



Contents lists available at ScienceDirect

Journal for Nature Conservation

journal homepage: www.elsevier.de/jnc

European Policy

How to identify future sampling areas when information is biased and scarce: An example using predictive models for species richness of Iberian water beetles

David Sánchez-Fernández^{a,*}, Jorge M. Lobo^b, Pedro Abellán^b, Andrés Millán^a^a Departamento de Ecología e Hidrología, Universidad de Murcia, Campus de Espinardo, 30100 Murcia, Spain^b Museo Nacional de Ciencias Naturales (CSIC), Madrid, Spain

ARTICLE INFO

Article history:

Received 4 August 2009

Accepted 22 May 2010

Keywords:

Sampling effort Assessment

Species richness

Iberian Peninsula

Water beetles

Distribution models

Conservation

ABSTRACT

Good distribution maps based on adequate sampling of a number of taxonomic groups are required to provide reliable conservation strategies. Nevertheless, it is common that inventories of many animal groups, particularly insects, are incomplete or nonexistent, with large gaps appearing once all available information of insects is mapped, especially when wide spatial scales are considered. Due to the lack of resources and manpower, the accomplishment of future field campaigns in these generally poorly surveyed areas should be directed so as to maximise the information obtained with the minimum survey effort. Using an exhaustive database of Iberian water beetles (which was relatively scarce and biased), we aim to prioritise areas to propose a planned survey design able to generate more accurate geographical representations of species distributions. For this, a prediction on the geographic distribution of the species richness of this group in the Iberian Peninsula was first obtained using the information coming from seven sets of well-surveyed grid cells determined by using progressively more exigent completeness values. Both observed and estimated by accumulation curves, species richness values of these different cell groups were subsequently used as the dependent variable in the modelling procedure. We used generalised linear models and 18 environmental variables as predictors. In this manner, 14 species richness predictions were obtained whose predictive power was assessed by a Jackknife procedure. The best model explained 57.5% of total deviance with a high mean Jackknife predictive error (29.9%). The overlay of these predictions with the survey effort map allows us to locate those areas where more sampling effort is necessary (areas of high predicted species richness that are not well inventoried).

© 2010 Elsevier GmbH. All rights reserved.

Introduction

Extensive biological databases are a primary tool in ecological and biogeographical research since good-quality distributional data are required for developing reliable designs and conservation strategies (Guralnick et al. 2007; Hortal et al. 2007; Prendergast et al. 1993; Soberón & Peterson, 2004). Only countries with a long-standing tradition of natural history and sufficient resources are able to produce good distribution maps based on adequate sampling of a number of taxonomic groups (Griffiths et al. 1999; Lawton et al. 1994). However, this is not the case in Mediterranean countries such as Spain, where inventories of many animal groups, particularly insects, are incomplete or nonexistent (Ramos et al. 2001). Due to the lack of resources and manpower in taxonomy, the accomplishment of future field campaigns in these generally poorly surveyed countries and taxonomical groups should be directed to maximise the information obtained with the minimum survey effort.

The discrimination and recognition of poorly surveyed areas as well as the study of the bias associated with the available information are key tasks necessary for the design of more efficient survey strategies. In the Iberian Peninsula, different studies examine these questions for the case of insects (see Lobo & Martín-Piera, 2002; Romo et al. 2006; Sánchez-Fernández et al. 2008b). This information can be used to design more effective survey campaigns but also to forecast the distribution of biodiversity attributes or individual species distributions in the remaining poorly surveyed territory (Ferrier 2002; Ferrier & Guisan, 2006; Hortal et al. 2001; Hortal et al. 2004; Lobo 2008a; Lobo & Martín-Piera, 2002). A first and necessary step for the identification of well-surveyed areas is estimating the degree to which the data represents reliable inventories on a given scale (Petersen & Meier, 2003), which can be carried out using a variety of statistical techniques (Colwell & Coddington, 1994; Gotelli & Colwell, 2001; Koellner et al. 2004; Rosenzweig et al. 2003; Soberón & Llorente, 1993). Although this procedure is not error-free (Hortal & Lobo, 2005), an area is usually considered adequately sampled when its ratio between the number of recorded species against those predicted (its completeness) is above a specific arbitrary threshold (see Nakamura & Soberón, 2009).

* Corresponding author. Tel.: +34 868884977; fax: +34 868883963.

E-mail address: davidsan@um.es (D. Sánchez-Fernández).

Freshwater ecosystems are subject to high rates of human alteration and, therefore, of biodiversity loss (Allan & Flecker, 1993; Darwall & Vié, 2005; Ricciardi & Rasmussen, 1999; Saunders et al. 2002). Furthermore, human pressures on freshwater resources are likely to increase in coming decades, putting yet more species at risk (Strayer 2006). This is particularly true in the Mediterranean Basin, which is considered one of Earth's biodiversity hotspots (Myers et al. 2000). In this study, Iberian water beetles were used as a focal group. From a taxonomic and biogeographical perspective, water beetles are perhaps one of the best known groups of invertebrates in the Iberian Peninsula and Balearic Islands (Ribera 2000; Ribera et al. 1998). These species groups are especially diversified in the Mediterranean region, where they inhabit almost every kind of fresh and brackish water habitat, from the smallest ponds to lagoons and wetlands and from streams to irrigation ditches and reservoirs (e.g. Millán et al. 2002; Ribera et al. 1998; Ribera 2000). Furthermore, they are thought to be a good indicator of the whole biological diversity existent in aquatic ecosystems (Bilton et al. 2006; Sánchez-Fernández et al. 2006) and have been successfully used in the selection of priority areas for conservation (Abellán et al. 2005; Sánchez-Fernández et al. 2004).

