Searching for a Predictive Model for Species Richness of Iberian Dung Beetle Based on Spatial and Environmental Variables

JORGE M. LOBO* AND FERMÍN MARTÍN-PIERA†

Departamento de Biodiversidad y Biología Evolutiva (Entomología), Museo Nacional de Ciencias Naturales,
c/o José Gutiérrez Abascal, 2; 28006, Madrid, SPAIN

Abstract: *In Mediterranean countries, inventories of many animal groups, particularly insects, are incomplete or nonexistent. Hence, a feasible spatial picture of unequally surveyed areas is required to ascertain which faunistic surveys are good enough to produce reliable estimates of species richness. We used generalized linear models to build a multiple-regression function through which we predicted the distribution of Iberian dung beetle species richness. Given the scarcity and unevenness of the species-richness spatial distribution, the number of records of a dung beetle database (BANDASCA), falling within each of the 50 × 50 km grid squares, was chosen as a measure of the sampling effort for that square. Examining the asymptotic relationship between the number of dung-beetle species and database records for each physioclimatic Iberian subregion, we found that 82 grid squares (32% of the total) were adequately sampled. Dung-beetle species richness was related in each of these 82 cells to 24 explanatory variables. Curvilinear functions, interaction terms, and the significant third-degree polynomial terms of latitude and longitude were included to model species-richness distribution. The final model accounted for 62.4% of the total deviance after we eliminated seven outlier squares, with maximum elevation, grassland area, land-use diversity, forest area, geological diversity, interaction of terrestrial area and maximum elevation, and interaction between calcareous rock and geological diversity and latitude being the most significant independent variables. The residuals of the function were not spatially autocorrelated, and we validated the final model by a jackknife procedure. Large and environmentally complex hotspots in the Iberian Central, Baetic, and Subbaetic mountain ranges stand out from the emerging map of species richness. Further detailed research is required to determine the complementarity of the faunas of these two main hotspots, the key question in conservation planning for a dung-feeding beetle.*

En Busca de un Modelo Predictivo para la Riqueza de Especies de Escarabajos Coprófilos Ibéricos Basado en Variables Espaciales y Ambientales

Resumen: *Los inventarios de muchos grupos de animales, particularmente insectos, son incompletos o totalmente ausentes en los países Mediterráneos. Por ello, se requiere de un panorama espacial viable de áreas poco estudiadas para determinar que estudios faunísticos son suficientemente buenos para producir estimaciones confiables de la riqueza de especies. Utilizamos modelos lineales generalizados para construir una función de regresión múltiple con la cual predijimos la distribución de la riqueza de especies de escarabajos coprófilos ibéricos. Dada la escasez y desigualdad de la distribución espacial de la riqueza de especies, el número de registros de una base de datos de escarabajos coprófilos (BANDASCA) en cada uno de los cuadrantes de 50 × 50 km fue seleccionado como una medida del esfuerzo de muestreo para ese cuadro. Al examinar la relación asintótica entre el número de escarabajos coprófilos y los registros en la base de datos para cada región fisioclimática, encontramos que 82 cuadros (32 % del total) fueron adecuadamente muestreados. La riqueza de especies de escarabajos coprófilos se relacionó con 24 variables explicativas en cada uno de estas 82 celdas. Se incluyeron funciones curvilíneas, términos de interacción y los significativos térmi-*

158

nos polinomiales de tercer grado de latitud y longitud en el modelo de distribución espacial de la riqueza de especies. El modelo final consideró el 62.4 % de la desviación total después de eliminar siete cuadros externos, elevación máxima, superficie de pastizal, diversidad de uso de suelo, superficie forestal, diversidad geológica, la interacción entre superficie terrestre y elevación máxima, la interacción entre roca calcárea y diversidad geológica y latitud, les variables independiente más significativa. Los residuos de la correlación no fueron autocorrelacionados espacialmente, y validamos el modelo final con un procedimiento de navaja. En el mapa emergente de la riqueza de especies sobresalen sitios conflictivos extensos y ambientalmente complejos en las cadenas montañosas de Iberia Central, Baética y Subaética. Se requiere de más investigación detallada para determinar la complementariedad de la fauna de estos sitios conflictivos principales como el punto central en la planeación de la conservación de escarabajos coprófagos.

Introduction

Species-distribution databases provide for biogeographical research (Lubchenco et al. 1991; Blackmore & Cutler 1996; Systematics Agenda–2000 1994) and can aid biodiversity conservation by helping to identify a taxon's species-rich hotspots (Myers 1990). Some authors claim, however, that unequal sampling effort is a complication plaguing many biodiversity studies (Williams & Gaston 1994; Gaston 2000). In fact, researchers should take into account the skewing of data due to variation in sampling intensity (generally, the greater the intensity, the higher the number of species recorded) when making use of species-richness maps derived from such databases (Prendergast et al. 1993), planning additional sampling, determining the distribution of a particular species, or modeling species richness (Margules et al. 1987; Nicholls 1989; Walter 1990; Osborne & Tigar 1992; Buckland & Elston 1993; Austin et al. 1996; Heikkinen 1996; Skov & Borchsenius 1997).

Prendergast et al. (1993) developed a sampling-effort correction, based on the rarefaction method, to estimate the number of species in an area. Several methodological approximations estimate the "true" number of species in an assemblage through abundance or presence-absence data (Colwell & Coddington 1995; Colwell 1997). These statistical methodologies, however, can be applied only to high-quality data from well-known taxonomic groups and when unequal sampling effort can be corrected across the entire territory investigated. If the largest part of the territory is surveyed poorly, as often occurs in large Mediterranean (and tropical) regions, and if taxonomic and/or faunistic knowledge of the majority of the biota (mostly insect groups) is not comprehensive, reliable alternative methodological approaches are needed.

In Mediterranean countries there is often little more than species descriptions, specimens, and a list of locations from which they have been collected. Therefore, a feasible spatial picture (mapping) of unequally surveyed areas is required to ascertain which faunistic surveys are good enough to produce reliable estimates of species richness. These inventories are needed to build models based on environmental variables to forecast species distribution for the whole territory investigated. The asymptotic relationship between the number of species added by increments of sampling effort can be used to ascertain well-surveyed areas (Soberón & Llorente 1993; Colwell & Coddington 1995; Fagan & Kareiva 1997), and generalized linear models can be used to develop statistical models to extrapolate species richness for unsampled areas (Nicholls 1989).

Dung beetles are a cornerstone of natural-grassland biomes in Mediterranean pasturelands maintained by traditional grazing coupled with low-intensity agriculture (Martín-Piera & Lobo 1995). They play a major ecological role in pastures and grassland biomes because they remove vertebrate feces (mainly domestic and wild ungulates), drain soils, and aide recycling of organic matter and nutrients. Consequently, pasture quality and productivity depends on the activity of these beetles (Bornemissza & Williams 1970; Fincher 1981; Rougon et al. 1988). They also effectively reduce the breeding sites of cattle pests, such as flies (e.g., Bornemissza 1970; MacQueen & Beirne 1975) and intestinal parasites (Fincher 1975). As for all orders of insects (Samways 1993, 1994; Martín-Piera 1997) concern over dung-beetle conservation has been increasing in the last decade (Klein 1989; Lumaret 1994; Martín-Piera & Lobo 1995; Barbero et al. 1999; Van Rensburg et al. 1999). The massive extension of monocultures, disappearance of domestic livestock from highly mechanized agricultural landscapes (Barbero et al. 1999), and antiparasitic treatments are all harmful to dung beetles (Lumaret 1986; Lumaret et al. 1993). Because any successful dung-beetle conservation program must be based on sound knowledge of its spatial-diversity patterns, quick and innovative efforts are required to correct the historical neglect of insect-species conservation in Mediterranean countries (Balleto & Casale 1991) in general and in Spain in particular.

We predicted the species-richness distribution of the dung beetle for the entire Iberian Peninsula. We selected species richness because it provides a reasonably useful measure of biodiversity which is frequently correlated with ecological, morphological, phylogenetic, and

functional diversity (Gaston 1996). As far as we know, data on dung beetles is exceptional among data derived from Iberian insect faunistic surveys in that an exhaustive database is available (BANDASCA, Lobo & Martín-Piera 1991). We analyzed the database to (1) delimit adequately sampled 50-km universal transversal mercator (UTM) grid squares, defined according to the asymptotic relationship between species richness and sampling effort; (2) forecast Iberian dung-beetle species richness according to a stepwise multiple-regression model with environmental predictor variables; and (3) map and describe the resulting species-richness distribution.

