



ELSEVIER

Ecological Indicators 3 (2003) 213–221

ECOLOGICAL
INDICATORS

This article is also available online at:
www.elsevier.com/locate/ecolind

Comparison of resolution of methods used in mapping biodiversity patterns from point-occurrence data

David Stockwell^{a,*}, A. Townsend Peterson^b

^a San Diego Supercomputer Center, University of California, 9500 Gilman Drive, La Jolla, CA 92093, USA

^b Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, KS 66045, USA

Accepted 15 April 2003

Abstract

This paper examines three methods of mapping of biodiversity using point-occurrence data for the birds of Mexico: aggregation of species occurrence records, vegetation surrogate, and individual species models. We compare the approaches from the perspective of achieving potential gains in spatial resolution with existing data. We found that mapping the diversity of Mexican birds using individual species models yielded results 400-fold more finely resolved, quantifiable errors, and greater flexibility for many applications. We show that the aggregation and surrogate methods are susceptible to tradeoffs between bias and resolution that can only be ameliorated through more intensive sampling. A theoretical error model and an empirical demonstration shows that higher spatial resolution in the individual species approach can be achieved by controlling the modeling approach by reducing bias and decreasing random error. The method is particularly applicable for large-scale biodiversity mapping, where intensive ground survey data are lacking.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Map; Biodiversity; Accuracy; Bias; Prediction; Scale; Species richness; Species composition; Birds; Mexico

1. Introduction

Loss of biodiversity is a crisis (Wilson, 1988), making efficient access to and analysis of high-quality information crucial. Information management systems such as geographic information systems (GIS) are widely used for mapping conservation efforts (Aspinall and Matthews, 1994; Miller, 1994). Using a GIS framework, maps of species distributions can be developed in a number of ways: aggregation (Peterson et al., 1998) surrogates (Gaston and Blackburn, 1995; Scott et al., 1996a), and increasingly, multivariate modeling (Nix, 1986; Austin et al., 1990; Walker and Cocks,

1991; Ferrier and Watson, 1996; Stockwell, 1999; Stockwell and Peters, 1999). The obvious extension to modeling individual species' distributions is to combine them to create maps of richness, endemism, and species composition (Stockwell and Peters, 1993).

The data on the locations of species used to map biodiversity are relatively scarce due to the vastness and inaccessibility of some areas and the often secretive nature of species. Museum collections data are increasingly appreciated as source of data largely due to internet accessibility, availability in poorly-known regions, and historical information is easily incorporated into analyses (Alberch, 1993; Peterson et al., 1998). But unlike data from a well designed survey, museum data suffers from the common problem of all ad hoc data collections of uneven sampling effort, or what is more generally known as a form of bias

* Corresponding author. Tel.: +1-619-8220942;

fax: +1-619-5345113.

E-mail address: davids@spsc.edu (D. Stockwell).

(Stockwell and Peterson, 2001). In this paper we evaluate the relationship of resolution and bias in three alternative approaches to obtaining maps of species richness (SR) using museum collections data. The conclusions of this study on SR should apply more generally in developing other forms of biodiversity information such as maps of endemism or ‘hotspots’, where individual species models serve as the elements from which the synthetic research products on patterns of biodiversity are derived.

2. Background

The following definitions of the main terms apply. The resolution of a map is the size of the pixel (or grid cell). The bias is the divergence in the data on frequency of occurrence of species across the map relative to the actual frequency of occurrence of the species. The accuracy (in predicting SR) at a grid cell, is the percentage of species predicted correctly, and average accuracy is the average of accuracy in those cells.

To address limitations of the different methods, consider the use of exhaustive ground-based surveys for mapping species richness. To provide precise estimates, the standard method uses patterns of accumulation of species to identify the point at which additional surveys would yield few or no new species (Soberon and Llorente, 1993; Colwell and Coddington, 1994). An alternative, potentially less intensive, approach is that of use of available information which can yield comparable samples under assumptions of equal accessibility and detectability of individuals (Peterson and Slade, 1998). While less-than-complete occurrence data sets can be aggregated across grids at coarser spatial scales to improve per-pixel sampling completeness (i.e. aggregation methods) this alters spatially dependent measures such as endemism (Stoms, 1994; Peterson and Watson, 1998).

Because of the labor-intensive nature of most field surveys, however, such information generally falls short of providing the fine resolution and comprehensive distributional information required for answering critical questions such as community composition (Lande, 1996) critical scales in fragmented landscapes (Keitt et al., 1997) and reserve system planning (Franklin, 1993). As a result, ecotypes or

vegetation types have been used frequently as “environmental surrogates” for species occurrences (Stoms and Estes, 1993) and indeed have figured importantly in US federal programs for biodiversity conservation prioritization (Scott et al., 1996b).

