Contributed Papers

A Successful Predictive Model of Species Richness Based on Indicator Species

RALPH MAC NALLY* AND ERICA FLEISHMAN†‡

*Australian Centre for Biodiversity: Analysis, Policy and Management, School of Biological Sciences, Monash University, Melbourne 3800 Australia

†Center for Conservation Biology, Department of Biological Sciences, Stanford University, Stanford, CA 94305–5020, U.S.A., email efleish@stanford.edu

Abstract: *Because complete species inventories are expensive and time-consuming, scientists and land managers seek techniques to alleviate logistic constraints on measuring species richness, especially over large spatial scales. We developed a method to identify indicators of species richness that is applicable to any taxonomic group or ecosystem. In an initial case study, we found that a model based on the occurrence of five indicator species explained 88% of the deviance of species richness of 56 butterflies in a mountain range in western North America. We validated model predictions and spatial transferability of the model using independent, newly collected data from another, nearby mountain range. Predicted and observed values of butterfly species richness were highly correlated with 93% of the observed values falling within the 95% credible intervals of the predictions. We used a Bayesian approach to update the initial model with both the model-building and model-validation data sets. In the updated model, the effectiveness of three of the five indicator species was similar, whereas the effectiveness of two species was reduced. The latter species had more erratic distributions in the validation data set than in the original model-building data set. This objective method for identifying indicators of species richness could substantially enhance our ability to conduct large-scale ecological assessments of any group of animals or plants in any geographic region and to make effective conservation decisions.*

Exitoso Modelo Predictivo de la Riqueza de Especies Basado en Especies Indicadoras

Resumen: *Debido a que los inventarios completos de especies son costosos y consumen tiempo, los científicos y gestores de tierras buscan técnicas para aliviar las constricciones logísticas para medir la riqueza de especies, especialmente en escalas espaciales grandes. Desarrollamos un método para identificar indicadores de riqueza de especies aplicable a cualquier grupo taxonómico o ecosistema. En un estudio inicial, encontramos que el modelo basado en la ocurrencia de cinco especies indicadoras pu du explicar el 88% de la anormalidad de la riqueza de especies de 56 mariposas en una cadena montañosa en el occidente de Norte América. Validamos las predicciones y la transferabilidad del modelo utilizando datos nuevos e independientes de otra cadena montañosa cercana. Los valores predichos y observados de la riqueza de especies de mariposas estuvieron altamente correlacionados, con 93% de los valores observados dentro de 95% de los intervalos confiables de las predicciones. Utilizamos un método Bayesiano para actualizar el modelo inicial con los conjuntos de datos tanto de construcción como de validación del modelo. En el modelo actualizado, la efectividad de tres de las cinco especies indicadoras fue similar, mientras que la efectividad de dos especies fue reducida. Esta especie tuvo distribuciones más erráticas en el conjunto de datos de validación que en el conjunto de datos originales. Este método objetivo para la identificación de indicadores de la riqueza de especies podría reforzar significativamente nuestra habilidad para llevar a cabo evaluaciones ecológicas, a gran escala, de cualquier grupo de animales o plantas en cualquier región geográfica y para tomar decisiones de conservación efectivas.*

Introduction

Identifying a limited suite of species that reflects the species richness of an entire biota has become a holy grail for ecologists, conservation biologists, and natural resource agencies (Heywood et al. 1995; Chapin et al. 2000; Mooney 2000). Whether the emphasis is on so-called hotspots, coldspots, or the full gradient of species richness values, species richness is a critical variable for biodiversity management that has been used for decision making and prioritization of conservation efforts (ECC 2000; Pimm et al. 2001; Roberts et al. 2002). It is too expensive and time consuming to measure species richness over extensive areas, however, especially for non-vascular plants and invertebrates and in tropical or marine ecosystems (Walters 1986; Pressey et al. 2000; Faith et al. 2001). As an alternative to conducting exhaustive species inventories, scientists and management practitioners hope that species richness can be predicted more cheaply and quickly through the use of "indicator" species (Pearson 1994; Scott 1998; Gustafsson 2000), defined here as species with occurrence patterns that are correlated with the species richness of a larger group of organisms. Although it now seems unlikely that indicator species from a single taxonomic group (e.g., birds) will provide information on the richness of an entire biota (all vertebrates, invertebrates, and plants) at scales meaningful for most land-use decisions (Prendergast et al. 1993; Mac Nally et al. 2002; Vessby et al. 2002), indicators still may be effective within limited taxonomic boundaries (Fleishman et al. 2000).

Effective maintenance of native biodiversity in the face of increasing human impacts requires accurate measurement of key trends and responses to disturbances and alternative management actions at landscape, regional, and continental scales (Chapin et al. 2000; Pressey et al. 2000). Developing a standardized method of measuring species richness is vital for international conservation efforts (ECC 2000; Pimm et al. 2001; Roberts et al. 2002). For example, Conservation International's recently launched Tropical Ecology, Assessment, and Monitoring (TEAM) Initiative plans to establish a network of more than 50 field stations worldwide to assess and monitor biodiversity with a standard protocol (da Fonseca et al. 2002). Similarly, the Millennium Ecosystem Assessment aims to provide decision makers with relevant data on biodiversity patterns, ecosystem processes, and underlying forces at spatial scales from local to global (Gewin 2002). Few tools are available, however, with which to accomplish these goals. Most existing biodiversity measures have been developed for much smaller spatial and temporal scales (Kareiva & Wennergren 1995). The validated method of biodiversity assessment we present here should increase our capacity to estimate large-scale biodiversity patterns and enhance our ability to conserve natural capital.