Using the information compiled by an exhaustive database on Iberian water beetles, it was previously demonstrated that only a quarter of the Iberian and Balearic 50 km × 50 km Universal Transverse Mercator grid cells can be considered well surveyed, and that recording effort was skewed by simple variables related with collector activity (Sánchez-Fernández et al. 2008b). In this paper, we aim to prioritise areas to propose a planned survey design able to generate more accurate geographical representations of species distributions for conservation purposes. Using the information coming from different sets of well-surveyed grid cells, we firstly carried out a simple prediction on the geographic distribution of the species richness of this group in the Iberian Peninsula to subsequently identify the most probable species-rich regions that at the same time are insufficiently surveyed.

Materials and methods

Study area

The study focuses on the Iberian Peninsula and Balearic Islands, extending over 585,644 km² (Fig. 1). The territory includes a variety of biomes, relief, climates, and soil types, where altitude ranges from sea level to 3483 m in the Sierra Nevada Mountains. This is one

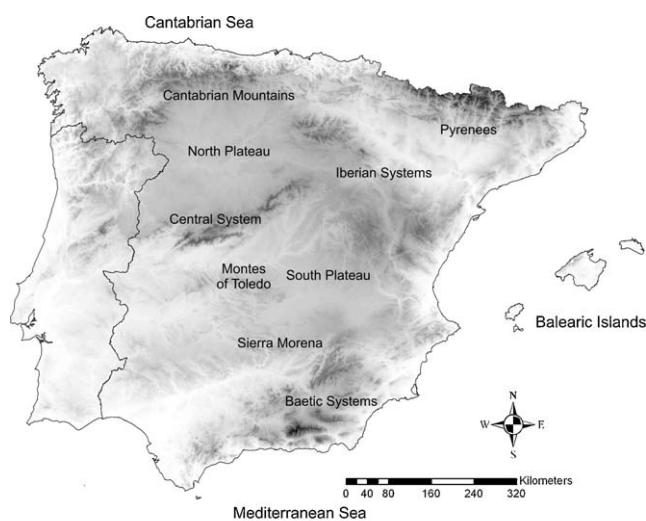


Fig. 1. Study area with some locations referred to in text highlighted.

of the richest European regions in terms of animal species diversity (Williams et al., 2000), and particularly in endemic water beetles (Ribera 2000; Ribera et al. 2003; Sánchez-Fernández et al. 2008a), and is characterised by a wide range of ecosystem types, some of which are rare on a European scale.

Source of biological data

An exhaustive database of records on Iberian water beetles (ESACIB "EScarabajos ACuáticos IBéricos") was used in this study. This database is the most complete source of information available for a major group of freshwater invertebrates in the study area. ESACIB contains over 50,000 records with associated location data for 510 species of water beetles, including all available geographical and biological data from the literature up to 2006, as well as from museum and private collections, PhD theses, and other unpublished sources. The database was originally referenced at a resolution of 100 km² (10 km × 10 km cells). However, for simplicity, and more importantly, due to the lack of exhaustiveness in the survey effort, 50 km × 50 km UTM cells were used here as geographical units ($n=257$) (see details in Sánchez-Fernández et al. 2008b).

Adequately surveyed cells

In order to assess the performance of species richness distribution models according to different completeness thresholds, completeness values were calculated using the asymptotic value of the Clench function on the accumulated number of database records (see Soberón & Llorente, 1993). Many studies suggest the appropriateness of using the number of database records as a surrogate of sampling effort (see Lobo 2008b). The ratio of recorded to asymptotic predicted species richness was used as a measure of the completeness of each cell inventory. The Clench species-accumulation function was chosen because their fit allows the estimation of relatively high asymptotic values and so more exigent degrees of completeness (Díaz-Francés & Soberón, 2005). Here, seven completeness thresholds were utilised (50%, 55%, 60%, 65%, 70%, 75%, and 80% of total Clench predicted values) in order to select those cells that could be considered adequately surveyed (see Fig. 2). With the thresholds mentioned above, the number of cells that can be considered as well surveyed were 121, 108, 92, 80, 56, 42 and 26, respectively. The number of species in each set of selected