The above discussion highlights that methods that make maximally efficient use of existing data are badly needed. A comparative study of surrogates, and multivariate logistic regression genetic algorithm models showed considerable efficiency for simple surrogate models in the use of occurrence data (Stockwell and Townsend Peterson, 2002). However, extant data sets contain biases that make inclusion of an inferential step greatly preferable, e.g. the concentration of specimen samples along routes of access, such as roads and trails (Ponder, 1992). Multivariate models have the potential to reduce bias because sampling is often unbiased or less biased in ecological dimensions than in geographic dimensions (Stockwell and Peterson, 2001). We explore the issue of bias on maps of SR based on multivariate models of ecological niches on a species-by-species basis (i.e. individual species modeling), and examine the implications for resolution achievable by mapping using the three methods.

3. Methods

The occurrence data were drawn from the Atlas of the Distribution of Mexican Birds: approximately 300,000 records of bird occurrences—the Mexican holdings of 43 scientific institutions (Peterson et al., 1998). The environmental data consisted of electronic versions of maps made available by the Comisión Nacional para el Uso y Conocimiento de la Biodiversidad (CONABIO): potential vegetation (10 classes), actual percentage forest cover, elevation (meters above sea level), average annual precipitation, average annual temperature, latitude, longitude, and distance from coast. These variables were originally provided with a grid cell size of 1' × 1' of latitude and longitude, but were reprocessed to combine nine adjacent cells into 3' × 3' grid cells to obtain smaller data layers for faster calculations. In preprocessing, values of the finer grid cells were averaged for continuous variables (elevation, precipitation, rainfall) while the dominant variable was used for categorical variables (e.g. vegetation). Inclusion of spatial variables (latitude,

longitude, distance from the coast) is to our knowledge a novel approach to incorporating spatial constraints into the framework of multivariate modeling.

The three methods of mapping biodiversity patterns were aggregation, vegetation surrogates, and individual species mapping (or averaging). In aggregation, a coarse grid is overlain on the map and occurrence of a species in a grid cell was counted to estimate SR. Small grid cells generally underestimate SR owing to undersampling; maps developed at appropriate scales, however, can provide accurate estimates of SR. Using the Mexican data, a scale of analysis that appears appropriate is $1^\circ \times 1^\circ$ (Peterson et al., 1998).

A second approach to mapping regional patterns of richness is to use surrogates of biodiversity, such as vegetation type (Stoms and Estes, 1993). Here, SR is calculated for each vegetation type. In the present example, we calculated a surrogate by counting numbers of species occurring in each of the 10 vegetation types, based on overlay of museum records on the vegetation map. Due to the relatively few number of types, and the large geographic area they consequently covered, the species composition of each vegetation type was well represented by the occurrence points.

For the individual species mapping approach, predictive models were developed from the museum occurrence data using the GARP modeling system (Stockwell and Noble, 1992; Stockwell and Peters, 1999; Stockwell, 1999). This system uses statistical sampling of the species occurrence and environmental data, and artificial intelligence modeling methods, to develop predictive models that consist of a set of rules that describe the species distribution in ecological space, which can then be projected onto landscapes to predict geographic distributions. The rules relate species' occurrence and environmental variables via diverse algorithms, including environmental ranges, logistic regression, and categorical rules. The rules are produced via a genetic algorithm through a random process of generation, testing, and selecting successful rules on independently resampled data sets. The program is designed to maximize significance and accuracy measures of rules while minimizing overfitting of the sample data.

GARP was used to predict distributions of all 1044 species of birds in the Mexican database as images with values representing (0) absence or presence (254) across the country. To develop SR maps, we

averaged the values for each $3' \times 3'$ pixel in the images. Normalization of the range of values resulted in an image with intensities ranging from 1 (no species) to 254 (all possible species).

4. Results and discussion

The three approaches used to map SR produced very different views of biodiversity patterns across Mexico (Figs. 1–3). The aggregation method showed variation in intensity at a coarse scale, and appears to intermix intensity of collection with true variation in species richness (Fig. 1). Indeed, the richest grid cells (e.g. corresponding to Chilpancingo, Guerrero) coincide with regions with good access for collectors over the past two centuries. Consequently, only the broadest patterns can be reliably extracted from the data via this approach, such as the trend towards increasing richness in the tropical lowland regions. Reduction of the grid cell size in an attempt to provide more detailed patterns would exacerbate the biases in the richness map beyond utility.