Methods

Origin of the Biological Information

The database of Iberian dung beetles (Coleoptera, Scarabaeinae), BANDASCA, includes biological and geographical information from museums and some private collections, theses, and published and unpublished data (see BANDASCA structure of Lobo & Martín-Piera 1991). This database has 15,740 records and 101,996 individuals of the 53 Iberian dung-beetle species (Martín-Piera 2000). We defined a database record as a pool of individuals of a single species with identical database field values: locality (UTM coordinates), elevation, date of capture (day/month/year), type of habitat, food resource, and others. Thus, any difference in any database field value gave rise to a new database record, regardless of the number of individuals. Because biological information comes from sources with heterogeneous sampling methodologies, we chose the number of database records in each 50-km UTM grid square as a surrogate of the sampling effort in it, assuming that the probability of species occurrence correlates positively with the number of database records. We examined the number of species and records from 255 UTM 50 × 50 km grid squares (briefly, squares) (Fig. 1), having added the records and species list from those squares more than 85% covered by water to the nearest and most environmentally similar neighbor.

Delimitation of Iberian Physioclimatic Subregions

We analyzed the species-effort relationship using the values of species number (response variable) and database records (predictor variable) in all the squares. Because the number of species per square depends on environmental conditions, the main physioclimatic subregions of the Iberian Peninsula were previously defined according to the values of eight variables in each square: annual mean temperature, annual temperature variation, annual days of sun, total annual precipitation, total summer precipitation, annual precipitation variation, mean elevation, and elevational range. After standardization of

all variables, we used cluster analysis to classify the squares, with Ward's method as the linkage rule. The squared Euclidean distance was chosen as the measure of similarity between squares to place progressively greater weight on grid squares that were farther apart. Using the groups of squares previously defined by cluster analysis, we performed a discriminant-function analysis to verify to which group a particular square belonged. Misclassified squares were reclassified according to the classification functions.

Definition of Adequately Sampled Grid Squares

The adequacy of sampling in each square was determined by a negative exponential function relating the number of species (S_r) to the number of database records (r). According to Soberón and Llorente (1993) and Colwell and Coddington (1995), this relationship is given by

$$S_r = S_{max}[1 - \exp(-br)], \qquad (1)$$

where S_{max}, the asymptote, is the estimated total number of species per square and b is a fitted constant that controls the shape of the curve. The curvilinear function was fitted by the quasi-Newton method. Because 100% richness requires an infinite number of database records, we calculated the number of records required for a rate of species increment of ≤0.001 (i.e., one added species for each 1000 records; $r_{0.001}$). According to Soberón and Llorente (1993),

$$r_{0.001} = 1/b \ln + (1 + b/0.001). \qquad (2)$$

We also calculated the number of records required to obtain species-increment rates of 0.005, one added species for every 200 records ($r_{0.005}$), and 0.01 added species for every 100 records ($r_{0.01}$). Of the 255 Iberian squares, 82 fell into the three categories and were considered adequately sampled.

Choice of Explanatory Variables

For the adequately sampled squares, we recorded 24 continuous variables: two spatial variables (central latitude and central longitude); two geographic variables (distance from Pyrenees and terrestrial area); three topographic variables (minimum, maximum, and mean elevation); two geological variables (calcareous and acid rock surface); six climatic variables (minimum and maximum monthly mean temperature, annual mean temperature, total annual and summer precipitation, and annual days of sun); four land-use variables (cultivated plus urban area, forest, scrub, and grassland area); and five variables that measured environmental diversity (elevational range, annual temperature and precipitation variation, land use, and geological diversity).

The topographic, climatic, and environmental diversity variables were chosen because elevation and productivity- and heterogeneity-related variables influence variation in number of species (e.g., Begon et al. 1990; Huston 1994; Gaston & Williams 1996; Wohlgemuth 1998). Land-use and geologic variables were chosen because the number of dung-beetle species is affected strongly by vegetation cover (Lumaret 1983; Lumaret & Kirk 1991) and because basic or acidic soils can also influence the composition of dung-beetle communities.

The climatic data for each square are courtesy of W. Cramer (CLIMATE database, version 2; available from
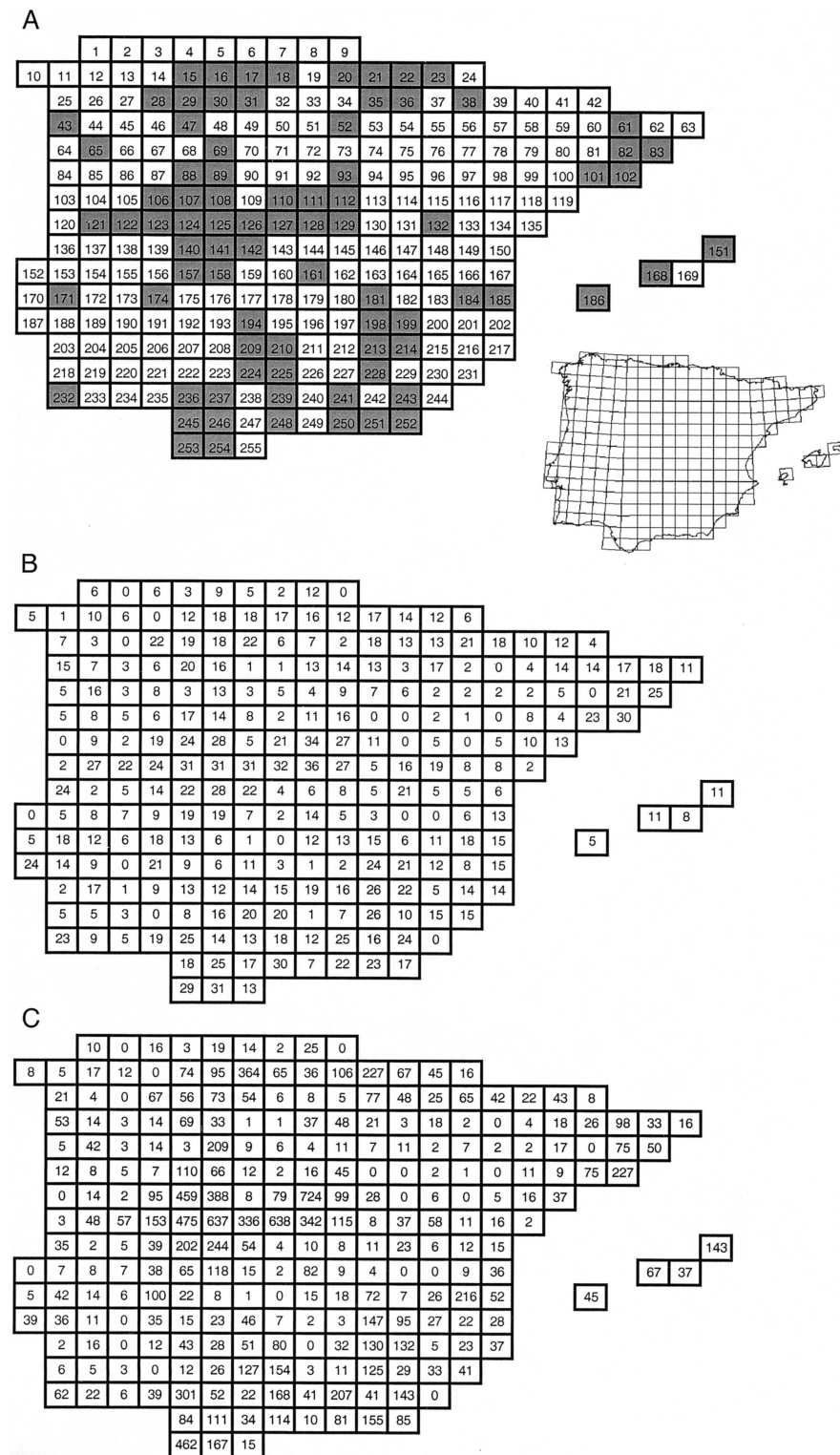
Figure 1. (a) Fifty-km UTM grid squares of the Iberian Peninsula numbered from 1 to 255. Shaded squares are those 82 grid squares judged adequately sampled according to the asymptotic relationship between species richness and sampling effort ($S_r = S_{max} [1 - exp (-br)]$) for each one of the six previously defined physioclimatic subregions of the Iberian Peninsula (Fig. 2). Adequately sampled grid squares were defined as those in which the number of database records is higher or equal to that required to obtain a rate of species increment per record of at least 1 species for every 100 records (see methods and Table 3). (b) Number of species recorded in each one of the 50-km UTM grid squares of the Iberian Peninsula. (c) Number of database records compiled in BANDASCA database of Iberian Scarabaeinae.

http://www.pik-potsdam.de/~cramer/climate.htm). Annual temperature and precipitation variation are the difference between the most extreme monthly scores. We selected the topographic and spatial data by overlaying a geographic information system (GIS: Idrisi 2.0; Clark Labs 1998) digital-elevation model of the Iberian Peninsula (grid square size of 1 km) with the polygons of squares. Land-use data, provided by the European Environment Agency, were extracted from the raster information (282-m resolution) on the 44 land-cover categories in Spain and Portugal (Corine Programme 1985–1990). This information was also overlaid with 50-km UTM polygons through the GIS. The different land-cover categories were grouped to obtain forest area (all types of forests), scrub and grassland surface area (either natural or artificial), and the surface of areas with strong anthropic influence (urban, industrial, and cultivation zones). We obtained geological data by digitizing, from an Iberian map (Instituto Geográfico Nacional 1995), soils on calcareous rocks, soils on acid rocks (siliceous), and soils on clay (spatial resolution of 1 km). Calcareous and acid rock surface in each grid square were estimated by overlaying the 50-km UTM polygons on this map by means of the GIS. We estimated land use and geological diversity in each grid square using the Shannon diversity index (Magurran 1988):

$$H' = -\sum p_i \log_2 p_i, \qquad (3)$$

where p_i is the relative frequency of each of the 44 landcover i categories and of each geological i category, respectively.