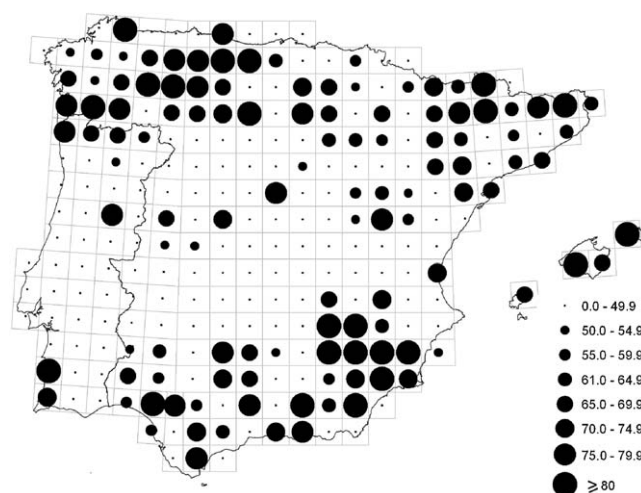


Fig. 2. Variation in completeness values (number of species recorded/number of species estimated by Clench function). Dot sizes indicate the different thresholds used to consider a cell as well surveyed.

cells as well as the values estimated by the asymptotic value of the Clench accumulation curves was used as a dependent variable in the subsequent modelling procedure (see Hortal et al. 2004).

Species richness modelling

Species richness predictions were obtained by using 18 environmental variables as potential predictors. For each cell (2500 km²), nine climatic variables (minimum and maximum monthly mean temperatures, mean annual temperature, total annual rainfall, summer precipitation, mean percentage of sunny hours per year, aridity, annual range of temperature variation and annual precipitation variation); four topographic variables (minimum, maximum and mean altitude, and elevation range), four lithological variables (percentage of area with clay, calcareous, and siliceous substrates, and lithological diversity) and one variable related with the total surface of water bodies in each cell were used. All these data come from the EDIT-Geoplatform (Sastre et al. 2009) which is freely accessible at <http://edit.csic.es/GISdownloads.html>. The scores of these variables at each 50 × 50 UTM square in the Iberian Peninsula were extracted and standardised (to 0 mean and 1 standard deviation) and then submitted to a Principal Component Analysis (PCA) in order to obtain uncorrelated factors (Varimax rotation). Five factors were selected with eigenvalues above 1 which accounted for 77% of total variability. Factor 1 (F1) was positively related with mean altitude (factor loading = 0.91) and negatively with minimum monthly mean temperature (−0.90); F2 is positively related with mean annual rainfall (0.89); F3 is negatively related with the percentage of calcareous soils (−0.76); F4 is related positively with maximum monthly mean temperature (0.59), and F5 positively with the percentage of clay soils (0.53). The scores of these five factors were extracted and subsequently used as predictors in a GLM modelling procedure (McCullagh & Nelder, 1989). A Poisson error distribution was assumed for the dependent variables (the number of species), related to the set of predictor variables via a logarithmic link function. To account for curvilinear relationships, the quadratic and cubic functions of each PCA factor were included in the regression model. Subsequently, a stepwise procedure was used to enter the factors into the model (Lobo & Martín-Piera, 2002; Nicholls 1989). First, the linear, quadratic or cubic function of the environmental factor that accounted for the most important change in deviance was entered. The remaining variables were added to the model sequentially according to their estimated explanatory capacity. The procedure was repeated iteratively until no more statistically significant explanatory variables remained ($p \leq 0.05$). At each step, the statistical significance of the terms already selected were tested by submitting the new model to a backward selection procedure. The terms that became non-significant in this step were then excluded. We also used Akaike's Information Criterion (AIC) to compare the model selected by this stepwise procedure with the other models composed by all possible subsets of predictors. The STATISTICA 6.1 package (StatSoft Inc. 2004) was used for all statistical computations.

The final model was evaluated by detecting outliers (cells with residual absolute values higher than mean ± standard deviation) and calculating the potential leverage (a measure of the distance of each observation from the centroid of the multi-dimensional space defined by the variables included in the model; Nicholls 1989). Thus, residuals were examined to determine whether they were due to erroneous data or to the environmental uniqueness of the cells.

The predictive power of the final model was estimated by a Jackknife procedure; model parameters were estimated as many times as the number of those cells considered to be adequately surveyed (n), deleting each cell once, and comparing the thus obtained predicted values against the values of the depen-

dent variable (observed richness or asymptotic species richness). The percentage of error for each cell value was subsequently calculated and the mean error percentage (MPE) for all the cells used to estimate the predictive error of competing model results.

Final model residuals were also checked for autocorrelation using Moran's I test (Legendre & Legendre, 1998). Whenever, as a result of this analysis, any spatial structure could be seen to remain in the residuals, such autocorrelation was taken to indicate the existence of at least one further variable not included in the analysis, with a spatially structured effect on species richness (Diniz-Filho et al. 2003).

Results

Predicting species richness

Model predictions depend on the completeness threshold used to discriminate the well-surveyed cells used as dependent variable. Explained deviance oscillates from 12.1% to 57.5%, the highest values being those generated when the data of the cells with higher completeness percentages (80%) and observed species richness values were used (Fig. 3a). Mean error percentages did not significantly differ between competing models (Fig. 3b), although the lower differences between observed and predicted Jackknife values also suggest that better model predictions are those obtained when the dependent variable is chosen by the most restrictive completeness threshold.