The indicator-species approach is a potentially efficient way to model species richness. Efforts to predict species distributions over extensive areas as functions of distributions of food or other habitat variables are often successful (Hanski 1999; Miller & Cale 2000), but obtaining these data frequently is more expensive and labor-intensive than direct inventories. Moreover, species-habitat models rarely can be developed for all the species in a given assemblage. If reliable indicators of species richness can be found, it may be much easier to measure the occurrence of indicators than to conduct comprehensive species inventories or habitat assessments (Gustafsson 2000). From a realistic, management-oriented perspective, it is also easier to train field biologists and other personnel to identify a limited set of species and to design monitoring plans for a few indicator species than to expect those personnel to recognize and track an entire fauna.

Indicator species are often selected according to ad hoc criteria, such as their charisma or legal protection status (Andelman & Fagan 2000). We argue that statistically based selection of potential indicators is better justified and likely to be more effective. Prediction of species richness based on the occurrence of indicator species should be regarded as a testable hypothesis. The hypothesis may be in the form of a statistical model—a function of the occurrence of indicator species—that should be confronted with new test data (Landres et al. 1998; Mac Nally et al. 2000).

We recently developed a statistical protocol to select potential indicators of species richness that are applicable over hundreds to thousands of square kilometers (Mac Nally & Fleishman 2002). These scales are typical of those over which much ecological and biogeographic research is conducted and many conservation and land-management decisions are made. Our objective at all stages of our work, and especially in this study, was to conduct a robust test of the effectiveness of our framework for identifying indicators of species richness. As an initial case study, we applied the method to butterfly assemblages in the central Great Basin of western North America. Butterflies are among the taxonomic groups most frequently suggested as indicators of species richness and ecosystem integrity (Kremen et al. 1993; New et al. 1995). Models were built from data collected in one mountain range in the late 1990s, whereas validation data specifically intended to facilitate a robust test of the model were gathered in a second, nearby mountain range (approximately 40 km away from the first) in the early 2000s. Mountain ranges are an appropriate scale for developing and testing the transferability (sensu Leftwich et al. 1997) of indicator-species models in this geographic region. Resource agencies in the Great Basin (U.S.A.) generally develop separate management plans for individual mountain ranges, yet there is little existing information on species distributions that can be used for ecological planning.

Methods

From 1996 to 1999, we used standard inventory methods to characterize assemblages of resident butterflies in 10 canyons in the Toquima Range (approximate north-south boundaries 39°17′50″ to 38°29′9″) (Fleishman et al. 2001). We divided canyons into multiple segments from base to crest (49 total). Each segment was 100 m wide and extended for approximately 100 m change in elevation. Mean segment length was >1 km, which is greater than the dispersal distances of virtually all resident butterflies (Fleishman et al. 1997).

We used well-established butterfly inventory methods described in detail by Fleishman et al. (1998). Field personnel were familiar with the regional butterfly fauna, and we restricted our inventories to weather most favorable for butterfly flight. It is reasonable to assume that a given butterfly species is absent in a given year if the area has been searched with these methods during the appropriate season and weather conditions (Pullin 1995; Reed 1996). Using these methods in the nearby Toiyabe Range, for example, we recorded 98% of the theoretical total number of resident species expected in the mountain range under a Michaelis-Menten model (Clench 1979; Raguso & Llorente-Bousquets 1990; Soberón & Llorente 1993). Thus, it is unlikely that we failed to detect more than a very few species that actually were present in a given site in a given year (Pollard & Yates 1993; Harding et al. 1995).

We recorded 56 resident species of butterflies from our study locations in the Toquima Range. We thought widespread species would not be useful for modeling variation in species richness and would thus have little potential to serve as indicators of species richness, whereas restricted species, which occur at relatively few sites, often have highly specific ecological requirements that are not shared with many other species. Therefore, we chose to consider as potential indicator species only the 22 species occurring in ≥30% and ≤70% of the 49 sites. Thus, we modeled species richness, including the potential indicator species, at each site as a function of the incidence of a set of indicator species drawn from those 22 species. Similar guidelines would be applicable to any taxonomic group.

We sought to identify one predictive model from among the 2^{22} combinations of predictor variables (i.e., all possible pairs, trios, and so forth of potential indicator species). The selected model should be the most statistically efficient, that is, the model that optimizes fitting error against model complexity (i.e., number of predictor variables) (Mac Nally 2000). A selection technique now advocated widely is Schwarz's Bayesian Information Criterion, (BIC) (Schwarz 1978), for which the model with a minimum value is sought. (For a general overview of Bayesian statistical methods and associated terminology, see Bergerud & Reed [1998] and references therein.)

The dependent variable, species richness, is likely to have a Poisson distribution because it must take relatively small, non-negative values (Crawley 1993). However, Poisson regression requires iterative, numerical procedures that are difficult to automate for millions of potential models. Therefore, we took a short-cut by modeling the logarithm of species richness against combinations of incidences of the 22 species with ordinary multiple linear regression. This is easier to automate because multiple linear regression involves matrix calculations rather than iterative fitting. The logarithmic transformation is appropriate for Poisson-distributed variables in generalized linear models (McCullagh & Nelder 1989; Cameron & Trivedi 1998).

