

Species distribution models that do not incorporate global data misrepresent potential distributions: a case study using Iberian diving beetles

David Sánchez-Fernández^{1*}, Jorge M. Lobo¹ and Olga Lucía Hernández-Manrique²

¹Museo Nacional de Ciencias Naturales, Departamento de Biodiversidad y Biología Evolutiva, José Abascal, 2. 28006 Madrid, Spain, ²Centro Iberoamericano de la Biodiversidad (CIBIO), Unidad de Diversidad y Comportamiento Animal, San Vicente del Raspeig, 03080 Alicante, Spain

ABSTRACT

Aim In this study, we (1) determine whether simple species distribution models based on regional data provide incomplete descriptions of potential distributions; (2) investigate whether underrepresented areas where potential distributions are estimated using only regional data are spatially and environmentally structured; and (3) examine why regional data may not adequately describe potential distributions.

Location Iberian Peninsula.

Methods We used a multidimensional envelope procedure to estimate the potential distributional areas of 73 species of Iberian diving beetles (Dytiscidae) using two data sets (Iberian data and data from the entire range). We used a Mann–Whitney *U*-test to compare the features (climate, number of database records and proportion of human-transformed land uses) of these underrepresented areas with those of the remaining Iberian territory.

Results By comparing species-richness estimates obtained by overlaying predicted species distributions modelled using either global or regional data, we found that some areas of species' potential distributions are underrepresented when only regional data are used. Incomplete estimates of potential distributions when using only Iberian data may be partly attributable to limited survey efforts combined with unique local climates, but none of the considered factors by itself seems to fully explain this underrepresentation.

Main conclusions Our results show that species data from regional inventories may provide an incomplete description of the environmental limits of most species, resulting in a biased description of species' niches. The results of distribution models based on partial information about the environmental niche of a species may be inaccurate. To minimize this error, we highlight the importance of considering all known populations of a given species or at least a sample of populations distributed across the whole range, to include environmental extremes of the distribution. We highlight some methodological and conceptual concerns that should be considered when attempting to infer potential distributions from occurrence data.

Keywords

Dytiscidae, Iberian Peninsula, potential distribution, species distribution models, water beetles.

*Correspondence: David Sánchez-Fernández, Museo Nacional de Ciencias Naturales, C/José Gutiérrez Abascal, 2, 28006 Madrid, Spain. E-mail: davidсан@mncn.csic.es

INTRODUCTION

The first studies designed to estimate the spatial distributions of species based on incomplete data appeared in the 1980s. Since

then, the number of publications dedicated to species distribution models (SDMs) has increased exponentially (Lobo *et al.*, 2010). Using a variety of modelling techniques (Guisan & Zimmermann, 2000; Elith & Leathwick, 2009), SDMs simulate

probable distributions by using species-occurrence data (primarily presences) and a set of predictive environmental variables that are thought to influence their distribution. When an SDM aims at geographically representing the environmental niche of a species, such a representation can be called a potential distribution (Jiménez-Valverde *et al.*, 2008; Soberón & Nakamura, 2009). A potential distribution reflects the areas in which a species may occur based on information about environmental variables in areas where it is known to occur.

To obtain a geographical representation of a species' niche, we must consider that dispersal limitations, biotic interactions (Colwell & Rangel, 2009; Soberón & Nakamura, 2009) and current environmental conditions (Jackson & Overpeck, 2000) may obscure the full spectrum of climate requirements that a species needs to maintain source populations (i.e. the realized distribution may provide only partial information on inhabitable conditions for a species). Therefore, if we do not include data representing the whole range of the species, a simulated potential distribution may provide a biased geographical representation. Although discussed previously (Pearson *et al.*, 2002; Thuiller *et al.*, 2004), the consequences of failing to include the full range of environmental conditions in potential-distribution simulations have still not been fully considered. This issue has been recently explored with regard to the spatial transferability of SDMs for invasive species (Beaumont *et al.*, 2009). Here, we extend such research by evaluating how model performance is affected by either using training data restricted to the studied region or including data from beyond that region.

We compare potential distributions estimated from Iberian versus global range-distribution data (i.e. occurrence records from the entire range of distribution, including data from both within and outside of the Iberian Peninsula) for 73 Iberian diving beetles (Coleoptera, Dytiscidae) to evaluate whether potential distributions estimated from Iberian data differ from those estimated from range-wide data. After identifying the areas that are underrepresented when global data are not incorporated, we examine whether these areas are spatially and environmentally structured. Next, we try to determine why the Iberian data do not fully represent the potential distributions by examining whether underrepresented areas have unique climatic conditions and/or whether they presently lack species data owing to survey biases or human activities.