Therefore, we selected the model that was able to explain the highest percentage of deviance which also has the lowest AIC value: the model built using those cells with the highest completeness values (80%) and the observed species richness values as dependent variable. The final model was: $S = \text{EXP}(0.11F1 - 0.01F1^2 - 0.03F2 + 0.04F3 + 0.05F3^2 - 0.07F4 + 0.05F4^2 - 0.08F5^2)$. This model was able to explain 57.5% of total species richness variability with a mean error percentage of 29.9% (95% confidence interval from 23.13% to 36.73%). Model residuals are not significantly autocorrelated at any one of the seven distance classes (lag = 80 km).

Species richness distribution pattern

When this simple final model was applied to the entire study area, the simulated geographic distribution pattern shows that high species richness areas can be found across the whole Iberian territory (Fig. 4a). According to the selected model, the richness hotspots of water beetles in the Iberian Peninsula are mainly concentrated in the eastern part, corresponding with areas of medium-high altitude far from the coast, but also in high-altitude mountain areas (Fig. 4a). Several main areas (see Figs. 1 and 4) can be distinguished: the Pyrenees; Central System; the Central System (mainly Sierra de Guadarrama); Sierra Morena Mountains; Montes de Toledo; and Baetic Systems, including the Sierra de Cazorla y Alcaraz, but excluding the highest zones of the Sierra Nevada Mountains.

The predicted species richness values differed greatly from the observed ones (Spearman rank correlation coefficient; $r_s = -0.04$; $p = 0.45$), and were also uncorrelated with the number of database records ($r_s = -0.09$; $p = 0.15$).

Identifying future survey areas

Future sampling efforts must be focused on the most probable species-rich regions; i.e., those comprised in the upper quartile of predicted species richness (Fig. 4a) that at the same time are insufficiently surveyed, i.e., those located in the lower quartile of raw number of records (Fig. 4b). According to our results, six main areas

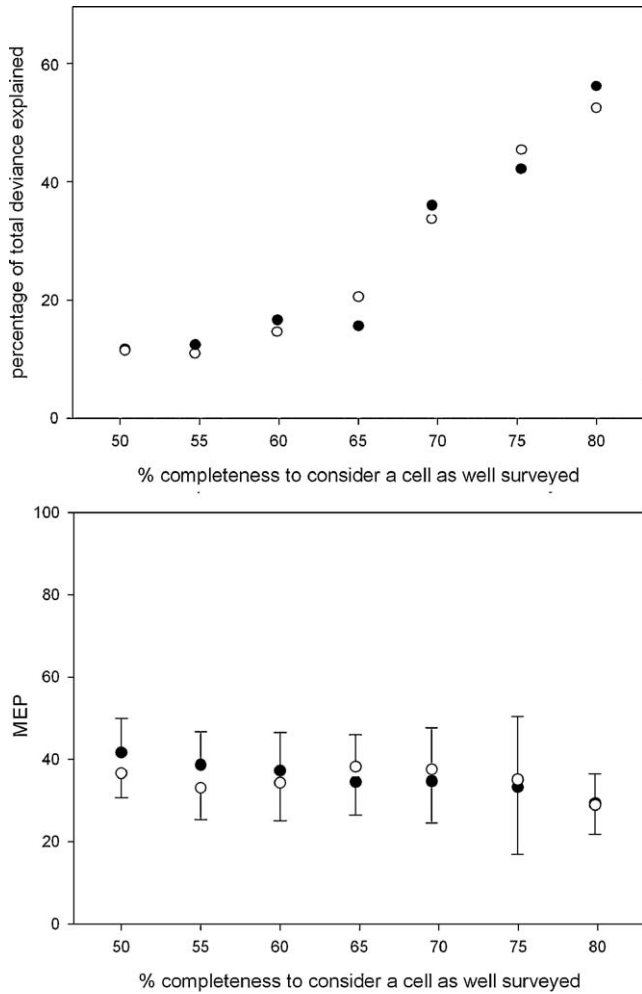


Fig. 3. (a) Differences in the deviance explained by different models depending on the completeness threshold selected to consider a cell as adequately surveyed and using observed richness (black circles) and estimated richness by accumulation curves (Clench function) as dependent variable (white circles). (b) Mean error percentage (MEP) and confidence interval (95%) depending on the threshold used to consider a cell as adequately surveyed and using observed richness (white circles) and estimated richness by accumulation curves (Clench function) as dependent variable (black circles).

can be identified (see Fig. 5) (A): a set of cells in the northern plateau including sites as Arribes del Duero, Riberas del Tormes, Hoces del Duratón and Río Arlanzón; (B) a set of cells in the southern Iberian plateau including the Montes of Toledo, the North face of Sierra Morena, and the Tietar and Alberche valleys; and four isolated cells including (C) Marismas of Santoña in the Cantabrian coast, (D) Alt Pallars in prePyrenees; (E) Montes of Arba and Río Zuera in the Ebro Valley; and (F) Sierras del Serral y Salinas in the Southeast Spain.

