How many studies are necessary to compare niche-based models for geographic distributions? Inductive reasoning may fail at the end

Terribile, LC.[a,b]*, Diniz-Filho, JAF.[b] and De Marco Jr., P.[b]

[a]Centro de Ciências Agrárias e Biológicas, Universidade Federal de Goiás – UFG,
Unidade Jatobá - Campus Jataí, Br 364, Km 192, CEP 75801-615, Jataí, GO, Brazil

[b]Departamento de Biologia Geral, Instituto de Ciências Biológicas – ICB,
Universidade Federal de Goiás – UFG,
CP 131, CEP 74001-970, Goiânia, GO, Brazil

*e-mail: levicarina@gmail.com

(With 2 figures)

Abstract

The use of ecological niche models (ENM) to generate potential geographic distributions of species has rapidly increased in ecology, conservation and evolutionary biology. Many methods are available and the most used are Maximum Entropy Method (MAXENT) and the Genetic Algorithm for Rule Set Production (GARP). Recent studies have shown that MAXENT perform better than GARP. Here we used the statistics methods of ROC – AUC (area under the Receiver Operating Characteristics curve) and bootstrap to evaluate the performance of GARP and MAXENT in generate potential distribution models for 39 species of New World coral snakes. We found that values of AUC for GARP ranged from 0.923 to 0.999, whereas those for MAXENT ranged from 0.877 to 0.999. On the whole, the differences in AUC were very small, but for 10 species GARP outperformed MAXENT. Means and standard deviations for 100 bootstrapped samples with sample sizes ranging from 3 to 30 species did not show any trends towards deviations from a zero difference in AUC values of GARP minus AUC values of MAXENT. Ours results suggest that further studies are still necessary to establish under which circumstances the statistical performance of the methods vary. However, it is also important to consider the possibility that this empirical inductive reasoning may fail in the end, because we almost certainly could not establish all potential scenarios generating variation in the relative performance of models.

Keywords: coral snakes, Elapidae, potential distribution models, GARP, MAXENT.

Quantos estudos ainda serão necessários para avaliar modelos de distribuição geográfica baseados em modelagem do nicho? Um pensamento indutivo pode estar fadado ao fracasso

Resumo

A utilização de modelos de nicho ecológico (ENM) para gerar distribuições geográficas potenciais de espécies tem aumentado rapidamente nas áreas de ecologia, biologia da conservação e biologia evolutiva. O Método de Máxima Entropia (MAXENT) e o Algoritmo Genético para Produção de Conjunto de Regras (GARP) estão entre os métodos mais utilizados, e estudos recentes têm atribuído ao MAXENT um melhor desempenho no processo de modelagem com relação ao GARP. Neste trabalho, foram utilizados os métodos estatísticos ROC – AUC (*area under the Receiver Operating Characteristics curve*) e de reamostragem (*bootstrap*) para avaliar o desempenho do GARP e MAXENT em gerar modelos de distribuição potencial para 39 espécies de cobras corais do Novo Mundo. Os resultados mostraram que os valores de AUC para o GARP variaram de 0,923 a 0,999, enquanto que para o MAXENT variaram de 0,877 a 0,999. Em geral, as diferenças de AUC entre os dois métodos foram pequenas, embora o GARP tenha apresentado melhor desempenho que o MAXENT para 10 espécies. Valores de média e desvio padrão de 100 amostras variando de 3 a 30 espécies não revelaram qualquer tendência de desvio em relação à diferença zero entre valores de AUC do GARP menos valores de AUC do MAXENT. Estes resultados sugerem que mais estudos serão necessários para determinar sob quais circunstâncias o desempenho estatístico dos modelos varia, embora seja importante considerar também a possibilidade de que argumentações empírico-indutivas em favor de um ou outro método podem falhar, já que é quase impossível estabelecer todos os cenários potenciais causadores de variação no desempenho dos modelos.

Palavras-chave: cobras corais, Elapidae, modelos de distribuição potencial, GARP, MAXENT.

