INTERPRETATION OF MODELS OF FUNDAMENTAL ECOLOGICAL NICHES AND SPECIES' DISTRIBUTIONAL AREAS

Jorge Soberón
*Comisión Nacional de Biodiversidad, México, and
Instituto de Ecología, UNAM, México*

AND

A. Townsend Peterson
*Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence,
Kansas 66045 USA*

Abstract.—Estimation of the dimensions of fundamental ecological niches of species to predict their geographic distributions is increasingly being attempted in systematics, ecology, conservation, public health, etc. This technique is often (of necessity) based on data comprising records of presences only. In recent years, modeling approaches have been devised to estimate these interrelated expressions of a species' ecology, distributional biology, and evolutionary history—nevertheless, a formal basis in ecological and evolutionary theory has largely been lacking. In this paper, we outline such a formal basis to clarify the use of techniques applied to the challenge of estimating 'ecological niches;' we analyze example situations that can be modeled using these techniques, and clarify interpretation of results.

Key words.—ecological niche, fundamental niche, realizad niche, geographic distribution

The fact that, at certain scales, climatic and physical factors affect profoundly the distributions of species has been known for a very long time. In the last two decades, mathematical techniques designed to estimate the geographic extent of the "fundamental ecological niche" (FN), or subsets of it, defined mostly in coarse-scale climatic dimensions [the "bioclimatic envelope" or climatic niche (Pearson & Dawson 2003)], have seen increasing use. Although of interest in itself regarding the evolutionary ecology of species, estimating the FN is often taken as an intermediate step towards estimating the geographic distribution of the species.

The FN has been estimated in two ways: (1) direct measurement or physical modeling of responses of individuals to temperature, humidity, and other physical parameters, and inferring from them fitness values of different combinations of physical variables. Then, using GIS technology, geographic regions of positive fitness can be displayed (Porter *et al.* 2000; Porter *et al.* 2002). This line of work has been referred to as the 'mechanistic approach' to niche modeling (Guisan & Zimmermann 2000). (2) Niches may be reconstructed by relating data on species'

occurrences with data sets summarizing climatic, topographic, edaphic, and other 'ecological' dimensions (in the form of GIS layers); combinations of environmental variables most closely associated with observed presences of species can then be identified and projected onto landscapes to identify appropriate regions, as above. The inferential steps in this manifestation of niche modeling have been achieved using diverse algorithms, including range rules (Nix 1986), DOMAIN (Carpenter *et al.* 1993), FloraMap (Jones & Gladkov 1999), multiple regression and other generalized linear and additive models (Guisan & Zimmermann 2000), neural networks (Pearson *et al.* 2002), and genetic algorithms (Stockwell 1999; Stockwell & Noble 1992; Stockwell & Peters 1999), and several others. All, in essence, extrapolate from associations between point occurrences and environmental data sets to identify areas of predicted presence on the map. These areas are (in some sense determined by the algorithm) *similar* ecologically to those where the species is known to occur, and this procedure can be termed as the 'correlative approach' to ecological niche modeling. Whether the result is interpreted as the species' distribution, the spatial

extent of its fundamental niche, or some other phraseology, it is important to remember that (strictly speaking) these extrapolation algorithms only find regions that "resemble," in terms of the layers provided, those where occurrence points are located. The rest of the process is interpretation.

Often (but not always), tools used to perform these niche estimations require information about absence of species. At certain spatio-temporal scales, and for certain taxa, this information may be available, perhaps in the form of well-surveyed sites at which the species was not detected. Alternatively, and at cost of restrictive assumptions, "absence data" may be generated in the form of pseudoabsences—areas without definite information regarding the species' occurrence, but assumed to be lacking the species for the purpose of providing statistical contrasts upon which analyses may be based.

The mechanistic approach, being based on direct measures of physiological variables, ignores biotic interactions, and indeed has little hope of taking them into account. The correlative approach, in contrast, is based on observations that already *include* effects of biotic interactions on distributions of species—here, of course, the challenge is *removing* the effects of those same interactions. The two approaches thus estimate quite-different phenomena, and should be interpreted carefully before being used interchangeably in applications. In what follows, we clarify the logic behind the correlative approach, and review its key limitations—our purpose is to provide a more formal discussion of what is being modeled in 'ecological niche modeling.' Pearson and Dawson (2003) discussed characteristics and limitations of these two methods as regards estimating niches and predicting distributions under scenarios of climate change.