Discussion

Unevenness in sampling effort may result in partial (and biased) descriptions of biodiversity variation (Dennis 2001), a common drawback that limits the usefulness of existing databases and/or atlases to accurately describe biodiversity patterns (Dennis & Shreeve, 2003; Hortal et al. 2007; Prendergast et al. 1993; Soberón et al. 2007). These drawbacks seriously compromise the usefulness of any distribution predictive model used as a basis for conservation purposes (Hortal et al. 2007; Lobo et al. 2007). However, such predictions might be used to select the location of future surveys

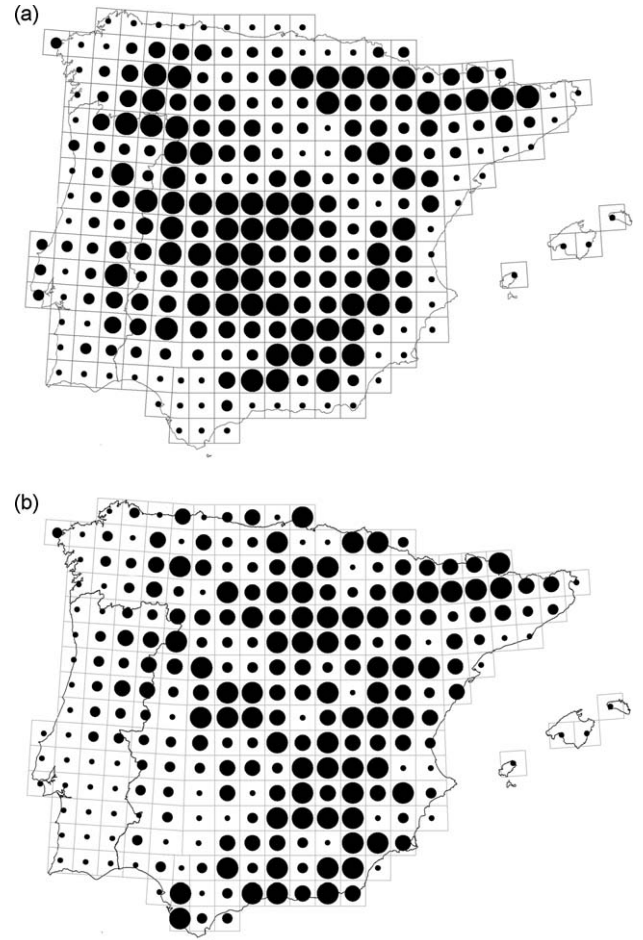


Fig. 4. Predicted species richness of water beetles in the Iberian Peninsula and Balearic Islands (a) and number of database records (b) in the 50 km × 50 km Iberian UTM cells. The varying diameter of symbols is proportional to the values of each variable on a scale of four categories (quartiles).

by prioritising those species rich-areas with a low level of sampling effort.

Selecting higher completeness values guarantees the appropriate choice of well-surveyed cells, but diminishes the number of cells eligible for use in a modelling procedure as well as their environmental coverage in the study area. According to the results of this study, we recommend to be rigorous in setting the thresh-

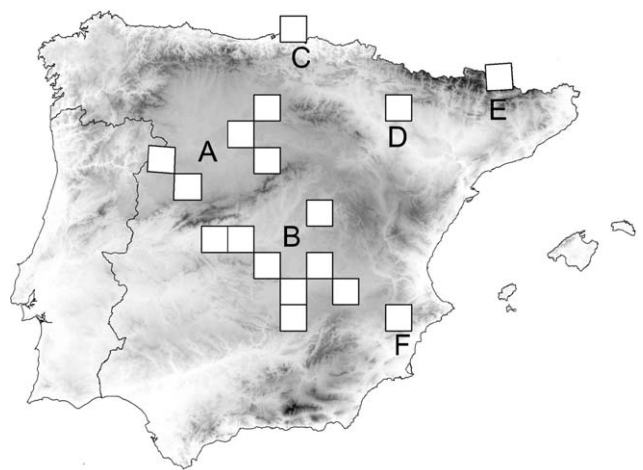


Fig. 5. Interesting areas to locate future sampling programmes.

olds for considering cells as well surveyed, thereby avoiding the inclusion of potentially poorly surveyed cells. However, the simulated species richness pattern generated by an exigent selection of well-surveyed cells should be interpreted with caution because of its lower geographical and environmental representativeness. This model is able to explain a moderate percentage of total species richness variability (57%) similar to that obtained in an analogous study performed on dung beetles (62%; see Lobo & Martín-Piera, 2002), the other Iberian insect group for which a similar modelling procedure has been carried out (see also Hortal et al. 2001; Hortal et al. 2004). Thus, the prediction of our species richness model is probably far from being accurate; it only considers the data from 26 cells out of 257 and, more importantly, it possesses a relatively high mean prediction error (30%) compared with that obtained for Iberian dung beetles (16%). All of these results suggest that this distributional proposal should be considered as a preliminary step towards a more definitive one, which will be made possible when a higher number of database records from some of the insufficiently surveyed cells is available. However, the value of our results is in pinpointing the areas in need of further surveys.

A former study (Sánchez-Fernández et al. 2008b) showed that cells with more than 70% of the theoretical species richness actually recorded are not evenly distributed across biogeographical and physioclimatic Iberian regions. Such bias, together with the limited predictive power of the species richness model developed here, highlight the imperative necessity of more reliable data before using such kind of modelling techniques on insect groups for conservation or biogeographical purposes, even at coarse resolutions (cells of 50 km × 50 km). The database used in this study does not seem to be so exhaustive; 50,000 database records on 510 species (around 100 records per species) does not seem to be enough to have an adequate number of training localities that are well-distributed across the whole spectrum of Iberian environmental conditions. Even at this wide resolution, approximately half of the Iberian territory remains characterised by a remarkable scarcity of water beetle records, with <50% of the predicted species recorded. Thus, a methodological conclusion of these results is that statistics cannot always efficiently overcome a data shortage. These modelling procedures might simply provide a preliminary approach for the majority of invertebrate groups, a higher survey effort being necessary to improve the accuracy of their predictions, even when all of the available information is compiled.

