Biodiversity and Conservation **9**: 1313–1322, 2000.

Preliminary distributional analysis of US endangered bird species

MANDALINE E. GODOWN and A. TOWNSEND PETERSON*
*Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, KS 66045, USA; *Author for correspondence (e-mail: town@ukans.edu; fax: +1-785-864-5335)*

Received 7 September 1999; accepted in revised form 20 December 1999

Abstract. A first exploration of applications of ecological niche modeling and geographic distributional prediction to endangered species protection is developed. Foci of richness of endangered bird species are identified in coastal California and along the southern fringe of the United States. Species included on the Endangered Species List on the basis of peripheral populations inflate these concentrations considerably. Species without protection in the US National Park System are focused particularly in peninsular Florida. Application of this methodology to additional taxa and regions holds promise for diverse conservation applications.

Key words: BIOCLIM, birds, conservation strategies, ecological niche modeling, endangered species, GARP, United States

Introduction

Countries worldwide are working to preserve species that have declined to the brink of extinction. The United States is no exception to this effort: the US Endangered Species Act (ESA) is arguably the strongest piece of environmental legislation in the country (Petit and Schmidt 1998). This federal mandate requires protection of endangered and threatened species and their critical habitats, and calls for development of site-specific recovery plans for each species (Congressional Register 1973, 16 U.S.C. §§ 1531–1544).

Nevertheless, endangered species in the United States exist over an extensive range of ecological and geographic features, making designing a conservation strategy to protect them an extremely complicated endeavor (Dunn et al. 1997; Ehrenfeld et al. 1997). Establishing regional conservation priorities within the United States for endangered species conservation based on quantitative criteria would be extremely valuable in landuse planning (Dobson et al. 1997). More generally, development of effective tools for identifying regions with concentrations of endangered species could, at the very least, highlight areas where preservation efforts are most needed.

It is now possible to build ecological models based on known occurrences of species that predict their geographic distributions with great precision (Peterson and

Cohoon 1999). With increased availability of biodiversity information over the Internet (Soberón 1999), computer modeling of ecological niches is feasible. These models can then be integrated in a GIS environment to visualize distributions and refine them into conservation strategies. In this way, the pieces of a complex conservation challenge are integrated into a coherent set of information and recommendations.

The purpose of this study is to apply these new methodologies to the United States' endangered bird species as a prototype application of new technologies to these challenges. The goal of this study is to identify at a regional level the areas most critical for preservation of endangered species, and to explore their importance in prioritization for protection. Most importantly, however, this study is designed as an exemplar for application to regions lacking detailed distributional data on their rare and endangered species.

Ecological niche modeling

The fundamental ecological niche of a species can be defined as the conjunction of ecological conditions within which it is able to maintain populations (MacArthur 1972); as such, it is a shape or shapes defined in multidimensional ecological/environmental space. The fundamental niche, the focus of our modeling efforts, must be distinguished carefully from the realized niche (that part which is actually occupied), to maintain focus on the ecological dimensions important to a particular species. Our applications of these approaches involve both physical variables and biotic variables, although the presence of the latter is limited to vegetative communities at present; future applications will involve individual species' distributions as potential predictors, and thus begin to take into account species' interactions as well. In many senses, all of the approaches used in distributional modeling approximate a niche, as they are defined in multidimensional ecological space; however, many algorithms intermingle modeling of the fundamental and realized niches, and therefore stray from a simple list of environmental factors defining a species' presence or absence.

Several approaches have been used to approximate species' fundamental ecological niches. The very simplest is BIOCLIM, which involves tallying species' occurrences in categories of each environmental dimension, trimming marginal portions of distributions, and taking the niche as the conjunction of the trimmed ranges (Nix 1986). Although easy to implement and conceptually attractive, BIOCLIM suffers reduced efficacy when many environmental dimensions are included: numbers of environmental combinations simply overwhelm most sampling protocols (Peterson and Cohoon 1999).