THE DISTRIBUTION OF A SPECIES

The distributional area of a species is a complex expression of its ecology and evolutionary history (Brown 1995), determined by diverse factors operating with different intensities at different scales (Gaston, 2003; Pearson & Dawson 2003). Four classes of factors determine areas in which a species is found:

1. Abiotic conditions, including aspects of climate, physical environment, edaphic conditions, etc., that impose physiological limits on species' ability to persist in an area.
2. Biotic factors–the set of interactions with other species that modify the species' ability to maintain populations. These interactions can be either positive (e.g., mutualists such as seed dispersors, pollinators, etc.) or negative (e.g., competitors, predators, diseases). By limiting or enhancing population processes, interactions can obviously affect distributions.
3. The regions that are accessible to dispersal by the species from some original area. This factor is extremely useful in distinguishing a species' actual distribution from its *potential* distribution, based on landscape configuration and the species' dispersal abilities.
4. The evolutionary capacity of populations of the species to adapt to new conditions. This factor, usually reserved from analyses or assumed negligible, is nevertheless an additional and important consideration in outlining the distributional possibilities of species. In theory (Holt 1996a, b; Holt & Gaines 1992; Holt & Gomulkiewicz 1996; Kawecki 1995), and in the limited experiments carried out to date (Etterson & Shaw 2001), effects of evolution in niche parameters over short periods of time appear minor.

These factors interact dynamically and with different strengths at different scales to produce the complex and fluid entity we call the geographic distribution of a species. Pulliam (2000) has developed a simulation for distributions of species based on quantitative descriptions of abiotic niches, dispersal, and competition as a sole type of biotic interaction. Pulliam (2000) used a simple graph to explore different relations between niche concepts and species' distributions. In Figure 1, we elaborate on Pulliam's (2002) graph, and present (as a heuristic tool), a static, non-mechanistic Venn diagram of the first 3 factors (evolution will be discussed later).

We assume that a species will be present at a given point where three conditions are met: (1) Abiotic conditions must be favorable (i.e., where density-independent fitness is positive), these conditions occur in region (**A**). (2) An appropriate suite of species is present (e.g., hosts, food plants, pollinators, seed-dispersors, mycorrhizae) and absent (e.g., strong competitors, diseases,

specialized predators), depicted in region (**B**). Finally, (3) the species will be present only in a region (**M**) that is reachable by the species from established distributional areas in ecological time (i.e., dispersal limitations are not a consideration).
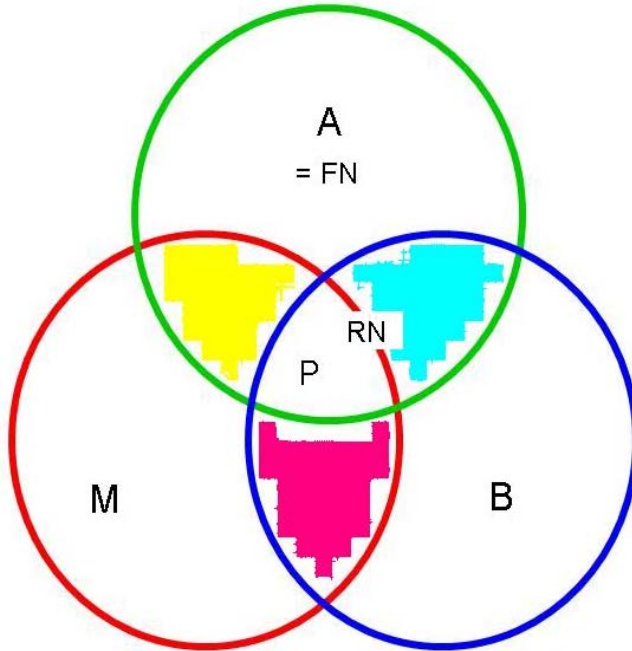

Figure 1. The green area **A** represents the geographic region with the appropriate set of abiotic factors for the species, and may be regarded as the geographic expression of the Fundamental Niche (FN). Area **B** (blue) is the region where the right combination of interacting species occurs, which may or may not overlap extensively with **A**. **A** ∩ **B** represents the geographic extent of the Realized Niche (RN) of the species. **M** (red) is composed of those parts of the world "accessible" to the species in some ecological sense, without barriers to movement and colonization. **A** ∩ **B** ∩ **M** = **P** is the region that has the right set of biotic and abiotic factors and that is accessible to the species, and is equivalent to the geographic distribution of the species.