The biogeographical diversity patterns for the majority of insects groups reflect the distribution of the areas investigated by entomologists (Dennis & Hardy, 1999), and, consequently, the richest areas may not necessarily be the most intensely surveyed ones. The results of this study show that the predicted richness patterns of Iberian water beetles differed greatly from the observed values, suggesting that in this case, entomologists' intuition could have failed in deciding where to focus their sampling efforts. As Sánchez-Fernández et al. (2008b) pointed out, researchers have tended to sample water beetles more intensely in accessible sites near their research centres, and have selected the study sites based on the presence of interesting species, and/or mountainous and protected areas, and ignoring other, less attractive regions with a high predicted richness. In this sense, these results provide a basis for the design of future sampling efforts, since they allow the identification of genuinely under-sampled and potentially species-rich regions. Even considering that Coleoptera are one of the best surveyed groups of freshwater biodiversity in the Iberian Peninsula, it is necessary to significantly increase the sampling effort in the detected areas to accurately describe the spatial distribution of these insects. Furthermore, new field work data from less-inventoried regions will allow us to validate the obtained model and continually improve the predicted figures (Hortal & Lobo, 2005; Lobo et al. 2004). The sampling effort to validate and improve this

model must be focused on the areas mentioned above (Fig. 5). The scarcity of water bodies in these areas, together with their perceived low level of attractiveness, could explain the low sampling effort invested in them to date (Sánchez-Fernández et al. 2008b).

Some conservation implications

Our results provide a preliminary picture of the distribution of water beetles species richness in the Iberian Peninsula and Balearic Islands, a region that harbours a high diversity and is also subjected to intense human pressures. Water beetles have been identified as excellent surrogates of wider inland water biodiversity (Bilton et al. 2006; Sánchez-Fernández et al. 2006). Therefore, species richness patterns from this study are likely to be similar to those of other less well-known groups of freshwater macroinvertebrates. Although preliminary, these results show that some of the species richness hotspots identified here are located in areas that are particularly threatened, especially those distributed at low and medium altitudes. These are the areas under the most intense human pressures which are subjected to frequent changes in land use (Gómez et al. 2005; Martínez-Fernández et al. 2000), via dredging and stream canalisation, drainage, urbanisation development, pollution, and loss of salinity (Millán et al. 2006; Velasco et al. 2006). The extent to which the current species richness could have been affected by these processes is unknown, as there is no suitable reference data. Thus, any intensification of survey efforts in these areas can provide lower species richness values than expected because of the relatively recent impoverishment of their natural assemblages. On the other hand, the predicted hotspots in mountainous areas are in general more well preserved, since they are under less intensive management and seem to be more appropriate for testing the validity of our conclusions.

It is clear that the sole reliance on maps of predicted species richness is not a sound strategy for identifying areas for conservation, even when models are apparently accurate (Ferrier 2002; Hortal et al. 2004). To decide where and how to locate protected areas, other biodiversity components (e.g. endemic, rarity, vulnerability) must also be estimated and used in combination with biogeographical and ecological information (Margules & Pressey, 2000). Nevertheless, if we aim to use the distributional information of hyperdiverse groups, such as insect species, to examine the effectiveness of current protected areas or to propose new additional areas for conservation (see Araújo et al. 2007), a major requisite is knowing where the important regions from a conservation point of view are located but not selected as conservation targets due to insufficient taxonomical information. In our case, the use of a species richness prediction map as a surrogate of the biodiversity relevance of each cell could be legitimated by the frequent positive correlation between water beetles species richness and rarity (Sánchez-Fernández et al. 2004), as well as by the lack of reliable information of the "true" composition of each cell. In Mediterranean countries this kind of information is not currently available for the vast majority of insects, although it is of great interest in further conservation work. Therefore, if we aim to develop more accurate conservation strategies by incorporating invertebrate data we imperatively need reliable information on the distributional patterns of these groups. Our message is that more focused sampling efforts need to be carried out, and predictive model techniques could be an effective and useful tool for designing the required sampling protocols.

Acknowledgements

This work was partially supported by projects CGL2006-04159 (MEC), O23/2007 (MMA) and CGL2006-10196/BOS (MEC). We

thank P. Aguilera, A. Castro, H. Fery, J. Fresneda, G. Foster, J. Garrido, C. Hernando, and L.F. Valladares for unpublished records of Iberian water beetles. We also thank the Ecología Acuática group (University of Murcia) for their help in the compilation of the data and I. Ribera, F. Picazo and J. Velasco for multiple comments on previous stages of the manuscript.

