

Predicting the richness of aquatic beetles and bugs in a semi-arid mediterranean region

Daniel Bruno^{1,*}, David Sánchez-Fernández^{1,2}, José Antonio Carbonell¹, Félix Picazo¹, Josefa Velasco¹ and Andrés Millán¹

¹ (Departamento de Ecología e Hidrología, Universidad de Murcia, 30100 Murcia, Spain).

² (Museo Nacional de Ciencias Naturales, C/José Gutiérrez Abascal, 2, 28006, Madrid, Spain).

* Corresponding author: dbrunocollados@um.es

Received: 20/10/2010

Accepted: 8/6/2011

ABSTRACT

Predicting the richness of aquatic beetles and bugs in a semi-arid mediterranean region

The southeastern Iberian Peninsula is a semi-arid region recognised as an area of high aquatic biodiversity. Water beetles (Coleoptera) and water bugs (Hemiptera) are two of the region's most thoroughly studied insect groups. An exhaustive database of aquatic beetles and bugs from the Region of Murcia (SE Spain) and a set of 33 variables were employed for two related purposes: a) to describe the sampling effort conducted in the study area on these groups in order to identify zones with reliable inventories and b) with this information as a base, to forecast the distribution of species richness for each group. Theoretical estimates were based on the Clench function applied to the accumulated number of samples and were constructed with general linear models (GLM). The following results were obtained: 1) the percentage of grid cells that can be considered to be well sampled in the study area (143 10 × 10 km UTM grid cells) is approximately 14 % for beetles and 22 % for bugs; 2) well-surveyed cells are evenly distributed according to the previously defined physioclimatic subareas, which enable the use of these grid cells for modelling the richness distribution; and 3) a more accurate picture of the actual richness patterns for both groups is shown by the analysis. These results were combined to identify key areas where future sampling efforts must be focused.

Key words: Sampling effort, predictive modelling, aquatic Coleoptera, aquatic Hemiptera, richness patterns.

RESUMEN

Prediciendo la riqueza de coleópteros y hemípteros acuáticos en una región semiárida mediterránea

El sureste de la Península Ibérica es una región semiárida con una alta biodiversidad acuática, siendo los coleópteros y hemípteros acuáticos dos de los grupos de insectos mejor estudiados. Una base de datos que recopila todas las citas existentes de ambos grupos en la Región de Murcia (SE de España) y un conjunto de 33 variables han sido utilizadas con los siguientes objetivos: a) describir el esfuerzo de muestreo llevado a cabo sobre estos grupos en el área de estudio para identificar zonas con inventarios fiables y usando como base esta información b) predecir la distribución de la riqueza para ambos grupos. Basándonos en estimaciones teóricas mediante el uso de la función de Clench sobre el número acumulado de muestras y modelos lineales generalizados (GLM), hemos encontrado que: 1) de las 143 unidades espaciales utilizadas (10 × 10 km UTM) en torno a un 14 % y un 22 % para coleópteros y hemípteros acuáticos, respectivamente, pueden considerarse bien muestreadas; 2) sin embargo, las celdas bien prospectadas se encuentran equitativamente distribuidas por las subregiones fisioclimáticas previamente definidas, lo que permite el uso de estas cuadrículas para modelar la distribución de la riqueza, 3) obteniéndose una imagen más cercana a los patrones de riqueza reales de ambos grupos. Combinando estos resultados, identificamos áreas clave donde deben ser localizados los futuros esfuerzos de muestreo.

Palabras clave: Esfuerzo de muestreo, modelo predictivo, coleópteros acuáticos, hemípteros acuáticos, patrones de riqueza.

INTRODUCTION

The importance of the Iberian Peninsula in terms of biodiversity conservation has been widely recognised (Domínguez-Lozano *et al.*, 1996; Williams *et al.*, 2000; Reyjol *et al.*, 2007). Approximately 80 % of its area is included in the 'Mediterranean hotspot list' (Myers *et al.*, 2000). The southeastern Iberian Peninsula is a region of particular interest due to its geographical location. It harbours species from Eurosiberian to northern African distributions. Therefore, although it is located in the most arid zone of Europe, it comprises a wide range of aquatic ecosystems, from freshwater streams, ponds and wetlands to hypersaline streams and continental and coastal salt pans (Vidal-Abarca *et al.*, 1992; Millán *et al.*, 1996). Many of these ecosystems are unique due to their ecological characteristics and the presence of rare and endemic species (Moreno *et al.*, 1997; Sánchez-Fernández *et al.*, 2004a; 2004b; Millán *et al.*, 2006). However, these aquatic ecosystems are subject to high rates of human alteration and to significant water shortages. Therefore, they are at risk for biodiversity loss (Allan & Flecker, 1993; Ricciardi & Rasmussen, 1999; Saunders *et al.*, 2002; Darwall & Vié, 2005).

To contribute to the effective conservation of these interesting and threatened ecosystems, it is desirable to rely on an exhaustive compilation of all faunal records, designed in a way that will allow the examination of the relationship between the distribution areas and their environmental and geographical determinants. Once the information is gathered, the assessment of the quality of the data is essential prior to any further analysis (Lobo & Martín-Piera, 2002; Hortal *et al.*, 2004; 2007).

Despite the general lack of survey-effort assessment (Hortal *et al.*, 2007), the quality of biodiversity databases from a variety of regions and covering a multitude of taxonomic groups is regularly evaluated (e.g., Dennis *et al.*, 1999; 2006; Lobo & Martín-Piera, 2002; Reddy & Dávalos, 2003; Romo *et al.*, 2006). To date, however, only one study of this type has focused on aquatic biodiversity (Sánchez Fernández *et*

al., 2008), probably due to the paucity of inventory data for freshwater systems (Lévêque *et al.*, 2005), especially in Mediterranean countries. Therefore, it is imperative to design sampling schemes that quickly and efficiently fill faunistic gaps in data that encompass the environmental and geographic variability of the study area (Funk *et al.*, 2005; Hortal & Lobo, 2005). These efforts are particularly important to identify poorly-surveyed areas and to study the bias associated with the available information. Complete inventories may be used to study the patterns and processes that have generated biodiversity and to develop reliable predictive distribution models. These models may aid in the selection of protected natural areas (Hortal *et al.*, 2001; Lobo & Martín-Piera, 2002; Hortal *et al.*, 2004; Ferrier & Guisan, 2006; Lobo, 2008) and facilitate the design of more efficient survey strategies (Sánchez-Fernández *et al.*, 2011).