Search for the Best Forecasting Model

We used generalized linear models (GLM; McCullagh & Nelder 1989) to summarize the relationship between the number of dung-beetle species and explanatory variables to obtain an appropriate forecasting model (e.g., Nicholls 1989; Austin et al. 1990, 1996; Heikkinen & Neuvonen 1997). A GLM can be used with a range of non-normal distributions for the random component and to relax the constraint imposed by a linear relation between the dependent and independent variables. We assumed a Poisson error distribution for the number of dung-beetle species, related to the set of predictor variables via a logarithmic link function (Crawley 1993).

In the first step, the total number of scarabaeine in the adequately sampled squares was related separately, one by one with each environmental variable. Because the relationship between species richness and environmental variables is often curvilinear (Austin 1980), we selected either a linear, quadratic, or cubic function of each environmental variable by comparing their reduction in deviance with that of a full model in which the number of parameters is equal to the total number of observations (Dobson 1999). The goodness-of-fit of the

competing functions was measured by the deviance statistic and the change in deviance was tested by the *F*-ratio test (McCullagh & Nelder 1989; Dobson 1999) with a 5% significance level. The percentage of explained deviance was also calculated for each model (Dobson 1999).

We used a forward stepwise procedure to enter the variables into the model (Nicholls 1989; Austin et al. 1996; Heikkinen & Neuvonen 1997). First, the linear, quadratic, or cubic function of the variable that accounts for the most important change in deviance was selected for inclusion. Next, all remaining variables were tested for significance by adding them sequentially one by one, and the procedure was repeated iteratively until no more explanatory statistically significant variables remained ($p \leq 0.05$). After each step of forward inclusion, the significance of the terms previously selected was also tested by submitting the new model to a backward-selection procedure to exclude those terms that had become nonsignificant.

Environmental heterogeneity variables have an important influence on variation in species' spatial diversity (Begon et al. 1990), but they probably are a combination of other more causal and primary environmental variables. Hence, primary environmental variables, which probably more directly influence physiological mechanisms, were included before environmental diversity variables were added to the model (Austin 1980).

Margules et al. (1987) demonstrated that interactions between variables are often more predictive than the same variables separately. Therefore, after checking which environmental variables should be included, we tested the significance of the interaction between the 24 spatial and environmental variables (278 interaction terms). All interaction terms were sequentially added one by one to the model. The interaction term that accounted for the most important change in deviance was selected for inclusion, and the procedure was repeated iteratively until no more significant interaction terms remained.

Finally, the spatial variables, which can incorporate the effects of otherwise ignored historical, biotic, or environmental variables (Legendre & Legendre 1998), were included in the model. To ensure that not only linear spatial patterns in the species-richness data were recovered, but also more complex spatial structures, the third-degree polynomial of the central latitude (LAT) and longitude (LON) of each square were used as independent variables (trend surface analysis; see Legendre 1993): $b_1\text{LAT} + b_2\text{LON} + b_3\text{LAT}^2 + b_4\text{LAT}\times\text{LON} + b_5\text{LON}^2 + b_6\text{LAT}^3 + b_7\text{LAT}^2\times\text{LON} + b_8\text{LAT}\times\text{LON}^2 + b_9\text{LON}^3$. A backward stepwise regression with the nine terms of the equation as predictor variables and the number of dung-beetle species as dependent variable was carried out to remove the non-significant spatial terms (Legendre & Legendre 1998). Latitude and longi-

tude were centered on their respective means prior to the backward selection procedure. Significant spatial terms ($p \leq 0.05$) contributing to a significant change in deviance were included in the model with environmental and interaction terms.

Environmental variables were standardized to eliminate the effect of differences in measurement scale. All statistical computations have been made with the STATISTICA package (StatSoft 1999).

Collinearity, Nonlinear Relationships, and Autocorrelation

When species-richness patterns are modeled, it is necessary to account for the co-variance between explanatory variables, their possibly nonlinear relationship with the dependent variable, and the spatial autocorrelation of variables (Margules et al. 1987; Austin et al. 1996; Gaston 1996; Gaston & Williams 1996). If the goal is to forecast a response variable, collinearity in explanatory variables is not a concern in maximizing the predictive power of the function (Legendre & Legendre 1998). Nonlinearity is incorporated in the model by the use of quadratic and cubic functions of explanatory variables (Margules et al. 1987; Austin et al. 1996; Heikkinen & Neuvonen 1997).

The autocorrelation of variables implies that observations at any given locality are influenced by the values of neighboring localities. Because on a geographical scale most environmental variables are autocorrelated, assessing the relationships between variables is complicated by the ineffectiveness of classical statistical tools (Legendre & Legendre 1998). When a variable is spatially au-

tocorrelated, its broad-scale spatial structure can be removed. If it is, however, the effect of those spatially structured factors that determine species richness are reduced, thus reducing the forecasting ability of the model. So autocorrelated data, an important source of information, must be taken into account as such to improve the predictive power of models (Smith 1994; Legendre & Legendre 1998). When autocorrelated variables are used, the spatial independence of errors is critical, meaning that residuals should not be spatially autocorrelated (Cliff & Ord 1981; Odland 1988). Indeed, we calculated the autocorrelation for the final model residuals (assuming independence) with Moran's I spatial autocorrelation statistic, choosing a lag distance of 50 km (Cliff & Ord 1981; Sawada 1999). Moran's I test was checked for significance with the Bonferroni-corrected significance level. Spatial autocorrelation in the residuals usually indicates that the model must be enlarged to incorporate spatially structured variables not otherwise accounted for (Cliff & Ord 1981; Odland 1988). Hence, addition of complex spatial terms (the third-degree polynomial of latitude and longitude) to the model of environmental variables and interaction terms can be expected to account for those otherwise ignored spatially structured variables.

Residuals and Model Validation

We determined the adequacy of the final model by identifying outliers that disproportionately influenced the fit of the model, examining standard errors of coefficients,
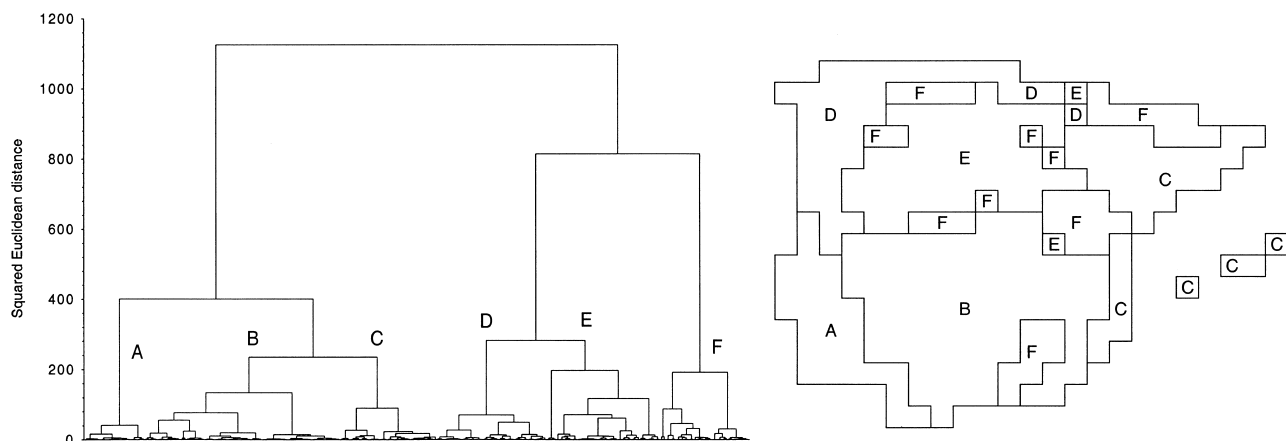

Figure 2. Dendrogram of Iberian 50-km UTM grid squares produced by a cluster analysis with Ward's method as the linkage rule and the squared Euclidean distance as the measure of similarity, and map of the physioclimatic subregions resulting from the cluster analysis and discriminant-function analysis. Cluster analysis was used to perform a preliminary classification of the UTM grid squares. Discriminant function analysis was then used to reclassify, by means of classification functions, those UTM grid squares misclassified by the cluster analysis (10 50-km UTM squares). The variables considered in each grid square are annual mean temperature and variation, annual days of sun, total annual and summer precipitation, annual precipitation variation, mean elevation, and elevation range. Codes: A, West Mediterranean subregion; B, South Plateau; C, East Mediterranean subregion; D, Eurosiberian subregion; E, North Plateau; F, Montane subregion.