References

- Abellán, P., Sánchez-Fernández, D., Velasco, J., & Millán, A. (2005). Conservation of freshwater biodiversity: A comparison of different area selection methods. *Biodiversity and Conservation*, *14*, 3457–3474.
- Allan, J. D., & Flecker, A. S. (1993). Biodiversity conservation in running waters. *Bioscience*, *43*, 32–43.
- Araújo, M. B., Lobo, J. M., & Moreno, J. C. (2007). The effectiveness of Iberian protected areas for conserving terrestrial biodiversity. *Conservation Biology*, *21*, 1423–1432.
- Bilton, D. T., McAbendroth, L., Bedford, A., & Ramsay, P. M. (2006). How wide to cast the net? Cross-taxon congruence of species richness, community similarity and indicator taxa in ponds. *Freshwater Biology*, *51*, 578–590.
- Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London B*, *354*, 101–118.
- Darwall, W. R. T., & Vié, J. C. (2005). Identifying important sites for conservation of freshwater biodiversity: Extending the species-based approach. *Fisheries Management and Ecology*, *12*, 287–293.
- Dennis, R. L. H. (2001). Progressive bias in species status is symptomatic of fine-grained mapping units subject to repeated sampling. *Biodiversity and Conservation*, *10*, 483–494.
- Dennis, R. L. H., & Hardy, P. B. (1999). Targeting cells for survey: Predicting species richness and incidence for a butterfly atlas. *Global Ecology and Biogeography*, *8*, 443–454.
- Dennis, R. L. H., & Shreeve, T. G. (2003). Gains and losses of French butterflies: Test of predictions, under-recording and regional extinction from data in a new atlas. *Biological Conservation*, *110*, 131–139.
- Díaz-Francés, E., & Soberón, J. (2005). Statistical estimation and model selection of species-accumulation functions. *Conservation Biology*, *19*, 569–573.
- Diniz-Filho, J. A. F., Bini, L. M., & Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, *12*, 53–64.
- Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? *Systematic Biology*, *51*, 331–363.
- Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, *43*, 393–404.
- Gómez, R., Hurtado, I., Suárez, M. L., & Vidal-Abarca, M. R. (2005). Ramblas in south-east Spain: Threatened and valuable ecosystems. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *15*, 387–402.
- Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, *4*, 379–391.
- Griffiths, G. H., Eversham, B. C., & Roy, D. B. (1999). Integrating species and habitat data for nature conservation in Great Britain: Data sources and methods. *Global Ecology and Biogeography*, *8*, 329–345.
- Guralnick, R. P., Hill, A. W., & Lane, M. (2007). Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters*, *10*, 663–672.
- Hortal, J., Garcia-Pereira, P., & García-Barros, E. (2004). Butterfly species richness in mainland Portugal: Predictive models of geographic distribution patterns. *Ecography*, *27*, 68–82.
- Hortal, J., & Lobo, J. M. (2005). An ED-based protocol for the optimal sampling of biodiversity. *Biodiversity and Conservation*, *14*, 2913–2947.
- Hortal, J., Lobo, J. M., & Jiménez-Valverde, A. (2007). Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife, Canary Islands. *Conservation Biology*, *21*, 853–863.
- Hortal, J., Lobo, J. M., & Martín-Piera, F. (2001). Forecasting insect species richness scores in poorly surveyed territories: The case of the Portuguese dung beetles (Col. Scarabaeinae). *Biodiversity and Conservation*, *10*, 1343–1367.
- Koellner, T., Hersperger, A. M., & Wohlgemuth, T. (2004). Rarefaction method for assessing plant species diversity on a regional scale. *Ecography*, *27*, 532–544.
- Lawton, J. H., Prendergast, J. R., & Eversham, B. C. (1994). The numbers and spatial distributions of species: Analyses of British data. In P. L. Forey, C. J. Humphries, & R. I. Vane-Wright (Eds.), *Systematics and Conservation Evaluation* (pp. 177–194). Oxford: Clarendon Press.
- Legendre, P., & Legendre, L. (1998). *Numerical ecology* (2nd ed.). Amsterdam: Elsevier Science.
- Lobo, J. M. (2008a). More complex distribution models or more representative data? *Biodiversity Informatics*, *5*, 15–19.
- Lobo, J. M. (2008b). Database records as a surrogate for sampling effort provide higher species richness estimations. *Biodiversity and Conservation*, *17*, 873–881.
- Lobo, J. M., Baselga, A., Hortal, J., Jiménez-Valverde, A., & Gómez, J. F. (2007). How does the knowledge about the spatial distribution of Iberian dung beetle species accumulate over time? *Diversity and Distributions*, *13*, 772–780.
- Lobo, J. M., Jay-Robert, P., & Lumaret, J. P. (2004). Modelling the species richness for French Aphodiidae (Coleoptera, Scarabaeoidea). *Ecography*, *27*, 145–156.
- Lobo, J. M., & Martín-Piera, F. (2002). Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conservation Biology*, *16*, 158–173.
- Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, *405*, 243–253.
