

0 Scaling of species distribution models across spatial resolutions and 53 5 extents along a biogeographic gradient. The case of the Iberian mole 55 *Talpa occidentalis*

Susana Suárez-Seoane, Emilio Virgós, Olga Terroba, Xosé Pardavila and Jose M. Barea-Azcón 60

10 *S. Suárez-Seoane (s.seoane@unileon.es), Área de Ecología, Depto de Biodiversidad y Gestión Ambiental, Facultad de Ciencias Biológicas y Ambientales, Univ. de León, Campus de Vegazana s/n, ES-24071 León (León), Spain. – E. Virgós, O. Terroba, X. Pardavila and J. M. Barea-Azcón, ESCET, Depto de Biología y Geología, Univ. Rey Juan Carlos, C/Tulipán s/n, ES-28933 Móstoles (Madrid), Spain.* 65

15 Scaling is a key process in modelling approaches since it allows for translating information from one scale to another. However, the success of this procedure may depend on ‘source’ and ‘target’ scales, but also on the biogeographic/ecological context of the study area. We aimed to quantify how performance and scaling of species distribution models (SDMs) varied across spatial resolution and extent along a biogeographic gradient using the Iberian mole as study case. We ran separate MaxEnt models at two extents (national and regional) using independent datasets (species locations and environmental predictors) collected at 10 km and 50 m resolutions respectively. Model performance and success of scaling SDMs were quantified on the basis of accuracy measures and spatial predictions. Complementarily, we calculated marginality and tolerance as indicators of habitat availability and niche truncation along the biogeographic gradient. Model performance increased with resolution and extent, as well as from North to South (mainly for high resolution models). When regional models were validated at different scales, their performance reduced severely, particularly in the case of coarse resolution models (some of them performed worse than random). However, when the 10 km-national model was downscaled within regions, it performed better (AUC_{test} : 0.82, 0.85 and 0.55 respectively for Galicia, Madrid and Granada) than models specifically calibrated within each region at 10 km (0.47, 0.65, 0.44). Indeed, it also had a better accuracy when projected at 50m (0.77, 0.91, 0.79) than models fitted at that resolution (0.62, 0.83, 0.96) in two of the three cases. The success of scaling model predictions decreased along the biogeographic gradient, being these differences associated to niche truncation. Models representing non-truncated niches were more successfully scaled across resolutions and extents (particularly in areas not offering all possible habitats for species), which has important implications for SDM applications. 70
75
80

30
35 During the last decade, species distribution models (SDMs) have been widely developed in ecology and biogeography (Guisan and Zimmerman 2000, Scott et al. 2002, Rushton et al. 2004, Elith et al. 2006, Franklin and Miller 2009). They are valuable tools for predicting species occurrence on the basis of the relationship between species locations and environmental features. Although still with a certain degree of uncertainty (Austin 2002, Guisan and Thuiller 2005, Araújo and Guisan 2006, Peterson 2006, Hirzel and Le Lay 2008), SDMs can be interpreted within the theoretical framework of ecological niche theory (Chase and Liebold 2003). For example, Kearney (2006) considers that correlative models based on distribution data are non-suitable to define niche per se, which could only be achieved by mechanistic models. Nevertheless, other authors (Franklin and Miller 2009) specify that SDMs cannot measure the fundamental niche from distribution data but only the realized niche, with the exception of sink populations (Pulliam 2000). Most niche-based SDMs are set at a large scale (when distributional limits matter), encompassing the full species’ geographic range or a section of this range (Pearson and Dawson 2003). Usually, they are developed on the basis of the current species distribution and a pool of abiotic variables describing the coarse-scale environmental requirements which allow population growth and persist. These models are interpreted as defining the ‘Grinnellian niche’ and contrasts with those characterizing the ‘Eltonian niche’, which are based on biotic interactions and define the functional niche at more detailed scales (Soberon 2007). However, this scale-dependent differentiation is probably misleading, because biotic and abiotic factors shape species distribution at all scales (Case et al. 2005, Morin and Lechowicz 2008, Soberón 2010). 90
95
100 Although the usefulness of SDMs has been demonstrated in many applications, there are still some methodological challenges that need to be addressed (Austin 2007). This is the case of scaling model predictions across spatial resolutions and extents. Spatial resolution (grain) refers to the ‘size’ of each observation and extent (domain) defines the study 105

0 area where the model is used to extrapolate from data (Elith and Leathwick 2009, Franklin and Miller 2009). Scaling is a
key process in modelling approaches since it allows for trans-
lating information from one scale to another (Wiens 1989,
King 1991, Wu 2004, Melbourne and Chesson 2005).
5 According to Lischke et al. (2007), this could be useful for:
1) making compatible ecological data gathered at different
scales; 2) translating model process from one scale to other;
and 3) building predictions on scales not accessible for direct
observation. Scaling can be done from small to large scales
10 (up-scaling) or viceversa (down-scaling). Upscaling can help
to understand complex ecological systems, as well as to iden-
tify source mechanisms relevant for target scale phenomena
(Lischke et al. 2007, Marcer et al. 2012). Downscaling can
be useful to analyze climate change effects on species distri-
15 bution shifts or to redefine conservation strategies at local
scale from the assessment of large scale distribution atlases
(Araújo et al. 2005, Keil et al. 2013). However, scaling can
be a difficult and problematic process (Bugmann et al. 2000)
since ecological systems are organized hierarchically (with
20 many feedback processes across scales) and also because they
are non-linear and spatially heterogeneous (as a consequence
of spatial variations in climate, edaphic properties and dis-
turbances). On the top of these constraints, choices on the
'source' and 'target' scales (Lischke et al. 2007) can also play
25 a relevant role in the performance of scaling exercises. The
selection of the 'source' scale generally depends on the avail-
able information more than on data adequacy for finding
causal relationships (Araújo and Guisan 2006, Austin 2007).
This fact is particularly problematic when the spatial resolu-
30 tion of the environmental variables does not match that of
the species data (Boyce 2006, McPherson et al. 2006), as it
may lead to erroneous predictions. Other limitation can be
arisen from the fact that model predictions are scale-depen-
dent (Trivedi et al. 2008), as the importance of predictors
35 for species may vary with the resolution and extent of the
analysis (Wiens 1995, Luoto et al. 2007, Menke et al. 2009).
Up to recent times, it has been widely accepted that, at large
scale, the major driver for species' distribution was climate
(Root 1988, Thuiller et al. 2005), while other abiotic (soils
40 or micro-habitat) or biotic factors were more relevant predi-
ctors at fine scale (Brown 1984, Wiens et al. 1987). However,
species distribution is actually explained by a combination of
abiotic and biotic predictors at all spatial scales (Case et al.
2005). For example, Normand et al. (2009), Meier et al.
45 (2011), Kissling et al. (2012) or Wisz et al. (2013), accord-
ing to the 'stress gradient hypothesis' (Menge and Sutherland
1987) recognized that species ranges are frequently limited
by abiotic constraints in scenarios of low resource availability
or stress. However, under more suitable environmental con-
50 ditions, they seem to be outcompeted by other species of the
same trophic level. This finding highlights the fact that biotic
interactions can also shape the range of species at large scale.
In spite of these questions, most modelling exercises consider
fine scale abiotic features (e.g. roughness or habitat edge) as
55 not very useful predictors at large spatial scales, where other
abiotic factors could overcome their importance. As a con-
sequence, large scale models may have problems to predict
suitable sites at small scales (when the importance of biotic
interactions or vegetation type increases). Furthermore, the
60 magnitude of all the above mentioned effects can vary along

biogeographic gradients due to differences in both environ- 61
mental suitability and habitat availability for species. As a
consequence, the ecological and biogeographic context of
the study area could affect performance (Osborne and
Suárez-Seoane 2002, Osborne et al. 2007) and scaling of 65
SDMs across spatial resolutions and extents. According to
Brown et al. (1995), species would tend to be more abun-
dant and generalists in the ecological centre of their distribu-
tion, where they would find a higher availability of suitable 70
habitats. In parallel, they would be rarer and more special-
ised towards the ecological boundaries, where they cope with
a reduction on habitat availability (Holt and Keitt 2000),
as well as on their fitness due to low habitat quality and
extreme environmental conditions (Brown 1984, Hoffman
and Blows 1994). Therefore, in ecologically marginal areas, 75
species would find their niche truncated (Braunisch et al.
2008). Niche truncation challenge current practices of scal-
ing in SDM approaches because models developed over small
extents can fail when extrapolated to larger extents (Austin
2007), as most of the suitable habitats found at large extent 80
are unavailable within smaller portions of this area (Pearson
and Dawson 2003, Thuiller et al. 2004).

In this paper we aim to assess how the performance and
success of scaling SDM predictions vary across spatial resolu-
tions and extents using the Iberian mole as study case. We 85
expect to find variations in these model parameters along
a biogeographic gradient, which be associated with differ-
ences in habitat availability (niche truncation) and spe-
cies responses. The particular questions to be explored are:
1) how does model performance change across scales along 90
the biogeographic gradient? We predict that SDMs will per-
form better at high spatial resolution and small extent, as
well as in marginal than in core ecological areas, because the
range of environmental conditions selected by the species
will be narrower (Brotons et al. 2004, Hernández et al. 2006, 95
Suárez-Seoane et al. 2008). 2) Do environmental responses
of species vary across scales and within the biogeographic
gradient? We hypothesize that the national model (large
extent) will be mainly driven by climate, while regional
models (small extent) should involve topography, vegetation 100
or human disturbances. 3) Does the success of scaling model
predictions across resolutions and extents changes along the
biogeographic gradient? Because scaling is probably context-
dependent, we cannot predict any a priori response along
the biogeographic gradient or among scales. 4) Could differ- 105
ences in performance and scaling of models among regions
be associated with changes in habitat availability? In the core
ecological area, the species should occupy a broader range
of suitable habitats and, therefore, niche would be non-
truncated. In ecologically marginal areas, niche would be 110
truncated because habitat availability would decrease.