Region **A** represents the geographic expression of the fundamental niche (FN) of the species (Hutchinson, 1957). If dispersing individuals or small populations of the species in question were to arrive or be introduced anywhere in areas fitting these conditions, the limiting factors (if any) to population growth would be biotic in nature. Region **B** represents the region in which obligate or very strong interspecific interactions necessary for a species' presence could be expected to be fulfilled. The intersection of **A** and **B** (hereafter denoted **A** ∩ **B**) is that part of the world in which (1) abiotic conditions are suitable for positive population growth of the species; and (2) required

mutualists are present and the suite of competitors, predators, and diseases present will not prevent positive fitness. **A** ∩ **B** is then the geographic expression of the realized niche (RN) of the species (Hutchinson, 1957). Notice that we include interactions other than competition in our conceptualization of the realized niche. Some recent treatments of niche concepts include both ecological requirements of species *and* their impacts on those niche factors (Chase and Leibold, 2003). By restricting the definition of **A** to factors that in the short term are independent of the species' population dynamics (climatic and topographic factors), we avoid complications arising from more dynamic concepts of niches.

Finally, region **M** represents that part of the world that is accessible to the species since its origin, or via anthropogenic or other means of introduction. Our heuristic scheme then states that stable populations of a species will be found only in the region of intersection of **M**, **B**, and **A** (**M** ∩ **B** ∩ **A**), which we term **P**. Sink populations may exist elsewhere (Pulliam, 2000; Gaston, 2003) in **M**, but outside of **A** ∩ **B**. Notice that this scheme implies a nested relationship between the geographic expressions of FN, RN, and the distributional area, since $\mathbf{A} \supseteq (\mathbf{A} \cap \mathbf{B}) \supseteq \mathbf{P}$.

Figure 1 thus provides a framework for discussing the main points of this paper. When an extrapolating algorithm (e.g., GARP, BIOCLIM, Domain, or FloraMap) is applied to a set of occurrence points and abiotic layers, how do we interpret the resulting geographic predictions? Under what circumstances can the result be interpreted as an estimation of FN (region **A**) versus the actual distribution (**P**)? Note that if occurrence data are obtained from source populations, they would come from **P**, which is a subset of **A**. However, in most examples in the literature, layers provided as input to extrapolating algorithms are abiotic in nature. In other words, extrapolating algorithms are generally not provided with information about biotic variables that determine **B**, since they are generally unavailable or complex to interpret. Therefore, the algorithm searches the map for regions resembling (in abiotic terms) occurrence points determined by a mixture of both biotic and abiotic factors. Note also that data points can come from sink populations (by mistake, carelessness, or simply because of lack of

information), within **M** but outside **A** ∩ **B** (the geographic expression of the RN). Given such complications, it is clear that such approaches require careful interpretation of results—we discuss some simple example cases below.

CASE I: A = B = M

The simplest (and most optimistic) case is where the three regions (**A**, **B**, **M**) overlap entirely (Figure 2, top). In this case, FN ≈ RN, and, since accessibility presents no restrictions, ecological potentials overlap broadly with the geographic distribution of the species **P**. An extrapolating algorithm would thus yield results equally interpretable as any of the three entities. Whether the estimate is good or bad depends (of course) on the extrapolating algorithm, and on how well occurrence points sample the abiotic variation in **A**, but these considerations always apply to questions of niche modeling. In other words, model quality depends on how well occurrence points cover the volume of the abiotic niche (which may or may be not related to the distribution of the points in geographic space), and results are interpretable as any of the spatial manifestations of a niche (fundamental niche, realized niche, actual distribution).

