REVIEW

Habitat suitability modelling and niche theory

Alexandre H. Hirzel\textsuperscript{1,2,*} and Gwenaëlle Le Lay\textsuperscript{1}

\textsuperscript{1}Department of Ecology and Evolution, Laboratory for Conservation Biology, University of Lausanne, CH-1015 Lausanne, Switzerland; and \textsuperscript{2}Zoological Institute, Division of Conservation Biology, University of Bern, CH-3012 Bern, Switzerland

Summary

1. The concept of the ecological niche relates a set of environmental variables to the fitness of species, while habitat suitability models (HSMs) relate environmental variables to the likelihood of occurrence of the species. In spite of this relationship, the concepts are weakly linked in the literature, and there is a strong need for better integration.

2. We selectively reviewed the literature for habitat suitability studies that directly addressed four common facets of niche theory: niche characteristics, niche interactions, community-wide processes and niche evolution.

3. We found that HSMs have mostly contributed to the study of niche characteristics, but the three other themes are gaining impetus. We discuss three issues that emerge from these studies: (i) commonly used environmental variables and their link with ecological niches; (ii) the causes of false absences and false presences in species data, and associated issues; (iii) the three axes of model generalization (interpolation and extrapolation): environmental, spatial and temporal. Finally, we propose a list of 12 recommendations to strengthen the use of HSMs for wildlife management.

4. Synthesis and applications. This selective review provides conservation biologists with a list of pointers to key niche-theory concepts and a wide palette of related HSM studies. It also brings together frameworks that are often separated: theoretical and applied ecology studies; botany, zoology and parasitology; and different HSM frameworks, such as Resource Selection Functions, Species Distribution Modelling, Ecological Niche Modelling, and Gradient Analysis. We hope that integration of all these slices of knowledge will improve the quality and reliability of HSM predictions.

Key-words: community ecology, conservation, fundamental and realized niche, Geographic Information Systems (GIS), generalization, niche evolution, Resource Selection Functions (RSF), spatial predictions

Introduction

That individual species only thrive within definite ranges of environmental conditions has stimulated one of the most fertile fields in ecology: the ecological niche theory (Chase & Leibold 2003). The requirement-based concept of the ecological niche (Grinnell 1917; Hutchinson 1957) defines it as a function that links the fitness of individuals to their environment. This quantitative definition allowed the niche theory to tackle a variety of issues, including evolutionary processes, competition and predation dynamics. In parallel, the last 15 years have seen the rise of habitat suitability modelling (Rushton, Ormerod & Kerby 2004), which aims to predict the likelihood of occurrence of species on the base of environmental variables (Franklin 1995; Guisan & Zimmermann 2000). Habitat suitability models (HSMs) can thus be seen as operational applications of the ecological niche, using environmental variables to predict the presence/absence or the abundance of a species throughout a study area (Fig. 1).

Despite the obvious relationship between niche theory and HSMs, they remain weakly linked and there is a strong need for more integration (Pulliam 2000; Austin 2002). We have explored the relationships between HSMs and niche theory to provide pointers to practical aspects of niche theory and to related HSM-based studies. We reviewed the literature, searching for articles that directly addressed common issues in niche theory by means of HSMs. We report herein only a selection of these studies. More comprehensive lists can be found elsewhere (Guisan & Zimmermann 2000; Peterson 2006).

*Correspondence author. E-mail: alexandre.hirzel@unil.ch
Several HSM frameworks have grown relatively independently and have been communicated in different groups of journals and conferences. Confusingly, techniques variously designated as Habitat Suitability/Selection Models (Akcakaya 2000; Hirzel et al. 2004), Habitat/Species Distribution Models (Guisan et al. 2000; Rushton et al. 2004), Resource Selection Functions (Boyce & McDonald 1999; Manly et al. 2002), Ecological Niche Models (Peterson 2006) or Gradient Analysis (Austin, Cunningham & Fleming 1984) actually address similar issues with different tools (regressions, envelope-modelling, classification trees, fuzzy logics, Bayesian models, artificial neural networks, factor analyses). We hope that this work will contribute to breaking down the ‘cultural’ barriers among these various HSM frameworks and terminologies.