Methods 115

The Iberian mole as the study system

The Iberian mole *Talpa occidentalis* is an endemic species of
the Iberian Peninsula (Loy 1999), where it is sympatric with
its European counterpart, the common mole *Talpa europaea*
(Kryštufek 1999). The Iberian species shows a wider 121

0 distribution, occupying the whole peninsular range, but
 1 the north-eastern corner, where only the common mole is
 2 present (Román 2007a, b). The Iberian mole is an IUCN
 3 Least Concern species (Fernandes and Herrero 2008) that
 4 may be found in a variety of temperate habitats, where soil is
 5 deep and moist all year-round, so long as it is not excessively
 6 stony, sandy or waterlogged. It prefers meadows, pastures and
 7 woodlands, as well as shrublands and rural gardens. In these
 8 habitats, moles find high availability of their preferred prey,
 9 earthworms (Funmilayo 1977, 1979, Niethammer 1990),
 10 but also good places for building their underground burrows
 11 (Macdonald et al. 1997). This species can be seen as a good
 12 study model to understand how aridity gradients constraint
 13 the distribution of temperate species in the Mediterranean
 14 region.

15

Setting the spatial scale of analysis and the biogeographic gradient

20 We defined two spatial scales with different resolution and
 21 extent: 1) a nationwide scale, covering the full Spanish
 22 range of the species, which was analysed at 10 km resolu-
 23 tion, as constrained by the Spanish Atlas of Mammals. 2)
 24 A regional scale, including three areas located along a bio-
 25 geographic (latitudinal, from north to south) gradient
 26 driven by temperature and precipitation: Galicia, Madrid
 27 and Granada (Fig. 1). Galicia holds the most suitable habi-
 28 tat and the highest mole abundance and, therefore, it can be
 29 considered as the core ecological area. Madrid and Granada
 30 support more extreme conditions for species survival and,
 31 then, both may be evaluated as ecologically marginal areas.
 32 All regions were mapped at 50 m resolution (i.e. pixel area
 33 was 2500 m²), since this value had a functional equiva-
 34 lency to European mole home range (estimated in 2324 m²
 35 by Macdonald et al. 1997), but also at 10 km resolution
 36 to assess model performance and scaling of models across
 37 resolutions.

38 Spanish topography, geology and climate are highly
 39 heterogeneous. Consequently, vegetation patterns are
 40 complex: from Atlantic Spain in the north and west,
 41 with extensive deciduous forests and rich grassy plains,
 42 to Mediterranean Spain in the south-east, with xero-
 43 phytic scrublands and sparse woodlands. In Galicia, the
 44 north-westernmost region, annual average rainfall exceeds
 45 900 mm and mean temperature is 13°C. Landscape is a
 46 fine-gained mosaic made of forests, crops, meadows and
 47 pasturelands. In order to maximize the environmental dif-
 48 ferences among study regions, we only considered those
 49 provinces with a typical Atlantic climate (A Coruña, Lugo
 50 and Pontevedra). Madrid lies in central Spain and holds
 51 a more diverse landscape, including xeric open habitats
 52 and deciduous woodlands. Winters are cold and sum-
 53 mers are hot. Annual average precipitation is below 500
 54 mm and mean temperature is 14°C. Granada is located
 55 in the southeasternmost corner of the country. It supports
 56 extreme variations in landscape and climate, associated
 57 with a sharp altitudinal gradient. The climate is hot and
 58 dry in summer and cold in winter. Annual average rain-
 59 fall is around 350 mm and mean temperature is 15°C.
 60 Landscape is dominated by dry and scattered scrublands.

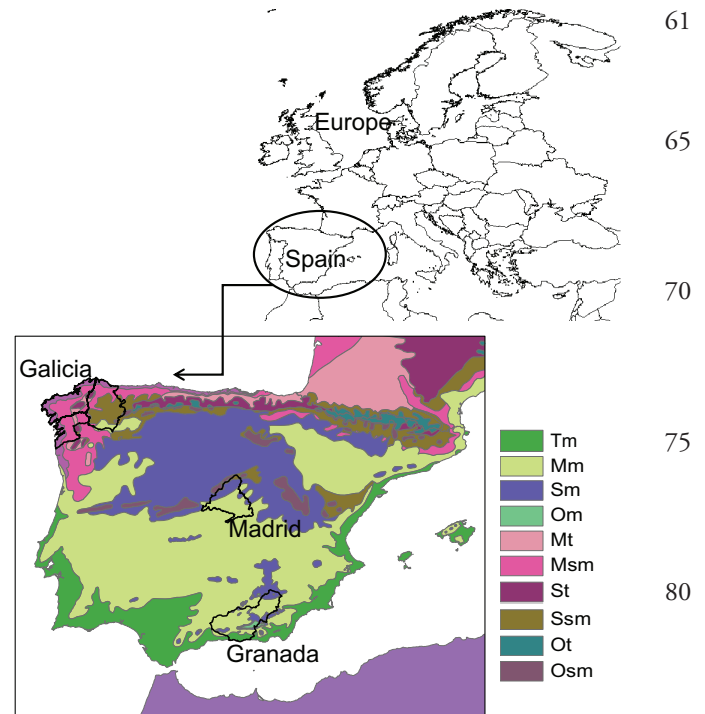


Figure 1. Location of the study area. Target regions are placed across a biogeographic gradient representative of different thermoclimatic belts. Mediterranean variants: Tm (Thermomediterranean), Mm (Mesomediterranean), Sm (Supramediterranean), Om (Oromediterranean). Temperate variants: Mt (Mesotemperate), Msm (Mesosubmediterranean), St (Supratemperate), Ssm (Suprasubmediterranean), Ot (Orotemperate), Osm (Orosubmediterranean). From Rivas-Martínez et al. (2004).

Mole data collection

An independent set of species occurrence data was collected at each scale. At the national scale, we used the Spanish Atlas and Red Book of Mammals (Román 2007a), which was built on the basis of direct observations of molehills (Mead-Briggs and Woods 1973, Edwards et al. 1999) for the period 1970–2000. We extracted the central coordinates of each 10 × 10 km square that was occupied by the species. Despite this coarse resolution, the relatively specialized mole habitat selection guarantees the suitability of the presences' dataset for modelling exercises. Since some of the reported absences could be imprecise because of insufficient sampling effort, we assume that absences are a mixture of true absences and pseudo-absences. At regional scale, we conducted field sampling in autumn–winter (October–December) of 2006–2008. During these months, males are more detectable since they show a high peak of activity associated with the construction of subterranean nests, which is coincident with a second mating period (Jiménez et al. 1990). Data were collected in 50 × 50 m plots, randomly located through each study area (196 in Galicia; 117 in Madrid and 114 in Granada). Plot size corresponds to the pixel size considered in further GIS analyses. Plots were walked by two-three people to exhaustively search for molehills or other signs of the presence (e.g. above-ground tunnels). Complementary we also extracted, for each region, the presence records from the

0 Atlas in order to model regional species distribution at coarse resolution (10 × 10 km).

5 **Environmental predictors**

10 We selected a pool of 11 environmental variables potentially affecting species distribution (climate, topography, hydrography, vegetation and human disturbances; Table 1, Supplementary material Appendix 1). Predictors were chosen according to previous knowledge on the habitat preferences of the Iberian mole (Loy 1999, Román 2007a) and other closely related species (Funmilayo 1979, Román 2007b). The number of predictors was relatively high to allow for flexibility when fitting models, since we expect important variations in habitat preferences across scales along the biogeographic gradient. Predictors were measured separately for each study area and scale. They were derived by applying identical GIS algorithms from different sources, when possible, to make them comparable. The size of the array (5 × 5 cells) used in the context analyses was selected according to the movement capacity of European moles at a local scale, which is a radius of 100 m around the external border of the territory (Macdonald et al. 1997). We used the normalized difference vegetation index (NDVI; Rouse et al. 1973, Pettorelli et al. 2005) as a proxy of aboveground net primary production (Goward et al. 1994, Paruelo et al. 1997, Osborne et al. 2001). This index has emerged as a powerful indicator for exploring the link between animal distribution and available resources (Pettorelli et al. 2006, 2011, Wiegand et al. 2008). It was estimated from a monthly series of cloud-free maximum value composite derived from NOAA-AVHRR imagery at 1 km resolution for the period 1983–1999 (see Suárez-Seoane et al. 2002, 2004 for details on the characteristics of the imagery). NDVI is based on the reflectance difference between visible and near infrared regions

of the electromagnetic spectrum in channels 1 and 2 of the AVHRR images. The range of NDVI values lies between -1 and +1. Only the positive values correspond with vegetated zones. Climatic variables were measured for the summer period, since this is the most limiting season for temperate species in the Mediterranean; Virgós and Tellería 1998, Virgós et al. 2004). To prevent multicollinearity effects, we verified that pairwise Spearman correlations among all predictors were under the threshold of 0.8 (Tabachnick and Fidell 1996).