In the study area, several authors studied species distributions, ecological traits and their use as biodiversity surrogates for both Coleoptera (Millán *et al.*, 1996; 2006; Moreno *et al.*, 1997; Sánchez-Fernández *et al.*, 2004a; 2004b; 2006; Abellán *et al.*, 2005; 2007) and Hemiptera (Millán *et al.*, 1988; Moreno *et al.*, 1997; Carbonell, 2008; 2010). The initial attempts to assess the sampling effort in the study area through use of the Clench function have suggested relatively complete inventories of water beetles (74 %) (Sánchez-Fernández *et al.*, 2004a) and water bugs (95 %) (Carbonell, 2010) on a regional scale. However, data on faunistics and information on occurrence and diversity patterns are probably incomplete and determined by the distribution of previous sampling efforts (Lobo & Martín-Piera, 2002; Hortal *et al.* 2007). The expansion of the knowledge of water beetle and bug distribution patterns enables their use as conservation tools. This principle is especially true for water beetles, which have been suggested by several authors as indicator taxa for monitoring population trends in other species and for identifying areas of high regional biodiversity (Davis *et al.*, 1987; Foster *et al.*, 1990; Foster & Eyre, 1992; Sánchez-Fernández *et al.*, 2004a; 2004b; Abellán *et al.*, 2005).

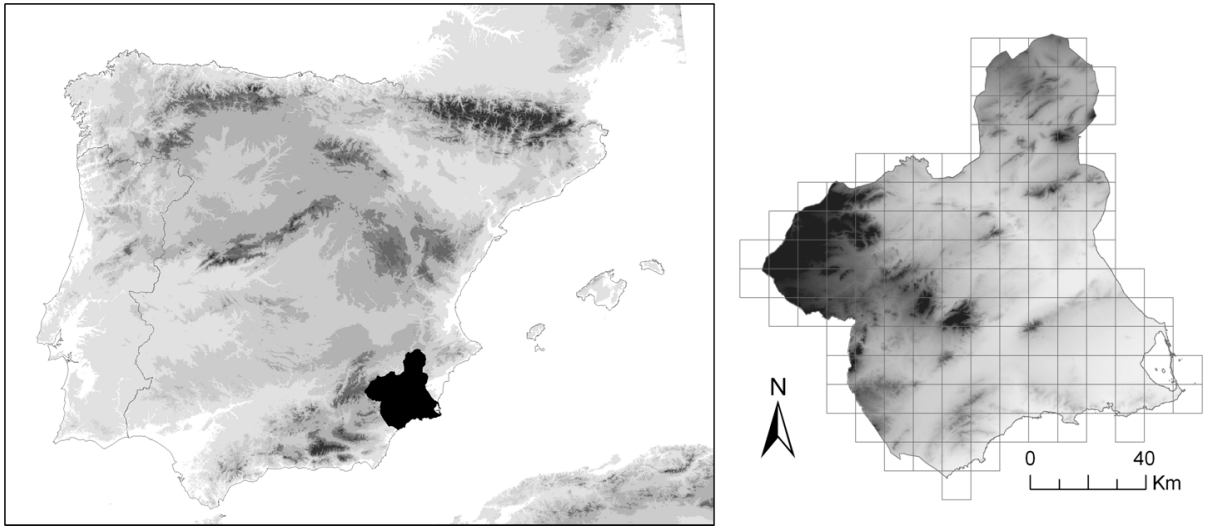


Figure 1. Study area (Region of Murcia) showing the 143 UTM 10×10 km squares studied. *Área de estudio (Región de Murcia), mostrando las 143 cuadrículas de 10×10 km UTM objeto de estudio.*

The main goal of this paper is to make simple predictions of the geographic distribution of the species richness of aquatic Coleoptera and Hemiptera in the southeastern Iberian Peninsula. To achieve this aim, an assessment of sampling effort was carried out in the study area. Only data for areas with reliable information were selected and the values estimated by the asymptotic value of the Clench accumulation curves were used as a dependent variable in the subsequent modelling procedure. Whether or not well-surveyed areas covered the environmental heterogeneity in the study area was also checked. By combining these results, we identified key areas where future sampling efforts must be focused.

METHODOLOGY

Study area

The study area corresponds to the 'Región de Murcia', located in the southeast of the Iberian Peninsula and encompassing an area of $11\,137\text{ km}^2$ (Fig. 1). This region has a Mediterranean climate, with an annual average rainfall of approximately 300 mm and an annual average temperature of 16°C . In a European context,

it is an area of high environmental heterogeneity and singularity. It includes a high diversity of bioclimatic stages, lithology and geology and consequently supports a wide range of aquatic ecosystems, from headwater streams to coastal wetlands. The various conservation statuses of these areas serve to increase the interest of the present study (Millán *et al.*, 2006). This paper focuses on the 143 UTM 10×10 km grid cells that overlap the 'Región de Murcia'. The defined cells are based on the Mercator Coordinate System.

Source of biological data

The records of aquatic bugs and beetles were compiled from an exhaustive database called BIODIVERSIDAD developed by the University of Murcia 'Ecología Acuática' group. This database is the most complete source of information available for aquatic invertebrates (water bugs and beetles at the specie level and the remaining aquatic macroinvertebrates at the family level) in the study area, including all available geographical and biological data from the literature up to 2009 and additional data from museum and private collections, PhD theses and other unpublished sources.

From this database, 3825 records for 162 species of water beetles (see species richness dis-

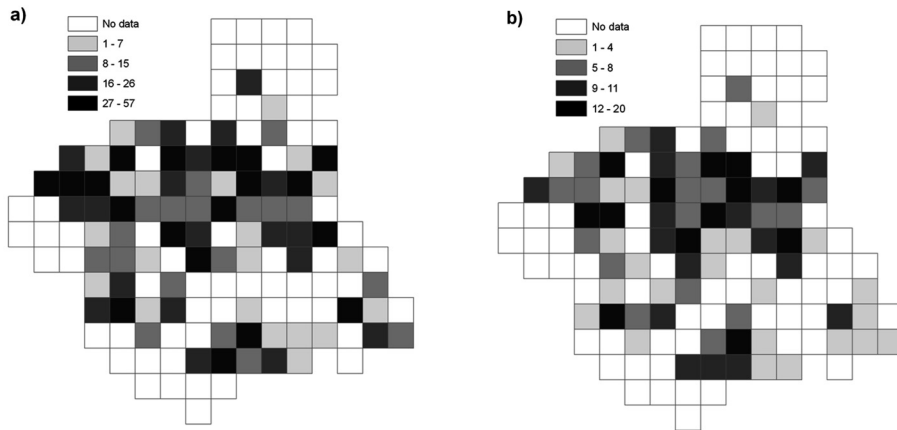


Figure 2. Observed richness of a) water beetles and b) water bugs. *Riqueza observada de a) coleópteros acuáticos y b) hemípteros acuáticos.*

tribution in figure 2a) and 1463 records for 33 species of water bugs (see species richness distribution in figure 2b) in the study area were collected. The data were referenced at a resolution of 100 km² (UTM 10 × 10 km squares) and these cells were used as geographical units ($n = 143$).