From this preliminary screening, we identified a set of five indicator species whose incidences produced the minimum BIC among models of species richness of butterflies in the Toquima Range: *Ochlodes sylvanoides*, *Everes amyntula*, *Euphilotes ancilla*, *Speyeria zerene*, and *Coenonympha tullia*. The five indicator species encapsulate a diversity of life-history characteristics found among the resident members of their taxonomic group in the biogeographic region. This may explain why this particular suite of species is so strongly associated with variation in species richness (Mac Nally & Fleishman 2002), and these characteristics of effective indicator species in the aggregate may be generalizable among taxonomic groups. In our butterfly case study, for example, the phenologies of flight activity of the species spanned the field season. In addition, the group of indicator species included taxa with varied larval host plants. Two species feed on different genera in the family Poaceae, and one each feeds on Polygonaceae, Fabaceae, and Violaceae. In addition, other work has demonstrated that although occurrence patterns of at least four of the five indicator species can be explained largely as functions of elevation (Fleishman et al. 2001), the species have different elevational distribution limits (Fleishman et al. 1998, unpublished data).

Having identified the five indicator species, we computed model coefficients—that is, regression coefficients associated with each indicator species—with Bayesian Poisson regression (Spiegelhalter et al. 2000). The model is

$$\log(\mu_i) = \alpha_0 + \sum_{k=1}^{Q} \alpha_k \delta_{ik} + \varepsilon;$$

$$Y_i \sim \text{Poisson}(\mu_i).$$

The αs are the regression coefficients to be estimated, μ_i is the estimated mean species richness at site i given the state of the predictor variables (i.e., presence or absence of the Q indicator species, δ_{ik}), δ_{ik} indicates whether species k is present or absent at site i, ε is model error, and Y_i is the observed species richness at site i, which is assumed to be distributed (∼) as a Poisson variable with mean μ_i. The Bayesian approach to estimating parameters

formally incorporates prior knowledge of the values of a parameter and produces a probability statement about the interval within which the parameter value lies. This means that each regression coefficient has a distribution of values rather than a single value. If one has virtually no prior knowledge of the values of a parameter, then it is appropriate to use a "noninformative" prior distribution for that parameter. Typically, one might use a flat uniform distribution (between two endpoints, ± 1000 perhaps) or a normal distribution with high variance (see below). Use of noninformative priors when there is little or no prior information means that the posterior probability distributions are dictated by the newly collected data (Lee 1989). In the model-building phase, we gave the αs noninformative normal priors ($\alpha \sim$ normal ($\mu = 0$, $\sigma^2 = 300$)), indicating that no prior information on their values was available. The fitted model accounted for 88% of the deviance in butterfly species richness (Mac Nally & Fleishman 2002).

To validate or test the efficacy of the model based on these indicator species—the model's ability to predict species richness patterns as well as to explain considerable deviation in species richness patterns in the data used to build the model—we wanted to use a data set that was roughly comparable in spatial and temporal extent to the model-building data set. Therefore, we conducted inventories of butterfly species in 29 new locations in 2000–2002 in the Shoshone Mountains, which are approximately 40 km west of the Toquima Range (approximate north-south boundaries $39°14'19''$ to $38°57'32''$) (Fleishman et al. 2001). The two mountain ranges are within the same biogeographic subregion (Austin & Murphy 1987). They have similar climates, a common biogeographic past and ancestral biota, and comparable land-use histories.

Bayesian calculations produce probability distributions for model parameters, such as regression coefficients. One often uses "95% Bayesian credible intervals" (Lee 1989) to characterize the range of values for the parameter that encompasses 95% of the probability mass for that parameter. To compute such credible intervals for predictions of species richness at newly inventoried locations, we used all possible ($2^5 = 32$) combinations of indicator-species incidence (e.g., [0, 0, 0, 0, 0], [0, 0, 0, 0, 1], . . . , [1, 1, 1, 1, 1]) in conjunction with the distributions of values for each regression coefficient computed from the Poisson-regression model. This formed a simple "look-up table" within which the observed combination of indicator species for a validation site was cross-referenced to a predicted value of species richness and associated credible intervals.

Ultimately, one of our primary objectives is to construct models that have low spatial decay rates—in other words, models useful for estimating species richness over extensive areas. There are several main steps in this process. First, the validity of the existing model must be assessed. To test the efficacy of our initial model, we constrained

the regression coefficients tightly so that the value of each coefficient was specified with high precision (i.e., very small standard deviation). This effectively prevents the "Bayesian updating" of the model by the new data and provides a test comparable to usual frequentist statistics.

If the initial model successfully predicts species richness patterns, as it did in this case, the model-building and model-validation data sets can be combined to yield an improved or "updated" model that more effectively represents species richness patterns over a large area. Therefore, in this second step we relaxed constraints on the regression coefficients by using the full regression-parameter distributions, including the computed precisions from the model-building phase (Mac Nally & Fleishman 2002). This relaxation allows full Bayesian updating to proceed so that the model parameters (both their means and their standard deviations) can be updated by exposure to the new data (see Hilborn & Mangel 1997). This approach melded our new validation data from the Shoshone Mountains with the existing model based on data from the Toquima Range to provide an updated model based on both data sets. The applicability of the updated model then can be tested by collecting new validation data over an even greater geographic extent.