Table 1. Spatial and environmental scores (mean ± SE) of the 50 × 50 km UTM grid squares belonging to each one of the physioclimatic subregions of the Iberian Peninsula (Fig. 2).[a]

Variable	Eurosiberian	Montane	North Plateau	South Plateau	West Mediterranean	East Mediterranean	$F_{(5,249)}$[b]	Post boc mean comparisons[c]
Longitude	−6.81 ± 0.36	−2.29 ± 0.39	−4.54 ±0.25	−4.28 ± 0.21	−8.03 ± 0.19	0.59 ± 0.25	93.79	1-2, 1-3, 1-4, 1-6, 2-3, 2-4, 2-5, 2-6, 3-5, 3-6, 4-5, 4-6, 5-6
Latitude	42.50 ± 0.18	40.90 ± 0.31	41.67 ± 0.11	38.63 ± 0.12	38.23 ± 0.22	40.78 ± 0.23	80.32	1-2, 1-3, 1-4, 1-5, 1-6, 2-3, 2-4, 2-5, 3-4, 3-5, 3-6, 4-6, 5-6
Mean elevation	427.32 ± 38.58	1211.98 ± 37.77	856.81 ± 24.19	563.20 ± 23.31	131.35 ± 12.76	360.33 ± 33.37	135.80	1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 2-6, 3-4, 3-5, 3-6, 4-5, 4-6, 5-6
Elevation range	1089.40 ± 61.41	1888.32 ± 99.65	818.93 ± 64.13	908.87 ± 51.06	360.61 ± 50.01	1048.06 ± 77.84	42.09	1-2, 1-5, 2-3, 2-4, 2-5, 2-6, 3-5, 4-5, 5-6
Annual mean temperature	14.70 ± 0.18	10.50 ± 0.33	12.93 ± 0.17	14.12 ± 0.17	15.98 ± 0.14	15.39 ± 0.30	64.53	1-2, 1-3, 1-5, 2-3, 2-4, 2-5, 2-6, 3-4, 3-5, 3-6, 4-5, 4-6
Annual temperature variation	13.75 ± 0.37	17.19 ± 0.22	18.25 ± 0.25	17.09 ± 0.16	12.70 ± 0.40	16.51 ± 0.24	60.18	1-2, 1-3, 1-4, 1-6, 2-3, 2-5, 3-4, 3-5, 3-6, 4-5, 5-6
Annual days of sun	47.14 ± 0.99	55.14 ± 0.72	54.37 ± 0.56	61.63 ± 0.21	63.89 ± 0.59	57.67 ± 0.51	102.52	1-2, 1-3, 1-4, 1-5, 1-6, 2-4, 2-5, 2-6, 3-4, 3-5, 3-6, 4-6, 5-6
Total annual precipitation	110.41 ± 5.53	70.28 ±3.29	57.68 ± 2.27	49.21 ± 1.54	61.02 ± 2.83	46.06 ± 1.71	63.35	1-2, 1-3, 1-4, 1-5, 1-6, 2-3, 2-4, 2-6, 3-6, 4-6, 5-6
Total summer precipitation	135.69 ± 7.78	151.38 ±11.67	91.83 ± 4.27	68.36 ± 3.15	38.08 ± 6.79	81.17 ± 4.99	39.05	1-3, 1-4, 1-5, 1-6, 2-3, 2-4, 2-5, 2-6, 3-4, 3-5, 4-5, 5-6
Annual precipitation variation	151.11 ± 11.39	61.24 ± 4.18	65.35 ±4.83	57.89 ± 2.50	102.46 ± 4.14	49.91 ± 2.70	49.18	1-2, 1-3, 1-4, 1-5, 1-6, 2-5, 3-5, 4-5, 5-6

[a]*Results of one-way analysis of variance for each variable and post boc mean comparisons by means of the Tukey test.*
[b]*$p \leq 0.001$.*
[c]*Pair numbers bigblight the statistically different pairwise comparisons between subregions at p = 0.05.*

plotting residuals against fitted values, and plotting residuals in a normal probability plot (Nicholls 1989; Dobson 1999). Using all adequately sampled squares ($n = 82$), we found a preliminary model in which eight squares had positive standard residuals higher than $+1.5$ (fitted values lower than observed) and six squares had negative standard residuals lower than -1.5 (fitted values greater than observed). Inspection of these 14 outliers showed that 7 were environmental and biological singular points lying outside the domain of the remaining squares (Nicholls 1989). The remaining 7 were located in the most extensively surveyed regions near Barcelona and Madrid (numbers 102, 111, 127, and 128 in Fig. 1) and in the Balearic Islands (151, 168, and 186). Historical oversampling for the positive residuals and insularity for the negative residuals explains these outliers. To build the final model, therefore, these 7 outliers were removed ($n = 75$).

To validate the final model, we applied a jackknife procedure. With a data set of 75 squares, we recalculated the model 75 times, leaving out one square in turn. Each one of the regression models based on the $n - 1$ grid squares was then applied to that excluded square in order to produce a predicted species-richness score for each one of the squares. We checked the predictive power of the model by examining the correlation between the recorded and the jackknife-predicted species richness and by estimating the percent error for each observation:

$$E_i = (|y_i - \overline{y_i}|/y_i) \times 100, \qquad (4)$$

where E_i is the predictive error for observation i; y_i is the observed species richness; and $\overline{y_i}$ the predicted value for y_i when this i observation is excluded. The mean of the 75 estimates of error provides a measure of the prediction error associated with the model and data used (Pascual & Iribarne 1993).

Results

Delimitation of Physioclimatic Subregions

Cluster analysis delimited six principal subregions in the Iberian Peninsula (Fig. 2) that were significantly different in their climate and topographic variables and in their spatial location. Results of one-way analysis of variance showed that all the variables significantly differed between groups, and post-hoc mean comparisons by Tukey test produced many significant differences between pairwise subregions (Table 1). The discriminant analysis showed that 96% of squares were well classified in the groups defined by the cluster analysis. The remaining 4% (10 squares) were assigned to the corresponding group according to classification functions. The number of squares that pertain to each physiocli-

Table 2. Number of 50 × 50 km UTM grid squares belonging to each Iberian subregion (Fig. 2), mean number of species, database records, and individuals per grid square in BANDASCA database (mean ± SE).

	Eurosiberian	Montane	North Plateau	South Plateau	West Mediterranean	East Mediterranean		Total
						continent	Balearic islands	
Squares	35	37	46	76	26	31	4	255
Species	7.9 ± 1.2	15.3 ± 1.5	11.5 ± 1.2	12.0 ± 0.9	12.0 ± 1.9	9.6 ± 1.5	8.8 ± 1.4	11.5 ± 0.5
Records	26.0 ± 7.2	115.6 ± 30.5	62.3 ± 16.3	51.3 ± 7.6	52.9 ± 20.7	27.2 ± 7.7	73.0 ± 24.2	56.7 ± 6.5
Individuals	187.3 ± 74.8	823.1 ± 290.8	495.4 ± 279.4	260.4 ± 56.0	544.4 ± 360.8	48.8 ± 14.1	195.5 ± 55.9	376.6 ± 78.4

Table 3. Asymptotic relationship between species richness and sampling effort ($S_r = S_{max} [1 - \exp(-br)]$) for the six physioclimatic subregions of the Iberian Peninsula (Fig. 2).