1. Introduction

Niche-based species distribution models, or eco-logical niche models (ENM), today play a central role in many areas of ecology, conservation and evolutionary biology, both because they can fill gaps in knowledge and allow a better estimate of multiple components of species diversity (Guisan and Zimmermann, 2000; Araújo and Guisan, 2006; Phillips et al., 2006). Also, they can be used, under certain assumptions, to predict the fate of biodiversity under ongoing climate change processes (Guisan and Thuiller, 2005; Araújo and New, 2007). Despite many challenges for the future (Araújo and Guisan, 2006), interest in these approaches is clearly growing, and, as a consequence, many different methods are now available to model a given species niche and, by extrapolation, their geographic ranges.

Many papers have compared the performance of several available ENM algorithms in the last few years (e.g., Segurado and Araújo, 2004; Elith et al., 2006; Pearson et al., 2007; Tsoar et al., 2007), trying to establish which of them are more statistically robust or appropriate to a given situation, in terms of scale of environmental tolerance of species, sampling patterns (presence/absence or presence-only data) and species dispersal ability generating non-equilibrium between geographic ranges and climate (Araújo and Pearson, 2005). Recently, McPherson and Jetz (2007) demonstrated that other ecological traits of species may also be correlated with ENM performance.

It is usually difficult to achieve a consensus about the performance of different algorithms, due to different reasons. First of all, new methods and algorithms arise continuously, and some current papers did not compare all available methods. Also, these studies are usually based on particular datasets, so it may be difficult to judge the generality of conclusions. Elith et al. (2006) recently did a particularly broad and general evaluation of ENM and showed that one of the most widely used methods, GARP (Genetic Algorithm for Rule Set Production) (Stockwell and Noble, 1992) performed poorly, whereas the recently developed MAXENT (Maximum Entropy Method) (Phillips et al., 2006) ranked among the best methods (together with novel methods like boosted regression trees [BRT], and regression- based methods [GAM, GLM and MARS], which were previously suggested as high-performance methods in most studies; see also Segurado and Araújo, 2004). Pearson et al. (2007) recently compared MAXENT and GARP to predict species distribution from small numbers of occurrence records and found that MAXENT was better than GARP when sample sizes were experimentally reduced to less than 10 presence-records.

However, it is difficult to judge at which point the conclusions by Elith et al. (2006) and from other studies are entirely independent of the particular characteristics of the datasets used, since different methods and algorithms may be vulnerable to different characteristics of data used. For instance, it would not be difficult to find other datasets for which their overall conclusions do not hold. Here we modeled the geographic distribution of 39 species of New World coral snakes (genus *Micrurus*, *Micruroides* and *Leptomicrurus*) using GARP and MAXENT and compared their results. Only these two methods were compared because just presence data were available, but hopefully our comparison will be enough for illustrating our point, since these two methods cover nearly the two extremes of the 'performance axis' established by Elith et al. (2006).

2. Material and Methods

2.1. Data

We modeled the geographic distribution of 39 species of New World coral snakes (genus *Micrurus*, *Micruroides* and *Leptomicrurus*, including 6 species from North America, 7 from Central America and 26 from South America) for which at least 5 occurrence records were available (Table 1). Occurrence data for the species were compiled based on voucher specimens held in North America (American Museum of Natural History – New York, Field Museum of Natural History – Chicago, Museum of Natural History – Los Angeles, Museum of Natural Science – Louisiana, Museum of Vertebrate Zoology – Berkeley, Smithsonian Institution – Washington, and Texas Memorial Museum – Austin) and South American museums (Coleção Herpetológica da Universidade de Brasília – Brasília, Colección Herpetológica Corrientes – Corrientes, Colección Herpetológica de la Fundación Miguel Lillo – Tucumán, Colección Herpetológica de Zoología de Vertebrados de la Universidad Nacional de Río Cuarto – Córdoba, Museo de Ciencias Naturales Bernardino Rivadavia – Buenos Aires, Instituto Butantan – São Paulo, Museo de História Natural Noel Kempff Mercado – Santa Cruz de la Sierra, Museo Nacional de Historia Natural del Paraguay – Asunción, Museu Paraense Emilio Goeldi – Pará, Museo de La Plata – La Plata, Museu Nacional – Rio de Janeiro, and Museu de Zoologia da Universidade de São Paulo – São Paulo). We supplemented our data sets with records that could be georeferenced from Campbell and Lamar (2004). The number of records for the 39 species studied here varied from 5 to 217.