A second class of approaches is based on logistic multiple regression, techniques aimed at predicting probability of 'yes' versus 'no' in the independent variable (Skidmore et al. 1996). This idea combines well with the concept of physiological tolerances determining species' presence along continuous climate dimensions, but does less well when categorical information (e.g., vegetation type, soil type) is also to be included. In effect, if a probability threshold is specified, logistic regression divides environmental space into two portions 'habitable' and 'uninhabitable'), an approach that may be useful under some circumstances. Recent implementations of this method have included important improvements, such as relaxation of distributional assumptions regarding distributions of errors in the regression.

Finally, the Genetic Algorithm for Rule-set Prediction (GARP) combines the methods described above with others in an iterative, artificial-intelligence-based approach (Stockwell and Noble 1991; http://biodi.sdsc.edu/). Here, individual algorithms are used to produce component 'rules' in the broader rule-set, and hence portions of the species' distribution are determined as within or without the niche based on different algorithms. As such, GARP represents a superset of other approaches, and should always have greater predictive ability than any one of them. Extensive testing of GARP has indicated excellent predictive ability (Ball et al., submitted) and insensitivity to BIOCLIM's problems with environmental data density (Peterson and Cohoon 1999).

Methods

Bird species used for this study were taken from the US Endangered Species List. Excluded from the study were species that do not breed in the United States (e.g., *Grus americana*), taxa found only in Alaska (e.g., *Branta canadensis leucopareia*), Hawaiian endemics (e.g., *Loxops coccineus coccineus*), endemics to San Clemente Island (*Amphispiza belli clementeae*, *Lanius ludovicianus mearnsi*), and extinct (*Numenius borealis, Campephilus principalis, Vermivora bachmanii*) and extirpated species (e.g., *Rhynchopsitta pachyrhyncha*). Two taxa on the list, *Sterna antillarum* and *S. a. browni*, the latter of which forms part of the former, were treated as a single unit.

Occurrence data were collected from the US Breeding Bird Survey (BBS) (http://www.mbr.nbs.gov/bbs/bbs.html), the scientific literature [Ridgway 1901–1919; Grinnell and Miller 1944; Peters 1934–1986; Bent 1922–1968; the Birds of North America series (Poole et al. 1992)], and The Species Analyst (http://chipotle.nhm.ukans.edu/nabin). Textual locality references were georeferenced to the nearest minute of latitude and longitude using the Geographic Names Information System of the US Geological Survey (http://mapping.usgs.gov/www/gnis).

Geographic distributional predictions were developed based on the GARP algorithm implemented in the Biodiversity Species Workshop (BSW) facility (http://biodi.sdsc.edu) developed by David Stockwell. The GARP algorithm works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection: first, a method is chosen from a set of possibilities (logistic regression, BIOCLIM, etc.), applied to the data, and a rule developed (logistic regression rules were taken as present when probability of presence exceeded 75%). Predictive accuracy is evaluated based on 1250 points resampled from the occurence point data and 1250 points sampled randomly from the study region as a whole (the lower 48 United States), and is calculated as the sum of points actually present predicted as present and actually absent predicted as absent, divided by 2500 (Stockwell and Noble 1991). Change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the model. The algorithm runs 1000 iterations, or until convergence. GARP thus combines the strengths of a number of algorithms, and should always have higher predictive efficiency than any of the individual algorithms, and should always have higher predictive efficiency than any of the individual algorithms (Stockwell and Noble 1991). Complete details and documentation of the algorithm are available at the facility website.

Geographic themes consisted of eight coverages drawn from the Global Ecosystems Database Project (http://www.ngdc.noaa.gov/seg/fliers/se- 2006.html). Themes used (and numbers of categories included) were as follows: Life Zones (6), Soil Class (7), Annual Temperature (8), Annual Precipitation (5), Vegetation Class (8), Vegetation Types (4), Wetlands (2), and World Ecosystems (12). Each coverage was generalized to a pixel resolution of 50 km. This set of base coverages was the subject of a previous intensive study of the effects of geographic data density on GARP's predictive efficiency (Peterson and Cohoon 1999), with the result that 4–5 of these coverages was sufficient to attain maximal predictive efficiency.