METHODS

Source of biological data

We studied 73 species and well-established subspecies (with five or more records on the Iberian Peninsula) of diving beetles from the family Dytiscidae. We chose this family because of the following characteristics: (1) taxonomically stable, (2) relatively well-known biology and life history and (3) component species showing a high variety of morphological, ecological and biogeographical features (body sizes,

distribution ranges, habitat requirements, environmental tolerances, etc.). Iberian occurrences come from an exhaustive database (ESACIB 'EScarabajos ACuáticos IBéricos') including all the available taxonomic and distributional data from Iberian water beetles (Sánchez-Fernández *et al.*, 2008). A total of 6388 Iberian database records were used (63.88 ± 68.69 ; mean number of database records per species \pm SD). As all considered species have distributions extending beyond the Iberian Peninsula, we compiled georeferenced localities to obtain data on climatic conditions at the environmental extremes of their whole range. Thus, a total of 1341 extra (out of Iberia) records from 1965 to 2009 were obtained from 25 specialized publications, unpublished data and private collections (mainly Palaearctic, but also Nearctic, Neotropical and Oriental regions) (13.06 ± 9.61 , mean number of database records per species \pm SD). Although these distributional records are incomplete and lack homogeneity, we assume that including these data will better represent the full spectrum of climatic conditions in which these species can inhabit. All biological information was georeferenced at a 100-km² spatial resolution using a Universal Transverse Mercator (UTM) projection.

Environmental variables

Environmental data were obtained from WORLDCLIM (version 1.3, <http://www.worldclim.org>; for details, see Hijmans *et al.*, 2005). WORLDCLIM contains climate data (i.e. monthly precipitation and monthly mean, minimum and maximum temperatures) at a spatial resolution of 30 arc-s (*c.* 1×1 -km resolution) obtained by the interpolation of climate-station records from 1950 to 2000. These data were used to derive biologically meaningful bioclimatic variables representing annual trends, seasonality and extreme conditions. In our study, we used 17 climate variables as predictors: annual mean temperature, isothermality, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, annual temperature range, mean temperature of the wettest quarter, mean temperature of the warmest quarter, mean temperature of the coldest quarter, annual precipitation, precipitation of the wettest month, precipitation of the driest month, precipitation seasonality, precipitation of the wettest quarter, precipitation of the driest quarter, precipitation of the warmest quarter and precipitation of the coldest quarter. The values of these explanatory variables were used at the same resolution as the biological data (UTM cells of 100 km²).

Modelling potential distributions

We used a simple multidimensional envelope procedure (MDE) to estimate potential species distributions because this procedure allowed us to represent regions with favourable conditions based on observed occurrences (see Jiménez-Valverde *et al.*, 2008; Lobo *et al.*, 2010 and Soberón, 2010 for the conceptual basis of this approach). First, we calculated

maximum and minimum scores (extreme values) for all relevant climatic variables in all observed presence cells, selecting as suitable grid squares all those with climatic values falling within that range and designating as unsuitable all cells outside it. Thus, we used the extreme values to derive a binary distributional hypothesis about the areas having climatically suitable conditions (potential distribution), assuming that recorded occurrences reflect the full spectrum of climatic conditions in which the species can survive and reproduce. Because distributional simulations obtained by MDE procedures are highly dependent on the number of selected predictors (Beaumont *et al.*, 2005), we estimated the minimum set of climatic variables needed to explain the occurrence of each species using ecological-niche factor analysis (ENFA; Hirzel *et al.*, 2002; Calenge & Basille, 2008) with the Iberian Peninsula as the background area. This procedure compares the climatic data of presence localities against the climatic conditions found throughout the study area, thereby computing uncorrelated factors that can explain both species marginality (the distance between the species optimum and the average climatic conditions in the study area) and specialization (the ratio of the ecological variance in the climate of the study area to that associated with the focal species). Factors were retained or discarded based on their eigenvalues relative to a broken-stick distribution (Hirzel *et al.*, 2002). Climatic variables selected as predictors were those showing the highest correlation values (factor scores > 0.30) with the retained ENFA factors. The number of climatic predictors ranged from 2 to 6, depending on the species. Once the relevant variables had been selected for each species based on Iberian data, we carried out an MDE to derive two different distributional hypotheses using data from either the Iberian Peninsula or the entire distributional range.

Differences between potential distributions

We produced two binary potential-distribution maps covering the extent of the Iberian Peninsula for each of the species considered: one using Iberian database records (I_{PD}) and the other using whole-distribution data (W_{PD}). The I_{PD} and W_{PD} maps for all species were overlapped to simulate potential species richness for the Iberian Peninsula and then self-subtracted ($W_{PD}-I_{PD}$) to estimate areas that would be underrepresented when using only Iberian data for the SDMs. We also estimated the 90th percentile of $W_{PD}-I_{PD}$ values (across a total of 731 cells) to identify the most underrepresented cells (UCs).

