

## No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation

Huijie Qiao<sup>1</sup>, Jorge Soberón<sup>2</sup> and Andrew Townsend Peterson<sup>2\*</sup>

<sup>1</sup>Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China; and <sup>2</sup>Biodiversity Institute, University of Kansas, 1345 Jayhawk Blvd., Lawrence, KS 66045, USA

### Summary

1. The field of ecological niche modelling or species distribution modelling has seen enormous activity and attention in recent years, in the light of exciting biological inferences that can be drawn from correlational models of species' environmental requirements (i.e. ecological niches) and inferences of potential geographic distributions. Among the many methods used in the field, one or two are in practice assumed to be 'best' and are used commonly, often without explicit testing.
2. We explore herein implications of the 'no free lunch' theorem, which suggests that no single optimization approach will prove to be best under all circumstances: we developed diverse virtual species with known niche and dispersal properties to test a suite of niche modelling algorithms designed to estimate potential areas of distribution.
3. The result was that (i) indeed, no single 'best' algorithm was found and (ii) different algorithms performing very different manners depending on the particularities of the virtual species.
4. The conclusion is that niche or distribution modelling studies should begin by testing a suite of algorithms for predictive ability under the particular circumstances of the study and choose an algorithm for a particular challenge based on the results of those tests. Studies that do not take this step may use algorithms that are not optimal for that particular challenge.

**Key-words:** conservation, niche modelling, population ecology, species distribution modelling and no free-lunch theorem, statistics

### Introduction

Empirical studies of geographic and environmental distributions of species have seen massive research interest in recent years (Lobo, Jiménez-Valverde & Hortal 2010), generally under the rubric of 'species distribution modelling' (SDM) or 'ecological niche modelling' (ENM), although conceptual confusions complicate the terminology (Peterson *et al.* 2011; Peterson & Soberón 2012). In general, these studies have had three objectives: (i) estimate current (occupied) distributions of species (termed  $G_o$ ), in effect a spatial interpolation of suitability in regions not yet sampled in enough detail; (ii) estimate ranges of tolerance to environmental conditions, or more specifically Grinnellian or scenopoetic dimensions of ecological niches; and (iii) transfer niche model rules over time or space to anticipate potential distributional responses to environmental change (Peterson *et al.* 2011). Although powerful alternative approaches exist (Porter *et al.* 2002; Kearney & Porter 2004; Buckley 2008; Kearney, Wintle & Porter

2010), large numbers of studies use correlative approaches derived from comparisons of primary species occurrence data with geospatial data summarizing environmental variation across landscapes (Peterson *et al.* 2011) – these correlative studies are the focus of this study.

For these correlative approaches, numerous mathematical and statistical algorithms have been developed to estimate ecological niches in highly dimensional environmental spaces and over complex geographic landscapes (Nix 1986; Hastie & Tibshirani 1990; Carpenter, Gillison & Winter 1993; Stockwell & Peters 1999; Breiman 2001; Rangel, Diniz-Filho & Araújo 2009; Thuiller *et al.* 2009; Elith *et al.* 2011). This richness of methods has encouraged many comparative studies (Guisan & Zimmermann 2000; Brotons *et al.* 2004; Segurado & Araújo 2004; Allouche, Tsoar & Kadmon 2006; Elith *et al.* 2006; Meynard & Quinn 2007; Ortega-Huerta & Peterson 2008), often attempting to find a 'best method' (Elith *et al.* 2006). Indeed, *de facto*, a maximum entropy algorithm 'MaxEnt' (Phillips, Anderson & Schapire 2006; Elith *et al.* 2011) has become the gold standard for correlative SDM/ENM development, based on the ease of its use, the detailed documentation that has been provided (Phillips, Anderson & Schapire 2006;

\*Correspondence author. E-mail: town@ku.edu

Elith *et al.* 2011), and strong performance in comparative studies (e.g. Elith *et al.* 2006).

Large numbers of empirical papers in SDM/ENM are produced each year. To examine in more detail the degree to which the field has ascribed to the idea of a single best method, we examined the 350 papers cited in Web of Science as using the terms 'ENM' or 'SDM' during 2010–2014 (search conducted 14 April 2014). Of the 350 listed in Web of Science, 59 were not easily accessed and 46 focused on methodology or were review papers, leaving 245 empirical papers for analysis. Of those papers, 196 (80%) used a single algorithm and did not even mention having tested other algorithms prior to final analyses; of the 196 single-algorithm analyses, 112 used MaxEnt, 24 used various GLM approaches and 9 used GAM approaches. Hence, even in the past 4 years, most researchers using SDM/ENM approaches used single algorithms without exploration of others, and the best part of the single-algorithm studies used MaxEnt. Some in the field, however, have pondered the possibility that different problems may require different methods and that no single 'best' algorithm may exist (Segurado & Araújo 2004); this idea is the focus of this study.

This idea can be explored more formally in the context of a well-known principle in econometrics, the 'no free lunch' theorem ('NFL'; Ho & Pepyne 2002). The NFL states that, considered over all possible optimization problems, no 'best' algorithm will exist: an algorithm may be ideal for one class of problems, but a poor performer for other classes of problems. An algorithm that indeed is best under all circumstances could be termed a 'silver bullet', after the mythical bullet that could kill both vampires and werewolves. Although the NFL theorem is a theoretical result, it does suggest that focusing on a single method may not be a good idea when a heterogeneity of problems exists.

In SDM/ENM, numerous classes of problems are faced: consider the three very different objectives listed above, diverse data types (presence only, presence-absence, presence-background and presence-pseudoabsence) and (most fundamentally) diverse aspects of the biological nature of the challenge (species with broad or narrow niches, species limited or not by dispersal, different shapes of response curves, etc.; Peterson *et al.* 2011), all of which may make for better performance by one inferential approach or another in particular cases. A recent analysis (Saupe *et al.* 2012) illustrated the vastly different outcomes that are obtained in SDM/ENM analyses under different configurations of a Venn diagram that summarizes abiotic requirements and movement considerations in structuring species' distributions (portions of the 'BAM' diagram; Soberón & Peterson 2005; Peterson *et al.* 2011). Moreover, in SDM/ENM, the distinction of what is actually the object of interest of the model is not always made. Is it the actual occupied region that is interesting (this inference will require information about absences or pseudoabsences to estimate) or is some 'potential distribution', henceforth called **A**, the object of the work (this area can be estimated with presences only, or presences/background data)? Adding pseudoabsences in a regression is not the same as adding background data in a maximum entropy approach. Quite simply, the analytical chal-

lenges involved in SDM/ENM are diverse, such that NFL issues are likely to enter powerfully into the picture.

In this contribution, we assess the role of NFL in considerations of 'best' algorithms in SDM/ENM. We use virtual (simulated) species with known niche and dispersal properties, and three BAM configurations (Soberón & Peterson 2005; Peterson *et al.* 2011; Saupe *et al.* 2012) to generate a series of biologically contrasting SDM/ENM estimation challenges; both **G<sub>o</sub>** and **A** are known precisely. We then measure and explore performance of a broad suite of inferential algorithms. The advantage of this approach is that we are able to compare our results against known truths in the context of explicit differences in causal factors. The outcome of these exercises is that, for the problems we pose and the methods we tested, no single algorithm emerged as best. Obviously, our results do not eliminate the possibility that a single 'best' algorithm might actually exist – however, as we presented several classes of common problems (narrow and broad niches; ranges constrained by dispersal, by niche or by both; estimating **G<sub>o</sub>** and estimating **A**) modelled explicitly using simplified scenarios, it is very unlikely that results with different data sets would be qualitatively different. Our results also illustrate further how different types of biological problems present harder or easier challenges for different types of algorithms in this field, although real-world complexities remain to be explored in this context.