- Martínez-Fernández, J., Esteve-Selma, M. A., & Calvo-Sendín, J. F. (2000). Environmental and socio-economic interaction in the evolution of traditional irrigate lands. A dynamic system model. *Human Ecology*, *28*, 279–299.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London: Chapman and Hall.
- Millán, A., Abellán, P., Ribera, I., Sánchez, D., & Velasco, J. (2006). The Hydradephaga (Coleoptera) of the Segura basin (SE Spain): Twenty five years studying water beetles. *Memorie della Società Entomologica Italiana*, *85*, 137–158.
- Millán, A., Moreno, J. L., & Velasco, J. (2002). *Estudio faunístico y ecológico de los coleópteros y heterópteros acuáticos y semiacuáticos de la provincia de Albacete*. Albacete: Instituto de Estudios Albacetenses.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–858.
- Nakamura, M., & Soberón, J. (2009). Use of approximate inference in an index of completeness of biological inventories. *Conservation Biology*, *23*, 469–474.
- Nicholls, A. O. (1989). How to make biological surveys go further with generalised linear models. *Biological Conservation*, *50*, 51–75.
- Petersen, F. T., & Meier, R. (2003). Testing species-richness estimation methods on single-sample collection data using the Danish Diptera. *Biodiversity and Conservation*, *12*, 667–686.
- Prendergast, J. R., Wood, S. N., Lawton, J. H., & Eversham, B. C. (1993). Correcting for variation in recording effort in analyses of diversity hotspots. *Biodiversity Letters*, *1*, 39–53.
- Ramos, M. A., Lobo, J. M., & Esteban, M. (2001). Ten years inventorying the Iberian fauna: Results and perspectives. *Biodiversity and Conservation*, *10*, 19–28.
- Ribera, I. (2000). Biogeography and conservation of Iberian water beetles. *Biological Conservation*, *92*, 131–150.
- Ribera, I., Foster, G. N., & Vogler, A. P. (2003). Does habitat use explain large scale diversity patterns in European water beetles? *Ecography*, *26*, 145–215.
- Ribera, I., Hernando, C., & Aguilera, P. (1998). An annotated checklist of the Iberian water beetles (Coleoptera). *Zapateri*, *8*, 43–111.
- Ricciardi, A., & Rasmussen, J. B. (1999). Extinction rates in North American freshwater fauna. *Conservation Biology*, *13*, 1220–1222.
- Romo, H., García-Barros, E., & Lobo, J. M. (2006). Identifying recorder-induced geographic bias in an Iberian butterfly database. *Ecography*, *29*, 873–885.
- Rosenzweig, M. L., Turner, W. R., Cox, J. G., & Ricketts, T. H. (2003). Estimating diversity in unsampled habitats of a biogeographical province. *Conservation Biology*, *17*, 864–874.
- Sánchez-Fernández, D., Abellán, P., Velasco, J., & Millán, A. (2004). Selecting areas to protect the biodiversity of aquatic ecosystems in a semi-arid Mediterranean region. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *14*, 465–479.
- Sánchez-Fernández, D., Abellán, P., Mellado, A., Velasco, J., & Millán, A. (2006). Are water beetles good indicators of biodiversity in Mediterranean aquatic ecosystems? The case of the Segura river basin (SE Spain). *Biodiversity and Conservation*, *15*, 4507–4520.
- Sánchez-Fernández, D., Bilton, D. T., Abellán, P., Velasco, J., Ribera, I., & Millán, A. (2008a). Are the endemic water beetles of the Iberian Peninsula and the Balearic Islands effectively protected? *Biological Conservation*, *141*, 1612–1627.
- Sánchez-Fernández, D., Lobo, J. M., Abellán, P., Ribera, I., & Millán, A. (2008b). Bias in freshwater biodiversity sampling: The case of Iberian water beetles. *Diversity and Distributions*, *14*, 754–762.
- Sastre, P., Roca, P., & Lobo, J. M. (2009). A Geoplatform for improving accessibility to environmental cartography. *Journal of Biogeography*, *36*, 568.
- Saunders, D. L., Meeuwig, J. J., & Vincent, A. C. J. (2002). Freshwater protected areas: Strategies for conservation. *Conservation Biology*, *16*, 30–41.
- Soberón, J., & Peterson, T. (2004). Biodiversity informatics: Managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society of London B*, *359*, 689–698.
- Soberón, J., Jiménez, R., Golubov, J., & Koleff, P. (2007). Assessing completeness of biodiversity databases at different spatial scales. *Ecography*, *30*, 152–160.
- Soberón, J., & Llorente, J. (1993). The use of species accumulation functions for the prediction of species richness. *Conservation Biology*, *7*, 480–488.
- StatSoft Inc. (2004). *STATISTICA (data analysis software system)*. Tulsa: StatSoft Inc.
- Strayer, D. L. (2006). Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society*, *25*, 271–287.
- Velasco, J., Millán, A., Hernández, J., Gutiérrez, C., Abellán, P., Sánchez, D., & Ruiz, M. (2006). Response of biotic communities to salinity changes in a Mediterranean hypersaline stream. *Saline Systems*, *2*, 1–15.
- Williams, P., Humphries, C., Araújo, M., Lampinen, R., Hagemeyer, W., Gasc, J.-P., & Mitchell-Jones, T. (2000). Endemism and important areas for conserving European biodiversity: a preliminary exploration of atlas data for plants and terrestrial vertebrates. *Belgian Journal of Entomology*, *2*, 21–46.