CASE II. A = M ≠ B

We now consider a complication (Figure 2, middle). Again, we assume that **M** overlaps almost entirely with **A** because the spatial scale is not too large, or because the biology of the species allows for efficient dispersal and the landscape does not include important barriers, making most of **A** accessible. However, in this situation, **B** overlaps with only a small part of the other two regions, because strong interactors (positive or negative) restrict use of the whole of the FN. In this case, the geographic extent of the FN (**A**) is larger than the distributional area. How good is the estimate of the FN depends on whether the occurrence points sample a large and representative section of **A**— that is, if interactions limit the species' distribution *in the dimensions in which* **A** *is being estimated*, then the estimate of **A** may also be suspect. Clearly, careful interpretation of the output of the extrapolating algorithm is needed, since underestimation of the FN may be expected (see Figure 3) in areas in which positive interactors are missing or negative interactors occur extensively.
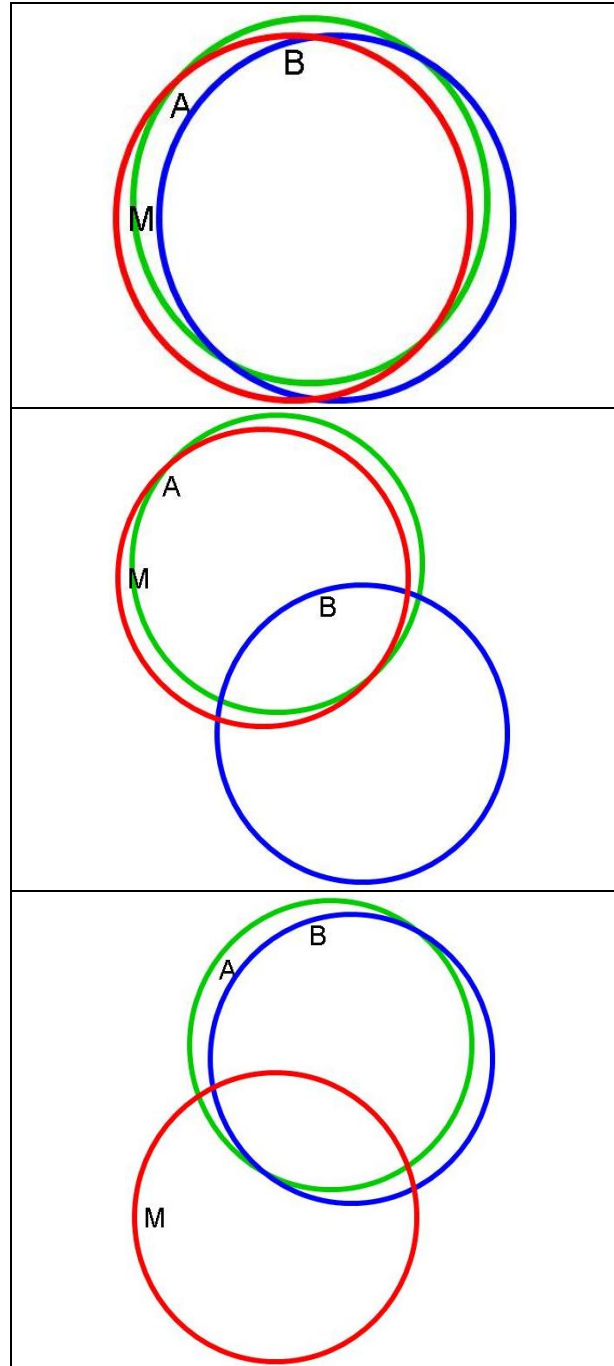

Figure 2. Three extreme cases of overlap between **A**, **B** and **M**. (Top) The three regions overlap almost completely and geographic range (FN and RN coincide). (Middle) All of FN is accessible, but some parts of this region either lack obligate mutualists or hold species that interact in a negative way with the species of interest in such a way as to reduce its fitness below zero. (Bottom) **A** overlaps almost entirely with **B**, but large areas remain inaccessible.

CASE III. A = B ≢ M

In this case (Figure 2, bottom), interactions are weak, diffuse, or non-specific, so FN overlaps broadly with the region in which interactions are appropriate (**A** ≈ **B**). However, given the large size of the region in relation to the dispersal capacities of the species, or owing to a complex geography of the area that generates barriers or promotes isolation of populations, the accessible region **M** overlaps with the other two in only a relatively small region. Hence, in this situation the FN will generally be larger than the actual distribution of the species (**P**).

ESTIMATION OF P

We have seen that, unless the three sets of factors affecting geographic distributions of species overlap entirely, most "bioclimatic" algorithms estimate the FN. Where interest in actual geographic distributions is strong, however, ancillary data and assumptions may be used to make possible its estimation. For example, Peterson *et al.* (2001) assumed that species under study were sampled well at the level of biogeographic provinces. Under this view, if an occurrence has been registered in a biogeographic province, the whole province is assumed accessible to the species; the geographic extent of the FN estimated by the algorithm on the basis of abiotic layers is then reduced to its intersection with a map of biogeographic provinces in which the species has been recorded (e.g., Figure 4). Other geographic features such as basins, drainages, or ecoregions, may be used instead of biogeographic provinces.