Using vegetation types as a surrogate for SR patterns in Mexico produced results that were even more difficult to interpret (Fig. 2). This map shows close correspondence, as expected, between SR and vegetation patterns, with homogeneous richness within polygons and abrupt changes between vegetation types. Hence, resolution was reduced even below that provided by the aggregation approach.

The third approach, based on averaging predicted distributions for all 1044 species (Figs. 3 and 4) produced a map resolved to $3' \times 3'$, 400 times smaller than the $1^\circ \times 1^\circ$ cells in the aggregation method, and far more detailed than those in the vegetation map. Fine-scale patterns visible in the averaging approach allow visualization of local SR patterns (Fig. 4). Indeed this map reflects variation in SR not reflected in the vegetation-based map: areas within the same vegetation type could be relatively rich or poor.

4.1. General discussion

4.1.1. Accuracy and error properties of averaging individual models

To examine error related to bias inherent in the three approaches, we explored a simple theoretical error

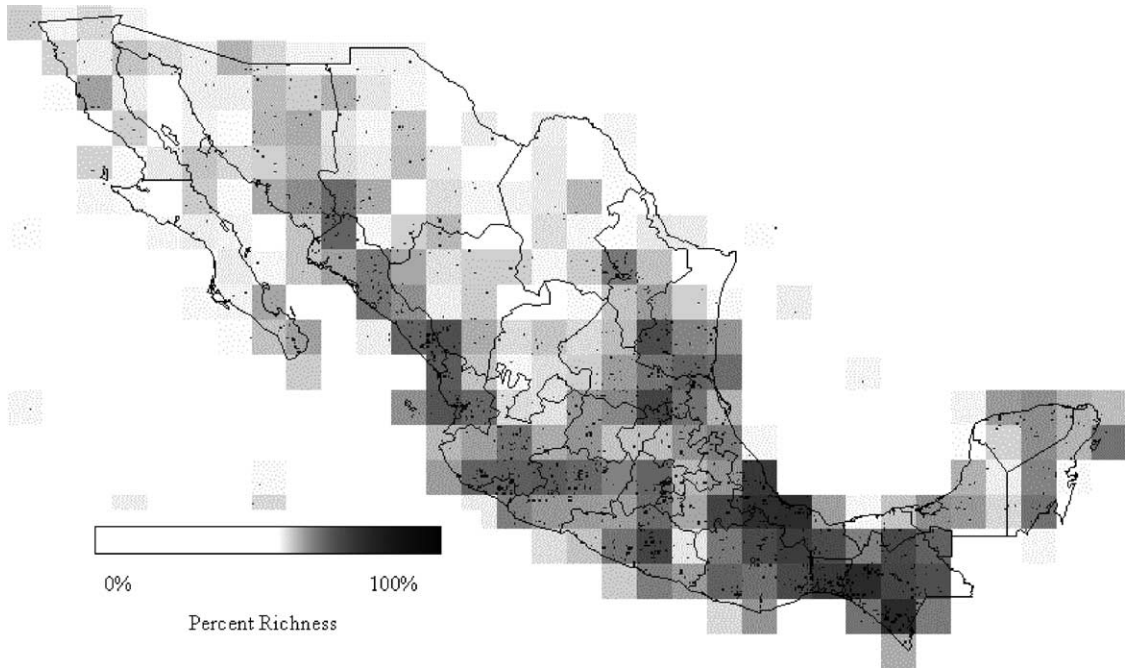


Fig. 1. Species richness of birds across Mexico based on the aggregation method, with richness tallied from known occurrences in $1^\circ \times 1^\circ$ grid cells. The lines represent state boundaries.