GARP results were exported from BSW as ASCII raster coverages, and imported into ArcView for further manipulation. Given that species' ecological niches are generally larger than actual distributions, predicted distributions (Peterson et al. 1999) were restricted to biogeographic regions from which particular species are known. The distribution of *Haliaeetus leucocephalus* was adjusted to include Florida and that of *Grus canadensis pulla* was counted as the two pixels that cover its limited range in Mississippi. Distributional information on protected areas within the US National Park System was obtained from *ArcAtlas* (ESRI 1998).

Results

Predicted geographic distributions of individual species (Figure 1) extended from two pixels (*Grus canadensis pulla*) to approximately half of the United States (e.g., *Haliaeetus leucocephalus*). The sum of individual species' distributions (Figure 2), representing the overall richness of endangered species, showed concentrations in Florida, along the Gulf Coast, in western and coastal California, and along the Mexican border to Texas and Arizona. Areas with the lowest richness of endangered species were in the northcentral portion of the United States.

The US Endangered Species List includes a substantial component of marginal populations of species broadly distributed elsewhere (Table 1) (Peterson, submitted). These species could potentially bias attention toward borderlands, particularly in the South, so we analyzed endangered species' distributions excluding peripheral

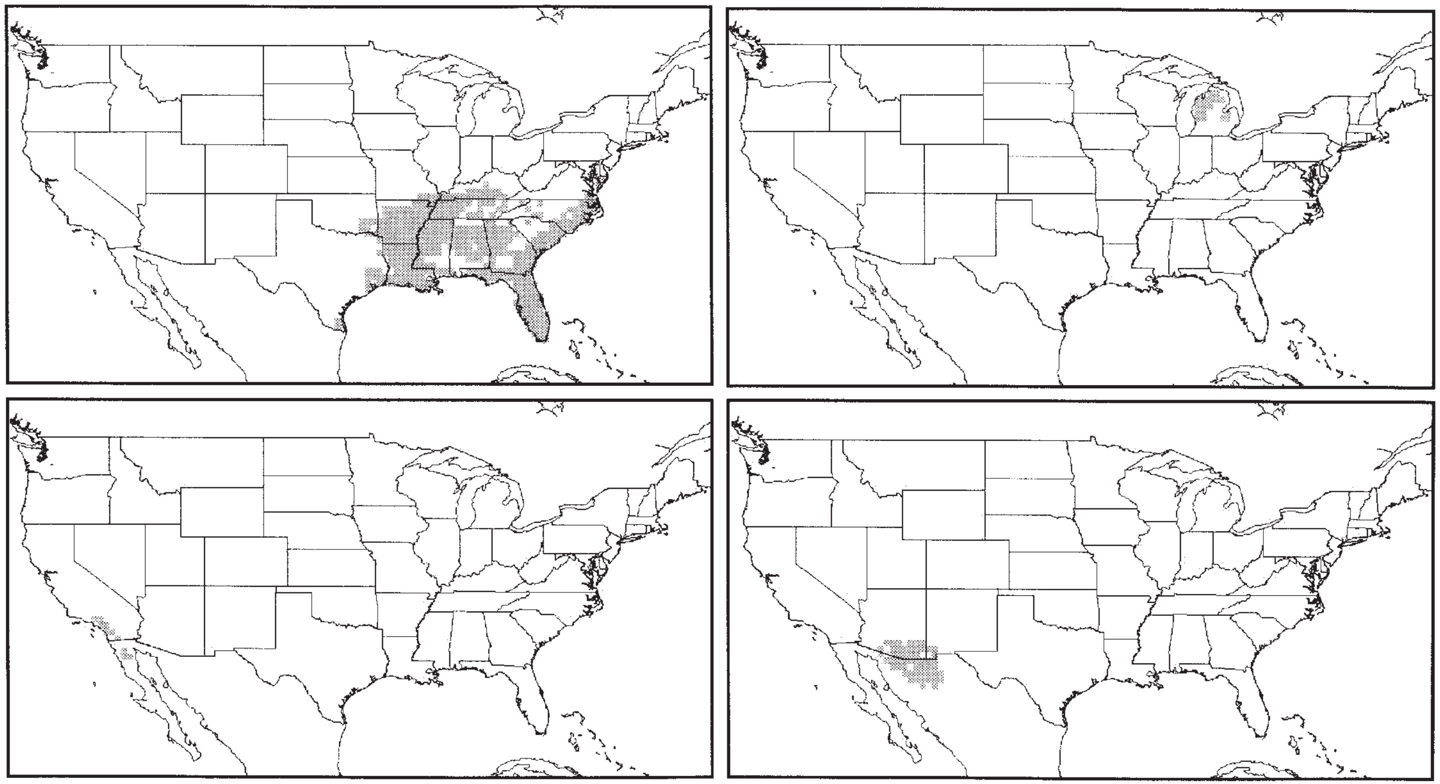

Figure 1. Predicted distributions for *Picoides borealis* (top left), *Dendroica kirtlandii* (top right), *Polioptila californica californica* (bottom left), and *Colinus virginianus ridgwayi* (bottom right).