Assessing potential explanations for expected differences

The Iberian Peninsula is traditionally divided into two main biogeographical regions: the Eurosiberian or Atlantic region (from north-western Portugal to the central Pyrenees) and the Mediterranean region (Rivas-Martínez, 1983). Thus, we divided the UCs into two groups according to their climatic

characteristics (UC₁ and UC₂) using a *k*-means cluster analysis. We used Mann–Whitney *U*-tests to compare the features (climate, number of database records and proportion of human-transformed land use areas) of each underrepresented region (UC₁ and UC₂) with those of the remaining Iberian territory. We calculated the proportion of human-transformed land use areas (urban, industrial and intensive cultivation; hereafter ‘transformed’) in the 100-km² grid cells of UC₁, UC₂ and the remaining Iberian territory. We obtained 250-m resolution land cover data from the Corine Land Cover 2000 database (<http://terrestrial.eionet.europa.eu/CLC2000>). Using simple and multiple linear regressions, we determined whether $W_{PD}-I_{PD}$ species-richness differences were related to the survey effort carried out in each cell (measured as the number of database records for these species available in an exhaustive database of Iberian water beetles; see Sánchez-Fernández *et al.*, 2008) or to climatic values. We also estimated the Mahalanobis distance (see Farber & Kadmon, 2003) between each grid cell and the average climatic conditions of the Iberian Peninsula and then regressed this distance against the $W_{PD}-I_{PD}$ values to determine whether cells with more biased estimations of potential distributions were climatically unique.

Finally, the percentage of underprediction ($UP = 1 - (I_{PD} \times 100 / W_{PD})$) was calculated for each species. This measure represents the percentage of cells of the potential distribution estimated using data from the whole distribution (W_{PD}) that was predicted using only Iberian database records (I_{PD}). Calculating the UP makes it possible to explore the proportion of species for which the percentage of underprediction (UP) is important. Also, we evaluated whether UP values were associated with three general geographical characteristics of species (northern limit, southern limit and latitudinal range extent, in degrees) and one ecological trait (adaptation to lentic-standing or lotic-running habitats), following Ribera & Vogler (2000).

RESULTS

The mean number of species per cell (59.6 ± 0.3 for W_{PD} and 52.1 ± 0.3 for I_{PD} ; mean \pm 95% confidence interval) in potential distributions differed between analyses using the Iberian and whole-range data (Mann–Whitney *U*-tests, $U = 11,200,352$, $P < 0.0001$). The species-richness values were higher when the potential distributions were estimated using whole-range data, except for 18 cells in which the W_{PD} - and I_{PD} -derived estimates were equal. The average difference between the W_{PD} and I_{PD} species-richness estimates ($W_{PD}-I_{PD}$) was 7.5 ± 0.1 ; the smallest differences generally appeared at high species-richness values (Fig. 1).

The geographical projection of $W_{PD}-I_{PD}$ values showed a spatially structured pattern with maximum values mainly distributed in two Iberian regions (Fig. 2a): a group of cells located along the Atlantic coast and another centred on the southern subsection of the Iberian inner plateau (Central Meseta). A *k*-means cluster analysis divided the UCs into two main regions (UC₁ and UC₂) that basically corresponded to