Six climatic variables were used for both GARP and MAXENT: annual mean temperature, temperature seasonality (coefficient of variation), mean temperature of driest quarter, annual precipitation, precipitation seasonality (coefficient of variation) and precipitation of warmest quarter, derived from the WorldClim interpolated map database (Hijmans et al., 2005), and three topographic variables (altitude, aspect and slope), derived from the U.S. Geological Survey's Hydro-1K data set (USGS, 2001). All variables were reduced to a grid resolution of 0.0417° for the analysis.

264

Braz. J. Biol., 2010, vol. 70, no. 2, p. 263-269

Table 1. List of 39 species of coral snakes ranked by number of records (in parentheses), and AUC values obtained for MAXENT and GARP.

Species (records)	MAXENT	GARP
M. frontalis (217) Duméril, Bibron and Duméril, 1854	0.9961	0.9969
M. altirostris (165) Cope, 1860	0.9957	0.9982
M. tener (150) Baird and Girard, 1853	0.9888	0.9884
M. nigricinctus (130) Girard, 1854	0.9889	0.9526
M. diastema (119) Duméril, Bibron and Duméril, 1854	0.9935	0.9878
M. lemniscatus lemniscatus (108) Linnaeus, 1758	0.9769	0.9416
M. pyrrhocryptus (108) Cope, 1862	0.9926	0.9905
M. surinamensis (107) Cuvier,1817	0.9832	0.9554
M. fulvius (103) Linnaeus, 1766	0.9947	0.9983
M. mipartitus (91) Duméril, Bibron and Duméril, 1854	0.9900	0.9578
M. spixii (84) Wagler, 1824	0.9853	0.9235
M. alleni (72) Schmidt, 1936	0.9973	0.9933
Micruroides euryxanthus (59) Kennicott, 1860	0.9972	0.9877
M. filiformis (52) Günther, 1859	0.9911	0.9676
M. corallinus (50) Merrem, 1820	0.9909	0.9853
M. annellatus (49) Peters, 1871	0.9900	0.9589
M. dumerilii (49) Jan, 1858	0.9897	0.9591
M. hemprichii (48) Jan, 1858	0.9839	0.9519
M. decoratus (42) Jan, 1858	0.9982	0.9982
M. lemniscatus carvalhoi (39) Roze, 1967	0.9943	0.9851
M. baliocoryphus (38) Cope, 1862	0.9982	0.9979
M. ibiboboca (35) Merrem, 1820	0.9902	0.9662
M. distans (32) Kennicott, 1860	0.9974	0.9891
M. multifasciatus (32) Jan, 1858	0.9837	0.9717
M. ancoralis (25) Jan, 1872	0.9939	0.9963
M. dissoleucus (25) Cope, 1860	0.9892	0.9511
M. isozonus (24) Cope, 1860	0.9953	0.9762
M. browni (22) Schmidt and Smith, 1943	0.9971	0.9946
M. laticollaris (22) Peters, 1870	0.9989	0.9971
M. tricolor (20) Hoge, 1956	0.9962	0.9954
L. scutiventris (16) Cope, 1869	0.9987	0.9911
M. brasiliensis (9) Roze, 1967	0.8955	0.9913
M. hippocrepis (9) Peters, 1862	0.8766	0.9866
M. putumayensis (9) Lancini, 1962	0.9973	0.9991
L. collaris (8) Schlegel, 1837	0.9700	0.9905
M. mertensi (8) Schmidt, 1936	0.9964	0.9932
M. bocourti (7) Jan, 1872	0.9966	0.9758
M. bernadi (6) Cope, 1887	0.9244	0.9968
M. peruvianus (5) Schmidt, 1936	0.9991	0.9995