Given the flexibility of the WinBUGS programming language (version 1.3; Spielgelhalter et al. 2000), we computed the proportion of the posterior probability masses (PPMs) of the updated distribution of values for each regression coefficient that fell within $\pm 10\%$ of the mean of the distribution of values for the regression coefficients in the initial model. WinBUGS is an iterative Markov chain Monte Carlo simulation model. Because the modeling process is iterative, one obtains many (typically several thousand) predictions that cover a range of values. The posterior probability mass refers to the probability distributions for the predicted values. If much of the PPM fell within the probability distribution of the regression parameters for the initial model (the "model-building distribution"), then the validation data confirmed that the indicator species had similar importance for predicting species richness in both the model-building and model-validation data sets. However, if the PPM distribution was shifted substantially upward (larger regression coefficient) or downward (smaller regression coefficient), then the validation data suggested that the relative importance of the corresponding indicator species in predicting species richness was different in the two data sets.

The efficacy of indicator species may be affected by differences in the spatial and temporal pattern of their occurrence between the model-building and model-validation data sets. Therefore, we analyzed changes in occurrence rates—the proportion of inventoried sites in which a species was recorded—and changes in "false-negative" rates between the model-building and model-validation data sets. A false-negative rate is a measure of the conditional likelihood that a species with an erratic temporal distribution will be present in any year at a site, given that

it is known to occupy the site at some point in time. In this context, false negatives represent real variation in species occurrence as opposed to observer error, which we assumed to be minimal, albeit not zero. In other words, a false negative can be obtained if a site can and often does support the species (i.e., if larval host plants and other necessary resources are available), but the species is not present at the site during a given year, perhaps because vegetation senesced unusually early or weather conditions were unfavorable.

Site occupancy can only be determined using inventory data. Accordingly, if a species was not recorded during our 3 years of inventories, we considered it absent. To compare false-negative rates among species, we calculated the false-negative rate as the number of annual absences divided by the number of years of inventory data for sites at which the species was present in 1 or more years. Thus, presence in only 1 of 3 years yields a false-negative rate of $(3-1)/3 \sim 0.67$, but presence in 2 years would produce a false-negative rate of $(3-2)/3 \sim 0.33$. To statistically assess differences in false-negative rates, we treated each site in either data set (model-building or validation) for a given species as a Bernoulli trial drawn from a common distribution with the same false-negative rate. The model is

$$\text{logit}(\rho_i) = \alpha + \beta\delta_i, \ \rho_i \sim \ \text{binomial}(P_i, N_i).$$

This models the false-negative rates (ρs) of all sites in the building and validation data sets together. Data sets were distinguished by the δs, which were set to 0 for the model-building data set and 1 for the validation data set. Thus, α is an estimate of the baseline false-negative rate and β is the additive difference associated with the validation data set. For example, if the computed $\alpha = 0.5$ and $\beta = 0.2$, then the false-negative rate for the model-building data set would be 0.5, but the value for the validation data set would be $0.5 + 0.2 = 0.7$. The ρs are assumed to be binomially distributed means, from which there are A_i absence years, in sites from which the species is known to occur, in N_i inventory years per site (3 for 25 validation sites and 2 for 7 validation sites). Noninformative normal priors were used for the regression coefficients (α and β). We again used WinBUGS to compute directly the distribution of the difference between the mean false-negative rates for the model-building and validation data sets and the PPM approach to evaluate that difference. In this case, we computed the proportion of the PPM above zero. If $\geq 90\%$ of the PPM exceeded zero then the coefficient was deemed to represent an important positive shift. If $\leq 10\%$ of the PPM exceeded zero, then this was regarded as an important negative change.

Results

We originally built the model with sites in the Toquima Range from which we had collected 3 years of inven-

tory data. We collected 3 years of inventory data for 22 of the validation sites in the Shoshone Mountains (2000–2002). For the remaining 7 validation sites, however, we had only 2 years of inventory data (2001–2002). For the former 22 validation sites, we conducted a linear regression in which we regressed species richness over 2 years (2001–2002) against species richness over 3 years (2000–2002). This allowed us to generate expected species richness values for the latter 7 validation sites. The projected 3-year richness function was $S_3 = 0.879$ (1.683 SE) $+ 1.068$ (0.069 SE) $\times S_2$, R^2 (adjusted) $= 0.920$, where S_3 is the projected 3-year species richness value and S_2 is the observed 2-year species richness value. The S_3 projected values were used to assess model fit for the 7 sites for which only 2 years of inventory data were available.

Twenty-seven of the 29 observed (or projected) species richness values fell within the 95% credible intervals (Fig. 1). Although the number of indicator species contributed to overall richness in the validation sites (maximum difference of four), this explained little of the difference between sites with the fewest and most species (32 species; Fig. 1). In other words, the number of indicator species present in a given site, which ranged from one to five, had little effect on the difference in species richness between the sites with lowest and highest species richness. Mean predicted values were highly correlated
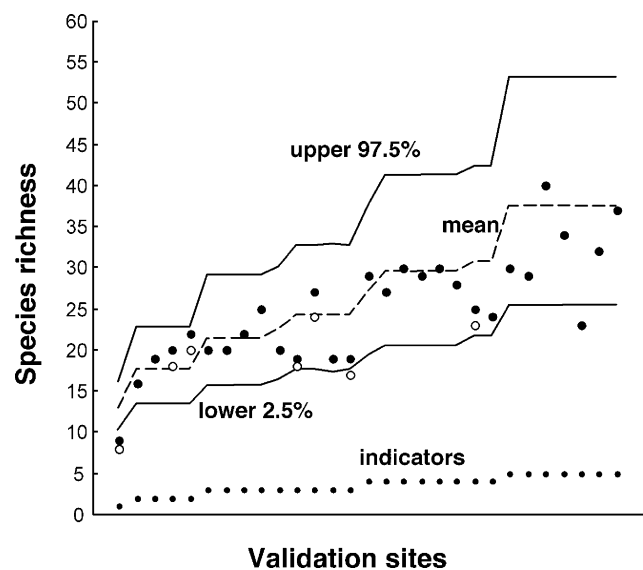

Figure 1. Mean predicted species richness (dashed line), upper and lower 2.5% credible limits, and observed or projected 3-year species richness values (large solid circles) for all model validation sites in the Shoshone Mountains. Small solid circles show the number of indicator species observed in each site. Observed 2-year species richness values for the seven sites with 2 rather than 3 years of inventory data are shown with open circles. The x-values are ordered by ascending mean predicted species richness.