Assessing sampling effort and identifying well-surveyed cells

To determine the adequacy of sampling (degree of completeness) in each cell, the asymptotic value of the Clench function for the accumulated number of database samples was employed. The Clench function (see Soberón & Llorente, 1993) estimates the cell-specific mean rate of species addition per unit of sampling effort, using here the number of database samples as a sampling effort surrogate. The ratio between the asymptotic value of the Clench function and the observed richness was used as a measure of the completeness of each cell's inventory. Complementarily, a non-parametric estimator (Jackknife 1) was used to guarantee the accuracy and confidence of the selection procedure. This estimator performs best at low sampling effort (Mao and Colwell, 2005) and fine grain (Hortal *et al.* 2006).

Previous studies have recommended using a 70 % threshold as the basis for considering an area well sampled (Jiménez-Valverde & Hortal, 2003). However, we used a threshold of 65 %,

the maximum value that allows us to conduct further biodiversity studies. We felt that this choice was justified for the following reasons: i) the two complementary richness estimators facilitate the confident selection of well-surveyed cells; and ii) the available information is limited by the evident scarcity and paucity of data on non-charismatic groups in Mediterranean countries (Ramos *et al.*, 2001) and especially in freshwater ecosystems (Lévêque *et al.*, 2005). In general, the most restrictive approach has been the Clench function. Accordingly, the values estimated by the asymptotic value of the Clench accumulation curves for these well-sampled cells were used as dependent variables in the subsequent modelling procedure (see Hortal *et al.*, 2004).

Environmental representation by the well-surveyed cells

Unevenness in sampling effort may result in partial (and biased) descriptions of biodiversity variation (Dennis, 2001). This common drawback limits the usefulness of existing databases and/or atlases to accurately describe biodiversity patterns (Hortal *et al.*, 2007). Although several assessments of environmental bias and environmental completeness have been proposed (e.g., Kadmon *et al.*, 2003; 2004; Hortal *et al.*, 2008), we checked whether or not the previously recognised well-sampled cells represented the envi-

ronmental heterogeneity of the study area on a broad scale. To perform this determination, we defined physioclimatic subregions with a set of 28 variables (the topographic and climatic variables) (Table 1). All the data were obtained from the EDIT-Geoplatform (Sastre *et al.*, 2009),

Table 1. List of the variables used. *Lista de las variables utilizadas.*

Type	Variable
Topographic variables	
A1	Altitudinal range (m) (A3-A2)
A2	Minimum altitude (m)
A3	Maximum altitude (m)
A4	Mean altitude (m)
Climatic variables	
B1	Annual mean temperature (°C)
B2	Max temperature of warmest month (°C)
B3	Min temperature of coldest month (°C)
B4	Mean temperature of wettest quarter (°C)
B5	Mean temperature of driest quarter (°C)
B6	Mean temperature of warmest quarter (°C)
B7	Mean temperature of coldest quarter (°C)
B8	Temperature mean diurnal range (Mean of monthly (max temp-min temp)) (°C)
B9	Temperature annual range (B2-B3) (°C)
B10	Temperature seasonality (standard deviation *10)
B11	Isothermality (B8/B9)(*10)
B12	Average annual precipitation (mm)
B13	Precipitation of wettest month (mm)
B14	Precipitation of driest month (mm)
B15	Precipitation seasonality (Coefficient of Variation)
B16	Precipitation of driest quarter (mm)
B17	Precipitation of wettest quarter (mm)
B18	Precipitation of warmest quarter (mm)
B19	Precipitation of coldest quarter (mm)
B20	Monthly mean solar radiation (Mj/m ² /day)
B21	Aridity (B12/B23)
B22	Real evapotranspiration (mm)
B23	Potential evapotranspiration (mm)
B24	Hydrologic balance (mm)
Spatial variables	
C1	Longitude (UTM X)
C2	Latitude (UTM Y)
Lithologic variables	
D1	Lithological diversity (Shannon Índice)
Water bodies availability and diversity	
E1	Total water surface (m ²)
E2	Water environment diversity (Shannon Índice)

which is freely accessible at <http://edit.csic.es/GISdownloads.html>. A Principal Component Analysis (PCA) (with the scores of these variables at each 10×10 UTM square standardised to a mean of 0 and a standard deviation of 1) was performed to obtain uncorrelated factors (Varimax rotation) that compress the environmental variability into new factors that show eigenvalues above 1. Subsequently, CLUSTER and SIMPROF analyses (using Primer 6) (Clarke, 1993) were applied to the new variables or axes obtained from the PCA to elaborate a classification of the main physioclimatic subregions. Finally, a Chi-square test was performed to determine whether the well-surveyed cells presented a balanced distribution in the different defined subregions.

Modelling species richness

Spearman's correlations were used to identify highly correlated variables within the complete set of potential predictors (Table 1) to avoid redundant information in the analysis. From among each group of highly correlated variables ($R > 0.95$), those that simultaneously showed a clear biological meaning were selected. As a result of this screening procedure, 18 environmental variables were removed due to high correlations with other variables. The longitude, latitude, mean altitude, altitudinal range, temperature annual range, temperature mean diurnal range, maximum temperature in the warmest month, mean temperature in the wettest quarter, isothermality, precipitation in the wettest month, precipitation in the coldest quarter, monthly mean solar radiation, total water surface, water environment diversity and lithological diversity were selected as the variables suitable for use as potential predictors of the target-group distribution patterns.