	Eurosiberian	Montane	North Plateau	South Plateau	West Mediterranean	East Mediterranean	Total
R^2 (%)[a]	82.4	84.9	83.7	81.7	90.4	69.7	
S_{max}[b]	18.8	27.4	22.2	22.6	26.3	19.9	
$r_{0.001}$[c]	91.7	160.7	108.9	114.5	100.6	101.3	
$r_{0.005}$[c]	54.4	83.7	62.4	64.9	58.6	58.9	
$r_{0.01}$[c]	39.9	56.6	44.8	46.2	42.5	42.7	
A[d]	2	6	6	12	3	2	31 (12.2%)
B[e]	3	5	6	9	2	3	28 (11.0%)
C[f]	3	6	6	4	1	3	23 (9.0%)
Total (A + B + C)	8	17	18	25	6	8	82
Total grid squares	35	37	46	76	26	35	255
Adequately sampled grid squares (%)	22.86	45.95	39.13	32.89	23.08	22.86	32.16

[a] Percentage of richness variation accounted for by the function.
[b] The asymptote is the estimated total number of species per square.
[c] The $r_{0.001}$, $r_{0.005}$, and $r_{0.01}$ are the number of records required to add one species for each 1000, 200, and 100 records, respectively (see Soberón & Llorente 1993).
[d] Number of grid squares with a rate of species increment of $\leq r_{0.001}$.
[e] Number of grid squares with a rate of species increment of $\leq r_{0.005}$ but $\geq r_{0.001}$.
[f] Number of grid squares with a rate of species increment of $\leq r_{0.01}$ but $\geq r_{0.005}$.

matic subregion (Fig. 2) varied between 26 (West Mediterranean subregion) and 76 (South Plateau subregion; Table 2). With these data, the mean number of species per square differed significantly only between the Eurosiberian and the Montane subregions (post-hoc Tukey test; $F_{(6,248)} = 2.81$; $p = 0.01$).

Relationship between Sampling Effort and Species Richness

The number of species and the number of database records differed greatly among the Iberian squares (Fig. 1b & 1c), but both values were spatially correlated (Spearman rank coefficient $r_s = 0.934$; $n = 255$; $p \leq 0.0001$) (i.e., the richest squares were the most surveyed).

The adequately sampled squares were mapped for each subregion (Fig. 1a) according to the species increment rates ($r_{0.001}$, $r_{0.005}$, and $r_{0.01}$) (Table 3). In all cases the percentage of explained variance of the regression (R_2) was high (between 69.7% and 90.4%). Hence, in total, 32.2% of Iberian squares were well inventoried (82 squares). The percentage of these squares differed moderately between subregions (between 22.9% and 45.9%), but a sufficient fraction of each subregion was sampled adequately (Table 3).

Forecasting Iberian Species Richness

Eight environmental variables were significant when tested separately as either linear, quadratic, or cubic functions (Table 4): the distance from the Pyrenees; minimum, maximum, and annual mean temperature; total annual precipitation; annual days of sun; grassland area; and land-use diversity. Only the linear term of annual days of sun (18.2%) and the quadratic function of minimum mean temperature (12.9%) explained more than 10% of the total deviance. With regard to the spatial variables, the backward stepwise regression of the third-degree polynomial terms of latitude and longitude selected the quadratic and cubic terms of latitude (Table 4).

Because the linear function of annual days of sun accounted for the most important change in deviance (18.1), this variable was the first one included in the model (Table 5). In the second step, the linear function of maximum elevation was selected. Next, the quadratic function of grassland area was selected, although only the quadratic term accounted for a significant change in deviance. None of the remaining primary variables were significant.

After the significant primary environmental variables were fit, the cubic function of land-use diversity was the only environmental diversity variable added to the model, although the quadratic term was deleted from the model (Table 5). Subsequently, the importance of interaction terms added one by one was tested. The interaction between forest area and geologic diversity accounted for the most significant change in deviance. Terrestrial area in grid square $\times$ maximum elevation, latitude $\times$ annual precipitation variation, and calcareous rocks $\times$ geologic diversity are the other significant interaction terms that enter in the model, but their inclusion removed the cubic term of land-use diversity (Table 5).

Finally, significant spatial terms—the quadratic and cubic terms of latitude—were added to the previous model, increasing the percentage of explained deviance by 2.3%, although the linear term of annual days of sun and the interaction between latitude and annual precipitation variation were removed during the selection procedure.

Table 4.   Environmental explanatory variables considered in the 75 adequately sampled 50 × 50 km UTM grid squares of the Iberian peninsula (Fig. 1).

Variable	Abbreviation	Selected terms	Deviance[a]	df	Change in deviance	F[b]	sign[c]
Full model			99.61	74			
Spatial[d]							
longitude	LON						
latitude	LAT	$LAT^2 + LAT^3$	74.16	72	25.45	24.71***	− −
Geography							
terrestrial area in grid square	A	A	98.46	73	1.15	0.85	
distance from Pyrenees	Dp	Dp	93.58	73	6.03	4.70*	+
Topography							
minimum elevation	e	e	99.45	73	0.16	0.12	
maximum elevation	E	E	97.47	73	2.15	1.61	
mean elevation	em	em	99.46	73	0.05	0.04	
Bedrock geology							
calcareous rocks	Cr	Cr	98.64	73	0.98	0.72	
acid rocks	Ar	Ar	98.16	73	1.45	1.08	
Climate							
minimum mean temperature	t	t	97.96	73	1.65	1.23	
		$t + t^2$	86.77	72	11.19	9.28**	− +
maximum mean temperature	T	T	93.71	73	5.90	4.60*	+
annual mean temperature	tm	tm	96.23	73	3.38	2.56	
		$tm + tm^2$	89.02	72	10.59	8.57**	− +
total annual precipitation	P	P	99.52	73	0.09	0.06	
		$P + P^2$	95.72	72	3.88	2.92	
		$P + P^2 + P^3$	93.01	71	6.60	5.04*	+ − +
total summer precipitation	Ps	Ps	99.29	73	0.22	0.34	
annual days of sun	Ds	Ds	81.52	73	18.10	16.21***	+
Land use							
cultivated and urban area	U	U	99.02	73	0.59	0.43	
forest area	F	F	99.60	73	0.01	0.01	
scrub area	S	S	99.21	73	0.40	0.30	
grassland area	G	G	96.13	73	3.48	2.84	
		$G + G^2$	94.55	72	5.06	3.86	
		$G + G^2 + G^3$	92.73	71	6.88	5.27*	− + −
Environmental diversity							
elevation range	AR	AR	97.71	73	1.91	1.42	
annual temperature variation	TV	TV	99.30	73	0.31	0.23	
annual precipitation variation	PV	PV	99.41	73	0.20	0.15	
land-use diversity	LUD	LUD	98.23	73	1.38	1.02	
		$LUD + LUD^2$	95.31	72	4.30	3.25	
		$LUD + LUD^2 + LUD^3$	91.25	71	8.36	6.51*	+ − −
geological diversity	GD	GD	99.00	73	0.61	0.45	

[a] *Deviance and change in deviance from a full model for total dung-beetle species number. The linear, quadratic, or cubic functions of each variable were selected if they accounted for a statistically significant change in the deviance with a probability lower than 0.05.*
[b] ****p < 0.001; **p < 0.01; *p < 0.05.*
[c] *Sign of the columns corresponds to sign of the term of each function.*
[d] *Spatial variables were chosen by removal of the nonsignificant terms from a third-degree polynomial equation of latitude and longitude by a backward stepwise selection.*

The final model was

$$S = \exp(c + E + G^2 + LUD + F \times GD + A \times E + Cr \times GD + LAT^2 + LAT^3), \quad (5)$$

where S is total number of dung beetles, c is the intercept, E is maximum elevation, G is grassland area, LUD is land-use diversity, F is forest area, GD is geological diversity, A is terrestrial area in the grid square, and Cr is calcareous rocks. This model explains 62.41% of the total deviance (Table 5).

Examination of residuals did not show special aberrant features. Residuals were normally distributed, the plot of residuals versus predicted values formed a homogeneous cloud around the center, and the standard errors of the coefficients were low (Table 5). Residuals were not spatially autocorrelated either. None of Moran's *I* values in the different lag classes were significant at a 0.05 level with the Bonferroni correction. However, plotting predicted versus observed values showed that the model tends to overestimate the num-

Table 5. Summary of the stepwise forward selection of variables to build a multiple-regression model for species richness of dung beetles.[a]

Step	Variable[b]	Deviance	df	Change in deviance[c]	F[d]	Explained deviance	Coefficients[e]	SE
Full model		99.61	74					
Step 1	Ds	81.52	73	18.10	16.21***	0.1817		
Step 2	E	71.45	72	10.07	10.14**	0.3177	0.062	0.026
Step 3	G	67.59	71	3.86	4.06*	0.3546		
	$+ G^2$	66.93	70	0.66	0.69	0.3608	0.014	0.006
	$- G$	67.11	71	0.18	0.19	0.3591		
Step 4	LUD	65.12	70	1.99	2.14	0.3782	0.114	0.041
	$+LUD^2$	62.70	69	2.42	2.66	0.4013		
	$+LUD^3$	59.27	68	3.43	3.93	0.4340		
	$-LUD^2$	59.54	69	0.27	0.31	0.4314		
Step 5	F × GD	54.13	68	5.41	6.80*	0.4831	−0.079	0.025
Step 6	A × E	49.19	67	4.94	6.73*	0.5303	0.093	0.026
	$-LUD^3$	49.21	68	0.02	0.03	0.5301		
Step 7	LAT × PV	45.30	67	3.91	5.78*	0.5674		
Step 8	Cr × GD	41.77	66	3.54	5.59*	0.6012	0.054	0.029
Step 9	$+LAT^2$	41.76	65	0.01	0.02	0.6013	−0.021	0.009
	$+LAT^3$	37.34	64	4.42	7.57**	0.6434	−0.009	0.002
	$-Ds$	37.40	65	0.06	0.11	0.6429		
	$-LAT × PV$	39.36	66	1.96	3.29	0.6241		
Intercept							2.940	0.050