Table 1. Original and updated model parameters for the Bayesian analysis of butterfly species richness.[*]

Parameter	Description	PPM distribution			
		mean ± SD	≥110%	mean ± 10%	≤90%
Original model parameters					
α_0	constant	2.352 ± 0.091	—	—	—
α_1	*Coenonympha tullia*	0.309 ± 0.074	—	—	—
α_2	*Euphilotes ancilla*	0.318 ± 0.083	—	—	—
α_3	*Everes amyntula*	0.236 ± 0.066	—	—	—
α_4	*Ochlodes sylvanoides*	0.193 ± 0.083	—	—	—
α_5	*Speyeria zerene*	0.195 ± 0.070	—	—	—
Updated model parameters					
α_0	constant	2.345 ± 0.082	0.000	1.000	0.000
α_1	*Coenonympha tullia*	0.315 ± 0.065	0.311	0.402	0.287
α_2	*Euphilotes ancilla*	0.249 ± 0.056	0.034	0.217	0.749
α_3	*Everes amyntula*	0.195 ± 0.047	0.120	0.292	0.588
α_4	*Ochlodes sylvanoides*	0.202 ± 0.055	0.424	0.258	0.318
α_5	*Speyeria zerene*	0.195 ± 0.062	0.410	0.242	0.348

[*]*Data are presented as posterior distributions. PPM indicates the proportion of the posterior probability mass for a parameter that was ≥110% of the original model mean value, within 10% (plus or minus) of the original model mean value and ≤90% of the original model mean value.*

with the 3-year observed values (sites from which 3 years of inventory data had been collected) or projected values (sites from which 2 years of inventory data had been collected) ($R_{Spearman} = 0.799$). The average absolute deviation was 3.6 species, but 4 sites contributed disproportionately to this deviation. Four sites had unusually large deviations (at least 6.7 species fewer than predicted). One of those sites had 14 fewer species than predicted. The average absolute deviation for the other 25 sites was 2.7 species.

In the posterior probability distributions for the Bayes-updated model, the constant was virtually unchanged, reflecting its very small standard deviation (0.091) relative to the mean (2.352) in the original model (Table 1). Coefficients for three of the five indicator species, *Coenonympha tullia*, *Ochlodes sylvanoides*, and *Speyeria zerene*, changed little when the model was updated with the validation data. This means that the presence of the latter three species provided a multiplicative factor (regression coefficient) in the updated model similar to that in the original model (i.e., e^{α}). However, the coefficients for the other two indicator species (*Euphilotes ancilla* [−22%] and *Everes amyntula* [−17%]) were shifted substantially downward, with much of the posterior probability mass <90% of the mean value in the original model (75% and 59% of the PPM, respectively; Table 1). Therefore, the effectiveness of *Euphilotes ancilla* and *Everes amyntula* as indicators of species richness was reduced, which may be why the predicted values of species richness overestimated the observed or projected values by approximately one species per site (Fig. 1).

We present occurrence rates of the indicator species without statistical tests of significance, as appropriate in the Bayesian context, although a test for identity of proportions (Sokal & Rohlf 1969) might be used in a classical frequentist context. Occurrence rates of two species, *Everes amyntula* and *Ochlodes sylvanoides*, differed lit-

tle between the building and validation data sets. However, occurrence rates for the other three species were substantially greater in the validation data set than in the model-building data set—38% greater for *Coenonympha tullia*, 147% greater for *Euphilotes ancilla*, and 137% greater for *Speyeria zerene* (Table 2).

False-negative rates in the two data sets were substantially different for four of the five indicator species (all but *Everes amyntula*), although the change for *Coenonympha tullia* (−0.06) fell just above the 0.10 decision criterion (recall that we considered a change in false-negative rate "substantial" if ≤0.10 or ≥0.90 of the PPM was >0). For four of the five species (all but *Euphilotes ancilla*), false-negative rates decreased markedly in the validation data set, indicating a substantially more consistent presence in that data set. That is, if a species was present at a location during our inventories, then it was more likely to be detected in more years (Table 2).

The updated model appeared to provide a consistent fit between the model-building (Toquima Range) and model-validation (Shoshone Mountains) data sets (Fig. 2). Locally weighted regression scatterplot smooths (Lowess; a smoothing method in which observations farther from the predicted x value are downweighted relative to observations closer to the predicted x value [Cleveland 1979; Quinn & Keough 2002]) for the two sets of data were similar. Therefore, the updated model seems to represent both sets of data well and will be used for future modeling.

Discussion

We developed an objective, statistically based method to identify indicators of species richness that should be applicable to any taxonomic group at landscape or regional scales (thousands of square kilometers) in any ecosystem.