#### DISTRIBUTIONAL ECOLOGY, VIRTUAL SPECIES AND IMPROVING METHODOLOGIES

Distributions of species are affected by diverse factors (Udvardy 1969; Brown & Lomolino 1998; Gaston 2003), including history, environments, interactions with other species, dispersal capacity and dispersal barriers (Peterson *et al.* 2011). The efficacy of algorithms designed to estimate the ecological niche also responds to numerous factors, which has been the subject of large numbers of analyses (e.g. Stockwell & Peterson 2002; Graham *et al.* 2007; Guisan *et al.* 2007a,b; Elith & Graham 2009). Following our previous explorations (Barve *et al.* 2011; Saupe *et al.* 2012; Owens *et al.* 2013) and those of others (Hirzel, Helfer & Metral 2001; Elith & Graham 2009; Godsoe 2010), here, we use the powerful tool of simulated, virtual species as a means of testing, developing and improving methodologies, with the advantage that the truth is known, rather than inferred, such that model performance can be measured precisely.

To focus on NFL issues, we simplified modelling challenges in the design of our virtual species to effects of two major factors only: environmental favourability and dispersal barriers. We excluded biotic interactions not because they are unimportant (see Leathwick 1998; Heikkinen *et al.* 2007; Kissling *et al.* 2012); rather, biotic interactions create dynamic complications that are complex and poorly understood (Gilman *et al.* 2010) and may be manifested at fine spatial resolutions without changing the coarse-resolution picture of species' distributions (the Eltonian Noise Hypothesis; Peterson *et al.* 2011). In the end, all of the objectives of this study – assessing whether any

single algorithm is likely to be ‘best’ algorithms under diverse circumstances – can be achieved in a world in which only abiotic niche dimensions and accessibility constrain species’ geographic distributions.

Virtual species have been used to study diverse problems in SDM/ENM, often methodological challenges of a technical nature, such as effects of prevalence, numbers of background points, relative weights, thresholding methods and so on (Meynard & Kaplan 2013). Here, we concentrate on fundamental niches and barriers to dispersal, **A** and **M** in the BAM diagram (Fig. 1; Soberón & Peterson 2005). We postulate fundamental niches ( $N_F$ ) that are convex in shape, in effect creating multi-dimensional ellipsoids (Norris, Jackson & Betancourt 2006): convex shapes are assumed as the appropriate approximation to fundamental niches based on theoretical expectations (Maguire 1973) and in the light of existing empirical evidence (e.g. Birch 1953; Green 1971). The intersection of  $N_F$  with the set of environmental combinations actually manifested across the accessible region **M** for a species determines the portions of the physiologically defined  $N_F$  that is actually available to the species (Jackson & Overpeck 2000). This reduced ‘niche’ is termed the *existing* fundamental niche, or  $N_F^*$  (Peterson *et al.* 2011), so we chose niches explicitly to vary in the degree to which they are represented completely, or only partially, in the existing environmental space. This point is crucial as different climatic environments may allow the same fundamental niche to be experienced in very different ways by species (Jackson & Overpeck 2000; Soberón & Peterson 2011). [Several additional factors may reduce the set of environments experienced by a species still further; e.g. negative interactions with other species can reduce existing fundamental niches to what Hutchinson (1957) classically called the realized niche, or  $N_R$ .]

For the above reasons, and assuming that species interactions are generally negative, the following set inequality

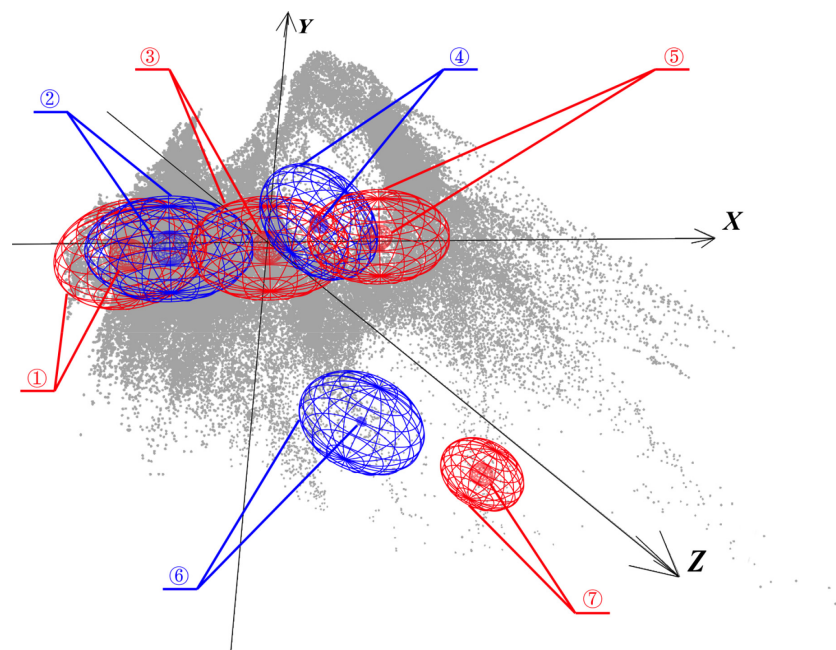
(notice subset symbol) probably holds (Soberón & Peterson 2011):

$$N_F \supseteq N_F^*(t) = N_F \cap \eta(\mathbf{M}, t) \supseteq N_R(t) = N_F^*(t) \cap \eta(\mathbf{B}, t) \quad (1)$$

where  $\eta(\mathbf{X}, t)$  indicates environments manifested in region **X** at time  $t$ . The inequality basically expresses the idea that environments accessible to a species in region **X** at time  $t$  reduce fundamental niches to existing niches and that environments associated with biotic interactions in region **X** at time  $t$  reduce existing niches further to realized niches. Particularly relevant is that these reductions of fundamental niches to existing or realized niches create rather complex shapes in environmental space. Niche modelling algorithms must thus be capable of dealing with complicated environmental spaces in a biologically intelligent manner. Projection of  $N_F$  into geographic space identifies areas putatively suitable for a species, termed **A** in the BAM diagram; **A** is the potential area of distribution, which may be contiguous or composed of disjoint regions not all of which are necessarily accessible to and inhabited by the species (Peterson 2003).

As a final step in setting up our tests, we made assumptions about dispersal ability of our virtual species over relevant time periods, to select one or more of the regions that comprise **A** as a hypothesis of an area actually occupied by the species ( $G_o$ ) but smaller than the whole suitable region. This step introduces the idea of distributional disequilibrium into our virtual species, which appears to be a very common feature of distributional ecology (Peterson 2003). This step follows Saupe *et al.* (2012), who showed that the degree to which dispersal considerations dominate the structure of species’ geographic distributions determines in large part the success or failure of correlational modelling efforts.

To summarize, we assembled scenarios that varied in niche position (central or peripheral within the environmental cloud), niche breadth (broad or narrow) and spatial position



**Fig. 1.** Illustration of example fundamental niches of virtual species used as the basis for this study, showing the distribution of environmental combinations with respect to the first three principal components of the environmental space across the study area (in grey). Niches are shown as ellipsoids, 2 per niche centre for seven centres. Red and blue colours are used to facilitate visualization. Details of virtual species are provided in the Supporting information.



of dispersal barriers. Our proposition is that this rich and diverse suite of virtual species offers an excellent test bed on which to examine the idea of whether silver bullets exist in correlative SDM/ENM. The results of these tests have important implications for the future of a field that is in an amazing growth phase, in which numerous ecologists, conservation biologists, biogeographers and systematists are adopting and exploring a new suite of methods.