ECOLOGICAL NICHE: AN OPERATIONAL DEFINITION

In the ecological literature, the term ‘niche’ is confusingly used to cover two distinct concepts (Vandermeer 1972; Chase et al. 2003): (i) the environmental requirements needed for a species to subsist without immigration (Grinnell 1917), and (ii) its relationships to other species (Elton 1927). These two concepts cannot be investigated with the same tools and they often apply to different scales (Soberon 2007). In this review, we consider only the ‘Grinnellian niche’.

In the absence of immigration, a species can only subsist if the local combination of environmental variables allows, on average, a positive population growth rate. These variables (also called predictors, covariates, or ecogeographical/independent variables) are the dimensions of the environmental space. The ecological niche is the volume in the environmental space that permits positive growth (Hutchinson 1957). The growth rate decreases from the niche optimum to the niche envelope where it is null (‘zero net growth isocline’, Tilman 1980), defining the ‘niche shape’. Further out, the growth rate is negative (Fig. 1).

The presence of a species in a location obeys three constraints (Soberon & Peterson 2005; Soberon 2007): (i) the local environment allows the population to grow (Grinnellian niche), (ii) the interactions with other local species (predation, competition, mutualism, etc.) allow the species to persist (Eltonian niche), and (iii) the location is actually accessible, given the dispersal abilities of the species. These constraints determine the geographical distribution of the species. It is thus theoretically possible to reconstruct a realized Grinnellian niche for a species from the environmental variables measured at the locations it occupies. This reconstruction and the spatial predictions derived from it are the goals of HSM. In practice, the strength of the distribution–niche link depends on the ecology of the species, local constraints and historical events (cf. Pulliam 2000).

Review

We focused on going beyond spatial predictions and specifically study four themes of niche theory: (i) individual niche characteristics, (ii) niche interactions, (iii) communities, and (iv) niche evolution.

NICHE CHARACTERISTICS

Published HSM studies have focused on three questions related to the niche of an individual species: (i) environmental variable selection, (ii) fitness response curves, and (iii) interactions between variables.
**Variable selection**

Identifying the key environmental variables that determine the niche is one of the most crucial HSM operations. The selection of candidate variables often relies on expert knowledge (Guisan et al. 2000; Manly et al. 2002). Then, most methods look for the smallest combination of variables that produces the best fit to the data (Johnson et al. 2006). Stepwise algorithms automate this selection/rejection process but risk rejecting variables more because of spurious correlations with other variables than for ecological reasons (Hirzel et al. 2002).

Other HSMs synthesize environmental variables into composite, uncorrelated variables—called factors—that summarize significant ecological information and can be used as new variables into HSMs. For instance, Guichon & Cassini (1999), using Principal Component Analysis, synthesized 26 variables into three factors that explained 60% of the information (see also Saab 1999). Rydgren, Ökland & Ökland (2003) found that the use of factors computed by a Detrended Correspondence Analysis smoothed the noise contained in the original variables and improved the predictions. While synthetic factors reduce system complexity, they may be difficult to interpret ecologically. The Ecological Niche Factor Analysis (ENFA, Hirzel et al. 2002; Braunisch et al. 2008) provides factors that are directly related to the niche: marginality and specialization (see also Calenge et al. 2008).

**Response curves**

Most theoretical models assume, mainly for mathematical convenience, that response curves are either sigmoid or Gaussian (Austin 1999). Although ecologically plausible, such a symmetry or continuity may be rare in real systems. Austin & Gaywood (1994a) argue that niches lying close to one extreme of a gradient are skewed away from it. Niche shapes may also depart from Gaussian because of interspecific interactions.

Regression-based HSMs fit curves ranging from parametric functions, like logistic regressions, Generalized Linear Models (GLM), Resource Selection Functions (RSF) to less-constrained shapes, like Generalized Additive Models (GAM), Fuzzy Envelope Models. For instance, Bio et al. (2002) modelled the response of some plants to nutrient concentrations with both GAMs and GLMs. The GAM functions were constrained to two degrees of freedom, and the GLM to two-degree polynomials. While conveying comparable complexity, GAM curves may take more flexible shapes (asymmetries or slope breaks) and thus provided a tighter fit than GLMs on 11 of the 18 species. Lehmann and colleagues (Lehmann 1998; Lehmann, Overton & Leathwick 2002) provided further striking examples of complex response curves.

