

Cross-scale predictions allow the identification of local conservation priorities from atlas data

P. Bombi¹, D. Salvi² & M. A. Bologna³

¹ SPACEnvironment, Rome, Italy

² CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Rua Padre Armando Quintas, Vairão, Portugal

³ Dipartimento di Biologia Ambientale, University of Roma Tre, Rome, Italy

Keywords

atlas maps; distribution models; downscaling; gap analysis; irreplaceability.

Correspondence

Pierluigi Bombi, SPACEnvironment, Via Maria Giudice 23, Rome 00135, Italy.
Tel: +393337012803
Email: pierluigi.bombi@gmail.com

Editor: Nathalie Pettorelli

Received 28 November 2011; accepted 18 January 2012

doi:10.1111/j.1469-1795.2012.00526.x

Abstract

For planning practical measures aimed at biodiversity protection, conservation priorities must be identified at a local scale. Unfortunately, identifying local conservation priorities requires high-resolution data on species distribution, and these are often unavailable. Atlases of species distribution provide data for several groups of organisms in many different areas but are often too coarse in resolution to provide valuable information. We explored the possibility of cross-scale modelling species distributions and we clarified, for the first time, its effect on prioritization exercises. We used different modelling techniques for scaling down atlas data for Sardinian reptiles, validated the outcomes with detailed, field-sampled data, and compared conservation priorities deriving from atlas maps and down-scaled models. Doing this, we obtained as a further result the identification of priority species and areas for future conservation strategies. Our results encourage us to experiment further with this approach. Through the downscaling procedure, we obtain high-resolution models with strong variations in predictive performances, although most of the models show satisfactory/excellent scores. This testifies that low-resolution data can be downscaled maintaining low rates of omission and commission errors. Increasing the resolution of distribution maps used for prioritization influences the spatial patterns of priority but does not modify the evaluation of species representation. Overall, we show that atlases can meet the large demand for distribution data by decision makers if appropriate downscaling procedures are adopted. In addition, we provide practical instruments for the conservation of reptiles in Sardinia by identifying priority species and areas that require strict management.

Introduction

Atlases of species distribution are common platforms to present locational data for a multitude of different organisms. Distribution atlases provide an excellent means for clarifying several aspects of species and community ecology (e.g. Turner, Lennon & Lawrenson, 1988; Dennis, 2003; Araújo, Densham & Williams, 2004) and represent a valuable source for data-thirsty activities, such as those involved in biodiversity conservation (Loiselle *et al.*, 2003; Carroll, 2010; Costa *et al.*, 2010). Unfortunately, decision makers often require species distribution information at a spatial resolution much higher than atlas data generally provide. Thus, setting up effective approaches for translating distribution information from one scale to another and making the most of coarse atlas data is a great challenge for conservation planning.

In cross-scale analyses of spatial data, species distribution models (SDMs) can be fundamental instruments. These models are generally based on the quantification of species–environment interactions (Guisan & Zimmermann, 2000) and are widely used for planning conservation measures (e.g. Guisan & Thuiller, 2005). Several studies explored the SDMs' capability of scaling down coarse atlas data into high-resolution predictions (e.g. Pearson, Dawson & Liu, 2004; Niamir *et al.*, 2011). Some of these papers experienced difficulties in producing realistic models (e.g. McPherson, Jetz & Rogers, 2006), but others obtained encouraging results (e.g. Pearson *et al.*, 2004). Furthermore, Araújo *et al.* (2005) evidenced that downscaled distributions can detect important environmental gradients unidentified by coarse resolution data. Recently, Bombi & D'Amen (2011) compared different approaches to cross-scale modelling, provided practical suggestions for optimizing this process and

analysed the theoretical framework of downscaling practices. To the best of our knowledge, no other study investigated the effect of downscaling on the identification of conservation priorities.

When spatial data are aggregated into larger areal units, data values may diverge from one scale to another, generating misleading outcomes (Gehlke & Biehl, 1934). This scale problem is a component of the modifiable areal unit problem (MAUP; Openshaw, 1984). MAUP may influence the production of high-resolution predictions from low-resolution data in two different ways: (1) it affects the aggregation of field records into coarse atlas' units on the one hand; (2) the calculation of predictors' values on the other. These two effects influence to varying degrees the various approaches to downscaling, which are based on different assumptions and have individual responses to MAUP (Bombi & D'Amen, 2011). Therefore, taking into account the uncertainty because of MAUP is essential in any downscaling exercise.

Sardinia is the second largest island in the Mediterranean basin, which is a global biodiversity hotspot (Mittermeier *et al.*, 2004), and the island is particularly important for biological diversity in the region (Blondel *et al.*, 2010). In particular, Sardinian reptiles have a high intrinsic conservation value, six of them being included in European Community and the International Union for Conservation of Nature (IUCN) conservation lists (*Euleptes europaea*, *Emys orbicularis*, *Testudo hermanni*, *T. graeca* and *T. marginata*: Annexes II–IV of the Habitats Directive 92/43/EEC; *Natrix natrix cetti*, Critically Endangered, IUCN, 2010). They encompass a broad spectrum of spatial distribution patterns (Salvi & Bombi, 2010), from very widespread to localized species, and most of them deserves the attention of conservationists.