## Methods

We created 14 virtual species for this project (Fig. 1 and Supporting information illustrate examples). These species are, in effect, ellipsoids in a three-dimensional environmental space, with or without barriers to dispersal in a corresponding geographic space. The species comprised seven pairs, one broad environmentally and the other with axes one-quarter the breadth (i.e. 64-fold smaller in volume), but with the same centroid and same directions of semi-axes. The study was based on a real-world landscape, within a polygon delimited by 65–145°E and 10–50°N: most of China, Mongolia, India, Pakistan, Kazakhstan, Kyrgyzstan, Tajikistan, Indochina, the two Koreas and Japan, plus parts of Russia, Uzbekistan, Afghanistan and the Philippines. Environmental dimensions included the 19 bioclimatic variables in the WorldClim archive (Hijmans *et al.* 2005); to avoid working in an overly dimensional space, we confined our analyses to the first three principal components of this space, which explained 80.7% of overall variance.

Once niches were created, we explored their projections into geographic space and created different scenarios of abiotic factors in relation to dispersal, following our methods in a previous study (Saupe *et al.* 2012). Specifically, we created a scenario in which all suitable conditions were accessible ('Hutchinson's dream', or HD), one in which distributional limitation was driven by dispersal/access exclusively and not by abiotic conditions ('Wallace's dream', or WD) and one in which both abiotic and dispersal considerations act to limit distributions in different areas ('classic BAM', or CB). For some of the species, not all BAM scenarios could be developed, because their extents in geographic space (**G**) or environmental space (**E**) were too small (i.e. fewer than 5 contiguous pixels, or so highly fragmented such that no WD areas of sufficient size could be created) – in all, then, we created 29 virtual species' distributions with different characteristics (see Fig. 1). Details of each virtual species are provided in the Supporting information.

For each species, we sampled random occurrence points comprising up to 10% of pixels within the species' **G<sub>o</sub>** (i.e. from within the area that is both habitable and accessible), but not exceeding 1000 points. To be able to include algorithms that require absence data explicitly in our tests, eliminate extra space points were sampled from the set of accessible pixels not known to hold presences via routines in BIOMOD (i.e. within its core programming). This step offers a non-occurrence data set against which to compare occurrence data in niche model calibration.

For each virtual species, we estimated ecological niches using 8 algorithms, which can be separated approximately into three functional groups (further development of this classification can be found in Peterson *et al.* 2011). First are algorithms that depend on the presence data only: BIOCLIM (Nix 1986) and Marble (Qiao 2010; Qiao *et al.* 2012). Second are algorithms that employ pseudoabsence or background data: GARP (Stockwell & Peters 1999) and MaxEnt (Phillips, Anderson & Schapire 2006). Finally (third) are algorithms that use both presence and true absence data: generalized additive models (GAM; Hastie & Tibshirani 1990), generalized boosted models (GBM; Ridgeway 1999), generalized linear models (GLM; McCullagh & Nelder 1989)

and random forests (RF; Breiman 2001). Details of implementation of each algorithm are provided in the Supporting information.

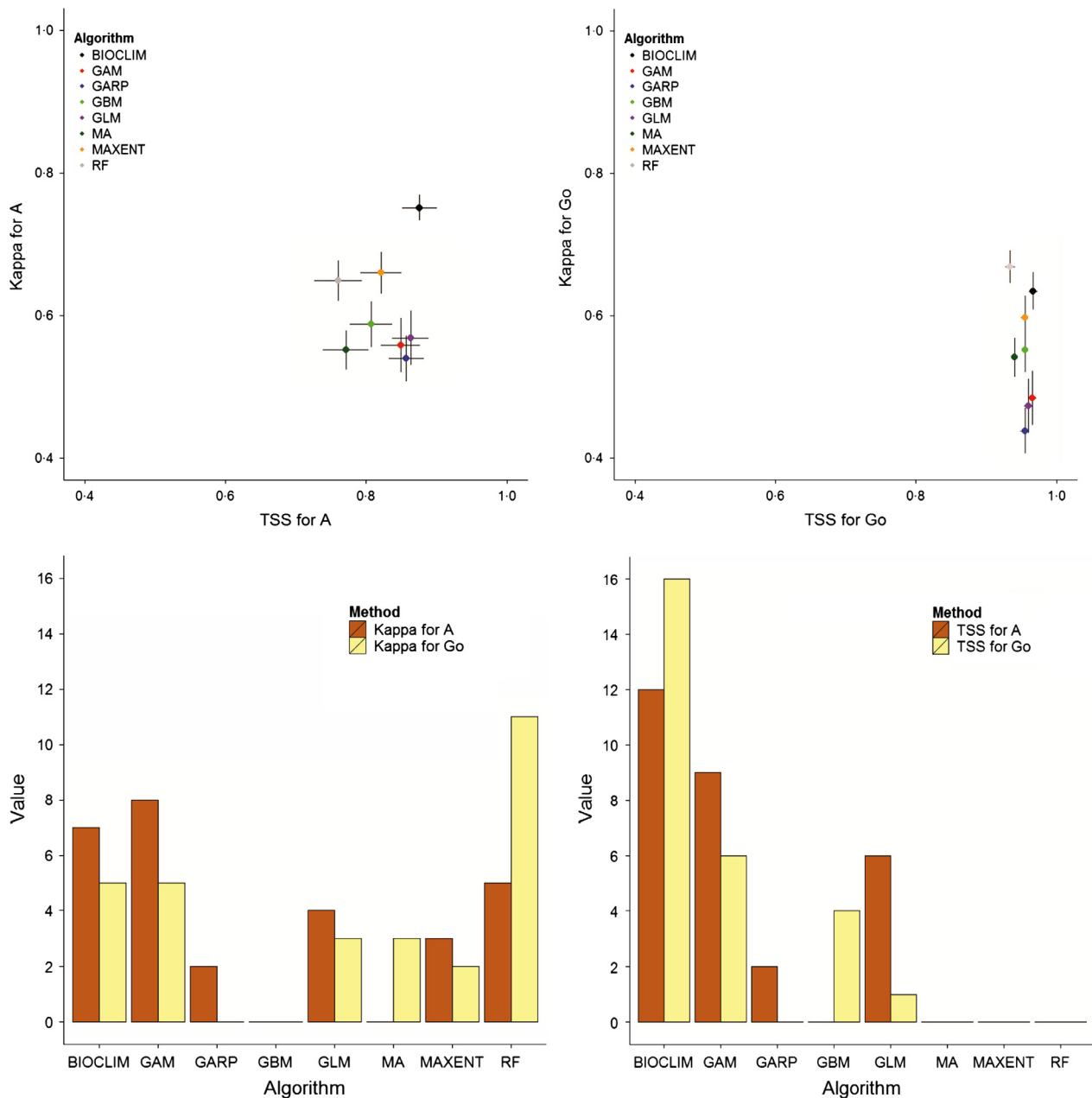
For each species × BAM scenario, we used the presence and non-presence points to calibrate models with each algorithm. We repeated random sampling of occurrence data 10 times to produce 10 replicate models for each species × BAM scenario, producing a total of 29 species × 8 algorithms × 10 random replicates = 2320 model results. We considered the error-free nature of the occurrence data (i.e. all presence points were sampled from **G<sub>o</sub>**, so all were from conditions within **N<sub>F</sub>**), and also the fact that different algorithms estimate different quantities, which will almost certainly not be comparable or even necessarily in the same units (Peterson *et al.* 2011). Hence, we thresholded each model result via minimum training presence thresholding (Liu *et al.* 2005; Pearson *et al.* 2007) to produce binary model predictions. This step creates a coarse, but comparable, scale for the different outputs.