2.2. Running GARP and MAXENT

GARP models were developed using the desktop version (see Pereira, 2002). GARP works with sets of rules of logic inference that indicate the presence or absence of a species in a region (Stockwell and Noble, 1992). Specifically, half of the data is randomly selected for the development of the rules (training data), whereas the other half is used to evaluate the accuracy of the rules (test data) (Peterson, 2001; Peterson et al., 2006). Through an iterative process of rule selection, evaluation, testing and incorporation or rejection, a method is chosen from a set of possibilities (e.g., logistic regression, bioclimatic rules, negated range rules) and applied to the training data to develop or evolve a rule (see Stockwell and Peters, 1999; Peterson, 2001 and Peterson et al., 2006, 2007 for details). For the analyses performed here, we implemented the best-subset model selection procedure (see Stockwell and Noble, 1992; Stockwell and Peters, 1999; Peterson, 2001). We generated 200 models, setting the convergence limit to 0.001, a 0% extrinsic omission error, 10% commission error, and 2000 maximum iterations. After that, we selected the 20 best models (i.e., the best subset) and summed them to make a composite GARP prediction.

Maxent is a machine learning method based on the principle of maximum entropy (see Phillips, 2006 and Phillips et al., 2006). It estimates the probability distribution of maximum entropy (i.e. that is closest to uniform) of each environmental variable across the study area. This distribution is calculated with the constraint imposed by the information available regarding the observed distribution of the species and environmental conditions across the study area (Phillips et al., 2006). Here we ran the iterative algorithm for 2000 rounds, or until the change in the objective function on a single round fell below 10^{-5}. For the regularisation parameter (β), we used 10^{-4}.

Both methods were compared using the area under the Receiver Operating Characteristics curve (ROC - AUC), an approach extensively used in species distribution modeling (see Allouche et al., 2006; Elith et al., 2006). To use this analytical approach without sample-records of true absence points, Phillips et al. (2006) generated a sample of 10,000 pseudo-absence points to join to the training sample and estimate AUC of the MAXENT procedure. We repeat the same procedure with the GARP predictions allowing the proper comparison of those methods. The AUC varies from 0 to 1, where a score of 1 indicates perfect discrimination between areas where a species is present, versus those where it is absent. Although AUC have recently been criticised (see Lobo et al., 2008; Peterson et al., 2008), it can provide at least a preliminary indication of the usefulness of the distribution models for identifying suitable areas of occurrence for particular species (Elith et al., 2006). We also recorded, for each species, the area of occurrence predicted by GARP and MAXENT, using the threshold generated to "cut" the potential distribution.

Differences between MAXENT and GARP AUC values were bootstrapped for different sample sizes (i.e., number of species studied) to evaluate how random combination of species, forming simulated 'studies' with increasing number of species will tend to support one or other method.

Braz. J. Biol., 2010, vol. 70, no. 2, p. 263-269

265

3. Results

Values of AUC for GARP ranged from 0.923 to 0.999 for the species analysed, whereas for MAXENT they ranged from 0.877 to 0.999 (Table 1), higher than those obtained by Elith et al. (2006). Taking into account the AUC values, we found that GARP does not work always worse than MAXENT, on the contrary that would be expected by considering the 'performance axis' suggested by Elith et al. (2006). Indeed, in 25.6% of the analyses (i.e., 10 species) GARP outperformed MAXENT, and in all cases differences in AUC were very small. We also found some relationship between the differences in the relative performance of the two methods and sample sizes. When sample-records are large (i.e., N > 140 records), the two models converge to similar situations, but when sample sizes are small (i.e., N < 10) or spatially clumped, GARP performed better than MAXENT in six out of eight species (see Table 1). In some intermediate situations (i.e., 15 < N < 100), MAXENT frequently achieved better results (AUC) than GARP (Figure 1a), but the difference was not as accentuated as illustrated in Elith et al. (2006).