[a] *Model employs a Poisson distribution of error assumption and a log link and uses the species-richness scores and environmental data of the 75 adequately sampled 50-km UTM grid squares (Fig. 1) after omitting seven outliers (see text).*
[b] *Variable codes defined in Table 4.*
[c] *The change in deviance after inclusion of a term in the model has been tested by an F-ratio test with a probability lower than 0.05.*
[d] ****p < 0.001; **p < 0.01; *p < 0.05.*
[e] *Only the coefficients and standard errors for parameters of the variables included in the final model are highlighted.*

ber of species for low species-richness squares and to underestimate it for high species-richness ones. This drawback of the model remained even when we eliminated all observations that could potentially be considered outliers. The results of a jackknife test on the final model showed strong correspondences between observed species-richness scores and those predicted by the jackknife procedure for the 75 squares (Fig. 3). The correlation between observed and predicted values was positive and significant (Pearson correlation coefficient: $r = 0.689$, $p = 0.0008$. Spearman rank correlation: $r_s = 0.688$; $p < 0.0001$). The mean jackknife predictive error of observations was 15.9%, with a 95% confidence interval between 13.2% and 18.7%. These results suggest that the predictive model was reasonably good in spite of its tendency to reduce the difference between the lowest and highest species-richness scores.
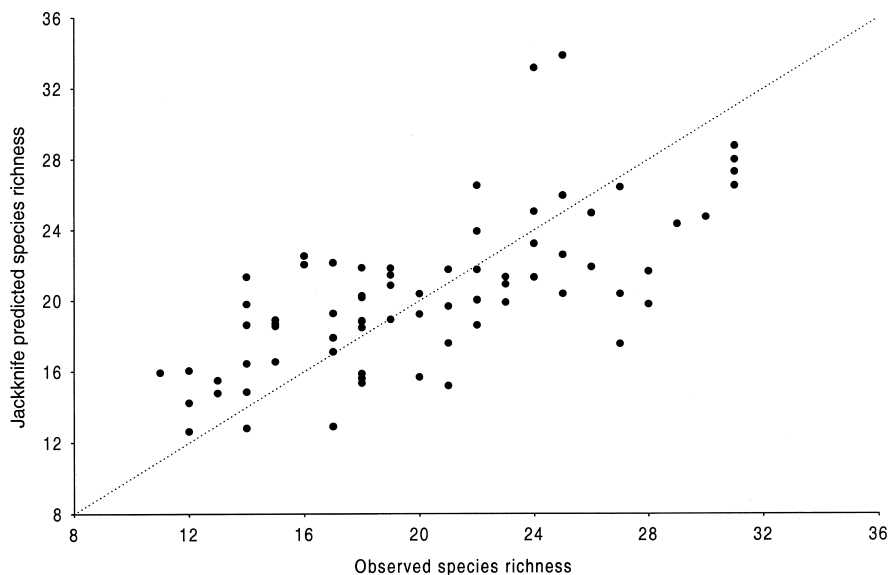

Figure 3. Relationship between the observed and jackknife-predicted species-richness scores. Using the 75 adequately sampled squares, we recalculated the coefficients of the final model (Table 5) 75 times, leaving out one square in turn. The regression model based on the n − 1 grid squares was then applied to that excluded square to obtain a predicted species-richness score for each one of the squares.

Iberian Distribution of Dung-Beetle Richness

When the final model was applied to the entire Iberian territory, the geographic distribution of species richness was forecasted (Fig. 4a). Species-richness hotspots of Iberian dung beetles were concentrated mainly in Mediterranean mountain regions, particularly in the Iberian Central System, the Baetic and Subbaetic mountain ranges, and the Iberian Cordillera (Fig. 4b & 4c). Few Pyrenees Mountain squares were predicted to be rich in dung-beetle species. The Iberian Central System hotspot extended to the southern slopes of the Galaico-Duriense massif in the northern half of the Portuguese Atlantic region, and the species richness of a few squares of the North and South Plateau of Central Spain was also high. Baetic and Subbaetic hotspots also extended westward to the Ibero-Atlantic southern corner. Shadow zones of species richness (North Ibero-Atlantic regions, North Plateau, the Ebro Valley, and the South Plateau) extended broadly eastward (Spanish Mediterranean Levante) and southwestward (Tajo Valley and southern Portugal) (Fig. 4).

The species richness of the physioclimatic subregions can be compared by taking into account the asymptotic value of the number of species as a function of the number of database records per square (S_{max} in Table 3). This comparison confirmed the general pattern of dung-beetle species-richness distribution predicted by the model. Eurosiberian and eastern Mediterranean subregions appeared the poorest (19 and 20 species per square, respectively), whereas 27 and 26 species per square occurred in the Montane and West Mediterranean subregions respectively, and 22 species occurred per square in North and South Plateau subregions.

Discussion

It is no longer a matter for discussion that decisions on the conservation of areas should be based on large, high-quality data sets. Nevertheless, except for some groups of vertebrates—mainly mammals and birds, higher plants, and a few groups of invertebrates (tiger beetles, butterflies) inhabiting temperate, adequately sampled areas—little accurate geographic information is available (Williams & Gaston 1994). It is known that biogeographic diversity patterns for the majority of the insect groups reflect the distribution of the areas investigated by entomologists (Dennis & Hardy 1999) and in Spain and Portugal the geographic distribution of the entomologists themselves, rather than the geographic distribution of organisms (Martín & Gurrea 1999). Some insights into the growth and evolution of entomological faunistic knowledge of the Iberian Peninsula emerge from a comparison of Figs. 1b and 4. Although the most intensely sampled areas would not necessarily have to be the richest, our

results confirm that Iberian dung-beetle species-richness predictions do not differ greatly from the observed geographic pattern (Pearson correlation coefficient: $r = 0.389$; $n = 251$; $p = 0.001$; Spearman rank correlation coefficient: $r_s = 0.302$; $n = 251$; $p \leq 0.0001$). This pattern suggests that entomologists have focused historically on those regions known to be richest, ignoring "poor" regions, and that positive feedback occurred at the beginning of inventorying species: most intensely sampled areas were those reputed to have more species, so more species tended to be recorded there. Model results did not correlate with observations in supposedly poor regions, where entomologist's intuition failed and sampling was not done. For example, there was a striking difference between observed and predicted species richness in the Iberian Cordillera (Fig. 4). Therefore, in poorly sampled areas the central question is how species-rich the supposedly poor regions are.

For the species-rich and poorly inventoried Mediterranean regions, much sampling must still be done to describe within a reasonable time the spatial distribution of insects, even if we are dealing with supposedly well-known groups such as dung-feeding beetles. This colossal task appears impractical (May 1990; Ehrlich 1992; Williams & Humphries 1996) because it is slow and expensive, and the financial support for taxonomic resources is continually declining (Cotterill 1995; Systematics Agenda–2000 1994; Blackmore & Cutler 1996; Martín-Piera 1997). The examination of an ex novo comprehensive biological database for Iberian dung beetles on a broad spatial scale (square = 2500 km^2) indicates that no more than 32% of the Iberian territory has been sampled adequately, and only 12% has a rate of species increment of fewer than one species per 1000 records (Table 3a). The situation is similar for French dung beetles, for which 21.4% of squares are well sampled (Lobo et al. 1997).