Austin and colleagues (Austin 1987; Austin et al. 1994a,b) directly tested the Gaussian shape hypothesis on several Eucalyptus species. They showed that response curves were often significantly asymmetric, with a tendency to skew away from extreme climatic values. Rydgren et al. (2003), by exploring the response curves of more than 1000 plants, found that long gradients and narrow niches increased the frequency of symmetric bell-shaped responses. Skewness, although slightly favoured by the optima for a species being near extremes of gradient, was rather uncommon.

**Interactions between variables**

Organisms usually respond to a complex of interdependent factors that consist of many environmental variables (Rydgren et al. 2003). For instance, a plant may survive heat or drought that occur separately, but not together (Carpenter, Gillison & Winter 1993). Similarly, opportunistic animals that are able to survive on several types of resource, often feed on the most common one (substitutable resources, Tilman 1980).

Some HSMs can detect and quantify interactions between variables. For instance, Bartlein, Prentice & Webb (1986) modelled two-dimensional climatic niches of several North American tree species on the basis of pollen percentages. They found several niche shapes (bell-shaped, planar, bimodal) with various degrees of interaction between the variables (see also Huntley, Bartlein & Prentice 1989; Prentice, Bartlein & Webb 1991). These plant studies showed that most of the investigated niches were affected by interactions between climatic variables.

**Niche interactions**

Species never live in isolation. Whenever several species co-exist, biotic interactions (like competition, predation, parasitism, mutualism) affect their fitness and behaviour and may drastically affect their niches (Pearson & Dawson 2003 and references therein). For instance, the presence of a superior competitor may prevent a species from occupying some part of its niche, leading to a truncated or even bimodal niche (Austin 1999). The interaction may be direct, such as through interference and predation, or indirect, by depleting a common resource or being preyed upon by a common predator. Hutchinson (1957) defined the realized niche as a subset of the fundamental niche a species was constrained to occupy because of interactions with other species.

A common opinion is that HSMs can only measure the realized niche (Guisan et al. 2000). However, the notion of equilibrium has become increasingly suspect because of species’ dispersal and environment variability (Alley 1982). Indeed, biotic interactions act at short distances, while most species have limited dispersal. This allows inferior competitors to evade negative interactions by settling in competitor-free locations (Harrison, Thomas & Lewinsohn 1995). This is particularly true for sessile organisms, or vagile species that live in metapopulations (Anderson, Peterson & Gómez-Laverde 2002; Peterson & Holt 2003). Thus, when spatial heterogeneity and dispersal limitations are accounted for, the differences between the realized and fundamental niches may be much smaller than expected from the spatially homogeneous environment assumed in most theoretical studies (Pulliam 2000). Moreover, wide-extent, heterogeneous areas favour a
high diversity of communities with varying equilibrium abundances of species (Huntley et al. 1989; Pearson & Dawson 2003). Accordingly, at such scales, abiotic variables may overcome biotic interactions. Indeed, many country- or continent-wide models achieve good accuracy of prediction when based on climatic variables only (Pearson et al. 2002; Vetaas 2002). For HSMs, it is preferable to distinguish between the Grinnellian and the Eltonian niche, rather than between the fundamental and the realized niche (Soberon 2007). The Grinnellian niche is based on broad-scale variables (climate) that are not affected by species density, while the Eltonian niche relates to fine-scale variables that may be consumed or modified by the species (nutrients).

### Niche comparisons

A simple way of comparing niches with HSMs is to model the habitat of different species independently in the same area and compare their characteristics. This may be done either in geographical space by comparing the predicted distributions of species (Anderson et al. 2002; Rice et al. 2003), or in the environmental space by measuring niche similarities (e.g. Reutter et al. 2003; Chefaoui, Hortal & Lobo 2005).