75 **Measuring performance and success of scaling model predictions across scales along the biogeographic gradient**

80 [AQ4] We fitted separate models to characterize mole distribution at different spatial resolution and extent using the Maximum Entropy algorithm (MaxEnt 3.3.3; Phillips et al. 2006, Phillips and Dudík 2008). We selected this method due to its good performance in comparison with other SMD techniques, particularly when species records are presence-only data derived from non-systematic biological surveys, as is the case for the Atlas data (Elith et al. 2006, 2010). In MaxEnt, the reliability of the predictions strongly depends on the number of ‘easy negatives’ in the background data used to represent the available environment, which lastly may affect prevalence. Then, selecting background points in unlike areas for the target species may result in over-prediction. On the other hand, many ‘easy positives’ in the background data have the opposite effect. See Chefaoui and Lobo (2008), Franklin and Miller (2009), Elith and Leathwick (2009) or Elith et al. (2010) for a discussion on this topic. To enhance model reliability and comparability, we created a sample of 10 000 randomly distributed background points for each model, avoiding points within highly

[AQ1] Table 1. Environmental variables used as predictors to train Maxent models (see Supplementary material Appendix 1 for more details).

| Code | Description of the variable | Source and input resolution of data | |
|---|--|--|--|
| | | National scale (10 × 10 km) | Regional scale (50 × 50 m) |
| Topography TOPOV5 | Topographic variability. It varies from ‘0’ (flat areas) to ‘1’ (maximal variability) | DEM 200 m | DEM 50 m |
| Hydrography RIVERDIST | Distance to the nearest river | Hydrography at 1:200 000 | Idem at 1:50 000 |
| Vegetation pattern PASTDENS WOODDENS | Percentage of meadows and pastures Percentage of broad-leaved and mixed woodlands | CORINE Land Cover Map (Heymann et al. 1994) at 1:100 000 | Regional habitat cartography of Madrid (CAM 1998), Galicia (SITGA 1998) and Andalucía (SINAMBA 1999) at 1:50 000 |
| Above ground net primary production NDVIMAX NDVIMIN | Annual maximum of NDVI Annual minimum of NDVI | NDVI from NOAA-AVHRR 1 km (monthly series, period 1983–1999) | |
| Human disturbances TOWNDIST ROADDIST | Distance to the nearest town Distance to the nearest road–trail | Towns and roads at 1:200 000 | Idem at 1:50 000 |
| Climate TMAXSUM PRECSUM | Maximal temperatures (°C) in summer (June to August) Mean precipitation (mm) in summer (June to August) | WorldClim data (Hijmans et al. 2005) (1965–2005) at 1 km | |

0 unsuitable areas (terrain slope higher than 60%; personal
 observations) and within a radius around the presence loca-
 tions used for model training (5×5 pixels for 50 m models,
 but only one pixel for 10 km models due to the reduced num-
 ber of pixels at this resolution). Each model was built using
 5 70% of the occurrence points (training dataset), retaining
 the remaining 30% for evaluation (testing dataset). Model
 performance was evaluated internally (within the same geo-
 graphic area and at the same spatial resolution at which it
 was built) by calculating the AUC (area under the receiver
 10 operating characteristic curve; Beck and Shultz 1986, Zweig
 and Campbell 1993, Fielding and Bell 1997), adapted for
 use with background samples, in the next three cases: 1) full
 models built from the complete training dataset and tested
 against calibration data with no independence (re-substi-
 tution test; Edwards et al. 2006); 2) 10-fold cross-validated
 15 models ran on the training dataset (Verbyla and Litvaitis 1989,
 van Houwelingen and Le Cessie 1990); 3) full models tested
 against the 30% set-aside independent data that were not used
 for model calibration. Complementarily, we also measured the
 20 Boyce index (Boyce et al. 2002, Hirzel et al. 2006) on the test-
 ing samples. This index measures the Spearman rank correla-
 tion coefficient between the frequency of presence locations
 within ranked classes of predicted habitat suitability (adjusted
 by the area) and the mean habitat suitability of these classes.
 25 It ranges from -1 to $+1$. AUC and Boyce index are threshold-
 independent SDM evaluation methods fairly insensitive to spe-
 cies prevalence (Franklin and Miller 2009). Full models were
 also externally evaluated with independent species data (testing
 datasets collected at a different resolution or extent, according
 30 to the case) through AUC and Boyce evaluators. To allow for reli-
 able estimates, we always compared pairs of models by using
 the same testing dataset. So, that the measured performance
 has the same meaning (Phillips pers. comm.).

The success of scaling model predictions across spa-
 tial resolutions and extents was quantified by applying an
 35 adaptation of the method of Randin et al. (2006), which
 was first used to measure model transferability across space.
 This approach was also applied by Tuanmu et al. (2011) to
 measure temporal transferability of wildlife habitat models.
 40 According to them, a successful scaling of models could be
 achieved if: 1) internal evaluation of models fitted in differ-
 ent frameworks is similar; 2) a model trained at a particu-
 lar spatial resolution/geographic area retains a comparable
 external evaluation when projected into other situation; and
 45 3) spatial predictions made by two models fitted in different
 frameworks match for each species location. Criteria 1 and
 2, based on AUC values, were assessed through Eq. 1; while
 the third criterion, based on spatial predictions, was evalu-
 ated with Eq. 2.

$$T1 = \frac{\frac{1}{2} \left(\left(1 - \frac{|AUC_{a \rightarrow a} - AUC_{a \rightarrow b}|}{0.5} \right) + \left(1 - \frac{|AUC_{b \rightarrow b} - AUC_{b \rightarrow a}|}{0.5} \right) \right)}{1 + \left| \frac{|AUC_{a \rightarrow a} - AUC_{a \rightarrow b}|}{0.5} - \frac{|AUC_{b \rightarrow b} - AUC_{b \rightarrow a}|}{0.5} \right|} \quad (1)$$

$$T1 = \frac{\sum_{i=1}^n HS_{ia} - \sum_{i=1}^n \min(HS_{ia}, HS_{ib})}{\sum_{i=1}^n HS_{ia}} \quad (2)$$

Equation 1 (T1) computes cross-scaling (a combination
 of up and downscaling estimates) of models across spatial
 resolutions using both internal and external measurements
 of validation (all based on testing samples). $AUC_{a \rightarrow a}$ and
 $AUC_{b \rightarrow b}$ are the internal validation of models that have been
 65 fitted and evaluated in a particular area and spatial resolu-
 tion. For example, $AUC_{a \rightarrow a}$ result from validating a model
 calibrated in Granada at 10km with the testing dataset cre-
 ated for that region at that resolution. $AUC_{a \rightarrow b}$ and $AUC_{b \rightarrow a}$
 are the external validation of models that have been fitted at
 70 other resolution. In the example, $AUC_{a \rightarrow b}$ result from vali-
 dating a model calibrated in Granada at 10km with the test-
 ing dataset created at 50 m for that region (i.e. downscaling),
 while $AUC_{b \rightarrow a}$ means the reverse situation (i.e. upscaling
 from 50 m to 10 km). Index values range from -1 (when
 75 internal validation reaches the maximum value, but external
 is null) to 1 (when internal and external validation values
 are maximal). Equation 2 (T2) is based on the Kulczynski's
 coefficient (Legendre and Legendre 2012) and quantifies the
 scaling of models across resolutions according to the distance
 80 between the spatial logistic predictions (habitat suitabil-
 ity index; HS_{ia} , HS_{ib}) made by two models fitted at differ-
 ent resolutions for each location of the testing sample. For
 example, the index may compare the predictions made by a
 model fitted in Madrid region at 10 km (situation a) with
 85 the predictions made for the same region at 50 m (situation
 b), using either the testing sample collected at 10 km or that
 at 50 m. Therefore, T2 can be used to estimate downscaling
 (T2a; if calculate the distance from logistic predictions made
 by a model set at low resolution with predictions made at
 90 high resolution) or up-scaling (T2b; the reverse situation).
 Index values varies from 0 (HS_{ia} and HS_{ib} are identical and
 scaling perform successfully) to 1 (differences between both
 sets of values are maximal and scaling perform badly).

To evaluate model scaling across extents (i.e. from national
 95 to regional level and viceversa), we also used Eq. 2. We only
 operated at 10 km resolution, since we had no data for the
 whole Spain at 50 m resolution.

Estimating the availability of suitable habitat along the biogeographic gradient

To understand the eventual relationships between perfor-
 mance and scaling of models with the availability of suitable
 habitat along the biogeographic gradient (i.e. niche trunca-
 105 tion), we calculated marginality and tolerance (Hirzel and
 Le Lay 2008) by using the ENFA analysis of BioMapper 4.0
 (Hirzel et al. 2002, 2007). Traditionally, marginality (M; Eq.
 3) has been understood in terms of how much the used habi-
 tat differs from the available (Hirzel et al. 2002, 2007), with
 110 values ranging from 0 (species living in average conditions
 throughout the study area) to 1 or higher (using extreme
 habitats) (Basille et al. 2008). Tolerance (To; Eq. 4) has been
 commonly considered as an index of specialization, which
 115 reports the ratio between the variance of available conditions
 and the variance of the used conditions. It ranges from 0
 (specialist or stenoecious species) to 1 (generalist or euriecious
 species). However, both marginality and tolerance could be
 best used as indicators of habitat availability and, therefore,
 of niche truncation. Marginality can be seen as a measure of
 121

0 how niche conditions are represented in a particular location. For example, if only a part of the niche is present in a geographic area, niche would be truncated (Austin 2002, Thuiller et al. 2004). In that case, species would not be able to select a habitat over the whole niche conditions offered at the whole distribution, but only over a narrow fraction. In such a situation, we could conclude an apparent specialization of the species, which is actually an artifact. Therefore, marginality can be re-interpreted as follows: values close to 0 indicate high availability of the whole niche conditions and values close to 1 (or higher) indicate a niche truncation. In the case of tolerance, values close to 0 can be related to niche truncation, whereas values close to 1 indicate that the whole niche conditions are available for the species.

$$15 \quad M = \text{sqrt} (\sum_{i=1,v} [M_i^2])/1.96 \quad (3)$$

$$20 \quad T_o = 1/(\text{sqrt} (\sum_{i=1,v} (\lambda_i)/V)) \quad (4)$$

Where sqrt is the square root function; M_i are the coefficients of the marginality factor; λ_i are the eigenvalues and V is the number of variables (Hirzel et al. 2002). The normality of each variable was previously verified according to a Kolmogorov–Smirnov test and, when necessary, variables were Box–Cox transformed.

25

Results

We collected 823 presence records at 10 km resolution from the Atlas for modelling Spanish distribution. From this dataset, 67 points were in Galicia, 28 in Madrid and 35 in Granada. By fieldwork we obtained 74 data at 50 m resolution in Galicia, 45 in Madrid and 35 in Granada.