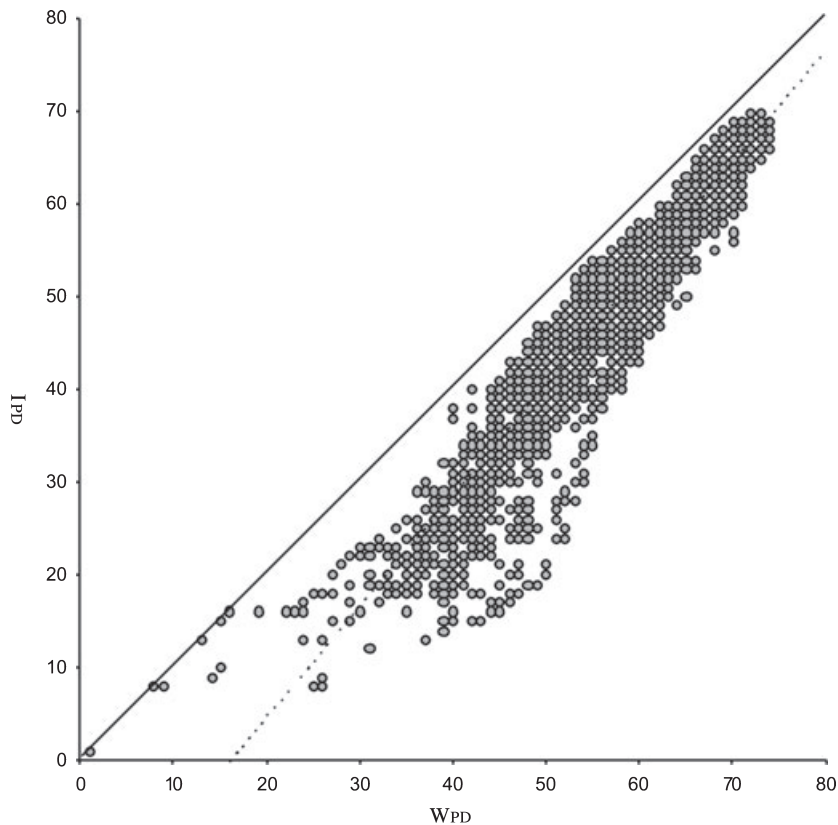


Figure 1 Relationship between species richness from potential distributions derived from Iberian data (I_{PD}) and whole-range distribution (W_{PD}) of 100 species of diving beetles in 10-km² UTM cells of the Iberian Peninsula. The slope of the linear regression is significantly different from one (1.20 ± 0.01 ; $\pm 95\%$ confidence limit), indicating that broader potential estimations are generated when using W_{PD} data.

the regions described above (Fig. 2b). These two regions differed from the rest of the Iberian territory in their climatic conditions and percentage of transformed land uses (Table 1). In general, UC_1 cells were characterized by higher precipitation, lower temperatures and higher proportions of transformed area, while UC_2 cells had higher temperatures, lower precipitation and lower proportions of transformed area (Table 1) than the remaining Iberian cells.

Neither UC_1 nor UC_2 cells differed from the remaining Iberian cells in the number of database records (Table 1). There was no relationship between W_{PD} - I_{PD} values and the number of database records ($F_{1,6135} = 0.58$; $P = 0.44$), and the 90th percentile of the W_{PD} - I_{PD} cell values (≥ 13 species; 731 UCs) did not have fewer database records (mean = 1.9) than the remaining Iberian cells (mean = 2.2) ($U = 1,940,253$; $P = 0.428$). Nevertheless, well-surveyed cells never had high W_{PD} - I_{PD} values (Fig. 3a).

There were relationships between W_{PD} - I_{PD} values and all climatic variables ($P < 0.001$) except annual precipitation, precipitation during the wettest month, precipitation during the coldest quarter and precipitation during the wettest quarter. However, the most important explanatory variable (maximum temperature of the warmest month) accounted for only 4.3% of the total variability ($adjR^2$). The complete model (including all climatic variables) explained 24.3% of the total variability in W_{PD} - I_{PD} values. The climatic conditions of the cells with the highest W_{PD} - I_{PD} values were similar to the average climatic conditions in the Iberian Peninsula (Fig. 3b).

In general, the percentages of UP were relatively high (mean UP = 18%). For more than a quarter of the species, the models built using regional data underpredicted more than 20% of their potential distributions as estimated using global data. Furthermore, only 26 species had UP values lower than 5%. These levels of UP were not statistically correlated with any of the general geographical characteristics of species (Spearman rank correlation coefficients), although models of the Iberian distributional ranges based on regional data seemed to be more underpredicted for lentic than for lotic species (Mann-Whitney U -tests, $U = 437$, $P = 0.03$).

DISCUSSION

Our results show that species data from regional inventories may provide an incomplete description of the environmental limits of most species, resulting in a biased description of species' niches. Consequently, the results of a distribution model based on partial information about the environmental niche of the species may be compromised (Hortal *et al.*, 2008), especially if the model is meant to geographically represent the fundamental niche of the species (Wolmarans *et al.*, 2010). Therefore, procedures requiring the estimation of potential distributional ranges, such as identifying additional survey sites (Guisan *et al.*, 2006), estimating niche conservatism (Peterson *et al.*, 1999; Martínez-Meyer & Peterson, 2006), assessing species-range shifts caused by climate change (Araújo & Rahbek, 2006; Thuiller *et al.*, 2006),