These correlative approaches can only show spatial relationships among species. Only an experimental approach (e.g. Connell 1983; Schoener 1983) can test the existence of a causal link. However, the spatial scales involved generally prevent experimentation. Some particular situations may offer a near-experimental design to HSM-based studies if they meet these three conditions: (i) the species occupy mostly distinct areas, thus providing a ‘natural, removal experiment’, (ii) their niches are partially overlapping, and (iii) the species occur in sympathy in some areas (Anderson et al. 2002).

### Competition

A basic tenet of ecology is Gause’s (1934) principle of competitive exclusion, which asserts that whenever two sympatric species have similar niches, one will out-compete the other and drive it to extinction (Tilman 1982). From there, most theoretical models assume a competitive equilibrium (i.e. sympatric species have evolved so as to minimise niche overlap and interspecific competition; Alley 1982).

Anderson et al. (2002) studied the mutual exclusion of two species of spiny pocket mice (*Heteromys* spp.). They built separate HSMs and identified regions of potential sympathy. This hinted towards potential asymmetry in competition for habitat use, and species displacement. Similarly, Leathwick & Austin (2001) showed that the tree *Nothofagus* was a strong competitor, able to displace other widespread species. The species experiencing the greater niche overlap with *Nothofagus* were the most strongly affected.

### Mutualism

Gutierrez et al. (2005) explored the mutualistic relationships between a butterfly and an ant. They computed three HSMs that differed by the predictors which included: (i) ant frequency, (ii) a set of environmental predictors, (iii) both. The results showed that probability of occurrence and abundance of the butterfly was higher in sites with high ant frequencies.

### Predation

Hebblewhite, Merrill & McDonald (2005) studied the relationship between environment and predation for the wolf–elk system. They built HSMs for all four predation stages: elk roaming, wolf searching, wolf/elk encounters, and kills. This allowed them to decompose the predation risk into probability of encounter given availability and probability of kill given encounter, allowing them to predict elk vulnerability in various types of environment. Similarly, Sanchez-Cordero & Martinez-Meyer (2000) used HSMs to assess the impact from various rodent species on several crop species.

### Parasitism / commensalism

Peterson and colleagues studied the potential of Chagas disease to spread by means of HSMs. The cycle of this parasitic protozoan involves mammal hosts and insects vectors. Hosts and vectors showed more than 90% habitat overlap (Peterson et al. 2002b). Another study explored how the expansion of the commensal house crow was favoured by human presence (Nyari, Ryall & Peterson 2006).

### Multi-species studies

Understanding the mechanisms that govern species richness and composition has stimulated much interest from ecology theoreticians (Klopfen & Macarthur 1961; Tilman 1994). In the context of biodiversity loss and global climate change, these issues are now of crucial importance for conservation biology and management (Loreau, Mouquet & Gonzalez 2003). Two opposing mechanisms have been proposed to explain species’ distribution patterns. First, niche theories argue that species composition is driven by environmental heterogeneity and adaptations of species (Bolliger, Kienast & Bugmann 2002; Tokesh & Schmid 2002). Second, neutralist theories argue that limited dispersal and stochastic extinction/speciation events are the driving factors, even if species have identical niches (MacArthur & Wilson 1967; Hubbell 2001). The synthesis of these contradictory viewpoints is beginning (Chave 2004; Ricklefs 2004; Tilman 2004), and HSMs no doubt have an important role to play.

Gilbert & Lechowicz (2004) elaborated a sampling design that eliminated covariance between geographical distance and environmental similarity. They used HSMs to relate the niches and dispersal patterns of six plant groups to various environmental and dispersal-related variables. They found that environment explained a larger proportion of the variance than dispersal.

Another approach studies global characteristics and spatial patterns of the species pool (e.g. biodiversity or species abundance) over wide areas, thus neglecting local species–
environment and most interspecific interactions. For instance, Zani, Lehmann & Overton (2002) identified fern biodiversity hotspots by summing single-species HSMs. Dominguez-Dominguez et al. (2006) performed a similar study of fish diversity. The effect of global climate change on the composition and distribution of species has recently attracted much attention. These analyses aim at assessing the extinction risk of various species by predicting habitat reduction and species turnover. They assume that niche characteristics cannot evolve fast enough to adapt to environmental changes. Thus, species must either track their suitable environment in space or die (Peterson, Soberon & Sanchez-Cordero 1999; Peterson & Vieglais 2001; Martinez-Meyer, Townsend Peterson & Hargrove 2004). For instance, Peterson and colleagues (2002a) used museum data to explore the consequences of climate change on the distribution of 1870 Mexican species. They tested three assumptions about dispersal abilities under two scenarios of climate change. Most of their predictions show a reduction in species’ distribution (with no differences between taxonomic groups), with a greater impact on endemic species. These changes also predict a high species turnover on a large part of the studied area.