Table 2. Comparison of occurrence rates (proportion of inventoried sites in which a species was recorded) and false-negative rates (see text) between the model-building (Toquima Range, 1996–1999) and model-validation (Shoshone Mountains, 2000–2002) data sets for the five indicator species.*

	Occurrence rates		False-negative rates			
Species	building	validation	building	validation	Change	PPM
Coenonympha tullia	0.673	0.931	0.17	0.11	−0.06	0.11
Euphilotes ancilla	0.265	0.655	0.37	0.47	0.10	0.90
Everes amyntula	0.367	0.379	0.57	0.50	−0.07	0.27
Ochlodes sylvanoides	0.592	0.621	0.54	0.41	−0.13	0.07
Speyeria zerene	0.408	0.966	0.55	0.08	−0.47	0

*PPM indicates the proportion of the posterior probability mass above zero with respect to the distribution of the difference between the false-negative rates for the two data sets; values ≤0.10 or ≥0.90 were deemed substantial for decreases and increases, respectively.

One of the most useful potential contributions of our work is that a standard method for identifying indicator species can be used for taxa and geographic locations beyond those employed to develop the method. The method's applicability is not restricted to butterflies, terrestrial invertebrates, or western North America. Thus, our framework for identifying indicator species may substantially enhance our ability to conduct large-scale ecological assessments and to make effective conservation decisions.

A model based on the occurrence of indicator species selected according to our method "explained" nearly 90% of the deviance of species richness in the data used for its construction (Mac Nally & Fleishman 2002). More important, from the perspective of our model's broader applicability both within and beyond our study system, is that validation tests indicated that the method is useful for estimating species richness at scales pertinent to contemporary land-use management. Predicted and observed values of species richness were highly correlated, and more than 90% of the observed values fell within the 95% credible intervals of the predictions. Over what distance will the model's accuracy remain high? Estimating the spatial "decay rate" of a model's usefulness is a critical stage in assessing the ecological reasons why the occurrence of indicator species is correlated with the richness of whole assemblages, and the confidence that conservation practitioners can have in statistical models of species richness. We expect that the model for butterflies in the Great Basin, and, by extension, models for other taxonomic groups and ecosystems to which our method is applied, will prove transferable over much larger areas, typically hundreds of kilometers or more, provided those areas have similar climate, pools of species, and land-cover and land-use patterns.

Pronounced environmental stochasticity often diminishes the predictive ability of correlative models. As explained below, we believe that the effects of precipitation extremes were detectable in the comparison between predicted and observed values of species richness in our case study. On the whole, however, the model was robust in the face of climatic variability. In our case study, four of the validation sites contributed disproportionately to the average absolute deviation between predicted and observed species richness. For the two sites with the greatest deviation, the low observed values of species richness (relative to the predictions) may reflect recent drought conditions in the study area. Running or standing water can function as a limiting resource for butterflies that obtain nutrients from mud puddles (Arms et al. 1974; Murphy & Wilcox 1986; Boggs & Jackson 1991). Presence of surface water and groundwater also delays senescence of larval host plants and adult nectar sources. Field measurements indicate that 2001 was the driest in the past 10 years in the central Great Basin and the fourth driest in the 70-year period of record (J. Korfmacher, personal communication). A hot, dry spring in 2002 similarly limited the amount of water, especially groundwater, available to plants and butterflies during the growing and flight seasons. The site with the largest deviation contains a spring that appeared virtually dry during the period in which we collected validation data. Similarly, the site with
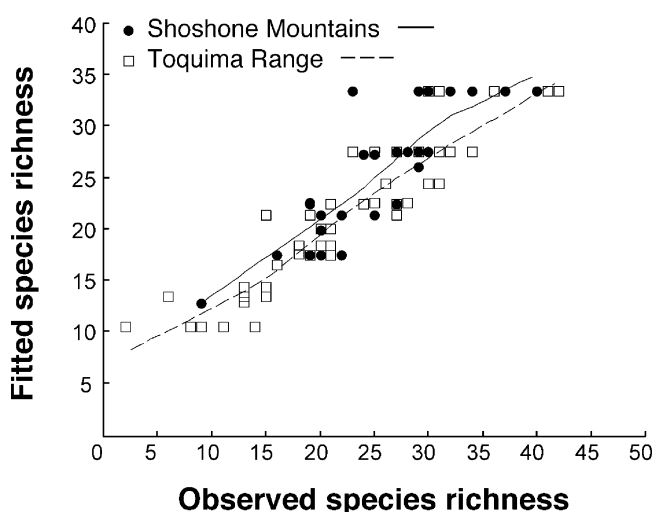

Figure 2. Fitted versus observed species richness for the Bayesian updated model. Data for the Shoshone Mountains and the Toquima Range are displayed separately. Lines are lowess (tension 0.66) smoothed curves.

the second largest deviation had the lowest mean distance to a stream channel of any of our validation sites, but the channel remained dry between 2000 and 2002. Although the third site did have standing and running water, it was heavily grazed by domestic cattle each year, which virtually eliminated food sources for adult butterflies.