The degree to which the three niche dimensions overlap in ways that fit the examples explored above is an empirical question that has received little attention. Biologically, **M** would be nearly equal to **A** if the region under consideration is small relative to the dispersal capacities of the species, and/or the region under study has a relatively homogeneous biogeographic and geological history. **B** would be expected to overlap broadly with the other two regions in those situations or at scales of resolution at which interactions among the species are diffuse or weak. Hence, we might expect **A** ≈ **M** at finest spatial scales, and **A** ≈ **B** at the broadest spatial scales, but again these conjectures await thorough empirical testing.

THE ROLE OF ABSENCE INFORMATION

Several algorithms used to estimate FNs require information on absences of species to be assessed across the area of analysis. If presence data can be scarce and difficult to come by, absence data are even more rare. Only in well-explored regions of the world do data sets exist with sufficiently high resolution and density that *absence* of records of a species can be interpreted as a true absence, rather than attributed to lack of data or to insufficient sampling.

Given this scarcity of absence data, researchers often resort to generating "pseudoabsences" by a number of methods (Fielding & Bell 1997; Stockwell & Peterson 2002), usually involving resampling information from some version of the set of points from which the species has *not* been recorded. Pseudoabsences are sometimes generated randomly over the entire region under study (including **P**, all of **A**, and indeed the entire region in question) as a simple approach. Some applications use this approach to generate populations of points that differ, *a priori*, in the probability of the species being present; these populations can then be fed into the predictive algorithms, even though some of the pseudoabsences turn out to represent actual presences (Stockwell & Peters 1999). Whereas this approach permits development of models under some circumstances, in general, it should be considered less acceptable. Rather, the universe for sampling pseudoabsences should be delineated carefully based on *a priori* hypotheses about the region and species to be modeled.