Another approach considers how the system as a whole responds to environmental changes. In this case, it is generally assumed that: (i) the system has reached an evolutionary equilibrium, and (ii) all species respond similarly to environment modifications. For instance, Haire et al. (2000) studied the impact of urban growth on bird communities. They showed that all species tended to decrease in abundance with the increase of human impact.

For functionally close species, expanding the niche concept to the whole group may be more relevant. Canonical Correspondence Analysis (CCA, Ter Braak 1986) is an HSM well-suited to study groups of species because it computes independent factors that emphasise patterns of species occurrence. These factors can be interpreted as composite environmental variables. The space defined by these factors allows delineation of groups of species that occur in similar habitats. For instance, Hill (1991) analysed jointly the niche and geographical distribution of birds and plants, summarizing 12 climatic, geological, topographic and spatial variables into four factors. Guisan, Weiss & Weiss (1999) assessed environmental factors relevant to a set of 10 tree species. However, as CCA expresses the global response of several species, its accuracy is generally lower than that of single-species models (Guisan et al. 1999).

### NICHE EVOLUTION

Although habitat use by individuals results from individual events at small time-scales, global patterns are ultimately driven by habitat-dependent fitness (Gilbert & Lechowicz 2004), that is, the fundamental niche is subjected to natural selection (Bradshaw, Zani & Holzapfel 2004). However, for a niche to evolve, the new conditions must not lie too far outside the ancestral niche. Thus, natural selection tends to act principally as a conservative force. (Holt & Gaines 1992). Accordingly, in the case of gradually, directionally changing environmental conditions, a species is condemned either to track its environment across space or to go extinct (Pease, Lande & Bull 1989).

Several HSM-based studies have contributed to niche evolution theory. For instance, Huntley et al. (1989) compared the climatic niche of two beechnut trees, one from Western Europe and the other from North-eastern America, which had been separated by continental drift. Despite this long separation, their niches still show greater similarity to one another than to those of any other taxon from their respective continents. The slowness of niche evolution was similarly confirmed by Peterson et al. (1999), who studied 37 sister-animal taxa pairs isolated on either side of a geographical barrier. Their results indicate strong niche conservatism between species pairs, but little at the family level. However, other taxa may show faster ecological differentiation (Peterson et al. 2003).

### Discussion

#### HSM AND NICHE THEORY

From our literature review, it emerges that HSM-based studies have traditionally addressed the niche issues of single species, with fewer studies about niche interactions, communities and evolution.

Most analyses addressing niche shape use one-dimensional response curves (mainly regression techniques). To improve on their limited ability to model interactions between variables, two- or three-dimensional polynomials may be used. Some methods also readily provide measures of niche geometry, such as niche breadth (e.g. Green 1971), overlap (e.g. Sattler et al. 2007), marginality and specialization (e.g. Hirzel et al. 2004) or skewness (e.g. Austin et al. 1994b). They generally allow visualization of the niche in one (e.g. Lehmann 1998; Bio et al. 2002), two (e.g. Bartlein et al. 1986; Hirzel & Arlettaz 2003; Peterson et al. 2006) or three dimensions (e.g. Prentice et al. 1991).

Comparatively few HSM studies have addressed niche interactions or comparisons, although they have proved useful to assess similarities among species, and to infer potential interactions such as competition, mutualism and parasitism. Here, HSMs can benefit from classical niche comparison studies (e.g. Colwell & Futuyma 1971). In particular, niche theory provides many niche overlap measures (e.g. Hurlbert 1978) that have still to be ported to the HSM world (but see Sattler et al. 2007).