35 Model performance and distribution patterns across scales along the biogeographic gradient

Table 2 summarizes the results of model validation across spatial resolutions and extents along the biogeographic gradient. At regional scale, the performance (AUC and Boyce values) of models set at high resolution (50 m) increased along the biogeographic gradient, from north to south. However, models set a coarse resolution (10 km) did not fit this trend, since Granada performed the worst. Looking at the internal evaluation, regional full models performed better at the most detailed resolution than at the coarsest one. When these models were cross-validated and tested against the 30% set-aside of data not used for calibration, AUC values reduced soundly, particularly in the case of coarse resolution models. In specific, values were lower than 0.5 in Galicia and Granada, indicating that 10 km-models performed worse than random in these regions. External validation across spatial resolutions showed that, when model predictions were downscaled, AUC and Boyce values were moderate and increased along the biogeographic gradient. However, when predictions were upscaled, AUC and Boyce values were much weaker (AUC was in all cases equal or lower than 0.5 and Boyce index showed negative values). In this case, the worst values were obtained in Granada. Regarding the effect of the extent, the national model had

a better internal evaluation than the regional models ran at the same resolution. When the 10 km-national model was downscaled across extents, AUC values (AUC_{test} : 0.82, 0.85 and 0.55 respectively for Galicia, Madrid and Granada) were higher than those of models fitted at the same resolution specifically for each region (AUC_{test} : 0.47, 0.65 and 0.44 respectively). Indeed, when the national model was downscaled across both resolutions and extents, results (AUC_{test} : 0.77, 0.91 and 0.79 respectively) were better than those obtained when fitted regional models at 50m (AUC_{test} : 0.62, 0.83 and 0.96 respectively) in two of the three cases, with the exception of Granada.

The spatial pattern of species occurrence at national scale was very close to the picture published by the Atlas (Fig. 2). Both identified two core areas (north-western square of the country and southernmost Betic Mountains) and other scattered and small marginal populations in eastern Spain. The model successfully excluded the area near Pyrenees, which is occupied by *Talpa europaea*. At regional level, important differences were detected when comparing models at both spatial resolutions (Fig. 3). Models ran at coarse resolution overestimated habitat suitability in comparison with models at detailed resolution. The range predicted by fine-resolution models strongly differed along the biogeographic gradient, becoming more restricted from north to south. At national scale, the most relevant driver of mole distribution was climate. At regional level, there were marked differences on the most contributing variables along the biogeographic gradient. In Galicia, species occurrence was linked to similar environmental factors as those identified at national level. In Madrid, the effect of climate was replaced by vegetation. In Granada, the main drivers were topography and human disturbances. Results were consistent across spatial resolutions, but in Granada (Table 3).

Species responses and habitat availability along the biogeographic gradient

Mole responses to environmental factors differed between spatial resolutions and along the biogeographic gradient (Fig. 4). Curves drawn at coarser resolution were broader than those built at detailed resolution. Galicia was the region where mole responses were more similar to those achieved in the Spanish model, while Granada was the most different. Marginality was lower, on average, on national than regional levels. It deeply changed along the biogeographic gradient, with the highest value in Granada and the lowest in Galicia. In the case of tolerance, the trend was the opposite. Tolerance was higher at national than regional scale and decreased from Galicia to Granada (Table 4). In summary, niche truncation increased (while habitat availability decreased) from national to regional scale, as well as from Galicia to Granada.

Scaling of models across resolutions and extents along the biogeographic gradient

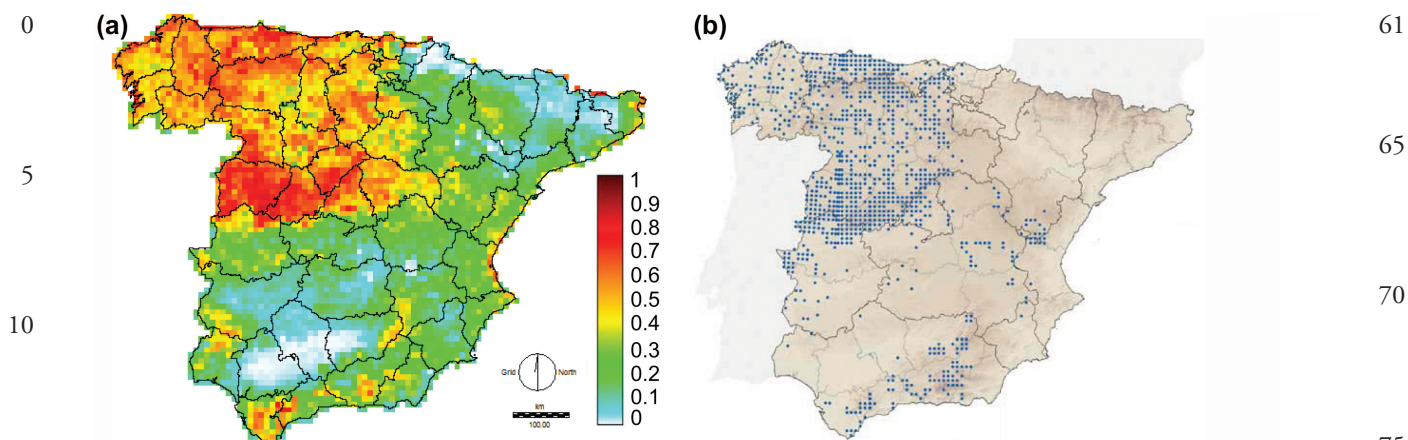
According to Eq. 1 and 2 (Table 5), the success of scaling model predictions across both spatial resolutions and extents decreased along the biogeographic gradient (from Galicia to Granada). The efficiency of downscaling practices (T2a; from coarse to fine resolution; from large to small extent)

Table 2. AUC and Boyce index values obtained from internal and external model evaluation. Models were fitted at national (Na) and regional scale (Ga: Galicia, M: Madrid, Gr: Granada) along a biogeographic gradient. They were evaluated internally (at the same spatial extent and resolution at which they were built) in three different ways: 1) by re-substitution of the full models ('full'); 2) by 10-fold cross-validation ('10cv'); and 3) by testing against the 30% set-aside of the original data not used for model calibration ('test'). Full models were as well evaluated externally ('test') with independent species datasets collected at a different spatial resolution/extent. To allow for reliable evaluations, we always compared models by using the same testing dataset. AUC values range from 0 to 1 and Boyce index from -1 to 1. Each row represents a different model; columns correspond to the test datasets the models were compared against.

| | Model validation | | | | | | |
|-------------------|--|---|---|--|--|---|---|
| | National 10 km | Galicia 10 km | Galicia 50 m | Madrid 10 km | Madrid 50 m | Granada 10 km | Granada 50 m |
| Model calibration | | | | | | | |
| National 10 km | <p>AUC_{full}Na10 km = 0.79</p> <p>AUC_{10cv}Na10 km = 0.77 ± 0.03</p> <p>AUC_{test}Na10 km → Na10 km = 0.78</p> <p>BoyCE_{test}Na10 km → Na10 km = 1.00</p> <p>AUC_{test}Ga10 km → Na10 km = 0.47</p> <p>BoyCE_{test}Ga10 km → Na10 km = 0.17</p> | <p>AUC_{test}Na10 km → Ga10 km = 0.82</p> <p>BoyCE_{test}Na10 km → Ga10 km = 0.24</p> <p>AUC_{full}Ga10 km = 0.72</p> <p>AUC_{10cv}Ga10 km = 0.55 ± 0.15</p> <p>AUC_{test}Ga10 km → Ga10 km = 0.47</p> <p>BoyCE_{test}Ga10 km → Ga10 km = 0.17</p> <p>AUC_{test}Ma10 km → Ga10 km = 0.17</p> <p>AUC_{test}Ga50 m → Ga10 km = 0.49</p> <p>BoyCE_{test}Ga50 m → Ga10 km = -0.30</p> | <p>AUC_{test}Na10 km → Ga50 m = 0.77</p> <p>BoyCE_{test}Na10 km → Ga50 m = 0.02</p> <p>AUC_{full}Ga50 m = 0.80</p> <p>AUC_{10cv}Ga50 m = 0.68 ± 0.09</p> <p>AUC_{test}Ga50 m → Ga50 m = 0.62</p> <p>BoyCE_{test}Ga50 m → Ga50 m = -0.15</p> | <p>AUC_{test}Na10 km → Ma10 km = 0.85</p> <p>BoyCE_{test}Na10 km → Ma10 km = 0.51</p> <p>AUC_{test}Ma10 km → Ma10 km = 0.61</p> <p>BoyCE_{test}Ma10 km → Ma10 km = 0.65</p> <p>AUC_{full}Ma50 m = 0.89</p> <p>AUC_{10cv}Ma50 m = 0.79 ± 0.19</p> <p>AUC_{test}Ma50 m → Ma50 m = 0.83</p> <p>BoyCE_{test}Ma50 m → Ma50 m = 0.08</p> | <p>AUC_{test}Na10 km → Ma50 m = 0.91</p> <p>BoyCE_{test}Na10 km → Ma50 m = 0.48</p> <p>AUC_{test}Ma10 km → Ma50 m = 0.61</p> <p>BoyCE_{test}Ma10 km → Ma50 m = 0.70</p> <p>AUC_{full}Ma50 m = 0.89</p> <p>AUC_{10cv}Ma50 m = 0.79 ± 0.19</p> <p>AUC_{test}Ma50 m → Ma50 m = 0.83</p> <p>BoyCE_{test}Ma50 m → Ma50 m = 0.08</p> | <p>AUC_{test}Na10 km → Gr10 km = 0.55</p> <p>BoyCE_{test}Na10 km → Gr10 km = -0.46</p> <p>AUC_{test}Ma10 km → Gr10 km = 0.61</p> <p>BoyCE_{test}Ma10 km → Gr10 km = 0.70</p> <p>AUC_{full}Gr50 m = 0.96</p> <p>AUC_{10cv}Gr50 m = 0.90 ± 0.15</p> <p>AUC_{test}Gr50 m → Gr50 m = 0.96</p> <p>BoyCE_{test}Gr50 m → Gr50 m = 0.70</p> | <p>AUC_{test}Na10 km → Gr10 km = 0.55</p> <p>BoyCE_{test}Na10 km → Gr10 km = -0.46</p> <p>AUC_{test}Ma10 km → Gr10 km = 0.61</p> <p>BoyCE_{test}Ma10 km → Gr10 km = 0.70</p> <p>AUC_{full}Gr50 m = 0.96</p> <p>AUC_{10cv}Gr50 m = 0.90 ± 0.15</p> <p>AUC_{test}Gr50 m → Gr50 m = 0.96</p> <p>BoyCE_{test}Gr50 m → Gr50 m = 0.70</p> |
| Galicia 10 km | | | | | | | |
| Galicia 50 m | | | | | | | |
| Madrid 10 km | | | | | | | |
| Madrid 50 m | | | | | | | |
| Granada 10 km | | | | | | | |
| Granada 50 m | | | | | | | |