To summarize results, we took advantage of the known truth nature of virtual species, following Saupe *et al.* (2012). Specifically, for each species × BAM scenario, we knew the true geographic extent of **A** (i.e. the potential distribution, equivalent to the spatial footprint of **N<sub>F</sub>**) and **G<sub>o</sub>** (i.e. the intersection of the potential distribution with the dispersal capabilities of the species). To assess correspondence between model outputs and known true configurations, we calculated Cohen's kappa, true skill statistics (TSS), sensitivity and specificity (Fielding & Bell 1997) with respect to both **A** and **G<sub>o</sub>**. TSS has significant advantages over kappa, in that it is not prevalence dependent (Allouche, Tsoar & Kadmon 2006), but both indices mix the contrasting behaviour of sensitivity and specificity. In an NFL context, we tallied the number of cases in which the algorithm 'won', that is, had the top performance overall among all competitors. We explored plots of sensitivity vs. specificity; dividing overall performance into these two components offers significant advantages in interpretation.

## Results

The general results were similar to those of other comparative modelling studies (e.g. Elith *et al.* 2006). That is, model predictions differed with respect to sensitivity and specificity and spread out broadly in performance space, indicating significant model-to-model variation in quality. In this sense, we are confident that the series of manipulations that we assembled (i.e. differences in niche position, niche breadth and BAM scenario) reconstruct at least some of the complexities of real-world challenges in estimating ecological niches of species. Exploring the frequency with which each algorithm performed 'best', in a more traditional view (Fig. 2a; cf. Elith *et al.* 2006), in terms of average performance, BIOCLIM and MaxEnt were clear winners (i.e. closest to the upper right corner in Fig. 2a). From a NFL viewpoint, however (Fig. 2b), it is immediately apparent that no best algorithm exists: each algorithm won under some circumstances, and even the metric used to define 'best' (i.e., Kappa vs. TSS) changes the result. The algorithm and challenge that yielded the highest number of 'wins' was BIOCLIM in anticipating **G<sub>o</sub>**, but several other algorithms proved best under some circumstances.

Dividing TSS values for species by BAM scenario (Fig. 3; parallel explorations of Kappa yielded similar results), it is immediately apparent that the WD scenario poses a difficult challenge for every method, paralleling previous results from our group (Saupe *et al.* 2012). Very simply, WD represents a



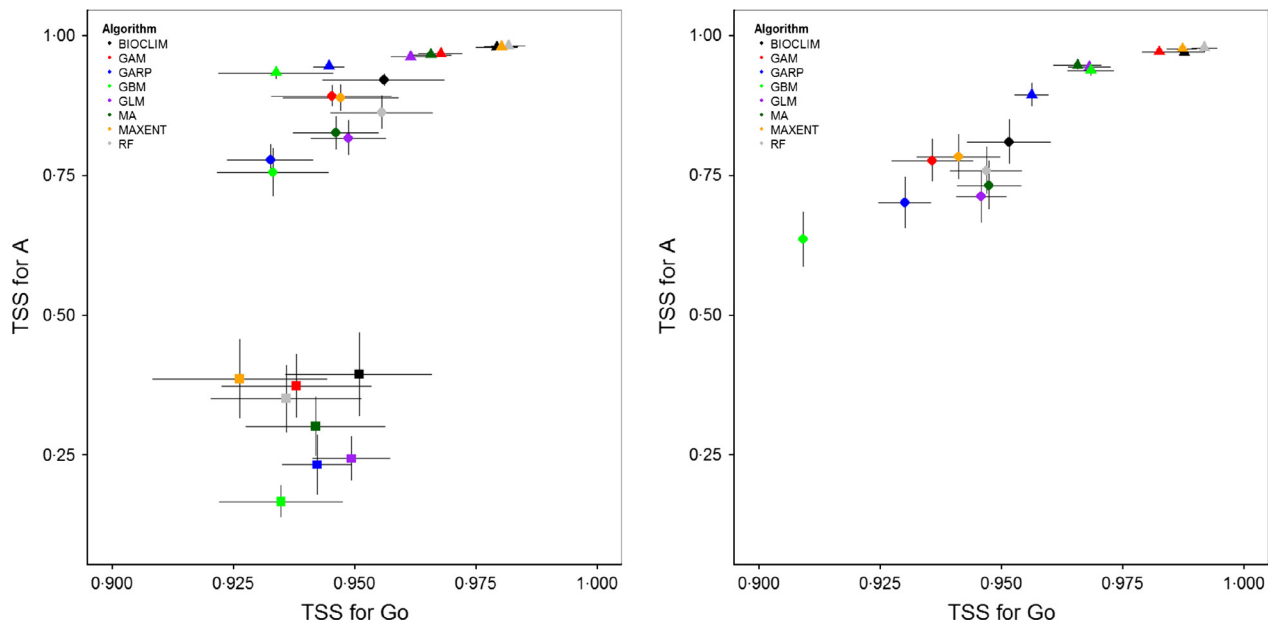
**Fig. 2.** Top row: average performance of 8 SDM/ENM algorithms in predicting areas **A** and **G<sub>o</sub>**, as measured via Cohen's kappa and TSS (average  $\pm$  95% confidence error). Bottom row: number of times that an algorithm produces the highest value of the kappa (orange bars) and the TSS (yellow bars) as an average of replicates per different challenge. Note the differences in the scaling of axes between the left and right panels in the top part of the figure. ENM, ecological niche modelling; SDM, species distribution modelling; TSS, true skill statistics.

situation in which distributional limits are set by dispersal limitations, rather than by the abiotic environment, such that appreciating the constraining role that the environment may play becomes very difficult. Niche breadth also played an important role in outcomes of niche modelling exercises (Fig. 3): broad fundamental niches were more difficult to model than narrow niches, with marked differences in performance for all algorithms.

The results discussed so far, however, mix two concepts within single statistics: proportion of suitable area (**A** or **G<sub>o</sub>**) correctly identified as such (sensitivity) and proportion of unsuitable area correctly identified as such (specificity). These

challenges are distinctly different from one another, so we assessed them separately. For the CB scenario (Fig. 4), two implications were immediately clear: (i) for most algorithms, capacity to predict absences correctly (specificity) is starkly different from capacity to predict presences (sensitivity); and (ii) niche breadth affects capacity to predict presences, particularly in RF, MaxEnt, GBM, MA and GLM. Results of predictive tests were better when the target was **G<sub>o</sub>** than when the target was **A**, in contrast to the results of [Saupe \*et al.\* \(2012\)](#).

The corresponding results in the HD and WD scenarios showed some differences (Fig. 5). HD is the simplest and most favourable scenario for niche modelling, and **A** and **G<sub>o</sub>** are



**Fig. 3.** True skill statistics contrasting model performance in predictions of **A** vs. **G<sub>0</sub>**. In the left-hand plot, circles, triangles and squares symbolize BAM scenarios (CB, HD, WD); in the right-hand plot, circles and triangles depict niche breadth (broad, narrow). Error bars represent 95% confidence intervals on means. HD, Hutchinson's dream; WD, Wallace's dream.

equivalent in this case. For WD, information was richest for species with broad niches, as narrow-niche WD species frequently had too few data for inclusion in analyses – for WD species, and particularly when the target was **A**, sensitivity performance was low in almost all cases. In essence, then, our explorations confirm what previous work had already shown: that SDM/ENM is a complex challenge, with numerous factors that influence model success or failure.