Model predictions represent falsifiable hypotheses that should be validated with new data before the model is used to guide management decisions. Thus, we strongly advocate that potential users of our method conduct a formal validation test of their own model and, ideally, "update" the model before using it for conservation planning. In our updated model, the influence of *Euphilotes ancilla* and *Everes amyntula* as indicators of species richness was reduced. *Euphilotes ancilla* was the only indicator species whose false-negative rate was greater in the validation data set than in the model-building data set. Temporal variation in its occurrence pattern may have rendered the species a less effective contributor to projected values of species richness. Similarly, although neither the occurrence rate nor the false-negative rate for *Everes amyntula* differed substantially between the building and validation data sets, the species had the highest false-negative rate (and the lowest occurrence rate) among the five indicator species in the validation data set. In other words, its occurrence was relatively infrequent and erratic. Again, in any model, fluctuation in predictor variables is likely to diminish predictive accuracy. As a result, in the context of our method for identifying indicators of species richness, it makes sense to focus particular attention during the model-validation stage on indicator species that have markedly irregular occurrence patterns.

Our understanding of the state of species and ecosystems around the world is poor. Assessing their current state and gauging how environmental changes may affect their future condition are vital to making decisions that will maximize protection of biological diversity and provision of ecosystem services (Gaston 2000; Margules & Pressey 2000). The ecological and socioeconomic complexities of conservation planning, including logistic and financial impediments to data collection, often force scientists and resource practitioners to seek short-cuts for developing management schemes. However, few organisms in any taxonomic group or ecosystem have been identified that are reliable and cost-effective indicators of variables at the community level or higher ecological levels (Scott 1998). We acknowledge that, depending on location and taxonomic group, an experienced observer would not need to spend considerably more time and money to conduct a comprehensive inventory of the focal taxonomic group than on a search aimed at just the indicator species. However, local field biologists potentially could collect information on the indicator species in the course of other work. This might be an efficient method, especially in large managed landscapes. Also, many taxonomic groups and ecosystems are not amenable to direct assessment. In the latter situations, our approach is likely to be much easier and more practical than attempting to measure the richness of all species in an array of sites distributed over an extensive area.

A demonstrably effective model of species richness as a function of indicator species is one of several tools that may help produce increasingly well-informed strategies for addressing diverse management objectives. Our approach has strategic relevance for large-scale ecological assessments because it can be applied in an identical way to any assemblage and ecosystem (da Fonseca et al. 2002). By developing and testing hypotheses to explain why a particular set of indicator species encompasses fundamental information about a larger biota, we may draw general and transferable inferences about the nature of ecological assemblages. Such models are tools that may substantially improve our ability to monitor and conserve native species and ecosystem functions in a rapidly changing world.

Acknowledgments

Literature Cited

Andelman, S. J., and W. F. Fagan. 2000. Umbrellas and flagships: efficient conservation surrogates, or expensive mistakes? Proceedings of the National Academy of Sciences of the United States of America **97:**5954–5959.

Arms, K., P. Feeney, and R. C. Lederhouse. 1974. Sodium: stimulus for puddling behavior by tiger swallowtail butterflies. Science **185:**372–374.

Austin, G. T., and D. D. Murphy. 1987. Zoogeography of Great Basin butterflies: patterns of distribution and differentiation. Great Basin Naturalist **47:**186–201.

Bergerud, W. A., and W. J. Reed. 1998. Bayesian statistical methods. Pages 89–104 in V. Sit and B. Taylor, editors. Statistical methods for adaptive management studies. Land management handbook 42. British Columbia Ministry of Forests, Forestry Division Services Branch, Victoria, Canada. (Also available from http://www.for.gov.bc.ca/hfd/pubs/docs/lmh/lmh42.htm).

Boggs, C. L., and L. A. Jackson. 1991. Mud puddling by butterflies is not a simple matter. Ecological Entomology **16:**123–127.

Cameron, A. C., and P. K. Trivedi. 1998. Regression analysis of count data. Cambridge University Press, Cambridge, United Kingdom.

Chapin, F. S., et al. 2000. Consequences of changing biodiversity. Nature **405:**234–242.

Clench, H. K. 1979. How to make regional lists of butterflies: some thoughts. Journal of the Lepidopterists' Society **33:**216–231.

Cleveland, W. S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association **74:**829–836.

Crawley, M. J. 1993. GLIM for ecologists. Blackwell Scientific Publications, Oxford, United Kingdom.

da Fonseca, G. A. B., C. Gascon, M. K. Steininger, T. Brooks, R. A. Mittermeier, and T. E. Lacher Jr. 2002. A global experiment under way. Science **295:**1835.

Environment Conservation Council (Victoria) (EEC). 2000. Box-ironbark: forest and woodlands investigation. EEC, Melbourne, Australia.

Faith, D. P., et al. 2001. The BioRap biodiversity assessment and planning study for Papua New Guinea. Pacific Conservation Biology **6:**279–288.

Fleishman, E., G. T. Austin, and D. D. Murphy. 1997. Natural history and biogeography of the butterflies of the Toiyabe Range, Nevada (Lepidoptera: Papilionoidea). Holarctic Lepidoptera **4:**1–18.

Fleishman, E., G. T. Austin, and A. D. Weiss. 1998. An empirical test of Rapoport's rule: elevational gradients in montane butterfly communities. Ecology **79:**2482–2493.

Fleishman, E., D. D. Murphy, and P. F. Brussard. 2000. A new method for selection of umbrella species for conservation planning. Ecological Applications **10:**569–579.

Fleishman, E., R. Mac Nally, J. P. Fay, and D. D. Murphy. 2001. Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. Conservation Biology **15:**1674–1685.

Fleishman, E., C. J. Betrus, R. B. Blair, R. Mac Nally, and D. D. Murphy. 2002. Nestedness analysis and conservation planning: the importance of place, environment, and life history across taxonomic groups. Oecologia **133:**78–89.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature **405:**220–227.