Similarly, few HSM studies have yet addressed issues of species assemblage, and they often make strong, dubious simplifications (e.g. equilibrium, niche conservatism). Within the current context of biodiversity erosion, this is a domain that needs urgent development and better coupling with complementary methods (such as experimentation).

Finally, HSMs have been quite successful at exploring niche evolution. As evolution is usually too slow to lend itself to experimentation, analytical techniques such as HSMs appear well-suited to address these questions. The limiting
factor is the availability of ancient data (fossils, pollen) and ancient environment (usually palaeoclimate). However, molecular and phylogenetic technologies have allowed HSMs to tackle the evolutionary issues indirectly, offering a promising research axis.

Overall, there is a strong bias in the studied organisms and ecosystems towards large, terrestrial species, which rarely involve complex life cycles such as parasitism, commensalism, multiple habitats, pluriomodal niches, etc. Our knowledge of the relationships between habitat selection and niche characteristics will remain incomplete as long as the diversity of biological systems is not further explored.

HABITAT VARIABLES AND NICHE DIMENSIONS

Grinnell (1917) listed the factors that potentially affect species’ distribution. These locally measured variables concerned vegetation, food, climate, soil, breeding and refuge sites, interspecific effects, and individual/racial preferences. The advent of remote sensing and geographical information system (GIS) technologies has made available a wider spectrum of spatial information, covering larger areas, which are known to influence the macro-distribution of species (such as climate, topography, land-cover and satellite imagery). An important contribution of HSMs to niche theory has been their ability to test the relevance of such variables for the species–environment relationship.

Climatic variables, and in particular temperature, are among the most important factors that drive species’ distribution (Grinnell 1917; Guisan et al. 2000), especially in large extents, as they have a direct influence on the behaviour and physiology of organisms. They are particularly important for plants, which cannot evade adverse weather by sheltering or migrating. Animals’ response to climatic variables is partly indirect, through correlations with the vegetation used as food or shelter. Bioclimatic variables (e.g. temperature of the coldest month, etc., WorldClim, Hijmans et al. 2004) might offer a better fit than the simple monthly or yearly averages.

Topography mostly affects species indirectly through its correlation with temperature and precipitation, but also through landscape diversity and configuration, soil and water dynamics. The variables derived from Digital Elevation Models (e.g. slope, aspect, convexity) are often crucial for plants, as they affect local conditions of light, wetness, daily temperature amplitude, soil stability and granulometry, etc. (Guisan, Theurillat & Kienast 1998). For animals, these factors may directly affect their locomotion but may also affect them indirectly through vegetation (Maggini, Guisan & Cherix 2002).

Land-cover data have the most diverse influence on ecological niches. They may represent food, shelter, breeding sites, disturbance, presence of competitors, etc. (e.g. Schadt et al. 2002; Sachot, Perrin & Neet 2003; Seoane, Bustamante & Díaz-Delgado 2004). However, land-cover maps are usually designed for land management and may not be well-suited to ecological purposes. Moreover, they often suffer from poor spatial accuracy and infrequent updates, which may preclude their use in fine-grained ecological models. Remote sensing data may sometimes substitute for them (Aspinall & Veitch 1993).

Some satellite sensors are providing ecologists with long time series of visible and infrared imagery. They allow computation of the Normalized Difference Vegetation Index (NDVI), a surrogate for plant biomass (Estrad-Peña, Venzal & Acedo 2006), or can be used directly (Zimmermann et al. 2007). They provide measures of seasonal vegetation and of contrasts within or between years.

Causal variables (i.e. those that directly affect fitness) may refer to fine-grain resources (Marchesi & Sergio 2005), to disturbances (Gibeaudeau al. 2002; Braunisch & Suchant 2007), to human activities (Le Lay 2002), to grazing (Stohlgren, Schell & Vanden Heulen 1999), to trampling (Cairns & Moen 2004), to landscape structure (Hokit, Stith & Branch 1999; With & King 2004) or to hunting sites (Titeux et al. 2007). Moreover, when a species is strongly affected by the presence of another organism (consumer, resource, mutualist), the predictors should include either its exhaustive distribution (e.g. Sachot et al. 2003) or its habitat suitability map (e.g. Zimmermann & Breitenmoser 2002). However, the variables that are closest to a species’ fitness are often measured ad hoc (Titeux et al. 2007), and are thus particular to the focal species and available at a limited number of sites. Although such variables are desirable for model accuracy, they may lower its applicability to wider areas. As often in modelling, there is a tight trade-off between accuracy and generality.