We compared differences in the area of occurrence predicted for the species, for each method (Figure 1b).
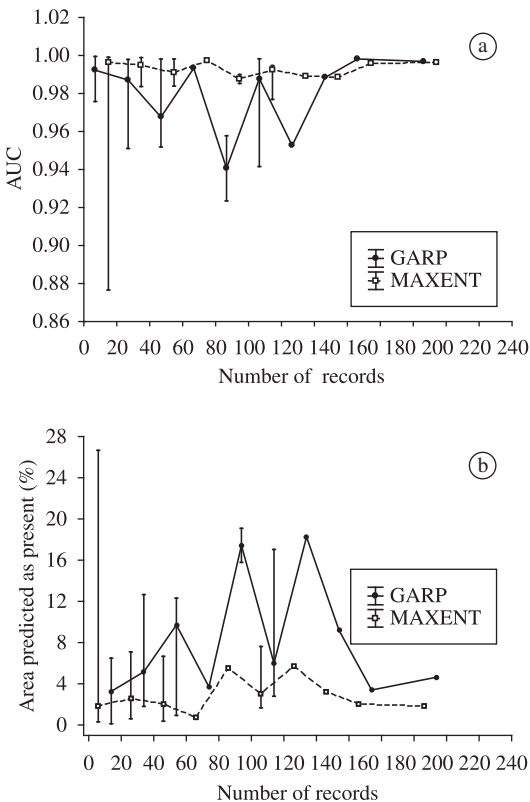
In general, MAXENT predicted a larger proportion of the area than GARP when the sample size was lower than 10 records. For large sample sizes (i.e. N > 20), GARP frequently predicts larger potential distributional areas when compared to MAXENT. When sample sizes were >150, the area predicted was similar for the two methods. In general, there is a significant correlation between the areas predicted by the two methods (r = 0.652; n = 37; P < 0.01), but only if two outlier species were removed (MAXENT generated very large extents for these two species with very small sample sizes).

Means and standard deviations for 100 bootstrapped samples with sample sizes ranging for 3 to 30 species (each one can be considered as a simulated 'study') did not show any trends towards deviations from a zero difference in AUC values between GARP and MAXENT when increasing the number of replications (i.e., species within studies) (Figure 2). Of course, it is possible to observe a decrease in the variance, revealing that all bootstrap samples are converging to the same zero difference at large sample sizes (i.e., 'studies' with more species). At the same time, this shows that, when studying only a few species, averages among 'studies' (i.e., combination of species) can vary a lot, so in some of them GARP will outperform MAXENT, whereas in some studies the other way around can be observed.

4. Discussion

What can we claim based on these results? The first and obvious issue to discuss is that the analysis of this particular dataset is not in agreement with Elith et al. (2006), regarding the performance of GARP and MAXENT. According to these authors, MAXENT belongs to the highest-performing group of methods, with
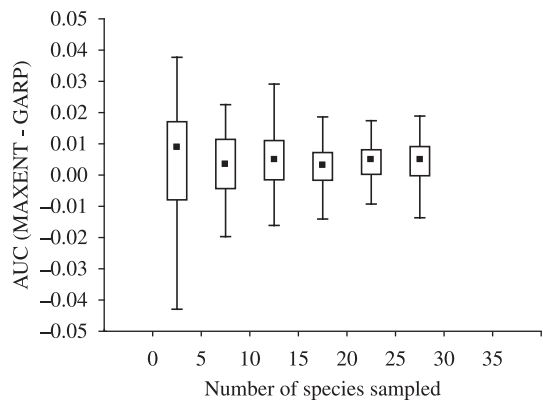

Figure 1. Means and standard deviations of AUC values for different classes of sample size (a), and percentage of area predicted as present vs sample size (b), comparing GARP and MAXENT.

Figure 2. Relationship between differences in AUC values (MAXENT – GARP) and accumulation of species studies, based on 100 bootstrapped samples (simulated 'studies') with sample sizes ranging from 3 to 30 species evaluated in each one. The differences between mean delta across samples sizes was not significant according to a one-way ANOVA using bootstrapped values as replications (F = 0.341; P = 0.915).