Gewin, V. 2002. The state of the planet. Nature **417:**112–113.

Gustafsson, L. 2000. Red-listed species and indicators: vascular plants in woodland key habitats and surrounding production forests in Sweden. Biological Conservation **92:**35–43.

Hanski, I. 1999. Metapopulation ecology. Oxford University Press, New York.

Harding, P. T., J. Asher, and T. J. Yates. 1995. Butterfly monitoring 1: recording the changes. Pages 3–22 in A. S. Pullin, editor. Ecology and conservation of butterflies. Chapman and Hall, London.

Heywood, V. H., I. Baste, and K. A. Gardner. 1995. Introduction. Pages 1–20 in V. H. Heywood and R. T. Watson, editors. Global biodiversity assessment. Cambridge University Press, Cambridge, United Kingdom.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey.

Kareiva, P., and U. Wennergren. 1995. Connecting landscape patterns to ecosystem and population processes. Nature **373:**299–302.

Kremen, C., R. K. Colwell, T. L. Erwin, D. D. Murphy, R. F. Noss, and M. A. Sanjayan. 1993. Terrestrial arthropod assemblages: their use in conservation planning. Conservation Biology **7:**796–808.

Landres, P. B., J. Verner, and J. W. Thomas. 1998. Ecological uses of vertebrate indicator species: a critique. Conservation Biology **2:**316–328.

Lee, P. M. 1989. Bayesian statistics. Edward Arnold, London.

Leftwich, K. N., P. L. Angermeier, and C. A. Dolloff. 1997. Factors influencing behaviour and transferability of habitat models for a benthic stream fish. Transactions of the American Fisheries Society **126:**725–734.

Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between—and reconciliation of—'predictive' and 'explanatory' models. Biodiversity and Conservation **9:**655–671.

Mac Nally, R., and E. Fleishman. 2002. Using 'indicator' species to model species richness: model development and predictions. Ecological Applications **12:**79–92.

Mac Nally, R., A. F. Bennett, and G. Horrocks. 2000. Forecasting the impacts of habitat fragmentation: evaluation of species-specific predictions of the impact of habitat fragmentation on birds in the box-ironbark forests of central Victoria , Australia. Biological Conservation **95:**7–29.

Mac Nally, R., A. F. Bennett, G. W. Brown, L. F. Lumsden, A. Yen, S. Hinkley, P. Lillywhite, and D. Ward. 2002. How well do ecosystem-based planning units cover different components of biodiversity? Ecological Applications **12:**900–912.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. 2nd edition. Chapman and Hall, London.

Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. Nature **405:**243–253.

Miller, J. R., and P. Cale. 2000. Behavioral mechanisms and habitat use by birds in a fragmented agricultural landscape. Ecological Applications **10:**1732–1748.

Mooney, H. A. 2000. Requiem for nature. Nature **403:**593–594.

Murphy, D. D., and B. A. Wilcox. 1986. Butterfly diversity in natural habitat fragments: a test of the validity of vertebrate-based management. Page 287–292 in J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife **2000:** modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison.

New, T. R., R. M. Pyle, J. A. Thomas, C. D. Thomas, and P. C. Hammond. 1995. Butterfly conservation management. Annual Review of Entomology **40:**57–83.

Pearson, D. L. 1994. Selecting indicator taxa for the quantitative assessment of biodiversity. Philosophical Transactions of the Royal Society of London Series B **345:**75–79.

Pimm, S. L., et al. 2001. Can we defy nature's end? Science **293:**2207–2208.

Pollard, E., and T. J. Yates. 1993. Monitoring butterflies for ecology and conservation. Chapman and Hall, London.

Prendergast, J. R., R. M. Quinn, J. H. Lawton, B. C. Eversham, and D. W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature **365:**335–337.

Pressey, R. L., T. C. Hager, K. M. Ryan, J. Schwarz, S. Wall, S. Ferrier, and P. M. Creaser. 2000. Using abiotic data for conservation assessments over extensive regions: quantitative methods applied across New South Wales, Australia. Biological Conservation **96:**55–82.

Pullin, A. S., editor. 1995. Ecology and conservation of butterflies. Chapman and Hall, London.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge, United Kingdom.

Raguso, R. A., and J. Llorente-Bousquets. 1990. The butterflies (Lepidoptera) of the Tuxlas Mts., Veracruz, Mexico, revisited: species-richness and habitat disturbance. Journal of Research on the Lepidoptera **29:**105–133.

Reed, J. M. 1996. Using statistical probability to increase confidence of inferring species extinction. Conservation Biology **10:**1283–1285.

Roberts, C. M., et al. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. Science **295:**1280–1284.

Schwarz, G. 1978. Estimating the dimension of a model. Annals of Statistics **6:**461–464.

Scott, C. T. 1998. Sampling methods for estimating change in forest resources. Ecological Applications **8:**228–233.

Soberón, J., and J. Llorente. 1993. The use of species accumulation functions for the prediction of species richness. Conservation Biology **7:**480–488.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman, San Francisco, California.

Spielgelhalter, D., A. Thomas, and N. Best. 2000. WinBUGS version 1.3. Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Institute for Public Health, Cambridge, United Kingdom.

Vessby, K., B. Söderström, A. Glimskär, and B. Svensson. 2002. Species-richness correlations of six different taxa in Swedish seminatural grasslands. Conservation Biology **16:**430–439.

Walters, C. J. 1986. Adaptive management of renewable resources. Collier Macmillan, New York.