The estimated species richness values calculated with the accumulation curves (only including well-surveyed cells) were regressed against the final selection of more representative environmental variables with general linear models (GLM, McCullagh & Nelder, 1989). A Poisson error distribution was assumed for the dependent variable (estimated number of species). The de-

pendent variable was related to the set of predictor variables via a logarithmic link function. To allow curvilinear relationships, a quadratic function of each factor was included in the regression model. A stepwise procedure was used to enter the factors into the model (Nicholls, 1989; Lobo & Martín-Piera, 2002). First, the linear or quadratic function of the environmental variable that accounted for the most important change in deviance was entered (a cubic term was included only for spatial variables). The remaining variables were added sequentially to the model according to their estimated explanatory capacity. The procedure was repeated until no more statistically significant explanatory variables remained ($p \leq 0.05$). The statistical significance of the terms that had already been selected was tested at each step by subjecting the new model to a backward selection procedure. The terms that became non-significant in this phase were then excluded. Akaike's Information Criterion (AIC) was also used to compare the model selected by this stepwise procedure with other models composed of all possible predictor subsets.

The final model was evaluated by detecting outliers (cells with residual absolute values higher than mean \pm 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, the residuals were examined to determine whether they were due to erroneous data or to the environmental uniqueness of the cells. Once the outliers had been removed and after having checked the balanced distribution of the final dataset, the model was rebuilt and applied. In addition, we identified the squares that remained with high residual values (values higher than mean \pm standard deviation).

The predictive power of the final model was estimated by a Jackknife procedure. The model parameters were estimated as many times as the number of cells considered to be adequately surveyed (n) by first deleting each cell once and then comparing the predicted values against the values of the dependent variables. The percentage of error for each cell value was subsequently

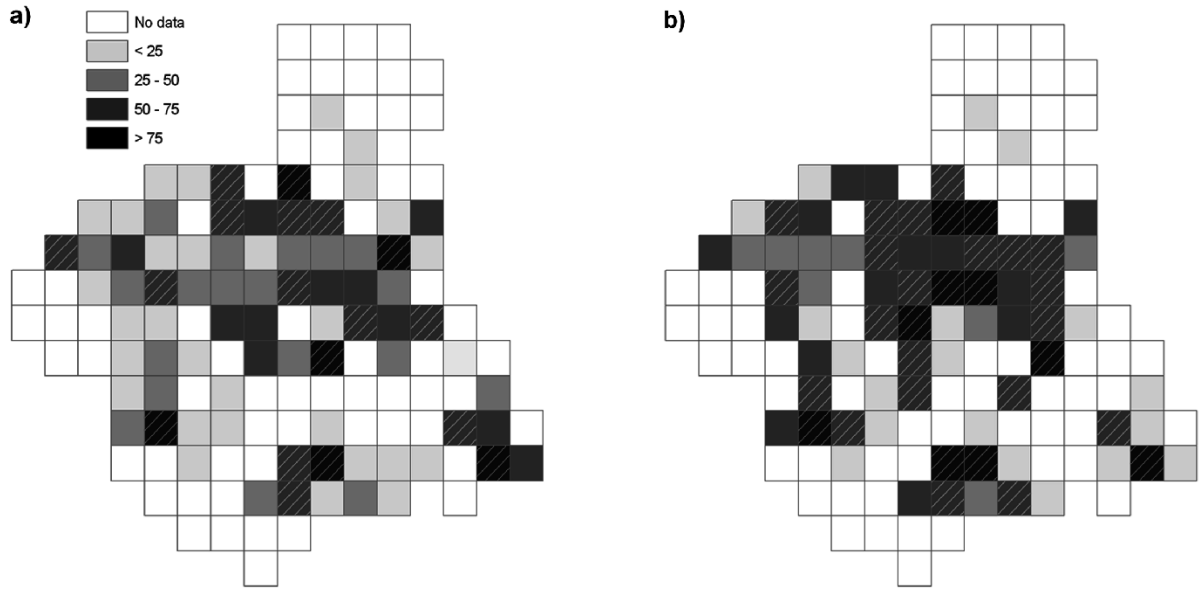


Figure 3. Completeness rate and well-surveyed cells (striped squares) for a) water beetles and b) water bugs. *Tasa de completitud y cuadrículas bien muestreadas (celdas con trama) de a) coleópteros acuáticos y b) hemípteros acuáticos.*

calculated to obtain the mean error percentage, which was used as a measure of model reliability (Lobo & Martín-Piera, 2002; Hortal *et al.*, 2004,

Sánchez-Fernández *et al.*, 2011). The STATISTICA 6.1 package (StatSoft Inc. 2004) was employed for all statistical computations.

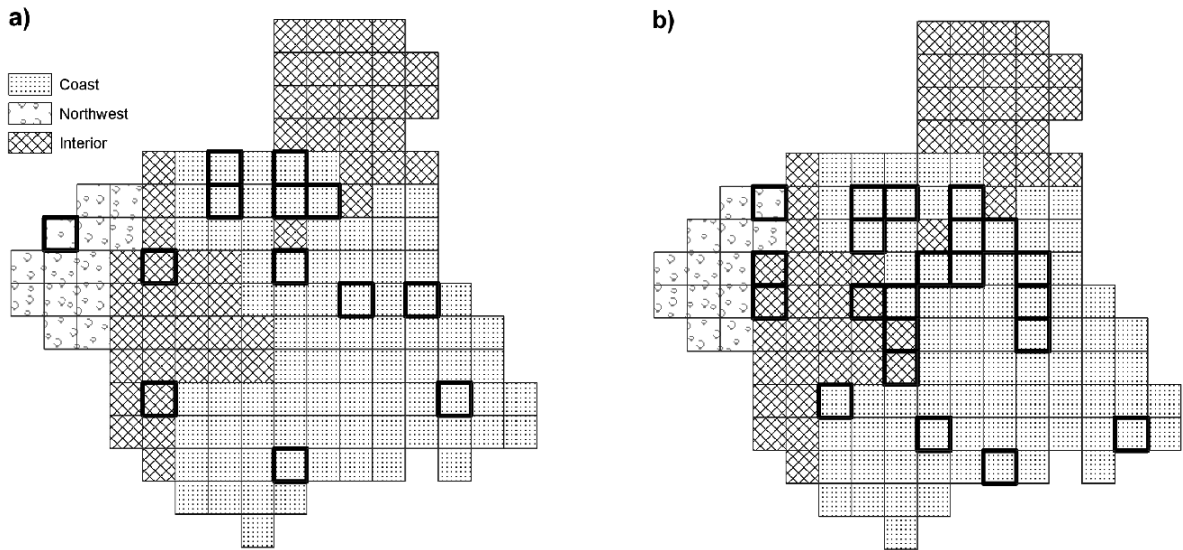


Figure 4. Distribution of squares used in the modelling procedure (marked cells) for the defined physioclimatic subregions for a) water beetles and b) water bugs. *Distribución de las celdas utilizadas para construir el modelo (celdas marcadas) en las subregiones fisioclimáticas definidas, para a) coleópteros acuáticos y b) hemípteros acuáticos.*

RESULTS

Well-surveyed cells

According to the Clench function, only 18 grid cells (10×10 km UTM) for aquatic beetles (Fig. 3a) can be considered to be well surveyed (23 % of the cells with data and 13 % of the whole study area). For bugs, 31 grid cells can be considered to be well surveyed (44 % of the cells with data and 22 % of the whole study area) (Fig. 3b).