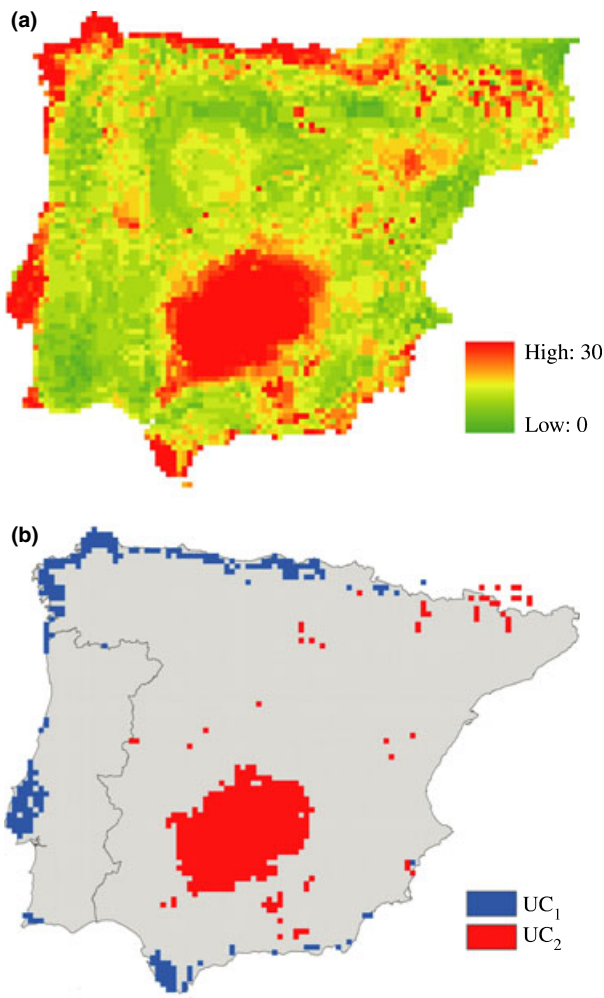


Figure 2 (a) Differences between the potential species richness ($W_{PD}-I_{PD}$) derived when using Iberian database records (I_{PD}) and whole-range data (W_{PD}). (b) Location of the 90th percentile of $W_{PD}-I_{PD}$ values or underrepresented cells (UCs), which were divided into two main regions according to their climatic characteristics by a *k*-means cluster analysis (UC_1 and UC_2).

identifying important areas for conservation (Cabeza *et al.*, 2004) or estimating invasion risk (Peterson, 2003; Thuiller *et al.*, 2005), may fail if only regional data are used. However, it is possible to estimate the locations where potential-distribution models may fail when only regional data are used by repeating the same analyses using data from outside the studied region. Further SDMs intended to represent the realized and potential distributions of Iberian diving beetles should consider these regions as areas of uncertainty (see Rocchini *et al.*, 2010).

The species we studied have all been observed in the target territory (Iberian Peninsula); therefore, underrepresented areas should be climatically unique, and areas where these Iberian species remain undetected are likely to be difficult to colonize or to have suffered species extinctions. The underrepresented areas that we identified are concentrated in two

climatically distinct regions, although these regions are not considered to have extreme climates when compared to mean Iberian climatic conditions. However, the location of underrepresented areas may depend on the target group used to derive predictions and on the climatic variability of the region. Further studies should be conducted to determine whether other species groups share the same underrepresented areas.

We cannot suggest a single explanation for the existence of underrepresented areas when using Iberian data alone. Even though there is a general lack of exhaustive data regarding invertebrate species (Sánchez-Fernández *et al.*, 2008), the survey effort is not particularly lower in underrepresented areas. Nevertheless, because well-surveyed sites did not have high $W_{PD}-I_{PD}$ values, we infer that a lack of survey effort may have played some role. Because the proportion of transformed land is higher in one underrepresented area and lower in the other, we cannot conclude that the underrepresentation in these areas is attributable to the extinction of former populations. Interestingly, the median dates of database records in underrepresented areas were earlier than those of the remaining Iberian Peninsula (data not shown). Therefore, we cannot conclude that recent surveys in these areas failed to detect extinct populations. Thus, anthropogenic disturbance may be a relevant but not a decisive factor. Finally, underrepresented areas in potential distributions derived from regional data may be characterized by their climatic uniqueness or inaccessibility (i.e. these areas may be difficult to colonize despite favourable climatic conditions). However, our results do not support this hypothesis. Hence, we cannot suggest a clear causal factor to explain underrepresentation of potential distributions when regional data are used. This phenomenon may be caused by a combination of factors, including scarcity of data (whatever its cause) combined with climatic uniqueness in some areas.

We can also examine whether the underrepresentation of Iberian data is attributable to a few particular species. The variation in the percentage of UPs suggests that the underrepresentation of potential distributional ranges by models calibrated from Iberian data alone is not attributable to a limited capacity to represent the climatic conditions of a few particular species. A notable number of species (more than a quarter) have highly underrepresented potential ranges when only regional data are used ($UP > 20\%$). Interestingly, the distributional-range size of a species does not seem to be associated with its underrepresentation. Thus, those species that are *a priori* capable of inhabiting a wider spectrum of climatic conditions do not show lower underrepresentation values, probably because their occurrence in some regions with climatic conditions similar to those of the Iberian Peninsula is a more decisive factor. However, habitat type partially explains the capacity of regional data to generate wider potential distributions for Iberian diving beetles. Lentic habitats are considered less predictable and ephemeral, requiring frequent immigration from neighbouring populations for their long-term survival (Ribera *et al.*, 2001). Our

Table 1 Comparison between UC₁ and UC₂ cell values for climatic variables, number of records, and percentage of anthropogenic land use and those present in the remaining (REM) Iberian cells ($nUC_1 = 245$, $nUC_2 = 486$, $nREM = 5406$). Significantly different climatic variables are shown (Mann–Whitney *U*-test, $P < 0.05$). Signs indicate when median values in UC₁ and UC₂ cells are higher (+), lower (–) or equal (=) than those of remaining Iberian cells.