SPECIES’ DISTRIBUTION AND NICHE INTERPRETATION

Existing HSMs rely on the premise that the observed geographical distribution of a species reflects its ecological requirements. Most HSMs assume that species are present in suitable habitats (true presences) and absent from unsuitable ones (true absences). In practice, however, the relationship between niche and distribution is more equivocal. A species may be absent from a suitable site or present in an unsuitable site for several reasons.

Unrecorded presences stem from these four causes: (i) incomplete sampling, (ii) selective sampling (e.g. lek or nests sites) (iii) cryptic or rare species, or (iv) faulty determinations. Whilst time, effort and skill can sometimes minimize this problem, it is often unavoidable. This random error becomes a bias when the probability of not recording a presence is larger in unsuitable than in suitable sites (MP2 > MP1 in Fig. 2), which may often be the case.

Fallacious absences (FA) stem from at least five causes, namely: (i) limited dispersal: geographical barriers or slow dispersal prevent a species from occupying some parts of its potential distribution; (ii) local extinction: environmental or demographic stochasticity has momentarily driven a local population to extinction; (iii) patch size: the area of suitable habitat is too small to harbour a viable population; (iv) alternative habitats: a generalist species may use several types of habitat, possibly at different periods; (v) biotic interactions (e.g. succession stage, competition, predation). FAs are

particularly problematic to HSMs because, contrary to common belief, they are unavoidable. When FA are suspected, presence-only HSMs should be preferred (Hirzel, Helfer & Métral 2001; Pearson et al. 2006; Chefaoui & Lobo 2008).

Fallacious presences (FP) stem from these two main causes: (i) source/sink dynamics: crowding in suitable areas induces individuals to settle in surrounding unsuitable sites (Pulliam 2000) (ii) high vagility: individuals are moving or migrating over a wide expanse of unsuitable habitats (Hirzel et al. 2004). Using fitness-related species data (e.g. nest sites instead of sightings) reduces the number of FPs.

The niche characteristics must be interpreted carefully, especially in conservation applications. Models usually offer only a partial view of the reality.

INTERPOLATION AND EXTRAPOLATION

In practice, HSMs are usually calibrated on a number of sites from the study area and thus, technically, are only valid for them. However, models may be generalized to wider contexts along three axes: (i) environmental generalization (i.e. making predictions for environments that were not found in the calibration set; Fig. 3); (ii) spatial generalization (i.e. making predictions for sites that were not part of the calibration set); (iii) temporal generalization (i.e. making predictions for times before or after the date of the calibration set). For each axis, the generalization can consist of application if it predicts only for the values found in the calibration set, of interpolation if the range of predicted sites stays within the calibration range, and of extrapolation if it goes beyond (Table 1). Generally, the degree of generalization is inversely proportional to sampling size.

Whilst environmental interpolation is acceptable and generally unavoidable, extrapolation is hazardous and should be avoided (Pearson et al. 2006; Peterson, Papes & Eaton 2007). Space and time are more amenable to generalization (Peterson 2003; Randin et al. 2006). The main risks are changes in the relationships among the variables or among the interacting species. Temporal extrapolations largely depend on the availability of climate models.

Extrapolation and interpolation generally require different types of species and environmental data, as well as different modelling techniques. Interpolation often aims to delineate precisely the actual distribution of a species. The environmental variables need more to be fine-grained and accurate than to be causal and widely available. Presence/absence techniques (e.g. generalized regression, classification trees) may provide better results. By contrast, extrapolation requires a more causal HSM. The larger extent of such studies may necessitate coarser but widely available variables. Presence-only techniques may be preferable (e.g. GARP, ENFA, RSF).