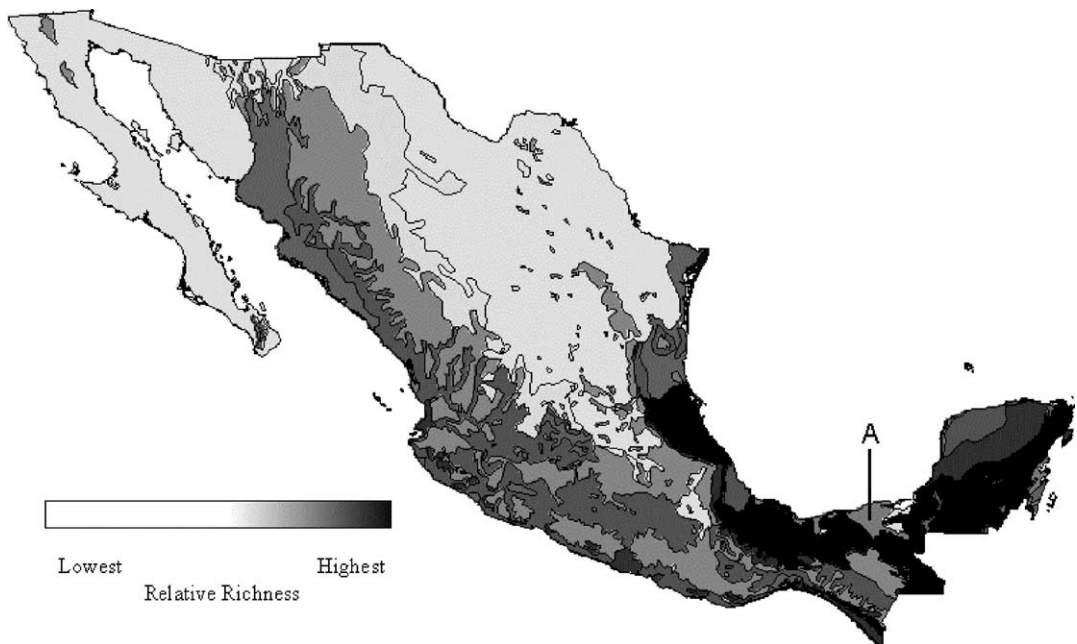


Fig. 2. Species richness of birds across Mexico based on a vegetation surrogate, with richness tallied for each vegetation type. The lines represent boundaries to vegetation types.

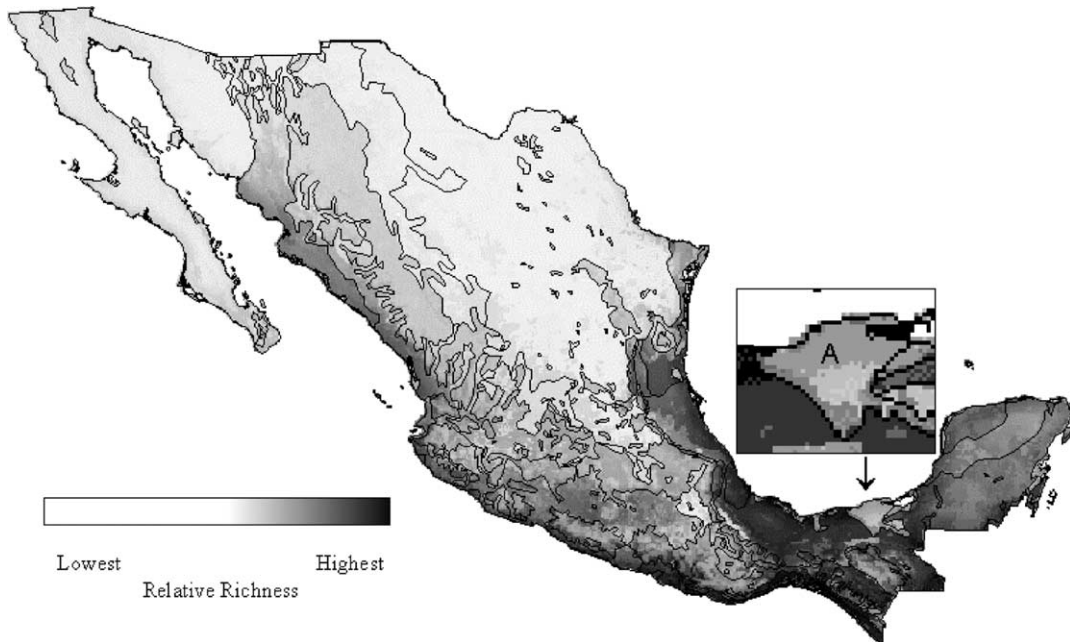


Fig. 3. Species richness of birds across Mexico based on averaging individual species maps developed using multivariate models—whole country. The lines represent boundaries to vegetation types.

model, in which prediction error is assumed to spring from random and systematic sources. Random error stems from mistakes in predicting presence of a particular species in a particular grid cell. Systematic error, in contrast, affects probabilities of predicting overall species numbers at particular locations.

The potential error, e_i , in predicting occurrence of the i th species in a particular grid cell can be decomposed into two components: systematic error or bias a_i and random error b_i as $e_i = a_i + b_i$ where a_i and b_i is a probability. The expected error over n species can be evaluating by averaging e_i across i from 1 to n . Assuming that random error is equally distributed across species, averaging maps across species makes each error component at a particular grid cell converge on its expected value, $E(a)$ and $E(b)$. The random component $E(b)$ has a mean of zero whereas bias $E(a)$ has a nonzero value at a biased grid cell (by definition). Hence, as n becomes large:

$$\sum_{i=0}^n \frac{e_i}{n} = \sum_{i=0}^n \frac{a_i}{n} + \sum_{i=0}^n \frac{b_i}{n}$$

$$\Rightarrow E(a) + E(b) = E(a) \neq 0$$

As the above shows, averaging maps over a number of species diminishes random error, but maintains systematic biases such as consistent over- or under-prediction of species richness in certain areas.