	UC ₁			UC ₂		
	U	P		U	P	
Annual precipitation	2,87,325	0.000	+	9,92,680	0.000	–
Annual mean temperature	4,45,962	0.000	+	9,02,806	0.000	+
Isothermality	4,07,255	0.000	+	8,47,672	0.000	–
Precipitation of the coldest quarter	2,38,858	0.000	+	12,66,155	0.186	=
Precipitation of the driest month	6,17,071	0.071	=	6,63,654	0.000	–
Precipitation of the driest quarter	6,03,814	0.019	+	6,57,266	0.000	–
Precipitation seasonality	4,65,489	0.000	+	8,96,802	0.000	+
Precipitation of the warmest quarter	5,83,268	0.002	+	5,99,375	0.000	–
Precipitation of the wettest month	2,17,081	0.000	+	10,02,577	0.000	–
Precipitation of the wettest quarter	2,20,367	0.000	+	10,53,445	0.000	–
Maximum temperature of the warmest month	2,48,789	0.000	–	3,35,974	0.000	+
Minimum temperature of the coldest month	1,49,623	0.000	+	12,93,529	0.575	=
Mean temperature of the coldest quarter	2,20,783	0.000	+	11,99,085	0.001	+
Mean temperature of the warmest quarter	5,06,650	0.000	–	5,31,287	0.000	+
Mean temperature of wettest quarter	5,40,990	0.000	+	12,26,261	0.015	+
Temperature seasonality	44,797	0.000	–	1,70,831	0.000	+
Temperature annual range	35,807	0.000	–	1,78,082	0.000	+
Number of database records	65,283	0.706	=	12,87,422	0.465	=
Percentage of anthropogenic land use	5,38,910	0.000	+	11,73,718	0.000	–

results suggest that species adapted to stagnant water bodies are more likely to have their potential distributions underpredicted when the pool of distributional information does not cover the entire global range. We hypothesize that the higher risk of local (Iberian) extinctions for these species (as consequence of the disappearance of these fragmented and scarce habitats under mediterranean conditions) and their higher tendency to include vagrant individuals may help to explain the predictions generated when using global occurrences.

Other factors may have influenced our results. For instance, other environmental variables that we did not include, such as edaphic characteristics, may help to explain the potential distribution of species. Similarly, we did not consider the phenotypic plasticity of a species across its entire range (Scheiner & Goodnight, 1984; Addo-Bediako *et al.*, 2000; Terblanche *et al.*, 2006); Iberian populations may not be able to inhabit the full range of climatic conditions to which other populations of a species are adapted. Furthermore, those interested in modelling species' distributions should note that our results are dependent on the modelling approach employed. We used simple multi-dimensional envelopes and presence information to obtain a geographical representation of an area where climatic conditions are potentially favourable to the species based on its observed occurrences. However, a variety of correlative techniques, predictors and types of biological data (presence versus presence/absence) may provide alternative

maps that represent some unknown point along the gradient between the realized and potential distributions (Jiménez-Valverde *et al.*, 2008; Soberón, 2010). While there is still disagreement on the best way to model the potential distribution of a species based only on occurrence data, we recommend the use of simple procedures like the one applied in this study.

Additionally, we should not forget that the estimation of niche conditions from distributional data and their geographical translation are compromised exercises. Undetected or regional species data may generate potential distributions that underestimate the 'true' inhabitable area, while the lack of knowledge about phenotypic or physiological variation among populations leads us to question predictions meant to represent potential distributions. Understanding the environmental limits of a species is fundamental to predicting its potential distributional area (Thuiller *et al.*, 2004; Soberón & Peterson, 2005; Araújo & Guisan, 2006). Nevertheless, the estimation of these limits is not free of error; even if all current populations of a given species can be included, thus capturing all environmental extremes, these limits may differ from the true limits of the species or may not be applicable to all populations. As a precaution, correlative SDMs should consider data from as many populations as possible or at least a sample of populations that are well distributed across the entire spectrum of climatic conditions throughout the range, including environmental extremes. This procedure may lead to