The preceding discussion argues that 'ecological niche modeling' algorithms generally produce estimates of the FN (or at least something more general than just the distribution). Estimating the actual distribution requires explicit assumptions about degrees of overlap of **A, B,** and **M**. If modeling FN is the goal of the study, "absence" information would be drawn most appropriately from an *a priori* hypothesis about the area outside of **A**, rather than outside the more complex **P**. In other words, investigators should take care not to sample pseudoabsences for this type of analysis from areas that are *within the niche* but not within the dispersal possibilities of the species, as this approach confuses presence of the niche conditions (but not the species) with real absence of the species and its niche.
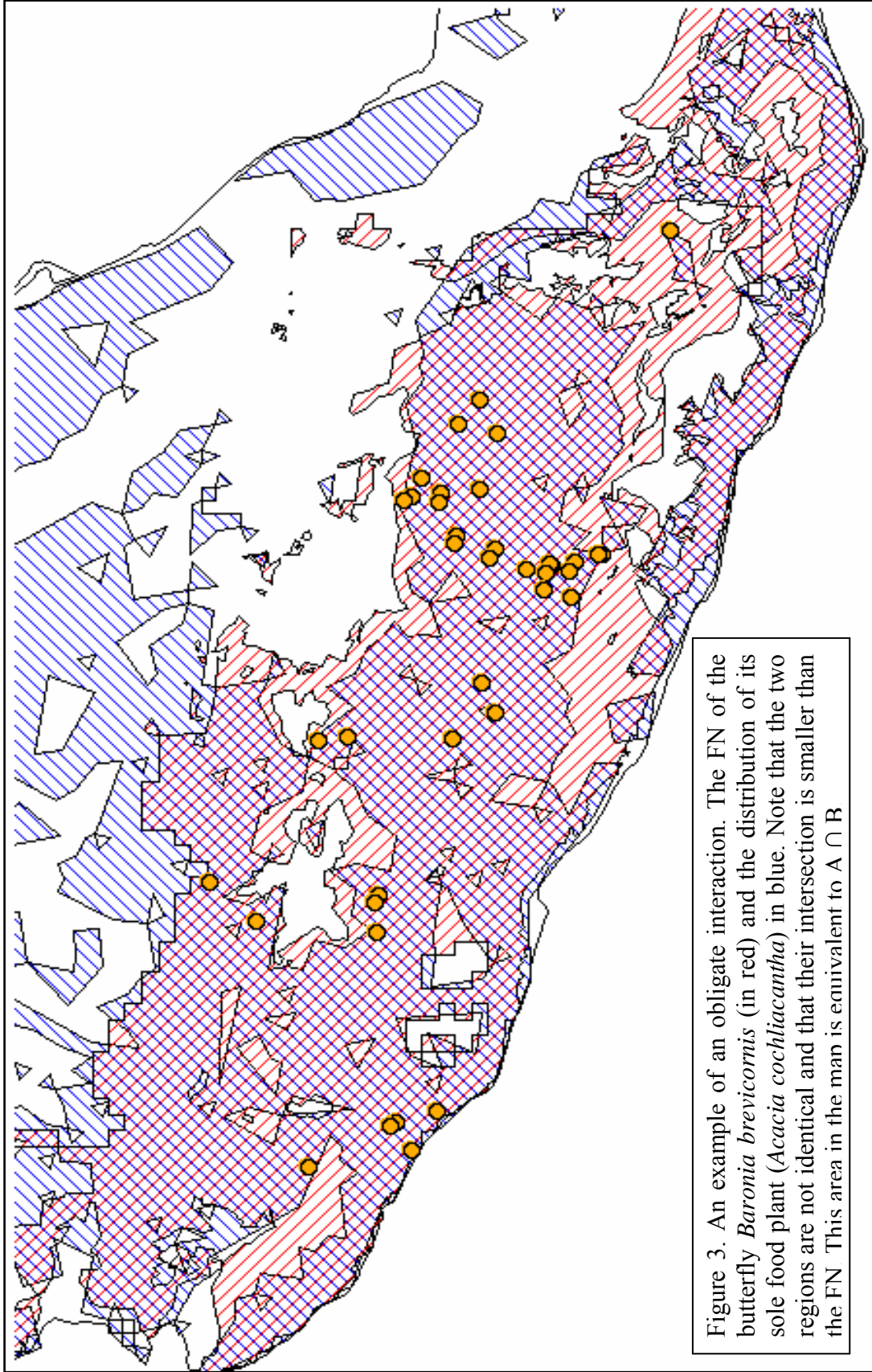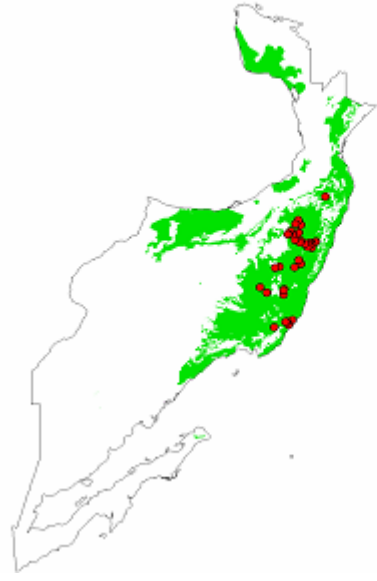
Figure 3. An example of an obligate interaction. The FN of the butterfly *Baronia brevicornis* (in red) and the distribution of its sole food plant (*Acacia cochliacantha*) in blue. Note that the two regions are not identical and that their intersection is smaller than the FN. This area in the map is equivalent to A ∩ B

Biogeographic Provinces (= M)
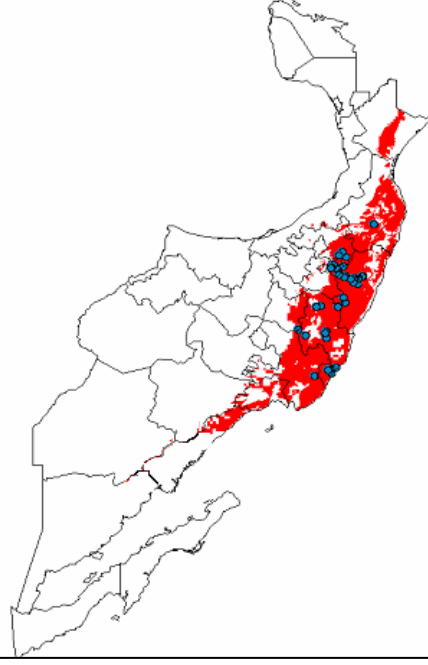
Fundamental Niche (= A)

Figure 4. An example of reduction of the FN area based on arguments of accessibility for understanding the distribution of *Baronia brevicornis*, a Mexican butterfly. It is assumed that every part of those biogeographic provinces in which the species has been recorded can be reached by potential colonists of this species; more explicitly, those biogeographic provinces from which the species has been recorded are within M, and as such the intersection of those provinces with the FN (A ∩ M) represents the portion of the FN that is accessible to the species. The presence of the species' coevolved host food plant—the third factor that would be equivalent to area B—is not shown.