AUC values usually near 1.0 (i.e., a "very good" ability of a model to discriminate between sites where a species is present, versus those where it is absent), whereas GARP belongs to the group of models that performed relatively poorly, with low AUC values. Here we have found that GARP did not differ, on average, from MAXENT, and both methods have high AUC values for most species. Our results also differ from the recent ones by Pearson et al. (2007), since they found that results for GARP at very low sample sizes (i.e., n < 10) were worse than MAXENT.

These conflicting results are in accordance to the recent discussions in the literature about the relative power and predictive ability of GARP (McNyset, 2005; McNyset and Blackburn, 2006; Stockman et al., 2006; White and Kerr, 2006; Fitzpatrick et al., 2007; Tsoar et al., 2007). Also, as recently pointed out by Peterson et al. (2008), the tests by Elith et al. (2006) were actually performed using high-quality data and designed to evaluate ENMs in a situation of fine-scale modeling of species distribution details. However, in some circumstances it is necessary to project distributions into large and unknown regions in which samples are sparse or non-existent (i.e., transferability). Because of the relatively poor sampling in the Neotropics and the large extent of the domain analysed here, our results strongly support the conclusions by Peterson et al. (2007), i.e., that GARP was more successful in predicting species distributions in broad unsampled regions (as evidenced here by the percentage of area predicted as present) than MAXENT.

However, perhaps a more general discussion is to establish how one can provide guidelines on the statistical performance of the ENM algorithms generalising from particular situations (i.e., a given taxonomic group, with variable life-histories, dispersal abilities or distributed in different ways in environmental space, subjected to different historical stochastic phenomena). The question that follows is: how many studies are necessary to show that a given method is 'robust', or behaves 'better' than others? As pointed out by Pickett et al. (1994), "...the relative youth of ecology suggests that many patterns are not yet to be discerned from. Accumulating and evaluating the generality of patterns by quantitative, statistical and inductive process is still a major need in ecology". However, this empirical and inductive approach and its future success are strictly conditioned to some important issues.

Comparative studies can provide guidelines for using the methods in the current time, but the problem is that they cannot be understood as final solutions to the problem if there are conflicting results, as we show here when comparing GARP and MAXENT (see also Peterson et al., 2008). First of all, the basic idea of accumulating results from multiple studies is that they must be directly comparable, both in terms of ENMs used and their evaluation criterion. Lobo et al. (2008) and Peterson et al. (2008), for instance, discussed many problems with ROC curves and AUC statistics used by Elith et al. (2006, and many other studies, including this one) that could partially take into account variations in performance rank among methods. So, a first critical step would be establishing an accurate and unequivocal rank of relative performance of ENMs.

More importantly, the empirical inductive approach will work only if clear trends towards convergence in model performance will appear as long as future studies are accumulated. The bootstrap did not reveal any trends towards deviations from a zero difference in AUC values between GARP and MAXENT when increasing the number of species within studies, although there are trends in variances. This same approach can be easily extended for comparison among multiple studies in the future, by using species or studies as replicates, and quantitative "neo-inductive" approaches, such as Bayesian statistics, to better evaluate if there are trends in the accumulation of information in time. Notice that this argument also holds for any other statistic obtained in empirical studies, including for example the relationship between ENM performance rank and dataset (or species) characteristics, such as sample sizes and their spatial configuration. Indeed, if relative rank between ENM varies in different studies, a promising research line would be to show in which particular situations different methods may perform differently. Here, for instance, we found variations in sample sizes (i.e., GARP outperforms MAXTENT when modelling restricted distributed species - but see Pearson et al., 2007, for an opposite result), and future research may also include comparisons of the predictive success to model fine details of species distributions across landscapes (see Peterson et al., 2007).

Still following the inductive reasoning, it is then important to stress that our main message here is not to directly criticise the recent paper by Elith et al. (2006), or other previous papers comparing niche-based methods (Segurado and Araújo, 2004; Phillips et al., 2006; Pearson et al., 2007), or even emphasise which of the two methods compared is better. Instead, we reinforce that results like ours suggest that many more studies, using standardised model evaluation criterion and ENM algorithms, would be necessary to try establishing under which circumstances the rank of statistical performance of the methods vary, so one could safely advise the use of a particular modelling technique, or at least of a group of techniques. Simulation studies using known species distributions can only partially account for this problem since it is very difficult to envision all possible realistic scenarios for modeling species geographic distribution, although they can still be very useful to test model performance under controlled, "quasi"-experimental, situations.