Sampling-effort bias has to be reduced to manageable proportions (Williams & Gaston 1994; Gaston 1996) to avoid erroneous predictions of hotspot richness and recommendations for conservation. But the species-number distribution of Iberian dung beetles (last species described in 1979) cannot be corrected for unequal sampling effort because of the lack of reliable biological information (many squares with zero or few records). To address this issue we propose building comprehensive databases of bibliographic, museum, and field data (e.g., Harding & Sheail 1992) and using forecasting models to obtain testable geographic patterns of biodiversity-related variables such as species richness. The number of database records, bringing together all the available taxonomic, ecological, and distributional specimen information, can be used as a sampling-effort surrogate to plan future sampling of inadequately or unexplored areas and, most important, to ascertain adequately inventoried regions. A model to forecast species-richness dis-

A

```
         11 12 10 10 11  9  8 10 10
   12 13 13 15 18 16 16 16 18 13 13 14 13 13  7
   16 17 17 20 19 20 20 18 17 16 15 15 17 18 22 16 24  8
   16 22 21 22 19 15 10 12 15 18 22 21 15 19 19 22 21 24 17 14 15
   19 22 21 20 20 15 11 15 16 19 21 25 23 16 15 15 19 17 16 23
   19 22 23 20 19 16 13 16 20 22 19 21 22 22 15 18 19 21 19
   19 26 23 21 27 20 18 20 27 21 19 18 21 22 22 21 17
   20 27 26 25 27 29 28 23 18 19 19 21 22 22 17
   19 28 24 20 19 22 20 20 18 18 16 20 19 22 20
   21 21 16 21 22 21 22 24 25 21 19 17 18 21 20 17
   20 20 12 16 18 21 22 21 21 21 17 18 19 20 19 19
   22 18 16 15 17 19 17 16 19 19 19 21 22 20 22 20
   17 16 16 18 19 18 20 19 20 17 22 22 22 20 20
   19 14 15 20 16 18 16 20 18 24 25 22 23 22
   20 18 23 22 21 15 15 19 21 32 26 23 19
         21 26 29 25 20 23 22 21
         28 29 17
```

B

C

CC
PC
NP
EV
ICS
IC
TV
SP
GB
BMR

Figure 4. (a) Plot of predicted species richness of Iberian dung beetles for 255 UTM 50 × 50 km grid squares (Fig. 1): white, 0–10 species; light grey, 11–20 species; dark grey, more than 21 species. (b) Interpolated surface of the total number of dung-beetle species from the predicted species-richness distribution per 50-km UTM squares. The interpolation model was designed with the Interpol module of Idrisi 2.0 (Clark Labs 1998). (c) Main physiographic regions of the Iberian Peninsula: BMR, Baetic and Subbaetic Mountain Ranges; CC, Cantabrian Cordillera; EV, Ebro Valley; PC, Pirineos Cordillera; GB, Guadalquivir Bassin; ICS, Iberian Central system; IC, Iberian Cordillera; SP, South Plateau; NP, North Plateau; TV, Tajo Valley. Tones of gray indicate elevation.

tribution, after delimiting the adequately sampled squares, would more rapidly produce an estimate of the spatial distribution of species richness within a reasonable time and could be validated and improved continually by adding new empirical data.

Because conservationists must quickly and cheaply identify areas of greatest diversity, it is more reasonable to design easy-to-use predictive models based on environmental variables. Several methodologies can be used to predict the geographic distribution of biodiversity (e.g., Nicholls 1989; Osborne & Tigar 1992; Huntley et al. 1995; Skov & Borchsenius 1997; Iverson & Prasad 1998; Wohlgemuth 1998). We used generalized linear models to estimate dung-beetle species-richness distribution patterns, including interaction terms (Margules et al. 1987) and significant spatial terms from a third-degree latitude and longitude polynomial. Although spatial terms are commonly omitted in the building of predictive models, we suggest adding them after contribution of the environmental variables and interaction terms, because they minimize residual autocorrelation, improving the predictive power of the model by including the influence of otherwise ignored spatially structured variables (Legendre & Legendre 1998).

The results of the regression-diagnostic methods and the lack of autocorrelation in the residual scores suggest that the assumptions of the model are reasonable and that no environmentally important variables were ignored. The main drawback of the model is its tendency to overestimate species richness in low-diversity squares and to underestimate it in high-diversity ones. This weakness, however, does not essentially modify the predictive species-richness map because the areas detected as coldspots are probably still poorest, whereas those detected as hotspots are still richest, thus providing a piece of evidence that more refined versions must be investigated.

To produce an accurate map of the probable geographic distribution of Iberian dung-beetle species richness, it is necessary to surpass several limitations, the main one being improvement of the accuracy of the dependent variable, which might require additional faunistic data. To achieve this goal, it is necessary to (1) increase the adequately sampled squares, thus enlarging the number of cases used in elaboration of the model; (2) improve the reliability of the procedure to select well-sampled squares; and (3) possess a more even spread of observations across the entire geographical range and in the full multivariate range of the environmental space. Moreover, if we aim to build new models based on smaller square size and to increase the organism's sensitivity to registered environmental variables, a still greater sampling effort will be required. The predictive utility of a model is constrained by limitations of the data on which the parameters have been estimated (Nicholls 1989). Indeed, our model appears credible

with the available species-richness data and the set of 24 explanatory variables we considered.

The predicted hotspots, the Iberian Central System and Baetic and Subbaetic mountain ranges, are environmentally heterogeneous and located in the Iberian Mediterranean region (Fig. 4). An ecologically heterogeneous faunistic pool of Iberian dung-beetle species of different phylogenetic lineages exists in river basins of middle and high mountains in these hotspots (Martín-Piera et al. 1992; Martín-Piera 2000). Hence, more detailed research on the wellsprings of Iberian biodiversity are needed to discriminate among the contributions of environmental, geographic, and historical variables. The question of faunistic complementarity in the species richness of the two main hotspots of Iberian dung beetles, important for conservation policy, remains to be addressed. Biodiversity variables such as phylogenetic composition, rarity, and endemism also need to be addressed for use in reserve selection (Reyers et al. 2000).

Acknowledgments

Literature Cited

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. Vegetatio **42:**11–21.

Austin, M. P., A. O. Nicholls, and C. R. Margules. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. Ecological Monographs **60:**161–177.

Austin, M. P., J. G. Pausas, and A. O. Nicholls. 1996. Patterns of tree species richness in relation to environment in south-eastern New South Wales, Australia. Australian Journal of Ecology **21:**154–164.

Balleto, E., and A. Casale. 1991. Mediterranean insect conservation. Pages 121–140 in N. M. Collins and J. A. Thomas, editors. The conservation of insects and their habitats. Academic Press, London.

Barbero, E., C. Palestrini, and A. Rolando. 1999. Dung beetle conservation: effects of habitat and resource selection (Coleoptera: Scarabaeoidea). Journal of Insect Conservation **3:**75–84.

Begon, M., J. L. Harper, and C. R. Townsend. 1990. Ecology: individuals, population and communities. Blackwell Scientific, Osney Mead, Oxford, United Kingdom.

Blackmore, S., and D. Cutler, editors. 1996. Systematics agenda 2000. The challenge for Europe. Proceedings of a workshop organised by the European Science Foundation Systematic Biology Network, the Linnean Society of London, the Rijksherbarium/Hortus Botanicus, Leiden University, and the Systematics Association. Samara Publishing, Cardingan, United Kingdom.

Bornemissza, G. F. 1970. Insectary studies on the control of dung breeding flies by the activity of the dung beetle *Onthophagus ga-*

zella F. (Coleoptera: Scarabaeidae). Journal of the Australian Entomological Society **9:**31–41.

Bornemissza, G. F., and C. H. Williams. 1970. An effect on dung beetle activity on plant yield. Pedobiologia **10:**1–7.

Buckland, S. T., and D. A. Elston. 1993. Empirical models for the spatial distribution of wildlife. Journal of Applied Ecology **30:**478–495.

Clark Labs. 1998. Idrisi 2.0. Clark University, Worcester, Massachusetts.

Cliff, A. D., and J. K. Ord. 1981. Spatial processes: models and applications. Pion, London.

Colwell, R. K. 1997. Estimates: statistical estimation of species richness and shared species from samples. Version 5. User's Guide and application available from http://viceroy.eeb.uconn.edu/estimates (accessed 7 March 2001).

Colwell, R. K., and J. A. Coddington. 1995. Estimating terrestrial biodiversity through extrapolation. Pages 101–118 in D. L. Hawksworth, editor. Biodiversity, measurement and estimation. Chapman & Hall, London.

Cotterill, F. P. D. 1995. Systematics, biological knowledge and environmental conservation. Biodiversity and Conservation **4:**183–205.

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

Dennis, R. L. H., and P. B. Hardy. 1999. Targeting squares for survey: predicting species richness and incidence of species for a butterfly atlas. Global Ecology and Biogeography **8:**443–454.

Dobson, A. J. 1999. An introduction to Generalized Linear Models. Chapman & Hall/CRC, Boca Raton, Florida.

Ehrlich, P. R. 1992. Population biology of checkerspot butterflies and the preservation of global biodiversity. Oikos **63:**6–12.

Fagan, W. F., and P. M. Kareiva. 1997. Using compiled species list to make biodiversity comparisons among regions: a test case using Oregon butterflies. Biological Conservation **80:**249–259.

Fincher, G. T. 1975. Effect on dung beetle activity on the number of nematode parasitoid required by grazing cattle. Journal of Parasitology **61:**759–762.

Fincher, G. T. 1981. The potential value of dung beetles in pasture ecosystems. Journal of the Georgia Entomological Society **16:**316–333.

