (Heikkinen *et al.*, 2007).

One of the main reasons for the growth of studies applying SDM methods is the increase in availability of world-wide abiotic environmental variables that are easily downloaded and incorporated into SDMs (Graham *et al.*, 2004; Peterson, 2006; Kozak *et al.*, 2008). On the other hand, biotic interactions are not easily mapped onto geographical space and very few studies have therefore been able to incorporate them into modelling. Many studies using only abiotic variables have nonetheless been able to provide good estimates of species distributions, and have helped address many questions in biogeography, evolution and conservation (Elith & Leathwick, 2009). Our study has important implications for SDM research. Our results, together with previous evidence, suggest that the use of biotic variables improves the resulting models. Our methods can be adapted to be used for models of different types of organisms, especially organisms that feed on particular plant species. The understanding of species geographical distributions and its implications for theoretical and applied research has gained considerable attention since the advent and development of SDM methods (Elith *et al.*, 2006; Kozak *et al.*, 2008; Elith & Leathwick, 2009). There are still many challenges facing the field, and factors such as sample size (Stockwell & Peterson, 2002; Hernandez *et al.*, 2006), spatial scale (Lassueur *et al.*, 2006; Guisan *et al.*, 2007; Trivedi *et al.*, 2008) and the nature of the environmental data set

(Parra *et al.*, 2004; Peterson & Nakazawa, 2008) have been identified as influences on the outcome of models produced. Among the myriad potential factors, the role of biotic interactions has been considered one of the major challenges, and despite many recent developments, has not been adequately explored (Pearson & Dawson, 2003; Elith & Leathwick, 2009; Zimmermann *et al.*, 2010). Recent studies have explored the conceptual and methodological issues regarding the role of biotic interactions in SDM (Boulangeat *et al.*, 2012; Godsoe & Harmon, 2012; Wisz *et al.*, 2013). This is not only a theoretical but also a practical problem for SDM, because the use of biotic variables in SDM requires complex biological phenomena to be mapped. This comes from the fact that biotic interactions may come in many forms. Our study explicitly incorporated a biotic variable (diet resource distribution) into SDMs, and we showed that we generally obtain improved models by doing so. This demonstrates that SDMs performed only with abiotic variables may depict only a partial representation of a species' geographical distribution. However, as seen in our results, there are species where the addition of biotic variables did not improve model performance, so we acknowledge that much work is still needed to provide a general theoretical and practical framework for the role of biotic interactions in predicting large-scale geographical distributions.

ACKNOWLEDGEMENTS

Don Shepard, Frank Burbrink and Rachel Costa provided useful comments on the manuscript. R. Caparroz, G. Brito, E. Lima, P. Silva and W. Nogueira kindly sent GPS localities of the parrots. CNPq (grant no. 563352/2010-8 to G.C.C.; grant no. 140121/2009-9 to C.B.A.) and CAPES provided funding for this work.

REFERENCES

Antas, P.T.Z., Carrara, L.A., Yabe, R.S., Ubaid, F.K., Oliveira-Junior, S.B., Vasques, E.R. & Ferreira, L.P. (2010) *A arara-azul na Reserva Particular do Patrimônio Natural SESC Pantanal*. SESC, Rio de Janeiro.

de Araújo, C.B. (2011) *Psitacídeos do Cerrado: sua alimentação, comunicação sonora e aspectos bióticos e abióticos de sua distribuição potencial*. PhD Thesis. Universidade Estadual de Campinas, Campinas, São Paulo, Brazil.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.

Araújo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.

de Araújo, C.B. & Marcondes-Machado, L.O. (2011) Diet and feeding behavior of the Yellow-faced Parrot (*Alipiopsitta xanthops*) in Brasília, Brazil. *Ornitología Neotropical*, **22**, 79–88.

de Araújo, C.B., Marcondes-Machado, L.O. & Vielliard, J.M.E. (2011) Vocal repertoire of the Yellow-faced Parrot (*Alipiopsitta xanthops*). *Wilson Journal of Ornithology*, **123**, 603–608.

Barbet-Massin, M. & Jiguet, F. (2011) Back from a predicted climatic extinction of an island endemic: a future for the Corsican Nuthatch. *PLoS ONE*, **6**, e18228.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Barbosa, A.M., Brown, J.A. & Real, R. (2013) *modEvA – an R package for model evaluation and analysis*. Available at: http://modtools.wordpress.com/ (accessed 15 August 2013).

Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584–593.

Brewer, A.M. & Gaston, K.J. (2003) The geographical range structure of the holly leaf-miner. II. Demographic rates. Journal of Animal Ecology, **72**, 82–93.

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.

CBRO (2010) *Listas das aves do Brasil*. Comitê Brasileiro de Registros Ornitológicos, São Paulo, Brazil. Available at: http://www.cbro.org.br/CBRO/listabr.htm (accessed July 2010).

Chave, J., Muller-Landau, H.C. & Levin, S.A. (2002) Comparing classical community models: theoretical consequences for patterns of diversity. *The American Naturalist*, **159**, 1–23.

Costa, G.C. & Schlupp, I. (2010) Biogeography of the Amazon molly: ecological niche and range limits of an asexual hybrid species. *Global Ecology and Biogeography*, **19**, 442–451.

Costa, G.C., Nogueira, C., Machado, R.B. & Colli, G.R. (2007) Squamate richness in the Brazilian Cerrado and its environmental–climatic associations. *Diversity and Distributions*, **13**, 714–724.

Costa, G.C., Wolfe, C., Shepard, D.B., Caldwell, J.P. & Vitt, L.J. (2008) Detecting the influence of climatic variables on species distributions: a test using GIS niche-based models along a steep longitudinal environmental gradient. *Journal of Biogeography*, **35**, 637–646.

Costa, G.C., Nogueira, C., Machado, R.B. & Colli, G.R. (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiversity and Conservation*, **19**, 883–899.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Endler, J.A. (1982) Problems in distinguishing historical from ecological factors in biogeography. *American Zoologist*, **22**, 441–452.

Feria, A., Patricia, T & Peterson, A.T. (2002) Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. *Diversity and Distributions*, **8**, 49–56.

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.

Gaston, K.J. (2009) Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1391–1393.

Godsoe, W. & Harmon, L.J. (2012) How do species interactions affect species distribution models? *Ecography*, **35**, 811–820.

Graham, R.W., Lundelius, E.L., Graham, M.A., Schroeder, E.K., Toomey, R.S., Anderson, E., Barnosky, A.D., Burns, J.A., Churcher, C.S., Grayson, D.K., Guthrie, R.D., Harington, C.R., Jefferson, G.T., Martin, L.D., McDonald, H.G., Morlan, R.E., Semken, H.A., Jr, Webb, S.D., Werdelin, L. & Wilson, M.C. (1996) Spatial response of mammals to late Quaternary environmental fluctuations. *Science*, **272**, 1601–1606.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, **19**, 497–503.

Guisan, A., Graham, C.H., Elith, J. & Huettmann, F. & NCEAS Species Distribution Modelling Group (2007) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, **13**, 332–340.

Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Körber, J.-H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773–785.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.-C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, **39**, 2163–2178.

Kozak, K.H., Graham, C.H. & Wiens, J.J. (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution*, **23**, 141–148.

Kulhanek, S.A., Leung, B. & Ricciardi, A. (2011) Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecological Applications*, **21**, 203–213.

Lassueur, T., Joost, S. & Randin, C.F. (2006) Very high resolution digital elevation models: do they improve models of plant species distribution? *Ecological Modelling*, **198**, 139–153.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*, 2nd edn. Elsevier, Amsterdam.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics* (ed. by P. Zarembka), pp. 105–142. Academic Press, New York.

Meier, E.S., Kienast, F., Pearman, P.B., Svenning, J.-C., Thuiller, W., Araújo, M.B., Guisan, A. & Zimmermann, N.E. (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038–1048.

Menard, S. (2000) Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, **54**, 17–24.

Muñoz, A.-R. & Real, R. (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. *Diversity and Distributions*, **12**, 656–665.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Oliver, T.H., Gillings, S., Girardello, M., Rapacciuolo, G., Brereton, T.M., Siriwardena, G.M., Roy, D.B., Pywell, R. & Fuller, R.J. (2012) Population density but not stability can be predicted from species distribution models. *Journal of Applied Ecology*, **49**, 581–590.