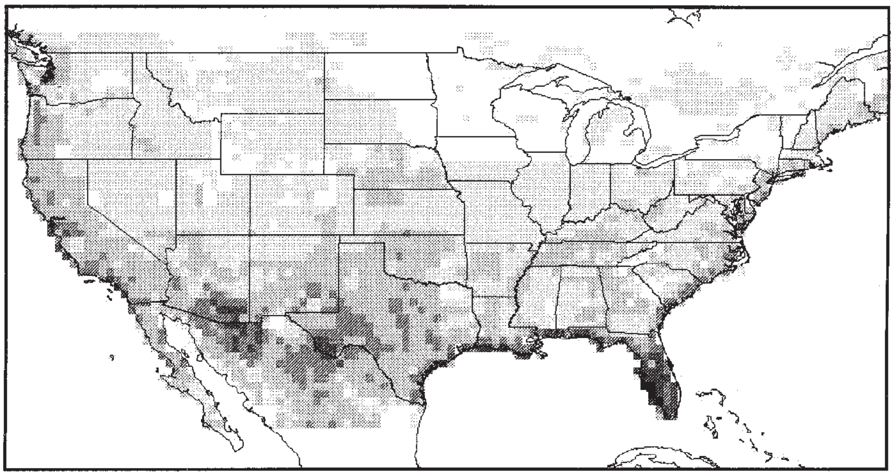

Figure 2. Overall predicted species richness for endangered bird species in the United States. White represents areas with zero species predicted, light gray one, soft gray 2–3, medium gray 4–5, dark gray 6–7, and black 8–9 species.

species (Peterson, submitted; Table 1) from consideration. Concentrations of peripheral species were in southern Arizona, southern New Mexico, southwestern Texas, along the Gulf Coast, the southern Atlantic coast, and much of Florida. Removal of these species from the analysis de-emphasized concentrations of endangered species in the southern United States, and made concentrations in Florida and California more prominent (Figure 3).

Predicted species' distributions were also compared with areas protected by the National Park Service to determine which species are currently protected by that system – a gap analysis of sorts (Figure 4). Of the 32 species being evaluated, all but nine (*Ammodramus maritimus mirabilis*, *Aphelocoma c. coerulescens*, *Dendroica chrysoparia*, *D. kirtlandii*, *Grus canadensis pulla*, *Polyborus plancus audubonii*, *Rallus longirostris obsoletus*, *Sterna dougallii dougallii*, *Tympanuchus cupido attwateri*) were afforded protection. The area with the highest concentration of unprotected endangered species was in peninsular Florida, where *Ammodramus maritimus mirabilis*, *Aphelocoma c. coerulescens*, and *Polyborus plancus audubonii* occur. Other areas holding unprotected species lay along the Gulf and Atlantic coasts, in Michigan, south-central Texas, and in central coastal California.