0
5
10
15
20
25
30
35
40
45
50
55
60

61
65
70
75
80
85
90
95
100
105
110
115
121



15 Figure 2. Spatial predictions of Iberian mole occurrence calculated at the national level at 10 km resolution. (a) Logistic continuous output. (b) Species distribution pattern from the Spanish atlas and Red book of mammals (Román, 2007a).

20 depended less on the biogeographic and ecological context of the target study area (index values were more similar along the gradient) than in the case of up-scaling (T2b; from fine to coarse resolution; from small to large extent).

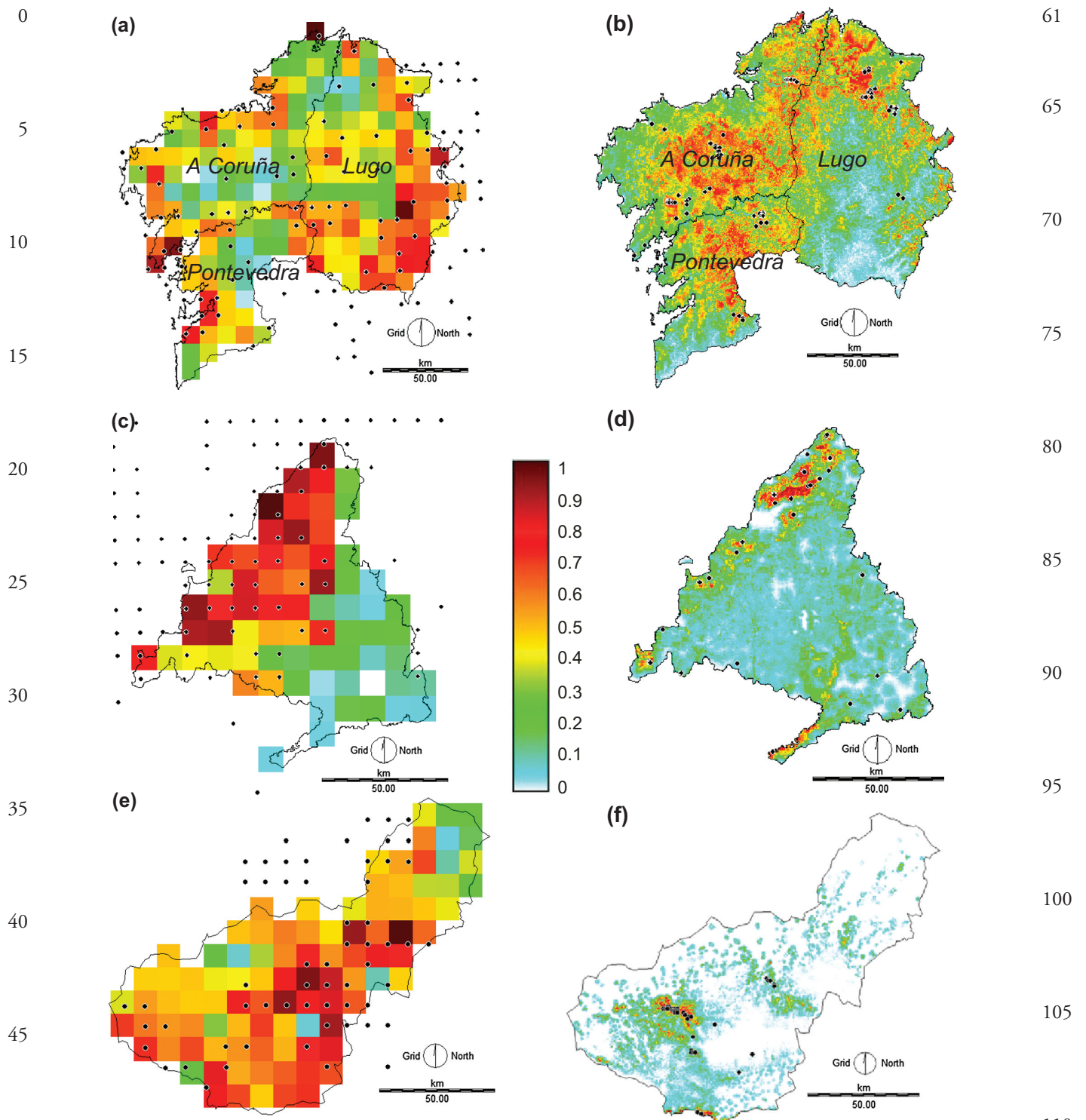
25 Discussion

Our results showed that model performance was dependent on spatial resolution and extent of the data. Indeed, both performance and scaling of models varied along the biogeographic gradient linked to differences in species environmental responses and also in the availability of suitable habitat which, in turn, is related to niche truncation. Models representing non-truncated niches were more successfully scaled across resolutions and extents (particularly in areas that did not offer all possible habitats for the target species), which has important implications for SDM applications, as those related to conservation planning under climate change (Thuiller et al. 2004, Seo et al. 2009).

40 As a trend, regional SDMs trained at coarse resolution performed worse than more detailed models. The weak performance of these models (see values of cross-validation and independent test with the 30% data-aside) could be, at least partially, explained by low sample size (Wisze et al. 2008) and/or sampling bias. Atlas data coverage across regions could be affected by different sampling effort and observer experience, which may imply either over or underestimation of the real distribution. Additionally, other authors as Luoto et al. (2007) or Menke et al. (2009) suggested that low resolution models may fail because they over-predict species occurrence since they tend to select climate variables against variables reflecting small-scale variations on vegetation or human disturbances. However, in our case, there were little differences between pairs of models trained for the same region at different resolution in terms of the nature of the key environmental predictors. Actually, differences were stronger between regions (Galicia was driven by climate, while vegetation, topography or human disturbances were more important in the other areas). Therefore, the need to consider variables accounting for vegetation or human disturbances

not necessarily depend on spatial resolution, but on other factors such as the location of the study area along environmental gradients, which may determine niche truncation, availability of suitable habitat, environmental responses (McPherson and Jetz 2007, Menke et al. 2009, Gottschalk et al. 2011) and particular life-history traits of target species. Specialist species (or species with low movement capacity) may need fine-grained variables, while generalists (or highly mobile species) do not respond so clearly to these variables (Suárez-Seoane and Baudry 2002). In Galicia (the core ecological area), moles showed broader environmental responses and were more widely distributed than in other regions of the biogeographic gradient. The explanation could be related to the availability of suitable habitat that decreased (as marginality and tolerance suggested) from North to South along the biogeographic gradient. Particularly, in the most marginal region (Granada), habitat suitability was very low and adverse climatic conditions, as well as a lack of plastic responses, could hamper the exploitation of alternative habitats. Consequently, it is of critical importance to explicitly consider the full ecological (both biological and environmental) and biogeographic context of target study areas when developing a model.

The way in which populations interact with environmental factors is also linked to ecophysiological stress, historical factors or dispersal opportunities (Costa et al. 2008), as well as interspecific interactions like competition that may drive local structure (Brown et al. 2000, Glen and Dickman 2005). The biogeographic gradient here analyzed is associated with a transition from the Atlantic to the Mediterranean region and it reflects critically diverse environmental conditions which, in a broad sense, relates to an increasing degree of aridity (higher air temperatures and lower water and food availability). Aridity is a key factor that determines the level of ecophysiological stress and, therefore, constrains species occurrence and abundance, especially for temperate species (Lara-Romero et al. 2012). It provides a proxy for the selection pressures experienced by animals (Mueller and Diamond 2001, Rezende et al. 2004), since it is directly related to primary production (Emberger 1955) which, in turn, may be considered as an index of food availability. In the case of



50 Figure 3. Comparison between model outputs obtained at both spatial resolutions for the three regions located along the biogeographic gradient: (a, b) Galicia at 10 km vs 50 m; (c, d) Madrid at 10 km vs 50 m; (e, f) Granada at 10 km vs 50 m. Points in (a), (c) and (e) are the Atlas centroids; points in (b), (d) and (f) are fieldwork data.