## Discussion

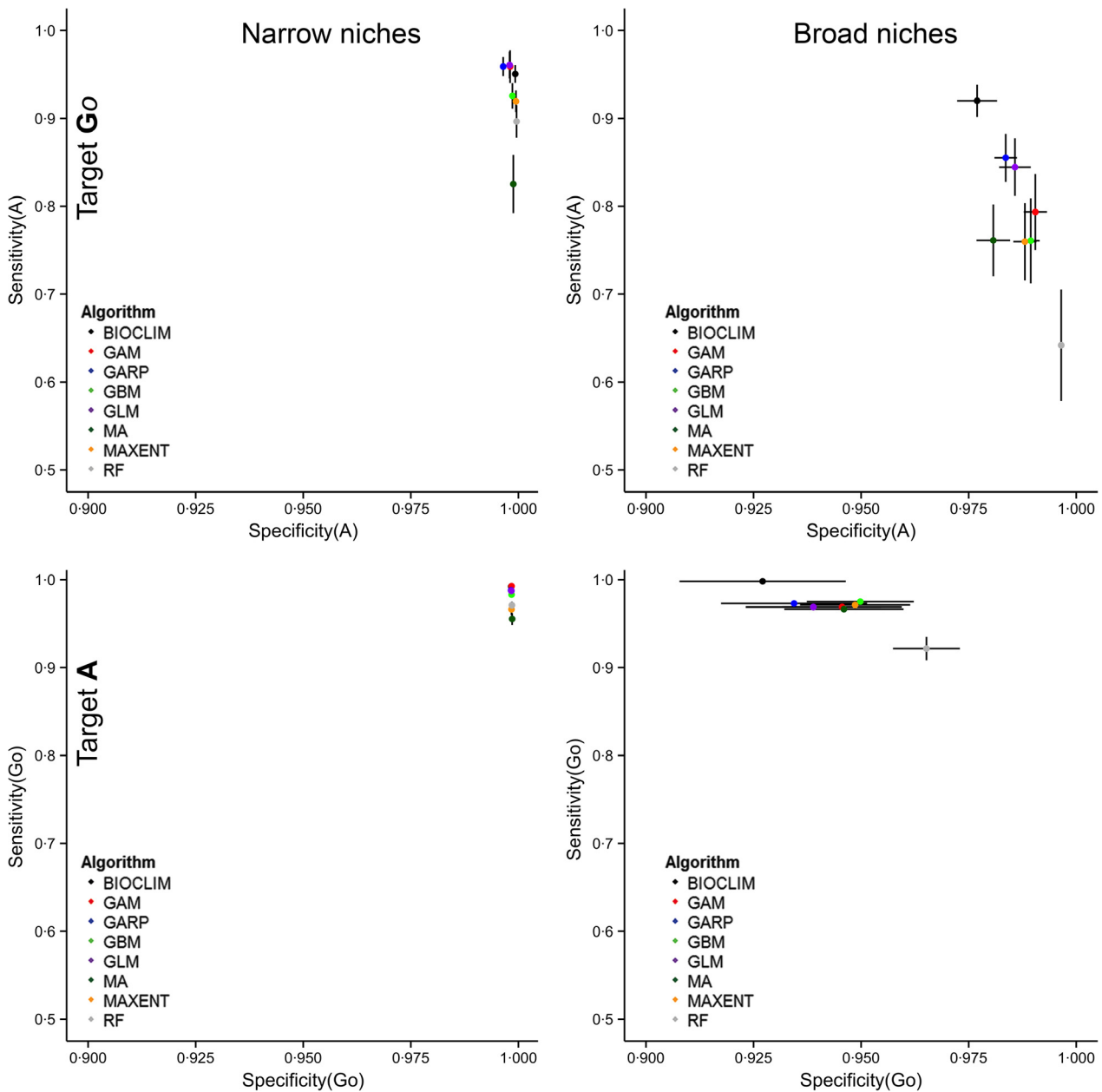
### NFL AND SILVER BULLETS IN SDM/ENM

This study in effect provides a different lens through which to view and contemplate comparative performance evaluations of niche and distribution model performance. The traditional approach has focused on average performance of algorithms (see fig. 3 in Elith *et al.* 2006), seeking consistent differences among algorithms. Here, we applied a distinct perspective to the idea of algorithm performance. We worked to create a suite of species and predictive challenges that was sufficiently complex as to create differences in performance among algorithms and then set out to reassess and revisit the reality of the idea of a best algorithm in niche and distribution modelling. NFL ideas suggest that, in theory, good performance under all possible circumstances is unattainable, which suggests that, when dealing with a variety of types of problems, performance of a particular algorithm may be good in some cases, but not as good under other circumstances. Our view of 'winner' and 'loser' algorithms bears this out: although BIOCLIM was most frequently a winner for TSS-measured challenges, followed by GAM and GLM, it was not always the 'best' algorithm. We hasten to note that our main point in this study is not that

BIOCLIM *per se* is best, as our simulated species do not capture the full complexity of real-world modelling challenges, but rather that each of 8 algorithms won under *some* set of circumstances. This view is the NFL reality: no algorithm will be optimal under all circumstances, because no single best approach is likely to exist.

A recent trend in the SDM/ENM field has been that of exploring ensemble approaches (Araújo & New 2007; Thuiller *et al.* 2009; Crimmins, Dobrowski & Mynsberge 2013), in which outputs of multiple algorithms are combined; indeed, some might argue that ensemble methods offer a means of dealing with an NFL world. In the simplest sense of ensembles, however, good and bad algorithm outputs get averaged together, and NFL would suggest that these approaches will produce less than optimal results. In more complex ensembles, however, algorithm outputs are weighted as to their contribution to the ensemble, based on some measure of performance. The NFL approach that we propose, in which a suite of algorithms is explored initially, algorithm performance evaluated, and high-performing algorithms used for final model development, is thus a form of ensemble approach that employs an extreme weighting scheme.

However, our main point would also apply to an ensemble method. Different classes of inferential challenges exist, because (i) the biology of the problems is different, either because species' distributions are constrained by dispersal or by lack of suitable environments, or because their niches are very narrow, or rather wide, or because their niches are on the periphery of the environmental space or in its centre (Saupe *et al.* 2012; Owens *et al.* 2013). (ii) The data will likely be different, as presence only and presence-absence situations pose substantially different mathematical problems. (iii) The statistics used to assess the degree of success of the algorithms



**Fig. 4.** Sensitivity vs. specificity for the CB scenario, when the target object is **A** (top row) and **G<sub>o</sub>** (lower row). The left column shows narrow niches, and the right column shows broad niches. Bars represent 95% confidence intervals on the means.