Physioclimate subregions

The PCA compressed the variables that defined the established physioclimate subregions. The compressed description showed that three factors accounted for 89.8 % of the total variability. Factor 1 was positively related to the annual mean temperature and negatively related to the annual precipitation. Factor 2 was positively related to altitudinal range and Factor 3 was negatively related to the temperature mean diurnal range. The CLUSTER and SIMPROF analysis revealed three main physioclimate subregions (Fig. 4a and 4b). These three areas represent the climatic

and topographic gradients (NW-SE) present in the ‘Región de Murcia’, from the high, wet and cold northwest to the low, dry and hot southeast. The distribution of the final set of cells used to construct the model after the detection of outliers was balanced across the three obtained physioclimate subregions for both the beetle data ($p \leq 0.01$ and 999 permutations) (Chi-square test $\chi^2 = 2.262$; $p = 0.323$; d.f. = 2) and the bug data ($p \leq 0.01$ and 999 permutations) (Chi-square test $\chi^2 = 2.108$; $p = 0.348$; d.f. = 2).

Modelling species richness

The model explaining the highest percentage of deviance and with the lowest AIC value for water beetles was: $S = \text{EXP}(0.16 \text{ Altitudinal range}^2 + 0.34 \text{ Altitudinal range} + 0.14 \text{ Precipitation of the coldest quarter} + 0.12 \text{ Water environments diversity} + 3.74)$. This model could explain 68.1 % of the total variability of species richness, with a mean error percentage of 32.4 %. The model for bugs was as follows: $S = \text{EXP}(-0.44 \text{ Mean Temperature of wettest quarter}^2 - 0.29 \text{ Mean Temperature of wettest quarter} + 0.22 \text{ Water environment diversity} - 0.37 \text{ Total water surface} +$

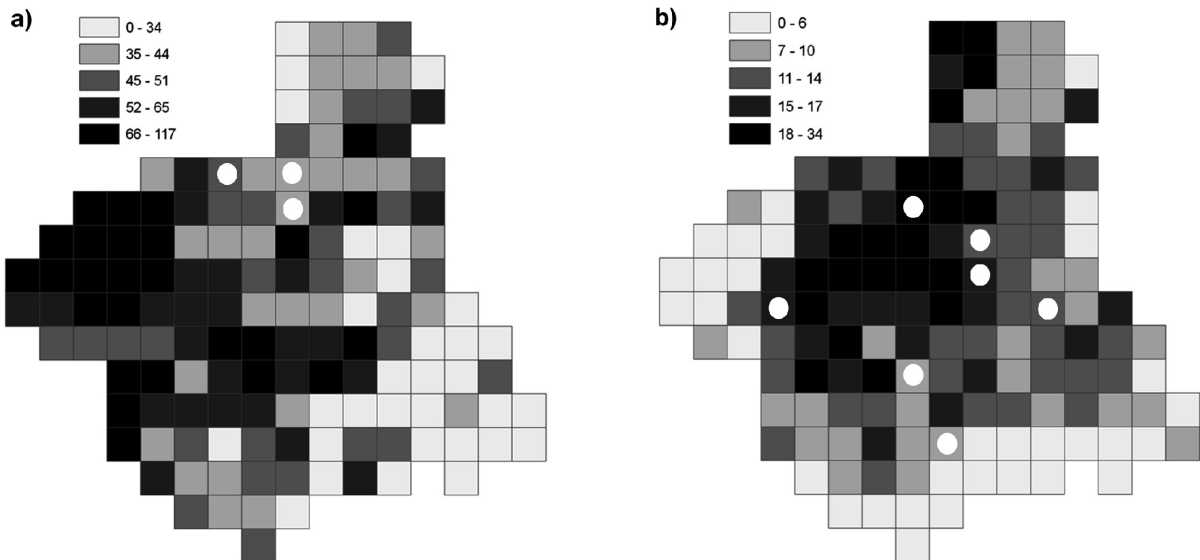


Figure 5. Predicted species richness and squares with high residual values (marked cells with white dots) for a) water beetles and b) water bugs. *Riqueza predicha de especies y cuadrículas con residuos altos (celdas marcadas con puntos blancos) para a) coleópteros acuáticos y b) hemípteros acuáticos.*

2.71). This model could explain 80.6 % of the total variability of species richness, with a mean error percentage of 38.2 %.

These simple final models were then applied to the entire study area. The simulated geographic distribution pattern showed a clear regional richness gradient, increasing from the southeast to the northwest for water beetles (Fig. 5a). The interior of the region was the most diverse area for water bugs (Fig. 5b). The squares that remained with high residual values are also shown in the figures.

DISCUSSION

Although both target groups are macroinvertebrates and are traditionally less well known, both groups showed reliable inventories on a regional scale. Nevertheless, these results suggest that less than a quarter (13 % for beetles and 23 % for water bugs) of the study area can presently be considered to be well surveyed (completeness values > 65 %). For similar levels of survey effort, the aquatic bug inventories are more complete due to their lower diversity. This lower diversity facilitates the construction of a good regional inventory based on fewer records. In this context, a general lack of complete and extensive inventory data for aquatic taxa (Lévêque *et al.*, 2005) at a fine scale (10 × 10 km squares) is apparent because it is probable that water beetles and bugs could be considered the best-studied groups of freshwater invertebrates in the Spanish Mediterranean region.

Based on the self-evident geographic pattern in the survey, it appears that additional effort is required in certain areas. The northeastern and northwestern parts of the region show an evident lack of surveys. Nevertheless, because only well-sampled cells were used to construct the model, this situation does not imply a lack of data reliability. These cells show a balanced distribution in relation to the environmental heterogeneity of the study area.