55 moles, earthworm abundance decreases with increasing aridity (Virgós et al. 2004), which may explain the reduced species occurrence in southern areas of the Iberian Peninsula. According to our results, temperature and rainfall were the most relevant predictors in Galicia, whereas production had a minimal contribution, probably because this is not a limiting factor in this region. We did not expect to find

115 climate as the main driver of species distribution in the core ecological region, but a probable explanation is that Galicia shows a strong gradient of temperature from the interior to the coast and, therefore, moles select particular climatic conditions. Complementarily, according to the 'stress gradient hypothesis' (Menge and Sutherland 1987), we should expect biotic interactions (competition) driving species distribution

| Variable | Spain | Galicia 10 km | Galicia 50 m | Madrid 10 km | Madrid 50 m | Granada 10 km | Granada 50 m |
|-----------|-------------|---------------|--------------|--------------|-------------|---------------|--------------|
| TOPO5 | 0.3 | 7 | 3.9 | 2.9 | 0.8 | 2.1 | 52.2 |
| RIVERDIST | 2.3 | 9.5 | 0.7 | 3.6 | 0.5 | 39.9 | 2.2 |
| PASTDENS | 12 | 0.9 | 0.1 | 1.4 | 0.7 | 0 | 0.4 |
| WOODDENS | 0.3 | 10.5 | 17.2 | 13.2 | 11.8 | 12.7 | 1.3 |
| PATCHSIZE | 1.3 | 8.6 | 0.5 | 70.2 | 28.7 | 1.2 | 0.7 |
| NDVIMAX | 18 | 3.5 | 0.2 | 5.1 | 2.9 | 2 | 2.5 |
| NDVIMIN | 2.4 | 13.1 | 1.8 | 2.9 | 39 | 5 | 1.1 |
| TOWNDIST | 1.3 | 1.5 | 3.7 | 0.5 | 5 | 2.1 | 24.0 |
| ROADDIST | 1.8 | 6.6 | 2.3 | 0 | 6.5 | 19.4 | 0.8 |
| TMAXSUM | 36 | 2.1 | 24.9 | 0.1 | 4.1 | 15.5 | 5.2 |
| PPSUM | 24.2 | 36.6 | 44.8 | 0 | 0 | 0.3 | 9.5 |

15 in Galicia (the core ecological area), while in Granada (the
 20 most ecologically marginal area), environmental constraints
 would explain better species responses. However, further
 research should be done in this respect. According to this
 idea, topography and human disturbances were actually the
 25 main drivers of mole distribution in Granada. In contrast
 with other regions, moles preferred here lower summer tem-
 perature (to avoid high physiological stress) and precipita-
 tion. The latter was unexpected, but may be due to the fact

that selected habitats by moles in this region include irri-
 gated crops and orchards that depend on the watering of
Populus plantations. Therefore, moisture and, as a result,
 earthworm availability do not depend so strongly on rainfall
 (García-López de Hierro et al. 2013). In Madrid, vegetation
 was the strongest explanatory variable, while climate was
 not an important constraint. Here, Iberian moles benefit
 from a good diversity of both climatic conditions and veg-
 etation types, as with other Eurosiberian species that live at

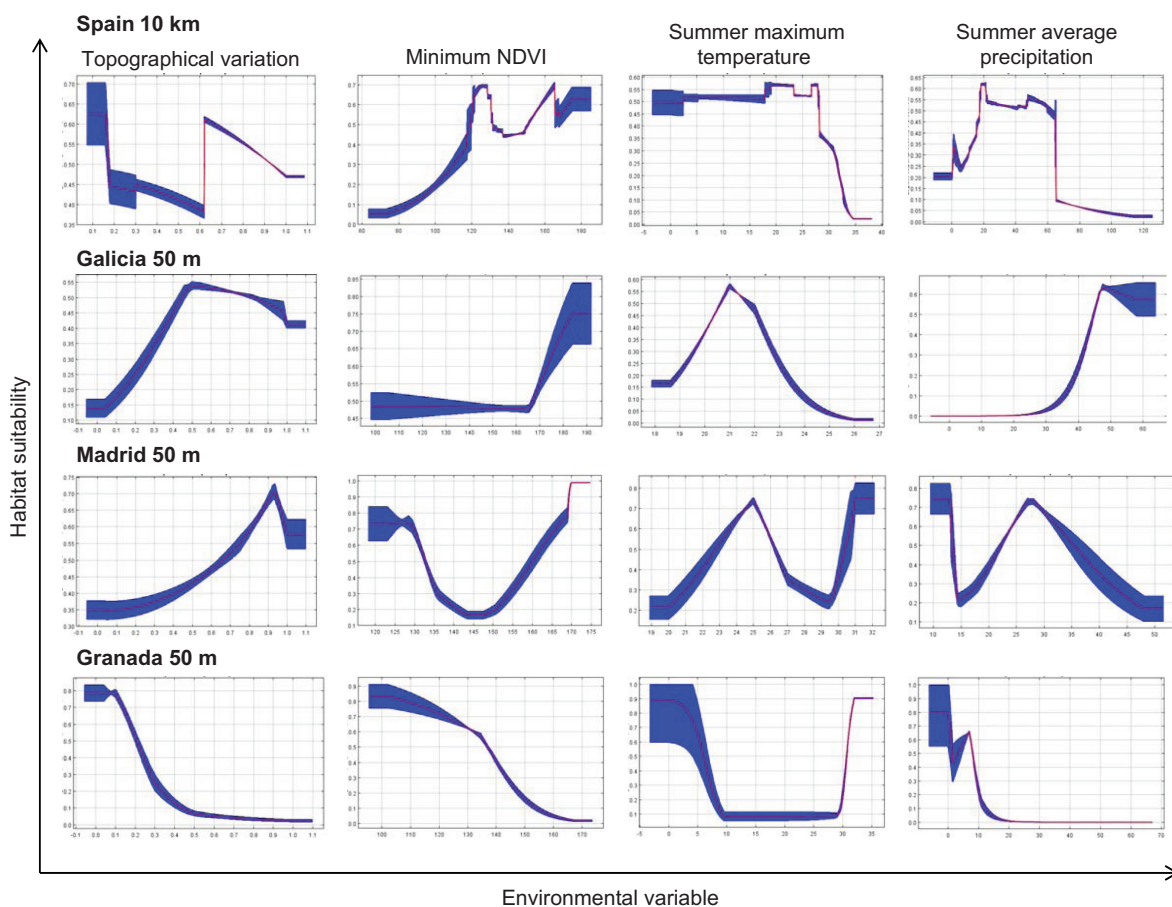


Figure 4. An example of response curves for some of the most relevant environmental variables when predicting Iberian mole distribution at different scales. Curves show how the logistic output (habitat suitability) change as each environmental variable is varied. Each curve represents a different model created using only the corresponding variable and shows the mean response of the 10 replicate MaxEnt runs (red) and the mean +/- one standard deviation (blue). NDVI original values (from -1 to 1) were stretched to 256 levels to maximize differences among values.

0 Table 4. Variations in marginality and tolerance across the biogeographic gradient calculated by using the ENFA analysis of BioMapper. A value of marginality close to 0 indicates a high availability of the whole niche conditions, whereas a value close to 1 (or higher) means niche truncation. Tolerance is bounded between 0 (niche truncation) to 1 (whole niche conditions are available for the species).

5

| | Spain* | Galicia* | Madrid** | Granada** |
|-------------|--------|----------|----------|-----------|
| Marginality | 0.419 | 0.406 | 0.790 | 1.439 |
| Tolerance | 0.897 | 0.651 | 0.369 | 0.298 |

10 *Patchsize was Box-Cox transformed.
 ** Patchsize, pastdens and wooddens were Box-Cox transformed.

15 relatively high densities in the mountains (Virgós and Tellería 1998, Virgós and Casanovas 1999). Mole prey reaches a high abundance associated with pasturelands and broad-leaved woodlands (Virgós et al. 2004).

20 Regarding the national model, it is highly remarkable that it performed better, in almost all cases, when validated (downscaled) within regions than models specifically trained at regional scale at 10 km and 50 m resolutions. Different explanations may be put forward to explain these results as a combination of biological, sampling and environmental features. A model with predictors covering a wider range in the training region (national model) is more likely to give accurate predictions when downscaled in a test area (regional models) than the reverse (Thuiller et al. 2004, Randin et al. 2006, Menke et al. 2009). In this sense, our national model incorporated a higher heterogeneity of environmental responses which allowed for detecting local variations more effectively (Osborne and Suárez-Seoane 2002, Barbosa et al. 2009). Therefore, this model fitted the non-truncated niche of the species much better than did any regional model. The success of downscaling across extents may also indicate a common pattern of species preferences across geographic ranges for some environmental conditions, such as rainfall or temperature, which were identified as important predictors regardless spatial scale. Indeed, rainfall and temperature can be considered as very good proxies of earthworm availability (Virgós et al. 2004) and this factor is probably very important to determine mole distribution at different extents and resolutions. Other explanation could be that sampling coverage at national scale was good and predictors selected at fine scale were also relevant at large scale (e.g. rainfall regime or temperature).

45 The success of scaling model predictions across resolutions and extents decreased from North to South along the

50 Table 5. Scaling of models across resolutions and extents along the biogeographic gradient. T1 quantifies cross-scaling of models on the basis of internal and external validation (AUC) on testing samples. Values range from -1 to 1, the highest values corresponding to the best performing scaling exercises. T2a measures down-scaling and T2b up-scaling, both on the basis of the distance between spatial logistic outputs. Values range from 0 to 1, higher values corresponding to poor performing scaling exercises.