Paranhos, S.J., de Araújo, C.B. & Marcondes-Machado, L.O. (2007) Comportamento alimentar do Periquito-de-encontro-amarelo (*Brotogeris chiriri*) no interior do estado de São Paulo, Brasil. *Revista Brasileira de Ornitologia*, **15**, 95–101.

Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.

Parra, J.L., Graham, C.C. & Freile, J.F. (2004) Evaluating alternative data sets for ecological niche models of birds in the Andes. *Ecography*, **27**, 350–360.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *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. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135–144.

Peterson, A.T., Papeş, M. & Soberón, J. (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, **213**, 63–72.

Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2002) Geographic variability of ecological niches of plant species: are competition and stress relevant? *Ecography*, **25**, 721–729.

Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.

R Development Core Team (2012) *R: a language and environment for statistical computing*. Version 2.15. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www. r-project.org/.

Raes, N. & ter Steege, H. (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727–736.

Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A. & Peterson, A.T. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, **426**, 837–841.

Real, R., Barbosa, A.M., Rodríguez, A., García, F.J., Vargas, J.M., Palomo, L.J. & Delibes, M. (2009) Conservation biogeography of ecologically interacting species: the case of the Iberian lynx and the European rabbit. *Diversity and Distributions*, **15**, 390–400.

Real, R., Romero, D., Olivero, J., Estrada, A. & Márquez, A.L. (2013) Estimating how inflated or obscured effects of climate affect forecasted species distribution. *PLoS ONE*, **8**, e53646.

Rissler, L.J. & Apodaca, J.J. (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, **56**, 924–942.

Roth, P. (1984) Repartição do habitat entre psitacídeos simpátricos no sul da Amazônia. *Acta Amazonica*, **14**, 175–221.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.

da Silva, J.M.C. (1995) Birds of the Cerrado Region, South America. *Steenstrupia*, **21**, 69–92.

da Silva, P.A. (2009) Orthopsittaca manilata *(Boddaert, 1783) (Aves: Psittacidae): abundância e atividade alimentar em relação à frutificação de* Mauritia flexuosa *L. f. (Arecaceae) numa vereda do Triângulo Mineiro*. MSc Thesis, Universidade Federal de Uberlândia, Uberlândia, Brazil.

Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.

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

Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.

Tôrres, N.M., De Marco, P., Santos, T., Silveira, L., de Almeida Jácomo, A.T. & Diniz-Filho, J.A.F. (2012) Can species distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. *Diversity and Distributions*, **18**, 615–627.

Trivedi, M.R., Berry, P.M., Morecroft, M.D. & Dawson, T.P. (2008) Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Global Change Biology*, **14**, 1089–1103.

VanDerWal, J., Shoo, L.P., Johnson, C.N. & Williams, S.E. (2009) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, **174**, 282–291.

Wiens, J.J., Graham, C.H., Moen, D.S., Smith, S.A. & Reeder, T.W. (2006) Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. *The American Naturalist*, **168**, 579–596.

Wisz, M.S., Pottier, J., Kissling, W.D. *et al.* (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, **88**, 15–30.

Zimmermann, N.E., Edwards, T.C., Graham, C.H., Pearman, P.B. & Svenning, J.-C. (2010) New trends in species distribution modelling. *Ecography*, **33**, 985–989.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Methods for selection of plant species, and methods to estimate parrot diet richness.

Appendix S2 Numerical results of the analysis shown in Fig. 1 and results of the analysis using as biotic variables the transformation of continuous scores into binary (presence/absence) maps.

Appendix S3 Supplementary figures: species distribution models for the 11 parrot species and 24 plant species studied.

BIOSKETCHES

Carlos Barros de Araújo is a postdoctoral researcher at the Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba. His research interests are the bioacoustics of several groups and the natural history of parrots.

Luiz Octavio Marcondes-Machado is an associate professor in the Department of Animal Biology, Institute of Biology of the Universidade Estadual de Campinas. His main research interests are the feeding and reproductive behaviour of Brazilian parrots.

Gabriel C. Costa is an assistant professor in the Department of Ecology of the Universidade Federal do Rio Grande do Norte, in the north-east of Brazil. His research interests are in macroecology, species distribution modelling and factors affecting species range limits.

Author contributions: C.B.A. and G.C.C. conceived the ideas; C.B.A. and L.O.M. collected the data; G.C.C. and C.B.A. analysed the data; and C.B.A. led the writing with contributions of both co-authors.

Editor: Miles Silman