Gaston, K. J. 1996. Species richness: measure and measurement. Pages 77–113 in K. J. Gaston, editor. Biodiversity: a biology of numbers and difference. Blackwell Science, Oxford, United Kingdom.

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

Gaston, K. J., and P. H. Williams. 1996. Spatial patterns in taxonomic diversity. Pages 202–229 in K. J. Gaston, editor. Biodiversity: a biology of numbers and difference. Blackwell Science, Oxford, United Kingdom.

Harding, P. T., and J. Sheail. 1992. The Biological Records Centre: a pioneer in data gathering and retrieval. Pages 5–19 in P. T. Harding, editor. Biological recording of changes in British wildlife. Institute of Terrestrial Ecology, London.

Heikkinen, R. K. 1996. Predicting patterns of vascular plant species richness with composite variables: a meso-scale study in Finnish Lapland. Vegetatio **126:**151–165.

Heikkinen, R. K., and S. Neuvonen. 1997. Species richness of vascular plants in the subarctic landscape of northern Finland: modelling relationships to the environment. Biodiversity and Conservation **6:**1181–1201.

Huntley, B., P. M. Berry, W. Cramer, and A. P. McDonald. 1995. Modelling present and potential future ranges of some European higher plants using climate response surface. Journal of Biogeography **22:**967–1001.

Huston, M. A. 1994. Biological diversity. Cambridge University Press, Cambridge, United Kingdom.

Instituto Geográfico Nacional. 1995. Atlas nacional de España. Volumes 1 and 2. Centro Nacional de Información, Madrid.

Iverson, L., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecological Monographs **68:**465–485.

Klein, B. C. 1989. Effects of forest fragmentation on dung and car-

rion beetle communities in Central Amazonia. Ecology **70:**1715–1725.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology **74:**1659–1673.

Legendre, P., and L. Legendre. 1998. Numerical ecology. 2nd edition. Elsevier, Amsterdam.

Lobo, J. M., and F. Martín-Piera. 1991. La creación de un banco de datos zoológico sobre los Scarabaeidae (Coleoptera: Scarabaeoidea) íbero-baleares: una experiencia piloto. Elytron **5:**31–38.

Lobo, J. M., J. P. Lumaret, and P. Jay-Robert. 1997. Les Atlas faunistiques comme outils d'nalyse spatiale de la biodiversité. Annales de la Société Entomologique de France **33:**129–138.

Lubchenco, J., et al. 1991. The sustainable biosphere initiative: an ecological research agenda. Ecology **72:**371–412.

Lumaret, J. P. 1983. Structure des peuplements de coprophages Scarabaeidae en région méditerranéenne française: relations entre les conditions écologiques et quelques paramètres biologiques des espèces. Bulletin de la Société Entomologique de France **88:**481–495.

Lumaret, J. P. 1986. Toxicité de certains helminthicides vis-a-vis des insectes coprophages et conséquences sur la disparition des excréments de la surface du sol. Acta Oecologica, Oecología Applicata **7:**313–324.

Lumaret, J. P. 1994. La conservation de l'entomofaune dans les aires naturelles protégées. Pages 57–65 in R. Jiménez-Peydró and M. A. Marcos-García, editors. Environmental management and arthropod conservation. Asociación Española de Entomología, Valencia.

Lumaret, J. P., and A. A. Kirk. 1991. South temperate dung beetles. Pages 97–115 in I. Hanski and Y. Cambefort, editors. Dung beetle ecology. Princeton University Press, Princeton, New Jersey.

Lumaret, J. P., E. Galante, C. Lumbreras, J. Mena, M. Bertrand, J. L. Bernal, J. L. Cooper, N. Kadiri, and D. Crowe. 1993. Fields effects of ivermectin residues on dung beetles. Journal of Applied Ecology **30:**428–436.

MacQueen, A., and B. P. Beirne. 1975. Influence of other insects on production of horn fly, *Haematobia irritans* (Diptera: Muscidae) from cattle dung in south-central British Columbia. Canadian Journal of Plant Science **107:**1255–1264.

Magurran, A. E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey.

Margules, C. R., A. O. Nicholls, and M. P. Austin. 1987. Diversity of *Eucalyptus* species predicted by a multi-variable environment gradient. Oecologia **71:**229–232.

Martín, J., and P. Gurrea. 1999. Areas de especiación en España y Portugal. Boletín de la Asociación Española de Entomología **23:**83–103.

Martín-Piera, F. 1997. Apuntes sobre biodiversidad y conservación de insectos: dilemas, ficciones y ¿soluciones? Boletín de la Sociedad Entomológica Aragonesa **20:**1–31.

Martín-Piera, F. 2000. Familia Scarabaeidae. Pages 215–446 in M. A. Ramos Sánchez, editor. Coleoptera, Scarabaeoidea I. Fauna Ibérica. Volume 14. Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas, Madrid.

Martín-Piera, F., and J. M. Lobo. 1995. Diversity and ecological role of dung beetles in Iberian grassland biomes. Pages 147-153 in D. I. McCracken, E. M. Bignal, and S. E. Wenlock, editors. Farming on the edge: the nature of traditional farmland in Europe. Joint Nature Conservation Committee, Peterborough, United Kingdom.

Martín-Piera, F., C. M. Veiga, and J. M. Lobo. 1992. Ecology and biogeography of dung beetle communities (Coleoptera, Scarabaeoidea) in an Iberian mountain range. Journal of Biogeography **19:**677–691.

May, R. 1990. Taxonomy as destiny. Nature **347:**129–130.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman & Hall, London.

Myers, N. 1990. The biodiversity challenge: expanded hot-spots analysis. The Environmentalist **10:**243–256.

Nicholls, A. O. 1989. How to make biological surveys go further with generalised linear models. Biological Conservation **50:**51–75.

Odland, J. 1988. Spatial autocorrelation. Sage Publications, Los Angeles.

Osborne, P. E., and B. J. Tigar. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. Journal of Applied Ecology **29:**55–62.

Pascual, M. A., and O. O. Iribarne. 1993. How good are empirical predictions of natural mortality? Fisheries Research **16:**17–24.

Prendergast, J. R., S. N. Wood, J. H. Lawton, and B. C. Eversham. 1993. Correcting for variation in recording effort in analyses of diversity hotspots. Biodiversity Letters **1:**39–53.

Reyers, B., A. S. van Jaarsveld, and M. Krüger. 2000. Complementarity as a biodiversity indicator strategy. Proceedings of the Royal Society of London, Series B **267:**505–513.

Rougon, D., C. Rougon, J. Trichet, and J. Levieux. 1988. Enrichissement en matière organique d'un sol sahélien au Niger par les insectes coprophages (Coléoptères: Scarabaeidae): implications agronomiques. Revue d'Ecologie et de Biologie du Sol **25:**413–434.

Samways, M. J. 1993. A spatial and process sub-regional framework for insect and biodiversity conservation research and management. Pages 1–28 in K. J. Gaston, T. R. New, and M. J. Samways, editors. Perspectives on insect conservation. Intercept, Andover, United Kingdom.

Samways, M. J. 1994. Insect conservation biology. Chapman & Hall, London.

Sawada, M. 1999. ROOKCASE: an Excel 97/2000 visual basic (VB) add-in for exploring global and local spatial autocorrelation. Bulletin of the Ecological Society of America **80:**231–234.

Skov, F., and F. Borchsenius. 1997. Predicting plant species distribution patterns using simple climatic parameters: a case study of Ecuadorian palms. Ecography **20:**347–355.

Smith, P. A. 1994. Autocorrelation in logistic regression modelling of species' distributions. Global Ecology and Biogeography Letters **4:** 47–61.

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

StatSoft, Inc. 1999. STATISTICA. Computer program manual for Windows. StatSoft, Inc., Tulsa, Oklahoma.

Systematics Agenda–2000. 1994. Systematics agenda 2000: charting the biosphere. Technical report. American Society of Plant Taxonomists, Society of Systematics Biologists, Willi Hennig Society, and Association of Systematics Collection, New York.

Van Rensburg, B., M. A. McGeoch, S. L. Chown, and A. S. Van Jaarsveld. 1999. Conservation of hetrogeneity among dung beetles in the Maputaland centre of endemism, South Africa. Biological Conservation **88:**145–153.

Walter, P. A. 1990. Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. Journal of Biogeography **17:**279–289.

Williams, P., and K. J. Gaston. 1994. Measuring more 'diversity': can higher taxon richness predict wholesale species richness? Biological Conservation **67:**211–217.

Williams, P., and C. J. Humphries. 1996. Comparing character diversity among biotas. Pages 54–76 in K. J. Gaston, editor. Biodiversity: a biology of numbers and difference. Blackwell Science, Oxford, United Kingdom.

Wohlgemuth, T. 1998. Modelling floristic species richness on a regional scale: a case study in Switzerland. Biodiversity and Conservation **7:**159–177.