The predictive model obtained in this study identified the northwest zone as the most diverse area for aquatic beetles and showed an increasing richness gradient from the coast to

this zone. Although the most diverse areas for aquatic bugs are located in the interior of the region, the obtained potential distribution map does not show a clear richness distribution pattern. These predicted high-richness areas differ noticeably in their environmental features. The northwest of Murcia shows a high altitude, a steep orography and a supra-Mediterranean climate with more frequent rainfall and lower temperatures than the interior of the region, which has a meso-Mediterranean climate with a high thermic contrast and a medium altitude. The conservation status of the two areas differs markedly. The interior of the region (especially the northeast) has experienced a profound land use transformation that has negatively affected the quality of aquatic ecosystems. In contrast, the northwest remains the most natural area and boasts well-preserved ecosystems. Such ecosystems are generally infrequent in Mediterranean zones (Gasith & Resh, 1999).

The observed differences in the distribution patterns of both groups can be explained by an ecological difference. The ecological requirements of the majority of aquatic beetle species are much narrower than those of generalist aquatic bugs (Polhemus & Polhemus, 2008; Carbonell, 2010). In fact, water beetles have been identified as excellent surrogates for inland water biodiversity due to their highly specific ecological requirements and the significant correlations between their species richness patterns and those of other groups (Bilton *et al.*, 2006; Sánchez-Fernández *et al.*, 2006). Thus, the use of a species richness prediction map as a surrogate for the biodiversity relevance of each cell could be justified by the frequent positive correlation between water beetle species richness and rarity (Sánchez-Fernández *et al.*, 2004), which makes this group an interesting conservation tool. Although aquatic bugs have more complete inventories, their distribution pattern is much more difficult to interpret. It is probable that this difficulty results from their broad ecological requirements and high dispersal ability. Nevertheless, aquatic bugs seem to prefer the central zones of the region because these areas offer a combination of lotic and standing waters with some degree of human

impact. Such areas constitute preferred habitats for this group of insects (Millán *et al.*, 1988; 2002).

Although it is probable that the forecast obtained from this species richness model is far from accurate (it has a mean prediction error of approximately 30 %), the resulting potential distribution map shows a picture that could closely approximate reality. Furthermore, the results can be used to determine whether more intense sampling effort has been focused on the most diverse areas. In this case, Spearman's correlations between the observed richness and the number of records are highly positive for beetles ($R = 0.96$, $p < 0.01$) and bugs ($R = 0.88$, $p < 0.01$). However, the predicted richness patterns of water beetles and bugs differed greatly from the actual sampling effort ($R = 0.1$, $p = 0.37$) for beetles and ($R = 0.1$, $p = 0.41$ for bugs). This result suggests that entomologists' intuitions could have failed to select the best focus for their sampling effort in this case.

Identifying future sampling areas

Because no differences in the distribution of well-surveyed cells were detected among the dif-

ferent physioclimatic subregions, we recommend that further sampling effort be located in the most probable species-rich regions that are also insufficiently surveyed. This method seeks a cost-effective selection of further survey areas that will increase the information to a considerable degree (because new records mean new species in the grid cell). In contrast, a complementary strategy involves increasing the number of well-surveyed areas to improve the richness distribution obtained by predictive modelling. To implement this second strategy, two different types of squares need further surveys. These surveys should include moderately surveyed squares with completeness values near the established threshold. The new information from these cells should be included in the construction of the model. The squares with high residual values should also be included because they represent areas where the model cannot be used as a good proxy of real richness and where further survey effort is therefore required. At the same time, a balanced distribution of well-surveyed cells in the defined physioclimatic subregions should be maintained.

The areas selected by the first strategy considered (the cost-effective method) are shown in

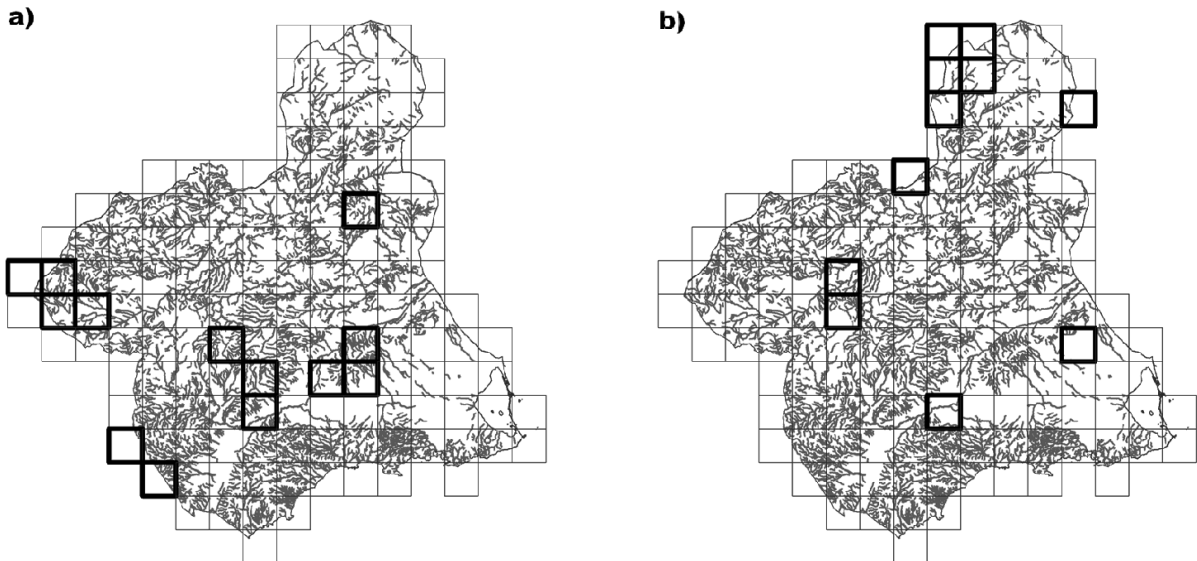


Figure 6. Sampling priority areas for a) water beetles and b) water bugs. *Áreas de muestreo prioritarias para a) coleópteros acuáticos y b) hemípteros acuáticos.*

figure 6a for beetles and figure 6b for bugs. According to these results and based on the self-evident geographic pattern in the survey, the priority sites to be sampled for aquatic beetles are located in several northwestern and central areas. For aquatic bugs, the most advantageous areas would be the northeastern parts of the region. Additional isolated areas distributed throughout the region should also be sampled for this insect group.