Discussion

Application of technology for modeling ecological niches and predicting geographic distributions is extremely useful in defining foci of species diversity and developing conservation strategies (Egbert et al. 1998; Peterson et al. 1999). These tools provide

Table 1. Birds on the US Endangered Species List, showing those included in this study, peripheral and protected status, and number of occurrence points ('sites') available to the authors for development of distributional predictions. 'Status' indicates endemic *versus* peripheral status of populations of that species in the United States; 'protected' refers to whether the species is predicted to be included within the present national parks system.

Taxa	Common name	Sites	Status	Protected
Ammodramus maritimus mirabilis	Cape sable seaside sparrow	7	Endemic or nearly so	No
Ammodramus savannarum floridanus	Florida grasshopper sparrow	4	Endemic or nearly so	Yes
Aphelocoma coerulescens coerulescens	Florida scrub-jay	3	Endemic or nearly so	No
Brachyramphus marmoratus marmoratus	Marbled murrelet	5	Endemic or nearly so	Yes
Charadrius alexandrinus nivosus	Western snowy plover	13	Substantial	Yes
Charadrius melodus	Piping plover	5	Substantial	Yes
Colinus virginianus ridgwayi	Masked bobwhite	8	Peripheral	Yes
Dendroica chrysoparia	Golden-cheeked warbler	7	Endemic or nearly so	No
Dendroica kirtlandii	Kirtland's warbler	3	Endemic or nearly so	No
Empidonax traillii extimus	Southwestern willow flycatcher	3	Endemic or nearly so	Yes
Falco femoralis septentrionalis	Northern aplomado falcon	10	Peripheral	Yes
Falco peregrinus anatum	American peregrine falcon	71	Substantial	Yes
Glaucidium brasilianum cactorum	Cactus Ferruginous pygmy-owl	4	Peripheral	Yes
Grus canadensis pulla	Mississippi sandhill crane	1	Endemic or nearly so	No
Haliaeetus leucocephalus	Bald eagle	46	Substantial	Yes
Mycteria americana	Wood stork	107	Peripheral	Yes
Pelecanus occidentalis	Brown pelican	47	Peripheral	Yes
Picoides borealis	Red-cockaded woodpecker	8	Endemic or nearly so	Yes
Pipilo crissalis eremophilus	Inyo California brown towhee	9	Endemic or nearly so	Yes
Polioptila californica californica	Coastal California gnatcatcher	13	Substantial	Yes
Polyborus plancus audubonii	Audubon's crested caracara	6	Peripheral	No
Rallus longirostris levipes	Light-footed clapper rail	6	Substantial	Yes
Rallus longirostris obsoletus	California clapper rail	9	Endemic or nearly so	No
Rallus longirostris yumanensis	Yuma clapper rail	6	Substantial	Yes
Rostrhamus sociabilis plumbeus	Everglade snail kite	4	Substantial	Yes
Sterna antillarum	Least tern	14	Substantial	Yes
Sterna antillarum browni	California least tern	16	Substantial	Yes
Sterna dougallii dougallii	Roseate tern	17	Substantial	No
Strix occidentalis caurina	Northern spotted owl	12	Substantial	Yes
Strix occidentalis lucida	Mexican spotted owl	3	Endemic or nearly so	Yes
Tympanuchus cupido attwateri	Attwater's greater prairie chicken	4	Endemic or nearly so	No
Vireo atricapillus	Black-capped vireo	11	Substantial	Yes
Vireo bellii pusillus	Least bell's vireo	8	Substantial	Yes