However, it is also important to consider the possibility that this empirical inductive reasoning may fail in the end, because almost certainly we could not establish all potential scenarios generating variation in the relative performance of models, especially considering

Braz. J. Biol., 2010, vol. 70, no. 2, p. 263-269

267

non-equilibrium (historical contingencies and dispersal processes) situations (see Araújo et al., 2004; Araújo and Pearson, 2005). The current lack of agreement among the relative few studies conducted so far suggests that this possibility must be seriously considered. If, as pointed out by Peterson et al. (2008; see also Levins 1966), precision must be sacrificed by generality when using ENM algorithms, perhaps the rank of algorithms will be strongly idiosyncratic and data dependent, and a final solution can never show up. So, we advise that, for forthcoming research projects, a given model rank may be better viewed as a working hypothesis to be tested as a prediction in another particular context, not easily generalised and thus not a final recommendation for practical enterprises.

Under this more "pessimistic" scenario, for any practical reason the choice among ENMs will probably be based on a statistical criterion (after a consensus about this is achieved) or, if all methods prove to work relatively well for particular datasets, choose the ENM algorithm by other subjective reasons (such as computational facility). Alternatively, perhaps a better approach for prediction would be the recent idea championed by Araújo and New (2007), by combining a large number of models and scenarios to produce robust ensembles of species distributions and deal with uncertainty at various hierarchical levels. Although more powerful and faster algorithms and computer skills would be necessary to achieve this endeavour, perhaps this is safer and will lead us out of the pitfalls of comparing niche-based methods without knowing the particular situations we can face in the future.

References

ALLOUCHE, O., TSOAR, A. and KADMON, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, vol. 43, no. 6, p. 1223-1232.

ARAÚJO, MB. and GUISAN, A., 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, vol. 33, no. 10, p. 1677-1688.

ARAÚJO, MB. and NEW, M., 2007. Ensemble forecasting of species' distributions. *Trends in Ecology and Evolution*, vol. 22, no. 1, p. 42-47.

ARAÚJO, MB. and PEARSON, RG., 2005. Equilibrium of species distributions with climate. *Ecography*, vol. 28, no. 5, p. 693-695.

ARAÚJO, MB., CABEZAS, M., THUILLER, W., HANNAH, L. and WILLIAMS, PH., 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, vol. 10, no. 9, p. 1618-1626.

CAMPBELL, JA. and LAMAR, WW., 2004. *The venomous reptiles of the Western Hemisphere*. Ithaca: *Cornell University Press* (vol. I e II)

ELITH, J., GRAHAM, CH., ANDERSON, RP., DUDÍK, M., FERRIER, S., GUISAN, A., HIJMANS, RJ., HUETTMANN, F., LEATHWICK, JR., LEHMANN, A., LI, J., LOHMANN, LG., LOISELLE, BA., MANION, G., MORITZ, C., NAKAMURA, M., NAKAZAWA, Y., OVERTON, JM., PETERSON, AT., PHILLIPS, SJ., RICHARDSON, K., SCACHETTI-PEREIRA, R., SCHAPIRE, RE., SOBERÓN, J., WILLIAMS, S., WISZ, MS. and ZIMMERMANN, NE., 2006. Novel methods improve prediction of species'distributions from occurrence data. *Ecography*, vol. 29, no. 2, p. 129-151.

FITZPATRICK, MC., WELTZIN, JF., SANDERS, NJ. and DUNN, RR., 2007. The biogeographyof prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, vol. 16, no. 1, p. 24-33.

GUISAN, A. and ZIMMERMANN, NE., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, vol. 135, no. 2, p. 147-186.