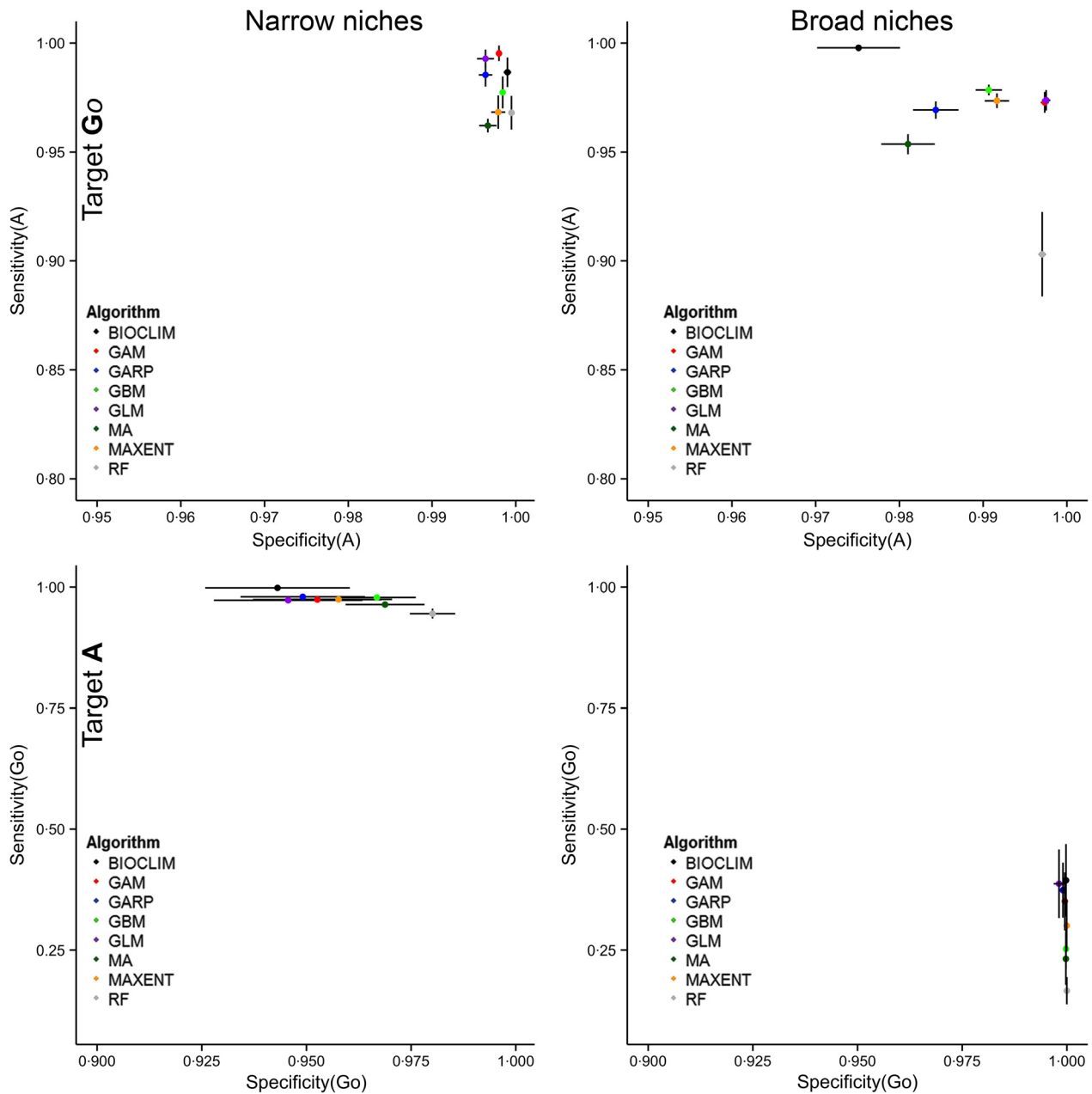
should be disaggregated. Finally, (iv) because the object being modelled may differ from one application to another, these different objects (e.g. occupied distributional area, potential distributional area) pose very different problems. Put most simply, the challenge is not one of number crunching, but rather a matter of understanding one's modelling challenge deeply.

#### AGGREGATED STATISTICS AND ALGORITHM PERFORMANCE

The NFL result presented above is simple, but perhaps does not capture all of the complexity and broad implications of our analyses. In particular, given our use of virtual species that

varied in a few, controlled dimensions, we were able to delve more deeply into causes of better or worse performance among the many tests that we conducted. As a consequence, we explore implications of these results beyond simple NFL ideas.

Using a single statistic, such as TSS or Kappa, can mask major differences in performance between algorithms (compare, e.g., Figs 3 and 4). By viewing model results in terms of sensitivity and specificity separately, consistent effects of BAM scenario or niche breadth on model outcomes can be appreciated. Generally speaking, we found marked differences in capacity to estimate areas related to specificity vs. sensitivity. Sensitivity of model predictions varied dramatically depending on whether the target was **A** or **G<sub>o</sub>**, as well as on the BAM scenario, niche breadth and algorithm. Mean sensitivity for some



**Fig. 5.** Top row, sensitivity vs. specificity for the Hutchinson's dream (HD) scenario. In this scenario,  $A = G_o$ . The left column is for narrow niches, and the right column is for broad niches. In the bottom row, we present the Wallace's dream (WD) scenario, for broad niches only: the left panel is for  $G_o$  as a target, and the right panel is for  $A$  as target. Bars represent 95% confidence intervals on the means.

algorithms was as low as 50%, with some individual values as low as 10% (Fig. 5). For specificity, in contrast, algorithms performed well, regardless of scenario and niche size, with most mean values >97%.

At least for some of the more popular algorithms (e.g. MaxEnt, GARP and GAM), errors with respect to sensitivity were apparent when the target of prediction was  $A$ , whereas errors regarding specificity were more pronounced when the target of modelling was  $G_o$  (in the HD scenario, all suitable areas are accessible, such that  $A = G_o$ , so this contrast does not appear). This asymmetry has not been noticed previously and further highlights the need to make targets of modelling exercises

explicit (Peterson & Soberón 2012). More generally, specificity and sensitivity are not equivalent in this process: indeed, correctly identifying the presence proves to be much more important in many situations than correctly identifying the absence in the light of differences in degree of confidence in data documenting presences and absences of species (Anderson, Lew & Peterson 2003; Peterson *et al.* 2011). It is thus worrisome that, in our simplified study, it is precisely sensitivity that is most prone to poor model performance.

Algorithms that performed well in general in terms of specificity were RF, GLM and GAM; all three of these algorithms can be regarded as regression-based methods. BIOCLIM per-



formed well in terms of sensitivity, probably because it is generous in terms of overpredicting across broad regions. In this sense, BIOCLIM may be capable of estimating correctly both **A** and **G<sub>o</sub>** simply by including them within relatively broad (artifactual) predictions. Whether a simplified algorithm (like BIOCLIM) or a more complex one is the best choice for a given situation needs to be decided on the basis of the question being asked, the biological nature of the problem, the type of data available and the specifics of the algorithm in question (Peterson 2006), as well as – the point of this study – on the results of tests developed at the outset of the study (see below). It is not just that many SDM/ENM users do not bother to understand the mechanics of the software packages they utilize (Joppa *et al.* 2013), but that the suitability of a software package may well change with the particularities of the estimation problem.