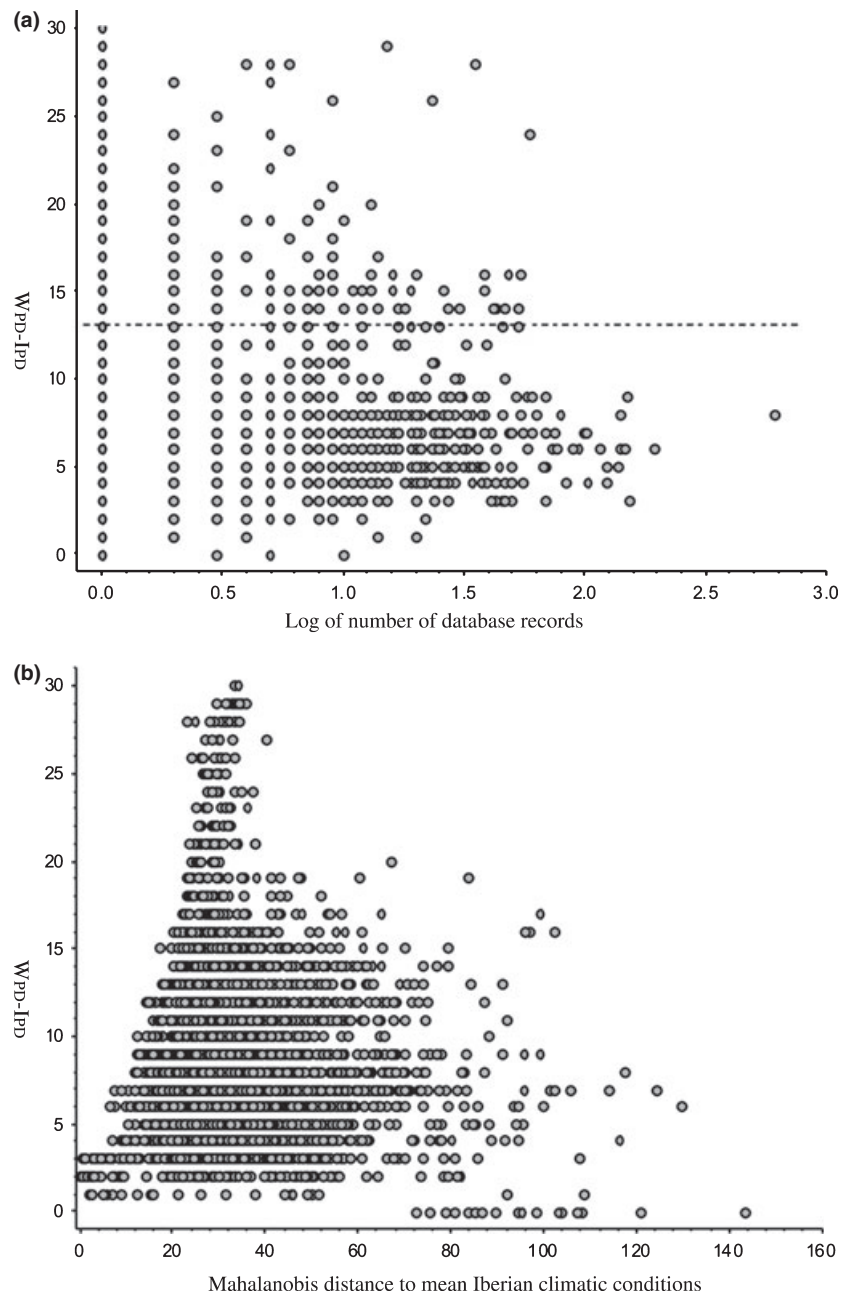


Figure 3 (a) Relationship between Iberian and whole range derived potential-distribution difference (W_{PD-IPD}) and survey effort, measured as the number of database records available in an exhaustive database of diving beetles, in 100-km² Iberian UTM cells (Sánchez-Fernández *et al.*, 2008). The dotted line represents the 90th percentile of W_{PD-IPD} cell values. (b) Relationship between W_{PD-IPD} values and the Mahalanobis distance for each grid cell compared to average climatic conditions of the Iberian Peninsula.

over-prediction of potential ranges, but considering the lack of reliable distributional data and our insufficient knowledge of species' responses to environmental change, a precautionary principle should guide us towards building broader risk maps.

ACKNOWLEDGEMENTS

We thank Pedro Abellán, Andrés Millán and Ignacio Ribera for their help in compiling data. We also thank Joaquín Hortal for their multiple comments on previous versions of the manuscript. This work was supported by funding from a post-doctoral grant to D.S-F. from the Fundación Séneca, as well as by the CGL2008-03878 project.