How are these errors manifested in SR maps? The aggregation map would show strong bias towards greater richness in 1° grid cells surveyed most intensively. That is, patterns of species richness under the aggregation approach are likely to reflect the pattern of survey intensity, with poorly sampled areas showing reduced richness. This effect is clearly seen in tropical areas of southeastern Mexico, where SR is known to be high (Escalante-Pliego et al., 1993), where the aggregated SR map shows uneven, and at times low richness in some cells (Fig. 1). In practice the ultimate resolution may be limited by the tolerance for bias error due to the paucity of sampling in these regions.

Maps developed using the surrogate approach would show error resulting from bias in intensity of sampling of particular vegetation types, which would be manifested as zones of rapid change in SR among adjacent vegetation types, whereas in reality changes

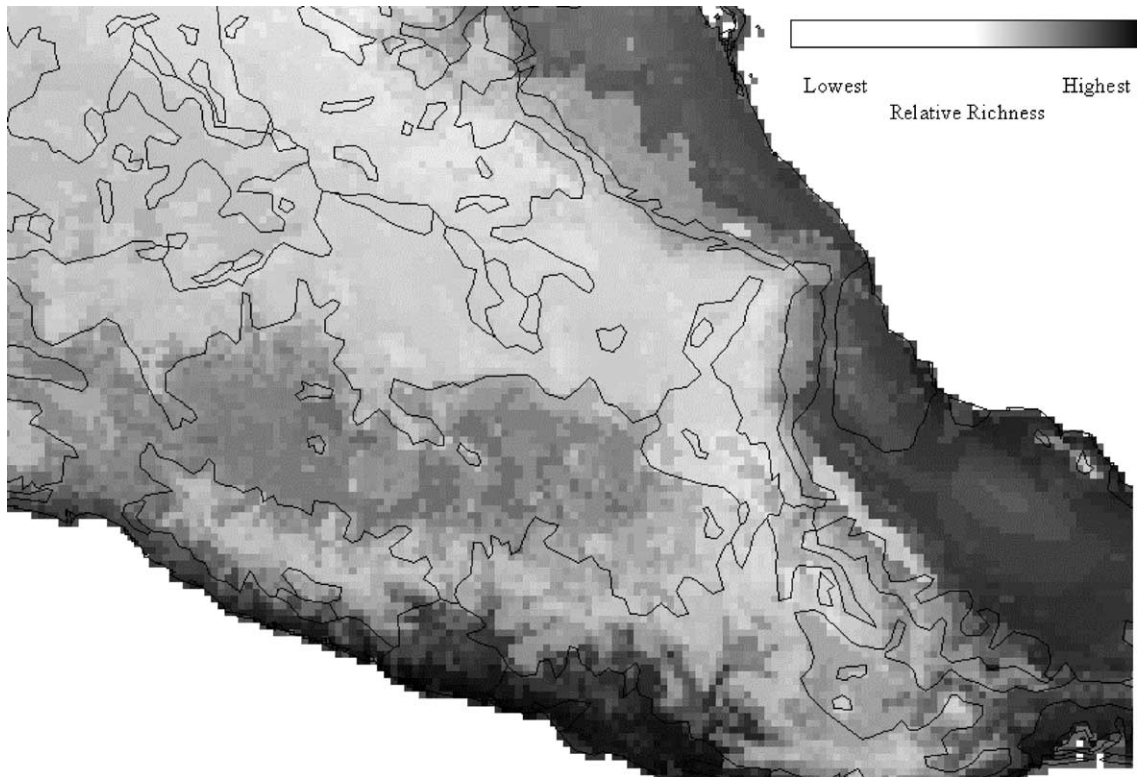


Fig. 4. Species richness of birds across Mexico based on averaging individual species models—detail of central Mexico. The lines represent boundaries to vegetation types.

are gradual. Indeed, area A in Fig. 2 is an area of lowland tropical aquatic vegetation with predicted SR strikingly lower than surrounding areas. In actuality, these areas have SR comparable to surrounding areas (Escalante-Pliego et al., 1993), but have not been sampled as intensively as other areas owing to difficulty of access: paucity of data for the area produces low SR for the vegetation type. In practice the ultimate resolution may be limited by the tolerance for bias error due to the paucity of sampling in these vegetation types.