One possible approach—as yet untested—would be selection of pseudoabsences with respect to an *a priori* estimate of region **A** based on an initial analysis with universal sampling of pseudoabsences. If, instead, pseudoabsences are placed at random outside an *a priori* estimate of **P**, as is often done in fine-scale studies, then the algorithm is forced to estimate **P** with the information about **B** and **M** being introduced indirectly and subjectively via the *a priori* knowledge of the expert—this approach is perilous, given that many sectors *within **A*** may not fall within **P** for reasons unrelated to abiotic features of the environment. A modeling technique rigorously designed to allow inclusion of *a priori* knowledge of the absences within **A** or **P** has been developed recently (Argaez *et al.*, in press)

DISCUSSION

The Introduction to this paper treated the two approaches that have been used to estimate FNs of species—the mechanistic approach and the correlative approach. Given the arguments developed above, several differences between the two can be appreciated. Whereas a mechanistic approach explicitly does not consider **B**, a correlative approach runs the risk of being biased by such interaction effects—still, it has the *potential* to take them into account explicitly via inclusion of the geography of other species in single-species models. Evolutionary effects, particularly in the form of geographic variation in niche characters (Peterson & Holt 2003), are explicitly excluded from the mechanistic approach, whereas they can be assessed in the correlative one via regional submodels within a broader species' distribution (Peterson & Holt 2003). In sum, the two approaches are complementary in the information that they provide; the relationships between the 'FN' estimated by each need to be clarified by direct empirical comparisons.

Based on the theoretical framework developed above, we can suggest some 'best practices'—suggestions for appropriate steps to be followed in development of predictive models for species' ecological niches. Key points are as follows.

- *Effects of scale*.—Interactions between spatial scales on which models are developed and the relative roles of **A**, **B**, and **M** in determining spatial and ecological 'behavior' of species lead to some clear lessons. For coarse-scale studies (continental or regional scales), **M** should prove particularly important—this point refers to the pervasive influence of landscape history, dispersal limitations, and other factors that are not strictly speaking 'ecological' in nature. On the other hand, in fine-scale studies, a different suite of factors becomes important—here, issues of historical heterogeneity and accessibility are less important, and considerations of the potentially more important role of interspecific interactions and metapopulation dynamics should be paid more attention.

- *Absence information*.—Information regarding absences of species should be used rigorously, if at all possible, given that 'absence' can have different meanings, and can prove positively misleading. That is, because **P** and **A** are probably only rarely coincident, it is very important to consider the referent for 'absence' information—absences in an ecological niche modeling context should be absences from **A,** but absences in a distributional modeling context should be absences from **P**. In this sense, when any sort of resampling for absence or pseudoabsence information is conducted, it is critical to consider carefully the goals and assumptions involved.

- *Model validation*.—Here again, it is important to consider carefully assumptions behind procedures being used. Niche models have universally been validated via comparisons with occurrences *in geographic space*, even though they are, strictly speaking, models of *ecological* niches and not geographic distributions. As such, large numbers of occurrence points distinct in geographic space *could* represent an artificial inflation of sample size, as they could all come from a single or a few environmental combinations. That is, if the model functions by predicting specific *environmental combinations* as 'present' or 'absent,' then this level is that at which we should be testing models, rather than projecting models onto geography and using geographic occurrences in pixels that could be environmentally redundant. This approach has seen little or no exploration—this gap appears to result from lack of effective tools for exploring, analyzing, and visualizing ecological niches in many-dimensional environmental space. Once such tools are developed, testing ecological niche models in ecological space will constitute an important step forward in improving the formal nature of ecological niche modeling.

The discussions developed in this paper, we hope, clarify the interplay among factors upon which ecological niche modeling applications can be built. The literature holds many variations in interpretation, and in the assumptions on which such models are founded, so we hope that this contribution serves to clarify dimensions to which this field is applicable. In essence, we analyzed the question of what *is* it that one models when one uses ecological niche modeling? The answer to this question appears to be that such algorithms mainly estimate the fundamental ecological niche (in the ecological dimensions used, often climatic in nature), which for comparatively small regions may coincide with the distributional area of a species. As we have seen however, this conclusion is subject to many caveats and conditions.