REFERENCES

- Addo-Bediako, A., Chown, S.L. & Gaston, K.J. (2000) Thermal tolerance, climatic variability and latitude. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 739–745.
- Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.
- Araújo, M.B. & Rahbek, C. (2006) How does climate change affect biodiversity? *Science*, **313**, 1396–1397.
- Beaumont, L.J., Hughes, L. & Poulsen, M. (2005) Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, **186**, 250–269.

- Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R. & Hughes, L. (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, **15**, 409–420.
- Cabeza, M., Araújo, M.B., Wilson, R.J., Thomas, C.D., Cowley, M. & Moilanen, A. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology*, **41**, 252–262.
- Chalenge, C. & Basille, M. (2008) A general framework for the statistical exploration of the ecological niche. *Journal of Theoretical Biology*, **252**, 674–685.
- Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences USA*, **106**, 19651–19658.
- Elith, J. & Leathwick, J. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, **40**, 677–697.
- Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modelling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. & Zimmermann, N.E. (2006) Using niche-based models to improve the sampling of rare species. *Conservation Biology*, **20**, 501–511.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hirzel, A.H., Hausser, J., Chessel, D. & Perrin, N. (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, **83**, 2027–2036.
- Hortal, J., Jiménez-Valverde, A., Gomez, J.F., Lobo, J.M. & Baselga, A. (2008) Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos*, **117**, 847–858.
- Jackson, S.T. & Overpeck, J.T. (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194–220.
- Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885–890.
- Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.
- Martínez-Meyer, E. & Peterson, A.T. (2006) Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. *Journal of Biogeography*, **33**, 1779–1789.
- Pearson, R.G., Dawson, T.P., Berry, P.M. & Harrison, P.A. (2002) SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecological Modelling*, **154**, 289–300.
- Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology*, **78**, 419–433.
- Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.
- Ribera, I. & Vogler, A.P. (2000) Habitat type as a determinant of species range sizes: the example of lotic-lentic differences in aquatic Coleoptera. *Biological Journal of the Linnean Society*, **71**, 35–52.
- Ribera, I., Barraclough, T.G. & Vogler, A.P. (2001) The effect of habitat type on speciation rates and range movements in aquatic beetles: inferences from species-level phylogenies. *Molecular Ecology*, **10**, 721–735.
- Rivas-Martínez, S. (1983) Pisos bioclimáticos de España. *Lazaroa*, **5**, 33–43.
- Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G. & Chiarucci, A. (2010) Uncertainty in species distribution mapping and the need for maps of ignorance. *Progress in Physical Geography* (in press).
- Sánchez-Fernández, D., Lobo, J.M., Abellán, P., Ribera, I. & Millán, A. (2008) Bias in freshwater biodiversity sampling: the case of Iberian water beetles. *Diversity and Distributions*, **14**, 754–762.
- Scheiner, S.M. & Goodnight, C.J. (1984) The comparison of phenotypic plasticity and genetic variation in populations of the grass *Danthonia spicata*. *Evolution*, **38**, 845–855.
- Soberón, J. (2010) Niche and area of distribution modelling: a population ecology perspective. *Ecography*, **33**, 159–167.
- Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences USA*, **106**, 19644–19650.
- Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distribution areas. *Biodiversity Informatics*, **2**, 1–10.
- Terblanche, J.S., Klok, C.J., Krafur, E.S. & Chown, S.L. (2006) Phenotypic plasticity and geographic variation in thermal tolerance and water loss of the tsetse *Glossina pallidipes* (Diptera: Glossinidae): implications for distribution modelling. *American Journal of Tropical Medicine and Hygiene*, **74**, 786–794.
- Thuiller, W., Brotons, L., Araújo, M.B. & Lavorel, S. (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 165–172.
- Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.
- Thuiller, W., Lavorel, S., Sykes, M.T. & Araújo, M.B. (2006) Using niche-based modelling to assess the impact of climate

change on tree functional diversity in Europe. *Diversity and Distributions*, **12**, 49–60.

Wolmarans, R., Robertson, M.P. & Van Rensburg, B.J. (2010) Predicting invasive alien plant distributions: how geographical bias in occurrence records influence models performance. *Journal of Biogeography*, **37**, 1797–1810.

BIOSKETCHES

David Sánchez-Fernández is a post-doctoral researcher at the Museo Nacional de Ciencias Naturales of Madrid. He is interested in biogeography and conservation of freshwater biodiversity.

Jorge M. Lobo is interested in the patterns and processes of species distributions from a macroecological perspective and in the management of biodiversity and conservation biogeography. He is a specialist in the biogeography and ecology of dung beetles (Scarabaeoidea).

Olga Lucía Hernández-Manrique is interested in the study of conservation status of threatened invertebrates.

Editor: Mark Robertson