Maps developed based on averages of individual species clearly show much finer resolution of patterns than those resulting from other approaches, generally without sharp differences between adjacent areas, due to the smoothly varying fine-scale independent variables used in the models. However, the models may be biased to the extent that individual variables correlate with sampling intensity. That is, if sampling

of vegetation types is biased, it can contribute to bias in the overall model. On the other hand, bias can be reduced by using other variables for which survey bias is negligible or by omitting the biasing variable (Stockwell and Peterson, 2001).

To demonstrate this possibility, we developed another SR map based on averaging individual species models built without the vegetation variable. This new map was then subtracted from the all-variable richness map (Fig. 3) to produce a difference map (Fig. 5) which identifies areas in which vegetation was particularly influential in predicting the presence or absence of species. Table 1 lists the species richness by vegetation type for the predicted richness with (Fig. 3) and without vegetation (Fig. 5). Here again, area A is an anomaly: it has a predicted SR of 22.3% of Mexican bird species in the all-variable analyses, but 42.3% when vegetation is excluded, suggesting that the low SR originally predicted for

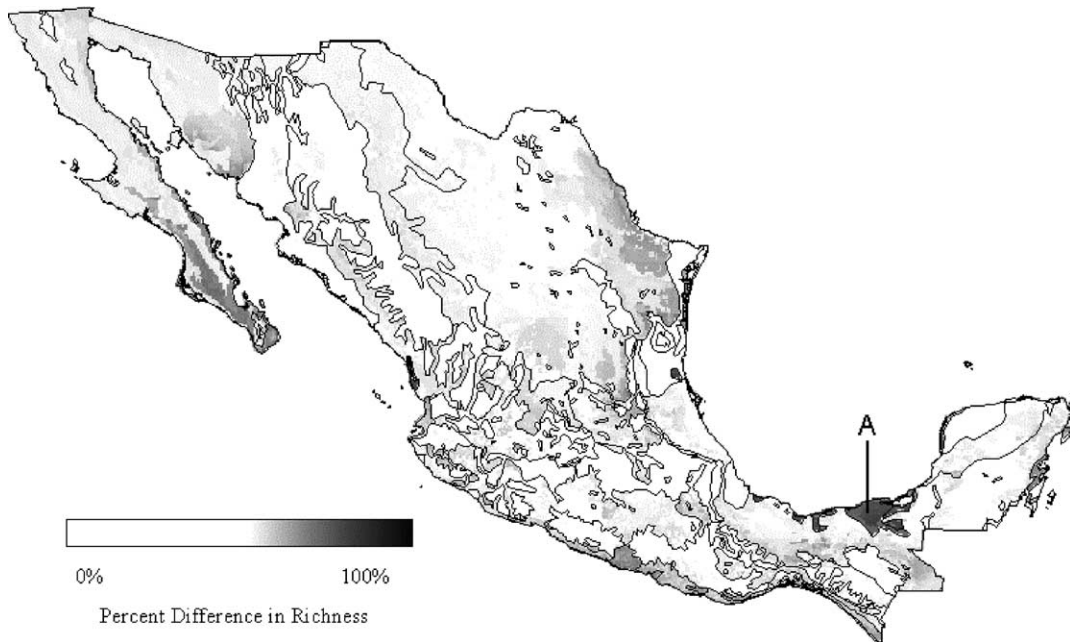


Fig. 5. Difference between the richness maps developed with multivariate models using all variables (Fig. 3) and with the vegetation variable removed from analyses.

this vegetation type represents accumulation of bias error in species distribution maps. Hence, omission of potentially biased variables may represent a useful exploratory tool for such analyses (Peterson and Cohoon, 1999; Stockwell and Peterson, 2001).

4.1.2. Comparison of methods

The goal of each of these methods is the production of fine resolution maps of species richness with

widespread coverage, without the need for costly new surveys of the birds of Mexico. Of the three methods: aggregation of data, use of surrogate variable, and averaging maps for individual species developed via multivariate modeling, the latter approach clearly outperformed the other two in detecting patterns at much finer resolutions using the existing data.