Interpolation and extrapolation must be used with caution. Increasing the sampling size and tightening the study area to the study’s goals may help to reduce the amount of generalization needed. Generalization should be particularly avoided when studying partial niches (Selander 1966) such as juvenile/adults, reproduction/migration, male/female (Lodé 1996), or seasonal variations (Martinez-Meyer, Peterson & Navarro-Siguenza 2004). Whilst partial niches provide a finer understanding of the system by focusing on critical processes (Sachot et al. 2003), they are difficult to generalize.

Table 1. Types and axes of model generalization from calibration sites to prediction sites

<table>
<thead>
<tr>
<th>Type</th>
<th>Environmental</th>
<th>Spatial</th>
<th>Temporal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Same value combination</td>
<td>Same sites</td>
<td>Same dates</td>
</tr>
<tr>
<td>Interpolation</td>
<td>Within calibration range</td>
<td>Within convex hull</td>
<td>Within calibration period</td>
</tr>
<tr>
<td>Extrapolation</td>
<td>Outside calibration range</td>
<td>Outside convex hull</td>
<td>Outside calibration period</td>
</tr>
</tbody>
</table>
Habitat suitability modelling and niche theory

MANAGEMENT IMPLICATIONS

Existing HSMs have already contributed greatly to conservation and management by allowing ecologists to delineate the ecological requirements of species and their limiting factors; to understand biogeography and dispersal barriers; to find unknown populations and new species; to identify reintroduction sites; to design conservation plans and reserves; to predict effects of habitat loss; to anticipate species invasions; and to predict climate change effects (review in Peterson 2006).

Insights from niche theory and related HSM studies allow the derivation of this set of recommendations for computing meaningful habitat suitability maps:

1. Determine early the goals of the study, in particular in regard to grain-size and generalization.
2. Pre-select environmental variables carefully; this is arguably the most critical point. Grouping the variables by theme (e.g. climate, resources, disturbances) and testing them separately is often insightful.
3. Carefully delineate the study area by removing all places where the species absence is trivial (e.g. lakes, roads, mountain tops).
4. Identify possible sources of unrecorded presences, fallacious absences and presences. If possible, use species data that are directly linked to fitness.
5. Select the HSM technique most relevant to the goal of the study and the characteristics of the species (see Guisan et al. 2000; Guisan & Thuiller 2005). If possible, try several approaches.
6. Know your HSM: its assumptions, its caveats, its strengths. Understand its inputs and outputs, and how it works.
7. Test model sensitivity to various inputs and parameters. If the predictions are not robust, understand why.
8. Interpret the HSM. Does it make ecological sense?
9. Evaluate the predictions. Assess predictive power and variance. Interpolation and extrapolation studies generally need different approaches (Hirzel et al. 2006; Peterson et al. 2007; Lobo, Jiménez-Valverde & Real 2008).
10. Provide a map of prediction confidence. In particular, indicate the areas where the model is applied, interpolated and extrapolated (Hirzel et al. 2008).
11. Reclassify the predictions into meaningful and honest values. The results of HSMs often imply a misleading precision. A map displaying four robust levels of suitability is more reliable than a rainbow of unsupported nuances (Hirzel et al. 2006). It is also easier to communicate to policy makers and the public.
12. Take home range size into account. Large species often require a wide expanse of suitable habitat to survive. Consider discarding suitable patches that are too small.

These 12 steps are essential to guarantee habitat suitability models reflecting at best the species–environment relationships occurring in the study area.

Acknowledgements

This work benefited greatly from discussion with A. Guisan, N. Perrin, J Hauser, P Vogel and the participants to the following workshops: ‘Resource Selection by Animals’ (Laramie, USA, 2003), ‘Predicting Species’ Distributions’ (NCEAS, Santa Barbara, USA, 2003), and ‘Generalized Regression Analyses’ (Riederlap, Switzerland, 2001, 2004). J Bonetti and Y Delorme helped us with the bibliography. D. Strickland, M. Boyce, A.T. Peterson, M. Hebblewhite and three anonymous referees provided helpful comments on earlier versions of the manuscript. This work was supported by the MAVA foundation and the Swiss National Foundation to A.H. (nos. 3100A-112511).

References


Habitat suitability modelling and niche theory


Received 31 May 2007; accepted 23 June 2008

Handling Editor: Mark Hebblewhite