Implications for conservation

In view of the high human pressure that threatens the ecological status of aquatic ecosystems, it is important to verify the actual distribution of the groups considered by this study and then to increase sampling effort in the areas previously mentioned to improve and validate the models. It is possible that any intensification of survey efforts in these areas would produce lower-than-expected species richness values because of the relatively recent impoverishment of the areas' natural assemblages.

If the ultimate goal is to develop more accurate conservation strategies by incorporating invertebrate data, there is a substantial need for reliable information on the distributional patterns of these groups. Although it should be noted that the database used here is the most complete source of information available for aquatic invertebrates in the SE of the Iberian Peninsula, additional sampling effort is undoubtedly needed. Predictive modelling techniques could be an effective and useful tool for designing the required sampling protocols and for contributing to a more realistic picture of the distribution of these groups.

ACKNOWLEDGEMENTS

We thank Cayetano Gutiérrez, Pedro Abellán, Oscar Belmar, Paula Arribas and the remainder of the 'Ecología Acuática' research group (Universidad de Murcia) for their wholehearted support. This work was supported by predoctoral grants (FPU) from MICINN to D.B. and J.A.C., predoctoral and postdoctoral grants from the Fundación Séneca to F.P. and D.S.-F, respectively, and projects CGL2006-04159, 023/2007 (A.M.).

We also thank the 'AIL' committee for recognising the poster based on this paper with an award as the best poster at the XV Congress of the 'AIL'.

REFERENCES

- ABELLÁN, P., D. SÁNCHEZ-FERNÁNDEZ, J. VELASCO & A. MILLÁN. 2005. Conservation of freshwater biodiversity: a comparison of different area selection methods. *Biodiversity and Conservation*, 14: 3457–3474.
- ABELLÁN, P., D. SÁNCHEZ-FERNÁNDEZ, J. VELASCO & A. MILLÁN. 2007. Effectiveness of protected area networks in representing freshwater biodiversity: the case of a Mediterranean river basin (SE Spain). *Aquatic Conservation: Marine and Freshwater Ecosystems*, 17: 361–374.
- ALLAN, J. D. & A. S. FLECKER. 1993. Biodiversity conservation in running waters. *Bioscience*, 43: 32–43.
- BILTON, D. T., L. MCABENDROTH, A. BEDFORD & P. M. RAMSAY. 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.
- CARBONELL, J. A. 2008. *Hemípteros acuáticos de la Región de Murcia. Estudio faunístico y estado de conservación*. Proyecto fin de carrera. Universidad de Murcia. 87 pp.
- CARBONELL, J. A. 2010. *Hemípteros acuáticos de la Cuenca del Río Segura. Estado del conocimiento, patrones de distribución y conservación*. Tesis de licenciatura. Universidad de Murcia. 119 pp.
- CLARKE, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, 18: 117–143.
- DARWALL, W. R. T. & J. C. VIÉ. 2005. Identifying important sites for conservation of freshwater biodiversity: Extending the species-based approach. *Fisheries Management and Ecology*, 12: 287–293.
- DAVIS, J. A., S. W. ROLLS & A. BALLA. 1987. The role of the Odonata and aquatic Coleoptera as indicators of environmental quality in wetlands. In: *The role of Invertebrates in Conservation and Biological Survey*. J. D. Majer (ed.): 31–42. Western Australian Department of Conservation and Land Management. Perth. Australia.
- 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., T. H. SPARKS & P. B. HARDY. 1999. Bias in butterfly distribution maps: the effects of sampling effort. *Journal of Insect Conservation*, 3: 33–42.
- DENNIS, R. L. H., T. G. SHREEVE, N. J. B. ISAAC, D. B. ROY, P. B. HARDY, R. FOX & J. ASHER. 2006. The effects of visual apparency on bias butterfly recording and monitoring. *Biological Conservation*, 128: 486–492.
- DOMÍNGUEZ-LOZANO, F., D. GALICIA-HERBADA, L. MORENO-RIVERO, J. M. MORENO-SAINZ, H. SAINZ-OLLERO. 1996. Threatened plants in Peninsular and Balearic Spain: a report based on the EU Habitats Directive. *Biological Conservation*, 76: 123–133.
- FERRIER, S. & A. GUISAN. 2006. Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43: 393–404.
- FOSTER, G. N. & M. D. EYRE. 1992. *Classification and ranking of water beetle communities*. UK Nature Conservation No 1. Joint Nature Conservation Committee. Peterborough. 110 pp.
- FOSTER, G. N., A. P. FOSTER, M. D. EYRE & D. T. BILTON. 1990. Classification of water beetle assemblages in arable fenland and ranking of sites in relation to conservation value. *Freshwater Biology*, 22: 343–354.
- FUNK, V. A., K. S. RICHARDSON & S. FERRIER. 2005. Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnean Society*, 85: 549–567.
- GASITH, A. & V. H. RESH. 1999. Streams in Mediterranean climate regions. Abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics*, 30: 51–81.
- HORTAL, J. & J. M. LOBO. 2005. An ED-based protocol for the optimal sampling of biodiversity. *Biodiversity and Conservation*, 14: 2913–2947.
- HORTAL, J., J. M. LOBO & F. MARTÍN-PIERA. 2001. Forecasting insect species richness scores in poorly surveyed territories: The case of the Portuguese dung beetles (Col. Scarabaeinæ). *Biodiversity and Conservation*, 10: 1343–1367.
- HORTAL, J., P. GARCÍA-PEREIRA & E. GARCÍA-BARROS. 2004. Butterfly species richness in mainland Portugal: Predictive models of geographic distribution patterns. *Ecography*, 27: 68–82.
- HORTAL, J., P. A. V. BORGES & C. GASPAR. 2006. Evaluating the performance of species richness estimators: sensitivity to sample grain size. *Journal of Animal Ecology*, 75: 274–287.
- HORTAL, J., J. M. LOBO & A. JIMÉNEZ-VALVERDE. 2007. Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. *Conservation Biology*, 21: 853–863.
- HORTAL, J., A. JIMÉNEZ-VALVERDE, J. F. GÓMEZ, J. M. LOBO & A. BASELGA. 2008. Historical bias in biodiversity inventories affects the observed realized niche of the species. *Oikos*, 117: 847–858.
- JIMÉNEZ-VALVERDE, A. & J. HORTAL. 2003. Las curvas de acumulación de especies y la necesidad de evaluar la calidad de los inventarios biológicos. *Revista Ibérica de Aracnología*, 8: 151–161.
- KADMON, R., O. FARBER & A. DANIN. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13: 853–867.
- KADMON, R., O. FARBER & A. DANIN. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, 14: 401–413.
- LÉVÊQUE, C., E. V. BALIAN & K. MARTENS. 2005. An assessment of animal species diversity in continental waters. *Hydrobiologia*, 542: 39–67.
- LOBO, J. M. 2008. More complex distribution models or more representative data? *Biodiversity Informatics*, 5: 15–19.
- LOBO, J. M. & F. MARTÍN-PIERA. 2002. Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conservation Biology*, 16: 158–173.
- MAO, C. X. & R. K. COLWELL. 2005. Estimation of species richness: mixture models, the role of rare species, and inferential challenges. *Ecology*, 86: 1143–1153.
- MCCULLAGH, P. & J. A. NELDER. 1989. *Generalized linear models* (2nd ed.). London: Chapman and Hall. London. United Kingdom.
- MILLÁN, A., J. VELASCO, & A. G. SOLER. 1988. Heterópteros acuáticos (Gerromorpha & Nepomorpha) de la cuenca del Río Segura. SE. de España. *Anales de Biología*, 15(4): 74–89.
- MILLÁN, A., J. VELASCO & L. RAMÍREZ-DÍAZ. 1996. Distribución espacial de los Adephaga acuá-