Mapping at finest scales possible is very important for biodiversity applications. As spatial resolution of

Table 1

Average difference in richness between aggregation and averaging maps by vegetation area, showing large difference in predicted richness in aquatic and sub-aquatic vegetation (area A)

Vegetation type	Land area (%)	SR % with all variables (Fig. 3)	SR % without vegetation	SR % difference (Fig. 5)
Coastal scrubland	1.4	40.2	38.9	1.3
Oak forests	18.7	20.3	16.5	3.8
Tropical thorn low forests	5.9	26.6	22.3	4.3
Cloud forest	0.86	34.8	33.9	0.9
Deciduous tropical forest	13.5	28.9	30.0	-1.1
Evergreen tropical forest	9.4	42.2	43.5	-1.3
Semi-deciduous tropical forest	2.8	34.4	39.4	-5.0
Arid and semi-arid scrublands	38.2	5.5	8.4	-2.9
Grassland	8.1	6.3	8.7	-2.4
Aquatic and sub-aquatic vegetation	1.2	22.3	42.3	-20

occurrence points is now possible using global positioning systems on the order of 1–10 m, predictions for individual species can be produced and averaged at scales finer than 100 m. Aggregating species occurrence data to grid cells, commonly done using cell sizes of $0.5' \times 0.5'$ or more (e.g. Peterson et al., 1998) involves considerable loss of resolution, explanatory capacity and information. Mapping biodiversity patterns using vegetation as a surrogate variable yielded poor resolution mapping in Mexico due to the coarse classification scheme in the available vegetation map (10 classes); however while a map with more classes would allow finer resolution, problems of bias with insufficient sampling across vegetation types would result in lower accuracy (Stockwell and Peterson, 2001).

The theoretical error model developed above predicts that the principal source of error in mapping SR is bias caused by the systematic under sampling of species in particular environments. All three methods were shown to be affected by bias: if a particular area or vegetation type is poorly sampled, then SR is underestimated. The only recourse of the aggregation and surrogate methods is to use more data (not possible under the assumption of using available data) or to increase the numbers of data in each cell or type by coarsening the resolution of the cells or the vegetation types to increase effective numbers of data points. Multivariate model generated maps inherit the finer resolution from the availability of finer-scale resolution environmental variables. As we have shown, like aggregation and surrogate methods, they are also susceptible to bias. However, unlike these methods, the strategy of selection of environmental variables is available to mitigate the effects of bias.

Does the averaging method suffer from the same tradeoff of resolution versus accuracy, effectively limiting the resolution of the approach? The principal limitations in resolution for this method is the accuracy inherent in the positioning of the occurrence and environmental data sets, and the fact that sampling is disproportionate across the region (bias). The limits of inherent accuracy are generally in the realm of 1 km worldwide, are usually below that required for most applications, although finer-scale applications may require more recent data with better spatial resolution. As shown by the theoretical model, the inherent in-

accuracy due to errors in the individual species maps diminishes to zero as the number of species averaged increases. As shown by the empirical demonstration, the bias can be reduced with appropriate choice of independent variables. Therefore, to the extent that predictions of species' distributions can be developed using variables that are not correlated with the sampling bias in the data, and the number of individual species models is sufficient to reduce the random error to a small value, such a tradeoff between resolution and accuracy should not exist in developing SR maps using the averaging approach.

This is the interesting result of the paper. Given the two forms of error, random and systematic, the random error is reduced by averaging more species, while the systematic error is reduced by removing the independent variables that exhibit bias. Therefore the obvious tradeoff that blocks the aggregation and surrogate methods from application to finer resolutions do not strictly apply to the averaging method, as there are strategies for dealing with error other than sampling more intensively.

Most importantly, from an environmental informatics perspective, the use of aggregate and surrogate methods loses critical species information potentially blocking further analyses. Hence, it is not possible to achieve more detailed studies of the composition of communities and areas finer than the resolution of the initial analyses. This limitation is particularly crippling when the objective is to design and optimize protected areas systems, yet the fundamental units on which such efforts focus (species) are out of reach. The approach used in this study to develop SR maps—mathematical operations such as averaging of maps of species distributions developed through multivariate modeling of the ecological niche—can potentially be applied to develop a wide range of synthetic views of biodiversity patterns including patterns of richness and rarity, maps of endemism or 'hotspots' and many others.

Acknowledgements

The environmental data layers for this study were made available by the Comision Nacional para el Uso y Conocimiento do la Biodiversidad (CONABIO). The Mexican Bird Distributional Atlas was developed

with the cooperation of Mexican and United States museums (listed in Peterson et al., 1998) by Adolfo Navarro. Funding for this work was provided in part by the National Science Foundation. The comments of the anonymous reviewers are gratefully acknowledged.