#### MANY TYPES OF NICHES

A further lesson of this study is the rather simple and obvious idea that specifying which ‘niche’ is being modelled is critical in such modelling exercises. Our study is based on creating virtual species defined by just two types of factors: the fundamental niche (i.e. tolerance limits with respect to environmental conditions) and dispersal constraints. **N<sub>F</sub>** has frequently been assumed to have a simple shape, probably convex and possibly ellipsoid (Birch 1953; Maguire 1973; Godsoe 2010), but estimating it robustly via correlational approaches is a significant challenge (Peterson *et al.* 2011). Correlative algorithms may estimate an object as circumscribed as the realized niche **N<sub>R</sub>** and the area **G<sub>o</sub>**, or the (likely broader) existing fundamental niche **N<sub>F</sub><sup>\*</sup>** if interactions can be disregarded (the Eltonian Noise Hypothesis of Peterson *et al.* 2011; Soberón & Peterson 2011), or something approximating **N<sub>F</sub>** if extrapolation and uncertainty are managed appropriately (Owens *et al.* 2013). Unfortunately, as **N<sub>R</sub>** and **N<sub>F</sub><sup>\*</sup>** depend on interactions among complicated factors (e.g. incomplete representation of combinations of conditions across real-world landscapes, biogeographic barriers, dispersal limitations, competition, etc.), their shape and size frequently are not related in any simple way to the original **N<sub>F</sub>**. In this sense, using estimates of **N<sub>i</sub>** or **N<sub>F</sub><sup>\*</sup>** to assess phylogenetic dimensions of niche evolution (e.g. Kozak & Wiens 2006; Eaton, Soberón & Peterson 2008), test ecological niche conservatism (e.g. Peterson, Soberón & Sánchez-Cordero 1999; Peterson 2011), or anticipate model behaviour when models are transferred (e.g., Thomas *et al.* 2004), is often a risky proposition (see Owens *et al.* 2013).

SDM/ENM algorithms estimate some region **N̂** in environmental space that in some sense falls in between **N<sub>R</sub>** and **N<sub>F</sub><sup>\*</sup>** or **N<sub>F</sub>** (Jiménez-Valverde, Lobo & Hortal 2008). The projection of **N<sub>F</sub><sup>\*</sup>** into geographic space is the area **A**, whereas the projection of **N<sub>R</sub>** includes the area **G<sub>o</sub>**, although it may include additional areas beyond the occupied area (Peterson *et al.* 2011). Our results suggest that, in many cases, and particularly when niches are narrow, **N̂** will estimate **G<sub>o</sub>** better than **A**. This result contrasts with that of a previous study from our group (Saupe *et al.* 2012), so this area and this set of conclusions merit

further research attention and assessment, perhaps best with more complex virtual species as a test bed.

#### VIRTUAL SPECIES

This study is based on 29 simulated or ‘virtual’ species, a design feature that offers significant advantages; however, their use also has disadvantages that must be weighed in any conclusions to be derived. On the positive side, virtual species approaches are most useful in that the truth is known. That is, one knows precisely the dimensions of **N<sub>F</sub>**, **N<sub>F</sub><sup>\*</sup>** and **N<sub>R</sub>** in environmental dimensions and the extents of **A** and **G<sub>o</sub>** in geographic dimensions, such that one can measure model outputs against known truths directly and without uncertainty. What is more, virtual species analyses can incorporate real-world complexity of environmental landscapes (e.g. this study is cast across Asia, using realistic environmental data) – this real-world dimension to simulated species gives at least some idea of the true complexity of species’ distributions.

However, on the negative side, some complexities are certainly lost with virtual species. Chief among these details are the shapes of fundamental niches – although considerable evidence from physiological ecological studies suggests that fundamental niches should be convex in multivariate space, variable-to-variable interactions in the real world may be more complex than those managed here. (Note that we did not explore at all how covariances among environmental variables affect outcomes of SDM/ENM explorations.) Similarly, the accessible areas that we have considered as **M** hypotheses for each species are clearly artificial – that is, we have simply drawn convenient lines and polygons on maps, which vastly underappreciate the complexity of real-world dispersal limitations and biogeographic barriers (Barve *et al.* 2011). In both cases, then, it must be borne in mind that virtual species do not replicate fully the complexity of real-world species, such that our results must be taken with a grain of salt, of size unknown. This simplicity may explain some of the success of relatively simple algorithms (e.g. BIOCLIM) in our tests.

#### Conclusions and Future Directions

The simplest results of this study are that we found no evidence of a single ‘best’ algorithm for development of correlative ecological niche models among the suite of algorithms that we tested. This outcome was expected on conceptual grounds (Ho & Pepyne 2002) and was corroborated in the results of this empirical test. The point is general: no single algorithm ought to be taken as a ‘no brainer’ or *de facto* choice in complex problems (indeed, NFL lessons should be contemplated in other fields of biology that require complex optimizations, such as phylogeny estimation). Instead, we suggest that researchers wishing to use correlative approaches to modelling ecological niches appreciate the need to ponder and assess algorithm choice carefully. Using a single algorithm simply because it has a convenient user interface, or because everyone else uses it, is not likely to be an ideal path (Joppa *et al.* 2013).

Our results suggest a significant modification to typical correlative SDM/ENM methodologies: a diverse suite of algorithms should be explored for each correlative SDM/ENM exercise that one carries out. One should develop an initial suite of model predictions from a variety of algorithms and under a diversity of parameter settings and initial conditions (Warren & Seifert 2011). These multiple solutions to the same problem can be challenged with a test – for example, a data-subsetting challenge over known distributional areas in the present time. Then, the algorithm that is most powerful in a given situation (species, landscape, sampling, etc.) can be used in the remainder of the study.

Such a methodology will be perhaps a bit scary for many in the field: most SDM/ENM users are accustomed to using one or a few algorithms, most commonly MaxEnt, which is available free of charge and has a convenient user interface. NFL thinking suggests, however, that studies begin with experimentation with and assessment of performance of several algorithms, which would take many (including at least one of the authors of this study) out of their comfort zones in terms of algorithm choice and use. This shift, however, is facilitated by the growing number of multiple algorithm platforms that has become available, including BIOMOD (<http://www.will.chez-alice.fr/Software.html>), OpenModeller (<http://openmodeller.sourceforge.net/>), Dismo (<http://www.inside-r.org/packages/cran/dismo>), mMWeb (<http://mmweb.animal.net.cn>) and BIOENSEMBLES (Rangel, Diniz-Filho & Araújo 2009).

Finally, several points are in considerable need of further exploration and documentation, so that this burgeoning field can continue to advance. In particular, many of these virtual species lessons need to be examined based on tests using real-world species: in particular, effects of BAM scenario (dispersal constrained vs. barrier-less species); effect of target object (A vs. G<sub>o</sub>, particularly as the results in this work appear to contradict results in Saupe *et al.* (2012); effects of niche breadth to understand where the ‘problem zone’ of overly broad niches begins; and others.

A more broad and critical point is estimation of fundamental niches, as opposed to existing or realized niches. Our entire conceptual analysis depends on how a physiologically determined set of environmental conditions within which populations of a species can persist (=the fundamental niche) is successively distorted and reduced by the existing conditions on Earth, by conditions within areas accessible to a species, and within areas available to the species given biological interactions and habitat limitations (these latter two complications were not explored in this study). This series of complicating phenomena means that conventional correlative approaches, particularly those that emphasize fitting of complex response surfaces (Elith *et al.* 2006), will frequently arrive at incorrect answers, particularly as regards modelling applications that require model transfers (Owens *et al.* 2013).

## Acknowledgments

We thank the members of the University of Kansas Ecological Niche Modeling Group for ideas and input throughout the development of this project. HQ was

supported by the National Natural Science Foundation of China (Project: *A New Method to Predict the Species Distributions*, grant number 31100390), JS was supported by NSF grant DEB-1208472, and JS and ATP were supported by a grant from Microsoft Research.

## Data accessibility

Species descriptions: uploaded as online supporting information.

Data deposited in the Dryad repository: <http://datadryad.org/resource/doi:10.5061/dryad.8g0v3>.