GUISAN, A. and THUILLER, W., 2005. Predicting species distributions: offering more than simple habitat models. *Ecology Letters*, vol. 8, no. 9, p. 993-1009.

HIJMANS, RJ., CAMERON, SE., PARRA, JL., JONES, PG. and JARVIS, A., 2005. Very highresolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, vol. 25, no. 15, p. 1965-1978.

LEVINS, R., 1966. The strategy of model building in population biology. *American Scientist*, vol. 54, no. 4, p. 421-431.

LOBO, JM., JIMÉNEZ-VALVERDE, A. and REAL R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, vol. 17, no. 2, p. 145-151.

MCNYSET, KM., 2005. Use of ecological niche modelling to predict distribution of freshwater fish species in Kansas. *Ecology of Freshwater Fish*, vol. 14, no. 3, p. 243-255.

MCNYSET, KM. and BLACKBURN, JK., 2006. Does GARP really fail miserably? A response to Stockman *et al*. *Diversity and Distribution*, vol. 12, no. 6, p. 782-786.

MCPHERSON, JM. and JETZ, W., 2007. Effects of species' ecology on the accuracy of distribution models. *Ecography*, vol. 30, no. 2, p. 135-151.

PEARSON, RG., RAXWORTHY, CJ., NAKAMURA, M. and PETERSON, AT., 2007. Predictingspecies distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, vol. 34, no. 1, p. 102-117.

PEREIRA, RS., 2002. *Desktop GARP*. Kansas: University of Kansas Center for Research, Inc. Available from: <http://www.nhm.ku.edu/desktopgarp/index.html>.

PETERSON, AT., 2001. Predicting species' geographic distributions based on ecological niche modeling. *Condor*, vol. 103, no. 3, p. 599-605.

PETERSON, AT., PAPES, M. and EATON, M., 2007. Transferibility and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, vol. 30, no. 4, p. 550-560.

PETERSON, AT., PAPES, M. and SOBERÓN, J., 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, vol. 213, no. 1, p. 63-72.

PETERSON, AT., SÁNCHEZ-CORDERO, V., MARTÍNEZ-MEYER, E. and NAVARRO-SIGÜENZA, AG., 2006. Tracking population extirpations via melding ecological niche modeling with land-cover information. *Ecological Modelling*, vol. 195, no. 3, p. 229-236.

PHILLIPS, SJ., 2006. Maxent software for species habitat modeling. Version 3.2.19. New Jersey: Princeton University. Available from: <http://www.cs.princeton.edu/~schapire/maxent/>.

PHILLIPS, SJ., ANDERSON, RP. and SCHAPIRE, RE., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, vol. 190, no. 3, p. 231-259.

PICKETT, STA., KOLASA, J. and JONES, CG., 1994. *Ecological understanding*: the nature of theory and the theory of nature. San Diego: Academic Press.

SEGURADO, P. and ARAÚJO, MB., 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31 no. 10, p. 1555-1568.

STOCKMAN, AK., BEAMER, DA. and BOND, JE., 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distribution*, vol. 12, no. 1, p. 81-89.

STOCKWELL, DRB. and NOBLE, IR., 1992. Induction of sets of rules from animal distributiondata: a robust and informative method of data analysis. *Mathematics and Computer Simulation*, vol. 33, no. 5, p. 385-390.

STOCKWELL, D. and PETERS, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, vol. 13, no. 2, p. 143-158.

TSOAR, A., STEINITZ, O., ROTEM, D. and KADMON, R., 2007. A comparative evaluation of presence only methods for modelling species distribution. *Diversity and Distribution*, vol. 13, no. 4, p. 397-405.

United States Geological Survey - USGS, 2001. *HYDRO1k Elevation Derivative Database*. South Dakota. Available from: <http://edcdaac.usgs.gov/gtopo30.hydro/>.

WHITE, P. and KERR, JT., 2006. Contrasting spatial and temporal global change impacts on butterfly species richness during the 20th century. *Ecography*, vol. 29, no. 6, p. 908-918.

Braz. J. Biol., 2010, vol. 70, no. 2, p. 263-269

269