- ticos (Coleoptera) en la cuenca del Río Segura (SE de la Península Ibérica). *Limnetica*, 12(2): 13–29.
- MILLÁN, A., J. L. MORENO & J. VELASCO. 2002. *Los coleópteros y heterópteros acuáticos y semiacuáticos de la provincia de Albacete. Catálogo faunístico y estudio ecológico*. Instituto de Estudios Albacetenses. Albacete. España. 180 pp.
- MILLÁN, A., P. ABELLÁN, I. RIBERA, D. SÁNCHEZ-FERNÁNDEZ & J. VELASCO. 2006. The Hydradephaga of the Segura basin (SE Spain): twenty five years studying water beetles. Monograph on Hydroadephaga. In memoriam of Prof. Franciscolo. *Memorie della Società entomologica italiana*, 85: 137–158.
- MORENO, J. L., A. MILLÁN, M. L. SUÁREZ, M. R. VIDAL-ABARCA & J. VELASCO. 1997. Aquatic Coleoptera and Heteroptera assemblages in waterbodies from ephemeral coastal streams (“ramblas”) of south-eastern Spain. *Archiv für Hydrobiologie*, 141: 93–107.
- MYERS, N., R. A. MITTERMEIER, C. G. MITTERMEIER, G. A. B. DA FONSECA & J. KENT. 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403: 853–858.
- NICHOLLS, A. O. 1989. How to make biological surveys go further with generalised linear models. *Biological Conservation*, 50: 51–75.
- POLHEMUS, J. T. & D. A. POLHEMUS. 2008. Global diversity of true bugs (Heteroptera; Insecta) in freshwater. *Hydrobiologia*, 595: 379–391.
- RAMOS, M., J. M. LOBO & M. ESTEBAN 2001. Ten years inventorying the Iberian Fauna: Results and perspectives. *Biodiversity and Conservation*, 10: 19–28.
- REDDY, S. & L. M. DÁVALOS. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, 30: 1719–1727.
- REYJOL, Y., B. HUGUENY, D. PONT, P. G. BIANCO, U. BEIER, N. CAIOLA, F. CASALS, I. COWX, A. ECONOMOU, T. FERREIRA, G. HAIDVOGL, R. NOBLE, A. DE SOSTOA, T. VIGNERON & T. VIRBICKAS. 2007. Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, 16: 65–75.
- RICCIARDI, A. & J. B. RASMUSSEN. 1999. Extinction rates in North American freshwater fauna. *Conservation Biology*, 13: 1220–1222.
- ROMO, H., E. GARCÍA-BARROS & J. M. LOBO. 2006. Identifying recorder-induced geographic bias in an Iberian butterfly database. *Ecography*, 29: 873–885.
- SÁNCHEZ-FERNÁNDEZ, D., P. ABELLÁN, J. VELASCO & A. MILLÁN. 2004a. Áreas prioritarias de conservación en la Cuenca del río Segura utilizando a los coleópteros acuáticos como indicadores. *Limnetica*, 23 (3-4): 209–228.
- SÁNCHEZ-FERNÁNDEZ, D., P. ABELLÁN, J. VELASCO & A. MILLÁN. 2004b. Selecting areas to protect the biodiversity of aquatic ecosystems in a semiarid Mediterranean region using water beetles. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 14: 465–479.
- SÁNCHEZ-FERNÁNDEZ, D., P. ABELLÁN, A. MELLADO, J. VELASCO & A. MILLÁN. 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., J. M. LOBO, P. ABELLÁN, I. RIBERA & A. MILLÁN. 2008. Bias in freshwater biodiversity sampling: The case of Iberian water beetles. *Diversity and Distributions*, 14: 754–762.
- SÁNCHEZ-FERNÁNDEZ, D., J. M. LOBO, P. ABELLÁN & A. MILLÁN. 2011. How to identify future sampling areas when information is biased and scarce: An example using predictive models for species richness of Iberian water beetles. *Journal for Nature Conservation*, 19: 54–59.
- SASTRE, P., P. ROCA & J. M. LOBO. 2009. A Geoplatform for improving accessibility to environmental cartography. *Journal of Biogeography*, 36: 568.
- SAUNDERS, D. L., J. J. MEEUWIGM & A. C. J. VINCENT. 2002. Freshwater protected areas: strategies for conservation. *Conservation Biology*, 16: 30–41.
- SOBERÓN, J. & J. LLORENTE. 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). StatSoft Inc. Tulsa
- VIDAL-ABARCA, M. R., M. L. SUÁREZ & L. RAMÍREZ-DÍAZ. 1992. Ecology of Spanish semiarid streams. In: *Limnology in Spain*, C. Montes, C. Duarte (eds): 151–160. Asociación Española de Limnología. Madrid. España.
- WILLIAMS, P. H., C. HUMPHRIES, M. B. ARÁUJO, R. LAMPINEN, W. HAGEMEIJER, J. P. GASC & T. MITCHELL-JONES. 2000. En-

demism and important areas for representing European biodiversity: a preliminary exploration of

atlas data for plants and terrestrial vertebrates. *Belgian Journal of Entomology*, 2:21–46.