## References

- Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.
- 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.
- Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.
- 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.
- Birch, L.C. (1953) Experimental background to the study of the distribution and abundance of insects: I. The influence of temperature, moisture and food on the innate capacity for increase of three grain beetles. *Ecology*, **34**, 698–711.
- Breiman, L. (2001) *Random Forests*; <http://oz.berkeley.edu/users/breiman/randomforest2001.pdf>. Unpublished Technical Report.
- Brotons, L., Thuiller, W., Araújo, M.B. & Hirzel, A.H. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.
- Brown, J.H. & Lomolino, M.V. (1998) *Biogeography*. Sinauer Associates, Sunderland, MA.
- Buckley, L. (2008) Linking traits to energetics and population dynamics to predict lizard ranges in changing environments. *The American Naturalist*, **171**, E1–E19.
- Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN: a flexible modeling procedure for mapping potential distributions of animals and plants. *Biodiversity and Conservation*, **2**, 667–680.
- Crimmins, S.M., Dobrowski, S.Z. & Mynsberge, A.R. (2013) Evaluating ensemble forecasts of plant species distributions under climate change. *Ecological Modelling*, **266**, 126–130.
- Eaton, M.D., Soberón, J. & Peterson, A.T. (2008) Phylogenetic perspective on ecological niche evolution in American blackbirds (Family Icteridae). *Biological Journal of the Linnean Society*, **94**, 869–878.
- Elith, J., Graham, C., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J. & Graham, C.H. (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., 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.
- 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. (2003) *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, Oxford.
- Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010) A framework for community interactions under climate change. *Trends in Ecology and Evolution*, **25**, 325–331.
- Godsoe, W. (2010) Regional variation exaggerates ecological divergence in niche models. *Systematic Biology*, **59**, 298–306.
- Graham, C.H., Elith, J., Hijmans, R., Guisan, A., Peterson, A.T., Loiselle, B.A. & NCEAS Species Distribution Modeling Group (2007) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239–247.
- Green, R.H. (1971) A multivariate statistical approach to the Hutchinsonian niche: bivalve molluscs of central Canada. *Ecology*, **52**, 544–556.
- Guisan, A., Elith, J., Graham, C.H., Phillips, S.J., Peterson, A.T. & Zimmermann, N.E. (2007a) What matters for predicting spatial distributions of trees:

- techniques, data, or species' characteristics? *Ecological Monographs*, **77**, 615–630.
- Guisan, A., Graham, C.H., Elith, J., Huetteman, F. & NCEAS Species Distribution Modelling Group (2007b) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, **13**, 332–340.
- Guisan, A. & Zimmermann, N. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hastie, T. & Tibshirani, R.J. (1990) *Generalized Additive Models*. Chapman & Hall, Boca Raton, FL.
- Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Korber, J.-H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.
- 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 Climatol*, **25**, 1965–1978.
- Hirzel, A.H., Helfer, V. & Metral, F. (2001) Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, **145**, 111–121.
- Ho, Y.C. & Pepyne, D.L. (2002) Simple explanation of the no-free-lunch theorem and its implications. *Journal of Optimization Theory and Applications*, **115**, 549–570.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.
- Jackson, S.T. & Overpeck, J.T. (2000) Responses of plant populations and communities to environmental changes of the Late Quaternary. *Paleobiology*, **26**, 194–220.
- 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.
- Joppa, L.N., McInerney, G., Harper, R., Salido, L., Takeda, K., O'Hara, K., Gavaghan, D. & Emmott, S. (2013) Troubling trends in scientific software use. *Science*, **340**, 814–815.
- Kearney, M. & Porter, W.P. (2004) Mapping the fundamental niche: physiology, climate, and the distribution of a nocturnal lizard. *Ecology*, **85**, 3119–3131.
- Kearney, M.R., Wintle, B.A. & Porter, W.P. (2010) Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. *Conservation Letters*, **3**, 203–213.
- Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerney, G.J. *et al.* (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, **39**, 2163–2178.
- Kozak, K.H. & Wiens, J.J. (2006) Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, **60**, 2604–2621.
- Leathwick, J.R. (1998) Are New Zealand's *Nothofagus* species in equilibrium with their environment? *Journal of Vegetation Science*, **9**, 719–732.
- Liu, C., Berry, P., Dawson, T. & Pearson, R. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.
- 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.
- MacCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, FL.
- Maguire, B. (1973) Niche response structure and the analytical potentials of its relationship to the habitat. *The American Naturalist*, **107**, 213–246.
- Meynard, C.N. & Kaplan, D.M. (2013) Using virtual species to study species distributions and model performance. *Journal of Biogeography*, **40**, 1–8.
- Meynard, C.N. & Quinn, J.F. (2007) Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, **34**, 1455–1469.
- Nix, H.A. (1986) A biogeographic analysis of Australian elapid snakes. *Atlas of Elapid Snakes of Australia* (ed R. Longmore), pp. 4–15. Australian Government Publishing Service, Canberra, ACT.
- Norris, J.R., Jackson, S.T. & Betancourt, J. (2006) Classification tree and minimum-volume ellipsoid analyses of the distribution of ponderosa pine in the western USA. *Journal of Biogeography*, **33**, 342–360.
- 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.
- Owens, H.L., Campbell, L.P., Dornak, L., Saupe, E.E., Barve, N., Soberón, J. *et al.* (2013) Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, **263**, 10–18.
- Pearson, R.G., Raxworthy, C., Nakamura, M. & Peterson, A.T. (2007) Predicting species' distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.
- Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology*, **78**, 419–433.
- Peterson, A.T. (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, **3**, 59–72.
- Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817–827.
- Peterson, A.T. & Soberón, J. (2012) Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, **10**, 1–6.
- 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, NJ.
- Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Porter, W., Sabo, J., Tracy, C., Reichman, O. & Ramankutty, N. (2002) Physiology on a landscape scale: plant–animal interactions. *Integrative and Comparative Biology*, **42**, 431–453.
- Qiao, H. (2010) *Construction and Comparison of Species' Potential Distribution Models*. PhD thesis. Graduate University of Chinese Academy of Sciences, Beijing.
- Qiao, H., Lin, C., Ji, L. & Jiang, Z. (2012) mWeb – an online platform for employing multiple ecological niche modeling algorithms. *PLoS ONE*, **7**, e43327.
- Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Araújo, M.B. (2009) *BIOENSEMBLES 1.0. Software for Computer Intensive Ensemble Forecasting of Species Distributions Under Climate Change*. Universidade de Goiás, Goiás.
- Ridgeway, G. (1999) The state of boosting. *Computing Science and Statistics*, **31**, 172–181.
- Saupe, E., Barve, V., Myers, C., Soberón, J., Barve, N., Hensz, C., Peterson, A.T., Owens, H.L. & Lira-Noriega, A. (2012) Variation in niche and distribution model performance: the need for a *a priori* assessment of key causal factors. *Ecological Modelling*, **237–238**, 11–22.
- Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.
- Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**, 1–10.
- Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**, 1348–1353.
- 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.
- Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C. *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W., Lafourcade, B., R., E. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.
- Udvardy, M.D.F. (1969) *Dynamic Zoogeography*. Van Nostrand Reinhold Company, New York, NY.
- Warren, D.L. & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335–342.

Received 3 December 2014; accepted 13 April 2015

Handling Editor: Darren Kriticos

## Supporting Information

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

**Data S1.** Details of methodology employed in fitting ecological niche models.

**Data S2.** The details about the characteristics of the 29 virtual species.