REFERENCES

Argaez, J., A. Christen, M. Nakamura, and J. Soberón. In press. Prediction of high potential areas of habitat for monitored species. Ecological Statistics.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago.

Carpenter, G., A. N. Gillison, and J. Winter. 1993. DOMAIN: A flexible modeling procedure for mapping potential distributions of animals and plants. Biodiversity and Conservation **2**:667-680.

Chase, J. M. and M. A. Liebold. 2003. Ecological Niches. Linking Classical and Contemporary Approaches. University of Chicago Press, Chicago.

Etterson, J. R., and R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science **294**:151-153.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation **24**:38-49.

Gaston, K. 2003. The Structure and Dynamics of Geographic Ranges. Oxford University Press, Oxford.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.

Holt, R. D. 1996a. Adaptive evolution in source-sink environments: Direct and indirect effects of density-dependence on niche evolution. Oikos **75**:182-192.

Holt, R. D. 1996b. Demographic constraints in evolution: Towards unifying the evolutionary theories of senescence and niche conservatism. Evolutionary Ecology **10**:1-11.

Holt, R. D., and M. S. Gaines. 1992. Analysis of adaptation in heterogeneous landscapes: Implications for the evolution of fundamental niches. Evolutionary Ecology **6**:433-447.

Holt, R. D., and R. Gomulkiewicz. 1996. The evolution of species' niches: A population dynamic perspective. Pages 25-50 in H. G. Othmer, F. R. Adler, M. A. Lewis, and J. C. Dallon, editors. Case Studies in Mathematical Modeling: Ecology, Physiology and Cell Biology. Prentice-Hall, Saddle River, N.J.

Hutchinson, G. E. 1957. Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology. 22: 415-42

Jones, P. G., and A. Gladkov 1999. FloraMap: A computer tool for predicting the distribution of plants and other organisms in the wild. Centro Internacional de Agricultura Tropical, Cali, Colombia.

Kawecki, T. J. 1995. Demography of source-sink populations and the evolution of ecological niches. Evolutionary Ecology **9**:38-44.

Nix, H. A. 1986. A biogeographic analysis of Australian elapid snakes. Pages 4-15 in R. Longmore, editor. Atlas of elapid snakes of Australia. Australian Government Publishing Service, Canberra.

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

Pearson, R. G., T. P. Dawson, P. M. Berry, and P. A. Harrison. 2002. SPECIES: A spatial evaluation of climate impact on the envelope of species. Ecological Modelling **154**:289-300.

Peterson, A. T., and R. D. Holt. 2003. Niche differentiation in Mexican birds: Using point occurrences to detect ecological innovation. Ecology Letters **6**:774-782.

Peterson, A. T., V. Sánchez-Cordero, J. Soberón, J. Bartley, R. H. Buddemeier, and A. G. Navarro-Sigüenza. 2001. Effects of global climate change on geographic distributions of Mexican Cracidae. Ecological Modelling **144**:21-30.

Porter, W. P., S. Budaraju, W. E. Stewart, and N. Ramankutty. 2000. Calculating climate effects on birds and mammals: Impacts on biodiversity, conservation, population parameters, and global community structure. American Zoologist **40**:597-630.

Porter, W. P., J. L. Sabo, C. R. Tracy, O. J. Reichman, and N. Ramankutty. 2002. Physiology on a landscape scale: Plant-animal interactions. Integrative and Comparative Biology **42**:431-453.

Pulliam, R. 2000. On the relationship between niche and distribution. Ecology Letters 3:349-361

Stockwell, D. R. B. 1999. Genetic algorithms II. Pages 123-144 in A. H. Fielding, editor. Machine learning

methods for ecological applications. Kluwer Academic Publishers, Boston.

Stockwell, D. R. B., and I. R. Noble. 1992. Induction of sets of rules from animal distribution data: A robust and informative method of analysis. Mathematics and Computers in Simulation **33**:385-390.

Stockwell, D. R. B., and D. P. Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems **13**:143-158.

Stockwell, D. R. B., and A. T. Peterson. 2002. Controlling bias in biodiversity data. Pages 537-546 in J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting Species Occurrences: Issues of Scale and Accuracy. Island Press, Washington, D.C.