References

- Alberch, P., 1993. Museums, collections and biodiversity inventories. *Trends Ecol. Evol.* 8 (10), 372–374.
- Aspinall, R., Matthews, K., 1994. Climate change impact on distribution and abundance of wildlife species: an analytical approach using GIS. *Environ. Pollut.* 86, 217–223.
- Austin, M.P., Nicholls, A.O., Margules, C.R., 1990. Measurement of the realized qualitative niche: environmental niches of five eucalyptus species. *Ecol. Monogr.* 60 (2), 161–177.
- Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. Roy. Soc. London B* 335, 101–118.
- Escalante-Pliego, P., Navarro-Siguenza, A.G., Peterson, A.T., 1993. A geographic, historical, and ecological analysis of avian diversity in Mexico. In: Ramamoorthy, T.P., Bye, R., Lot, A., Fa, J. (Eds.), *The Biological Diversity of Mexico, Part 1. Pac. Coast Avif.* 29, 1–202.
- Ferrier, S., Watson, G., 1996. An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity. Consultancy report to NSW National Parks and Wildlife Service, June 1996.
- Franklin, J.F., 1993. Preserving biodiversity: species, ecosystems, or landscapes? (in forum: preserving biodiversity). *Ecol. Appl.* 3, 202–205.
- Gaston, K.J., Blackburn, T.M., 1995. Mapping biodiversity using surrogates for species richness: macro-scales and New World birds. *Proc. Roy. Soc. London B* 262, 335–341.
- Keitt, T.H., Urban, D.L., Milne, B.T., 1997. Detecting critical scales in fragmented landscapes. *Conserv. Ecol.* (online) 1 (1), 4. Available from the Internet. URL: <http://www.consecol.org/vol1/iss1/art4>.
- Lande, R., 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76, 5–13 (Copenhagen).
- Miller, R.I. (Ed.), 1994. *Mapping the Diversity of Nature*. Chapman & Hall, London, 218 pp.
- Nix, H.A., 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore, R. (Ed.), *Atlas of Elapid Snakes of Australia*. Australian Flora and Fauna Series Number 7. Australian Government Publishing Service, Canberra, pp. 4–15.
- Peterson, A.T., Cohoon, K.P., 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecol. Model.* 117, 159–164.
- Peterson, A.T., Slade, N.A., 1998. Extrapolating inventory results into biodiversity estimates and the importance of stopping rules. *Diversity Distributions* 4, 95–105.
- Peterson, A.T., Watson, D.M., 1998. Problems with areal definitions of endemism: the effects of spatial scaling. *Diversity Distributions* 4, 189–194.
- Peterson, A.T., Navarro-Siguenza, A.G., Benitez-Diez, H., 1998. The need for continued collecting a geographic analysis of Mexican bird specimens. *IBIS* 140, 288–294.
- Ponder, W.F., 1992. Bias and biodiversity. *Aust. Zool.* 28, 47–51.
- Scott, J.M., Jennings, M., Wright, R.G., Csuti, B., 1996a. Landscape approaches to mapping biodiversity. *BioScience* 46, 77–78.
- Scott, J.M., Tear, T.H., Davis, F.W. (Eds.), 1996b. *Gap Analysis: A landscape approach to biodiversity planning*. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland. ISBN-1-57083-03603.
- Soberon, J., Llorente, J., 1993. The use of species accumulation functions for the prediction of species richness. *Conserv. Biol.* 7 (3), 480–488.
- Stockwell, D.R.B., 1999. Genetic Algorithms II. In: Fielding, A.H. (Ed.), *Machine Learning Methods for Ecological Applications*. Kluwer Academic Publishers, Boston, pp. 123–144.
- Stockwell, D.R.B., Noble, I.R., 1992. Induction of sets of rules from animal distribution data: a robust and informative method of analysis. *Math. Comput. Simul.* 33, 385–390.
- Stockwell, D.R.B., Peters, D.P., 1993. ERINYES, 18:8. Newsletter of the Environmental Information Resources Network.
- Stockwell, D.R.B., Peters, D.P., 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *Int. J. Geogr. Inform. Syst.* 13 (2), 143–158.
- Stockwell, D.R.B., Peterson, A.T., 2001. Controlling bias during predictive modeling with museum data. In: Scott J.M., Heglund P.J., Morrison M., Raphael M., Hauffer J., Wall B., Samson F. (Eds.), *Predicting Species Occurrences: Issues of Scale and Accuracy*. Island Press, Covelo, CA.
- Stockwell, D.R.B., Townsend Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.* 148, 1–13.
- Stoms, D.M., 1994. Scale dependence of species richness maps. *Professional Geogr.* 46, 346–358.
- Stoms, D.M., Estes, J.E., 1993. A remote sensing research agenda for mapping and monitoring biodiversity, *Int. J. Remote Sens.* 1839–1860.
- Walker, P.A., Cocks, K.D., 1991. *Global Ecol. Biogeogr. Lett.* 1, 108–118.
- Wilson, E.O., 1988. *Biodiversity*. National Academy Press, Washington, DC.