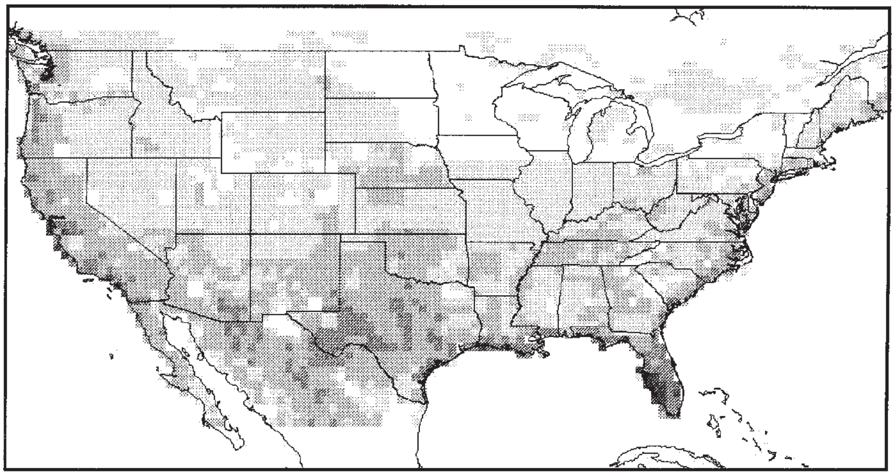

Figure 3. Endangered bird species richness in the United States with marginal species eliminated from the analysis. White represents area with zero species predicted, light gray one, soft gray 2–3, medium gray 4–5, and dark gray 6–7 species.

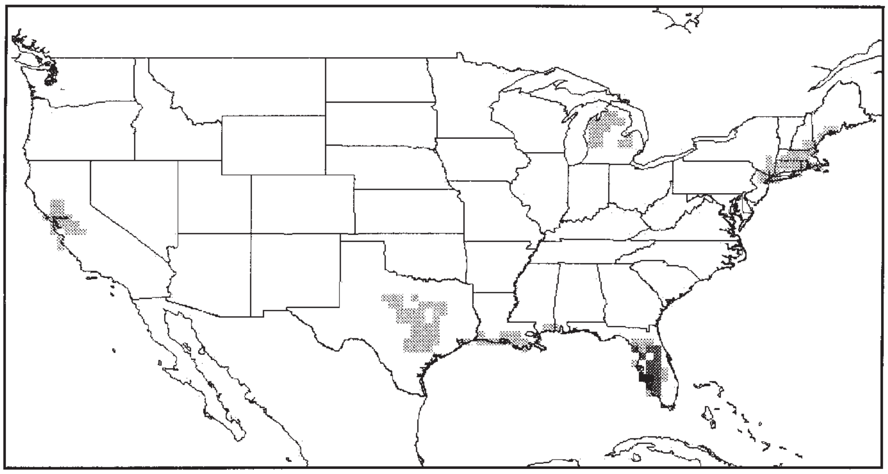

Figure 4. Richness of endangered bird species with ranges completely excluded from the national park system. White represents areas with zero species predicted, light gray one, medium gray two, and black three species.

a practical framework; permit investigators to evaluate multiple, complex factors; and generate visual images that are readily analyzed.

Clearly, the results of such studies are only as good as the inferential algorithms on which they are based. Although GARP is a relatively recent development (Stockwell and Noble 1991), tests carried out to date indicate robustness to variation in geographic data density (Peterson and Cohoon 1999), as well as excellent predictive ability based on independent test data sets (Ball et al., submitted). Clearly, conservation applications should not rely upon such tools alone: field corroboration of computer-generated predictions is critical. These predictive approaches can serve to highlight broad-scale trends, and to focus later field studies on areas of particular promise.

Although much of the United States holds distributional areas for endangered species, this study confirms the results of an earlier study (Dobson et al. 1997) that showed clear concentrations of endangered species along the southern coastal regions and in the southweastern border states. The latter concentration results in no small part from the overemphasis of the US Endangered Species List on species with ranges that are only marginal in the United States (Peterson, submitted). The importance of preserving the outer perimeter of a species range is at best questionable, many of these species (e.g., *Polyborus plancus*) are common in the remainder of their geographic distributions. Inclusion of peripheral species in preservation efforts may give undue attention to areas that are otherwise not critical. Eliminating marginal species in this study, the southwestern borderlands were significantly deemphasized, likely providing a more realistic view of endangered species concentrations.