55

| | Resolution | | | Extent | |
|---------|------------|------|------|--------|------|
| | T1 | T2a | T2b | T2a | T2b |
| Galicia | 0.71 | 0.13 | 0.25 | 0.20 | 0.03 |
| Madrid | 0.40 | 0.19 | 0.55 | 0.20 | 0.04 |
| Granada | 0.41 | 0.20 | 1.00 | 0.23 | 0.23 |

60

61 biogeographic gradient, being down-scaling less dependent on the biogeographic and ecological context of the study area than up-scaling. This result could be explained by the quality of the training and testing data and also by variations in environmental responses and niche truncation. Incomplete and/or geographically biased occurrence data samples may fail to encompass the full range of environmental conditions present within a region. When such a data are used to train SDMs, outputs may not be transferable to other scales (Guisan and Zimmermann 2000, Pearson and Dawson 2003, Thuiller et al. 2004, Randin et al. 2006, Menke et al. 2009). In our case, we found particularly bad results in Granada. Here, we cope with an interesting situation, consisting of an ecologically marginal population living at the edge of the species geographic range, in a highly diverse and patchy landscape, where most of the available habitats are non-suitable. Under this scenario, modelling and sampling are especially demanding. The patchy and low population density of Iberian mole in Granada can difficult an efficient field sampling, because individuals become easily undetected under the common sampling effort in core populations. Only at a higher effort, this bias could be overcome or, at least, be mitigated. On the other side, data collected from Atlas can record highly localized populations because volunteers can sample the space in an efficient, although non-systematic, way that allow for recording most of available habitats in a region. Complementarily, large-scale data at 10 km summarize in only one record all the environmental information of this plot, allowing for a higher success when modelling presence at more fine scales. Regarding the effect of environmental responses and niche truncation, the fact that Galicia (the core ecological area) was the region with the highest success in scaling of models could be interpreted considering that SDM for species living within wide environmental limits have to be very general. As a consequence, they can be applied more widely than models representing sharply defined niches of species occurring within narrow environmental limits (Evangelista et al. 2008). All these facts highlight that model interpretation and application is complicated in ecologically marginal situations, where only under exhaustive sampling is possible to obtain consistent and symmetric modelling outputs across scales, a situation probably exacerbated in very specialized species, as it the case of the Iberian mole.

105 We conclude that scaling of models across extents and resolutions may be hampered by the truncation of species fundamental niches, which is a major source of uncertainty. Since many species probably have truncated niches, SDM would tend to incorrectly scale model predictions in scenarios that exceed the evaluated environmental limits (see also Thuiller et al. 2004, Wiens et al. 2009, Feeley and Silman 2010, Veloz et al. 2012). We advertise to SDM practitioners about the robustness of downscaling exercises across extents, as well as about the need of carefully evaluating the geographic and ecological modelling context of the study area before performing upscaling procedures. Indeed, if we are about to scale SDM predictions, it is crucial to ran 'source' models at such an extent that warranties a sample size big enough to account for the whole environmental complexity of the species range (Wisz et al. 2008, Braunisch and Suchant 2010). Assessing how species' niches extend into portions of

[AQ4]
121

- 0 environmental space not observed in study areas is currently
a key challenge in modelling approaches.
- 5 *Acknowledgements* – This study was supported by research projects
CGL2005-07681 and CGL2009-13013 from Spanish Ministry of
Education and Science. Thanks to S. Phillips, B. Jimenez-Alfaro
and J. M. Alvarez for helping in different questions related to the
elaboration of the manuscript; also to A. Davies for her help with
language. Many thanks to N. Zimmermann that make a great effort
to improve the quality of this manuscript.
- 10
- ## References
- 15 Araújo, M. and Guisan, A. 2006. Five (or so) challenges for species
distribution modelling. – *J. Biogeogr.* 33: 1677–1688.
- Araújo, M. B. et al. 2005. Downscaling European species atlas
distributions to a finer resolution: implications for conserva-
tion planning. – *Global Ecol. Biogeogr.* 14: 17–30.
- Austin, M. P. 2002. Spatial prediction of species distribution: an
interface between ecological theory and statistical modelling.
– *Ecol. Model.* 157: 101–118.
- 20 Austin, M. P. 2007. Species distribution models and ecological
theory: a critical assessment and some possible new approaches.
– *Ecol. Model.* 200: 1–19.
- Barbosa, A. M. et al. 2009. Transferability of environmental favour-
ability models in geographic space: the case of the Iberian des-
man (*Galemys pyrenaicus*) in Portugal and Spain. – *Ecol. Model.*
220: 747–754.
- 25 Basille, M. et al. 2008. Assessing habitat selection using multivari-
ate statistics: some refinements of the ecological-niche factors
analysis. – *Ecol. Model.* 211: 233–240.
- Beck, J. B. and Shultz, E. K. 1986. The use of relative operating
characteristic (ROC) curves in test performance evaluation.
– *Arch. Pathol. Lab. Med.* 110: 13–20.
- Boyce, M. S. 2006. Scale for resource selection functions. – *Divers.*
Distrib. 12: 269–276.
- Boyce, M. S. et al. 2002. Evaluating resource selection functions.
– *Ecol. Model.* 157: 281–300.
- 35 Braunisch, V. and Suchant, R. 2010. Predicting species distribu-
tions based on incomplete survey data: the trade-off between
precision and scale. – *Ecography* 33: 826–840.
- Braunisch, V. et al. 2008. Living on the edge-modelling habitat
suitability for species at the edge of their fundamental niche.
– *Ecol. Model.* 214: 153–167.
- 40 Brotons, L. et al. 2004. Presence-absence versus presence-only
modelling methods for predicting bird habitat suitability.
– *Ecography* 27: 437–448.
- Brown, J. H. 1984. On the relationship between distribution and
abundance of species. – *Am. Nat.* 124: 255–279.
- 45 Brown, J. H. et al. 1995. Spatial variation in abundance. – *Ecology*
76: 2028–2043.
- Brown, J. H. et al. 2000. Assembly rules: desert rodent communi-
ties are structured at scales from local to continental.
– *Am. Nat.* 156: 314–321.
- 50 Bugmann, H. et al. 2000. Scaling issues in forest succession mod-
elling. – *Clim. Change* 44: 265–289.
- CAM 1998. Plan regional de estrategia territorial de la comunidad
de Madrid. – Dirección General de Urbanismo y Planificación
Territorial.
- Case, T. J. et al. 2005. The community context of species' borders:
ecological and evolutionary perspectives. – *Oikos* 108: 28–46.
- 55 Chase, J. M. and Liebold, M. A. 2003. Ecological niches: linking
classical and contemporary approaches. – Univ. of Chicago
Press.
- Chefaoui, R. M. and Lobo, J. M. 2008. Assessing the effects of
pseudo-absences on predictive distribution model perform-
ance. – *Ecol. Model.* 210: 478–486.
- 60
- Costa, G. C. et al. 2008. Detecting the influence of climatic
variables on species distributions: a test using GIS niche-based
models along a steep longitudinal environmental gradient.
– *J. Biogeogr.* 35: 637–646.
- Edwards, G. R. et al. 1999. Factors influencing molehill distribu-
tion in grassland: implications for controlling the damage
caused by molehills. – *J. Appl. Ecol.* 36: 434–442.
- Edwards, T. C. Jr et al. 2006. Effects of sample survey design on
the accuracy of classification models in ecology. – *Ecol. Model.*
199: 132–141.
- Elith, J. and Leathwick, J. R. 2009. Species distribution models:
ecological explanation and prediction across space and time.
– *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697.
- Elith, J. et al. 2006. Novel methods improve prediction of
species' distributions from occurrence data. – *Ecography* 29:
129–151.
- Elith, J. et al. 2010. A statistical explanation of MaxEnt for ecol-
ogists. – *Divers. Distrib.* 17: 43–57.
- Emberger, L. 1955. Afrique du Nord-Ouest. – In: UNESCO (ed.),
Plant ecology: reviews of research. UNESCO, pp. 219–249.
- Evangelista, P. H. et al. 2008. Modelling invasion for a habitat
generalist and a specialist plant species. – *Divers. Distrib.* 14:
808–817.
- 80 Feeley, K. J. and Silman, M. R. 2010. Biotic attrition from tropical
forests correcting for truncated temperature niches. – *Global
Change Biol.* 16: 1830–1836.
- Fernandes, M. and Herrero, J. 2008. *Talpa occidentalis*. – IUCN
Red List of Threatened Species, ver. 2010.4, <www.iucnredlist.
org>.
- 85 Fielding, A. H. and Bell, J. F. 1997. A review of methods for the
assessment of prediction errors in conservation presence/
absence models. – *Environ. Conserv.* 24: 38–49.
- Franklin, J. and Miller, J. A. 2009. Mapping species distributions.
Spatial inference and prediction. – Cambridge Univ. Press.
- 90 Funmilayo, O. 1977. Distribution and abundance of moles
(*Talpa europaea* L.) in relation to physical habitat and food
supply. – *Oecologia* 30: 277–283.
- Funmilayo, O. 1979. Food consumption, preferences, and storage
in the mole. – *Acta Theriol.* 25: 379–389.
- 95 García-López de Hierro, L. et al. 2013. Positive and negative
unintended human induced effects on Iberian mole abundance
at the edge of its distribution area. – *Mamm. Biol.* 78:
276–282.
- Glen, A. S. and Dickman, C. R. 2005. Complex interactions
among mammalian carnivores in Australia, and their implica-
tions for wildlife management. – *Biol. Rev.* 80: 387–401.
- Gottschalk, T. K. et al. 2011. Influence of grain size on species-
habitat models. – *Ecol. Model.* 222: 3403–3412.
- Goward, S. N. et al. 1994. Ecological remote sensing at OTTER:
satellite macroscale observations. – *Ecol. Appl.* 4: 322–343.
- 100 Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat dis-
tribution models in ecology. – *Ecol. Model.* 135: 147–186.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution:
offering more than simple habitat models? – *Ecol. Lett.* 8:
993–1009.
- Hernández, P. A. et al. 2006. The effect of sample size and species
characteristics on performance of different species distribution
modeling methods. – *Ecography* 29: 773–785.
- Heymann, Y. et al. 1994. CORINE Land Cover. Technical guide.
– EUR12585 Luxembourg, Office for Official Publications of
the European Communities.
- 105 Hijmans, R. J. et al. 2005. Very high resolution interpolated cli-
mate surfaces for global land areas. – *Int. J. Climatol.* 25:
1965–1978.
- Hirzel, A. H. and Le Lay, G. 2008. Habitat suitability modelling
and niche theory. – *J. Appl. Ecol.* 45: 1372–1381.
- Hirzel, A. H. et al. 2002. Ecological-niche factor analysis: how
to compute habitat-suitability maps without absence data?
– *Ecology* 83: 2027–36.
- 121