Another approach to prioritization is to identify concentrations of endangered species that are not protected. These species should probably be placed first in conservation efforts because they are not federally protected. Under this view, peninsular Florida held the highest species density, which could represent a priority zone for endangered species conservation.

In this study, we have used endangered bird species as indicators for determining critical areas for endangered species preservation. Extension of these analyses to other taxa and to finer scales of geographic resolution would be particularly informative. From such a study, a more accurate picture of the United States' endangered species 'hot spots' (Dobson et al. 1997) could be determined, and a better prioritization for action developed. This approach would produce even more useful and novel results if applied to regions and faunas not as well known as those of the United States.

Acknowledgements

We thank David Stockwell for assistance with technical aspects of ecological niche modeling and Bruce Peterjohn of the US Breeding Bird Survey for providing data. Additional data were drawn from the distributed data network entitled The Species Analyst, a project of the North American Biodiversity Information Network, including data provided by the University of Kansas Natural History Museum, and the Museum of Vertebrate Zoology, University of California, Berkeley. Financial support was provided by a REU supplement to a grant from the National Science Foundation.

References

Ball LG, Peterson AT and Cohoon KC (submitted) Predicting distributions of tropical birds. Ibis.
Bent AC (1922–1968) Life Histories of North American Birds. US National Museum Bulletin
Dobson AP, Rodríguez JP, Roberts WM and Wilcove DS (1997) Geographic distribution of endangered species in the United States. Science 275: 550–553
Dunn CP, Bowles ML, Rabb GB and Jarantoski KS (1997) Endangered species 'hot spots'. Science 276: 513–515
Egbert SL, Peterson AT, Sánchez-Cordero V and Price KP (1998) Modeling conservation priorities in Veracruz, Mexico. In: Morain S (ed) GIS in Natural Resource Management: Balancing the Technical-Political Equation, pp 55–63. High Mountain Press, Santa Fe
Ehrenfeld D, Noss RF and Meffe GK (1997) Endangered species 'hot spots'. Science 276: 515–516
ESRI (1998) ArcAtlas. Environmental Systems Research Institute, Redlands, California
Friedman H and Ridgway R (1941–1946) The Birds of North and Middle America. US National Museum Bulletin.
Grinnell J and Miller AH (1944) Distribution of the Birds of California. Hollywood, California: Cooper Ornithological Club
MacArthur RA (1972) Geographical Ecology. Princeton University Press, Princeton, New Jersey
Nix HA (1986) A biogeographic analysis of Australian elapid snakes. In: Longmore R (ed) Atlas of Elapid Snakes of Australia, Australian Government Publishing Service, Canberra
Peters JL (1934–1986) Check-List of Birds of the World. Vol. 1–14. Museum of Comparative Zoology, Cambridge, Massachusetts
Peterson AT (submitted) Peripheral populations and the U.S. Endangered Species List. Endangered Species Bulletin
Peterson AT and Cohoon KP (1999) Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modeling 117: 159–164
Peterson AT, Egbert SL, Sánchez-Cordero V and Price KP (in press) Geographic analysis of conservation priorities using distributional modelling and complementarity: Endemic birds and mammals in Veracruz, Mexico. Biological Conservation
Petit DR and Schmidt PR (1998) Bird conservation: More than just chasing the tails. Endangered Species Bulletin 23: 4–5
Poole AF, Stettenheim P and Gill FB (1992) The Birds of North America. Birds of North America, Inc., Philadelphia, Pennsylvania
Ridgway R (1901–1919) The Birds of North and Middle America. US National Museum Bulletin
Skidmore AK, Gauld A and Walker P (1996) Classification of kangaroo distribution using three GIS models. International Journal of Geographical Information Systems 10: 441–454
Soberón J (1999) Linking biodiversity information sources. Trends in Ecology and Evolution 14: 291
Stockwell DRB and Noble IR (1991) Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. Mathematics and Computer Simulation 32: 249–254