- 0 Hirzel, A. H. et al. 2006. Evaluating the ability of habitat suitability models to predict species presences. – *Ecol. Model.* 199: 142–152.
- Hirzel, A. H. et al. 2007. Biomapper 4.0. – Laboratory for Conservation Biology, Dept of Ecology and Evolution, Univ. of Lausanne, Switzerland, <www2.unil.ch/biomapper>.
- 5 Hoffman, A. A. and Blows, M. W. 1994. Species borders: ecological and evolutionary perspectives. – *Trends Ecol. Evol.* 9: 223–227.
- Holt, R. D. and Keitt, T. H. 2000. Alternative causes for range limits: a metapopulation perspective. – *Ecol. Lett.* 3: 41–47.
- 10 Jiménez, R. et al. 1990. The reproductive cycle of *Talpa occidentalis* in the southeastern Iberian Peninsula. – *Acta Theriol.* 35: 165–169.
- Kearney, M. 2006. Habitat, environment and niche: what are we modelling? – *Oikos* 115: 186–191.
- 15 Keil, P. et al. 2013. Downscaling of species distribution models: a hierarchical approach. – *Methods Ecol. Evol.* 4: 82–94.
- King, A. W. 1991. Translating models across scales in the landscape. – *Ecol. Stud.* 82: 479–517.
- Kissling, W. D. et al. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. – *J. Biogeogr.* 39: 2163–2178.
- 20 Kryštufek, B. 1999. *Talpa europaea*. – In: Mitchell-Jones, A. J. et al. (eds), *The atlas of European mammals*. Academic Press, pp. 82–83.
- Lara-Romero, C. et al. 2012. Habitat selection by European badgers in Mediterranean semi-arid ecosystems. – *J. Arid Environ.* 76: 43–48.
- 25 Legendre, P. and Legendre, L. 2012. *Numerical ecology*. – Elsevier.
- Lischke, H. et al. 2007. Model up-scaling in landscape research. – In: Kienast, F. et al. (eds), *A changing World*. Springer, pp. 249–272.
- 30 Loy, A. 1999. *Talpa occidentalis*. – In: Mitchell-Jones, A. J. et al. (eds), *The atlas of European mammals*. Academic Press, pp. 80–81.
- Luoto, M. et al. 2007. The role of land cover in bioclimatic models depends on spatial resolution. – *Global Ecol. Biogeogr.* 16: 34–42.
- 35 Macdonald, D. W. et al. 1997. Spatial and temporal patterns in the activity of European moles. – *Oecologia* 109: 88–97.
- Marcer, A. et al. 2012. Modelling invasive alien species distributions from digital biodiversity atlases. Model upscaling as a means of reconciling data at different scales. – *Divers. Distrib.* 18: 1177–1189.
- 40 McPherson, J. M. and Jetz, W. 2007. Effects of species' ecology on the accuracy of distribution models. – *Ecography* 30: 135–151.
- McPherson, J. M. et al. 2006. Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. – *Ecol. Model.* 192: 499–522.
- 45 Mead-Briggs, A. R. and Woods, J. A. 1973. An index of activity to assess the reduction in mole numbers caused by control measures. – *J. Appl. Ecol.* 10: 837–845.
- Meier, E. S. et al. 2011. Co-occurrence patterns of trees along macroclimatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. – *J. Biogeogr.* 38: 371–382.
- Melbourne, B. A. and Chesson, P. 2005. Scaling up population dynamics: integrating theory and data. – *Oecologia* 145: 178–186.
- 55 Menge, B. A. and Sutherland, J. P. 1987. Community regulation: variation in disturbance, competition, and predation in relation to environmental stress and recruitment. – *Am. Nat.* 130: 30–757.
- Menke, S. B. et al. 2009. Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. – *Global Ecol. Biogeogr.* 18: 50–63.
- Morin, X. and Lechowicz, M. J. 2008. Contemporary perspectives on the niche that can improve models of species range shifts under climate change. – *Biol. Lett.* 4: 573–576.
- 65 Mueller, P. and Diamond, J. 2001. Metabolic rate and environmental productivity: well-provisioned animals evolved to run and idle fast. – *Proc. Natl Acad. Sci. USA* 98: 12550–12554.
- Niethammer, J. 1990. *Talpa occidentalis* – Spanischer Maulwurf. – In: Niethammer, J. and Krapp, F. (eds), *Handbuch der Säugetiere Europas, Band 3/I, Insectivora, Primates*, Aula Verlag, pp. 157–161.
- 70 Normand, S. et al. 2009. Importance of abiotic stress as a range limit determinant for European plants: insights from species responses to climatic gradients. – *Global Ecol. Biogeogr.* 18: 437–449.
- Osborne, P. E. and Suárez-Seoane, S. 2002. Should data be partitioned spatially before building large scale distribution models? – *Ecol. Model.* 157: 249–259.
- 75 Osborne, P. E. et al. 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. – *J. Appl. Ecol.* 38: 458–471.
- 80 Osborne, P. E. et al. 2007. Non-stationarity and local approaches to modelling the distributions of wildlife. – *Divers. Distrib.* 13: 313–323.
- Paruelo, J. M. et al. 1997. ANPP estimates from NDVI for the Central Grassland Region of the United States. – *Ecology* 78: 953–958.
- 85 Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. – *Biodivers. Inform.* 3: 59–72.
- 90 Pettorelli, N. et al. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. – *Trends Ecol. Evol.* 20: 503–510.
- Pettorelli, N. et al. 2006. Using a proxy of plant productivity (NDVI) to find key periods for animal performance: the case of roe deer. – *Oikos* 112: 565–572.
- 95 Pettorelli, N. et al. 2011. The normalized difference vegetation index (NDVI): unforeseen successes in animal ecology. – *Clim. Res.* 46: 15–27.
- Phillips, S. J. and Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. – *Ecography* 31: 161–175.
- Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – *Ecol. Model.* 190: 231–259.
- 100 Pulliam, H. R. 2000. On the relationship between niche and distribution. – *Ecol. Lett.* 3: 349–361.
- 105 Randin, C. F. et al. 2006. Are niche-based species distribution models transferable in space? – *J. Biogeogr.* 33: 1689–1703.
- Rezende, E. L. et al. 2004. Climatic adaptation and the evolution of basal and maximum rates of metabolism in rodents. – *Evolution* 58: 1361–1374.
- 110 Rivas-Martínez, S. et al. 2004. Bioclimatic map of Europe (thermoclimatic belts). – Univ. de León, Secretariado de Publicaciones.
- Román, J. 2007a. *Talpa occidentalis*. – In: Palomo, L. J. et al. (eds), *Atlas y libro rojo de los mamíferos terrestres de España*. SECEM, pp. 89–91.
- 115 Román, J. 2007b. *Talpa europaea*. – In: Palomo, L. J. et al. (eds), *Atlas y libro rojo de los mamíferos terrestres de España*. SECEM, pp. 86–88.
- Root, T. 1988. Environmental factors associated with avian distributional boundaries. – *Journal of Biogeography* 15: 489–505.
- 121

| | | |
|----|--|-----|
| 0 | Rouse, J. W. et al. 1973. Monitoring vegetation systems in the great plains with ERTS. – Proc. of the 3rd ERTS Symp. 1, NASA SP-351: 309–317. | 61 |
| | Rushton, S. P. et al. 2004. New paradigms for modelling species distributions? – J. Appl. Ecol. 41: 193–200. | |
| 5 | Scott, J. M. et al. 2002. Predicting species occurrences: issues of accuracy and scale. – Island Press. | 65 |
| | Seo, C. et al. 2009. Scale effects in species distribution models: implications for conservation planning under climate change. – Biol. Lett. 5: 39–43. | |
| 10 | SINAMBA 1999. Sistema de Información Ambiental. – Junta de Andalucía. | 70 |
| | SITGA 1998. Sistema de Información Territorial de Galicia. – Xunta de Galicia. | |
| | Soberon, J. 2007. Grinnellian and eltonian niches and geographic distributions of species. – Ecol. Lett. 10: 1115–1123. | |
| 15 | Soberón, J. M. 2010. Niche and area of distribution modeling: a population ecology perspective. – Ecography 33: 159–167. | 75 |
| | Suárez-Seoane, S. and Baudry, J. 2002. Scale dependence of spatial patterns and cartography on the detection of landscape change: relationships with species' perception. – Ecography 25: 499–511. | |
| 20 | Suárez-Seoane, S. et al. 2002. Large-scale habitat selection by agricultural steppe birds in Spain: identifying species-habitat responses using generalised additive models. – J. Appl. Ecol. 39: 755–771. | 80 |
| | Suárez-Seoane, S. et al. 2004. Can climate data from METEOSAT improve wildlife distribution models? – Ecography 27: 629–636. | |
| 25 | Suárez-Seoane, S. et al. 2008. Maximum entropy niche-based modelling of seasonal changes in little bustard distribution. – Ecol. Model. 219: 17–29. | 85 |
| | Tabachnick, B. and Fidell, L. 1996. Using multivariate statistics, 3rd ed. – Collins. | |
| 30 | Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – Ecography 27: 165–172. | 90 |
| | Thuiller, W. et al. 2005. Niche-based modeling as a tool for predicting the risk of alien plant invasions at a global scale. – Global Change Biol. 11: 2234–2250. | |
| 35 | Trivedi, M. R. et al. 2008. Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. – Global Change Biol. 14: 1089–1103. | 95 |
| | Tuanmu, M. N. et al. 2011. Temporal transferability of wildlife habitat models: implications for habitat monitoring. – J. Biogeogr. 38: 1510–1523. | 100 |
| 40 | | |
| | Supplementary material (Appendix ECOG-00077 at <www.oikosoffice.lu.se/appendix>). Appendix 1. | 105 |
| 45 | | |
| | | 110 |
| 50 | | |
| | | 115 |
| 55 | | |
| | | 121 |
| 60 | | |