The main aim of this paper was to investigate the potential contribution of downscaling procedures for prioritization exercises. More specifically, we focus on the differences of results deriving from atlas data and downscaled models in terms of: (1) spatial patterns of species richness and land priority for conservation; (2) relative importance of existing protected areas; (3) species representation in the current reserve network. For doing this, we utilize atlas data for Sardinian reptiles, validate downscaled models with an independent and high-resolution dataset, and perform gap and irreplaceability analyses. We are also interested in providing practical suggestions for optimizing future management activities for the conservation of reptiles in Sardinia. In this light, our study has both a theoretical interest in the outcomes comparison, and a practical importance in the identification of conservation priorities.

Methods

Species and environmental data

We based our analyses on atlas data from Salvi & Bombi (2010) for reptiles of Sardinia. These data indicate the occurrence of each species in the 312 cells of the grid (UTM, 10 × 10 km) that intersect the Sardinian territory. This dataset represents the most recent and fulfilling synthesis of distribution information for Sardinian reptiles [see Salvi & Bombi, (2010) for details]. We considered all the 17 species of reptiles occurring in Sardinia (Table 1). On this dataset, we fitted SDMs utilizing four groups of environmental predictors: (1) 19 climatic variables; (2) six land cover variables; (3) four land morphology descriptors; (4) one measure of human impact (see Supporting Information Table S1 for data sources and description). All the variables

Table 1 Validation scores, priority scores (see text for details), range extents (in number of cells), conservation targets (in number of cells) and percentages of target met for the two approaches

Species	AUC	Priority	Atlas data			Downscaled models		
			Range extent	Target	% target met	Range extent	Target	% target met
<i>Algyroides fitzingeri</i>	0.783	5	81	15	–	13 410	2478	–
<i>Archaeolacerta bedriagae</i>	0.930	5	25	9	–	5 164	1811	77.187
<i>Chalcides chalcides</i>	0.668	3	37	6	–	16 666	2630	–
<i>Chalcides ocellatus</i>	0.537	3	87	14	–	15 856	2595	–
<i>Emys orbicularis</i>	0.832	1	13	4	26.570	6 985	2022	69.724
<i>Euleptes europaea</i>	0.785	5	35	14	51.307	4 331	1688	78.779
<i>Hemorrhois hippocrepis</i>	0.983	2	17	10	77.541	727	441	28.331
<i>Hemidactylus turcicus</i>	0.834	3	44	6	–	21 961	2823	–
<i>Hierophis viridiflavus</i>	0.538	3	110	14	–	23 181	2860	–
<i>Natrix maura</i>	0.593	3	47	8	–	14 511	2533	–
<i>Natrix natrix</i>	0.704	5	23	6	–	8 821	2185	–
<i>Podarcis sicula</i>	0.516	2	108	14	–	20 723	2782	–
<i>Podarcis tiliguerta</i>	0.737	4	124	22	–	14 394	2528	–
<i>Tarentola mauritanica</i>	0.840	3	34	8	73.120	9 166	2212	88.104
<i>Testudo graeca</i>	0.833	1	14	8	13.059	2 250	1231	32.015
<i>Testudo hermanni</i>	0.934	1	12	5	0.000	3 572	1554	55.418
<i>Testudo marginata</i>	0.979	1	12	1	0.000	393	11	–

Total gap species are evidenced in bold and partial gap species that met < 50% of their target in italic. Cell size corresponds to 100 km² for atlas data and approximately 1 km² for downscaled models. In bold are evidenced those species not represented at all in the protected areas.

were obtained at the resolutions of 30 arc-s (approximately 1 km). In order to eliminate the effect of correlation between variables and to ensure that our set of predictors is not saturated (Blanchet, Legendre & Borcard, 2008), we calculated the variance inflation factor (VIF; Belsley, Kuh & Welsch, 1980) on the occurrence data in R (version 2.8.1, R Development Core Team, 2010). We removed variables until all VIFs were below 10 (Belsley *et al.*, 1980; Marx & Smith, 1990). Thus, we trained the models utilizing a subset of 13 predictors (Supporting Information Table S1).

Downscaling strategy

In order to produce high-resolution models (30 arc-s) from atlas data, for each species, we used as presence data 100 random points throughout the ensemble of the occupied squares. Similarly, we created absence data by generating 200 random points out of the species' known distribution. Thus, for each species we utilized 300 records (prevalence = 0.33) for sampling the environmental conditions at high resolution and for fitting the models. We imposed this low prevalence because it increases model specificity (e.g. Allouche, Tsoar & Kadmon, 2006) and counterbalances the inflation of commission errors due to downscaling (Bombi & D'Amen, 2011).

Modelling procedure

Because the choice of algorithm influences the performance of downscaling process (Bombi & D'Amen, 2011), we adopted an ensemble forecasting approach, which makes it possible to take into account the inter-models variation (e.g. Thuiller, 2004; Araújo & New, 2007; Marmion *et al.*, 2008). In the R-based package BIOMOD (Thuiller, 2003), we fitted on species presence/absence data at a spatial resolution of 30 arc-s: artificial neural networks (ANN; Ripley, 1996), classification tree analysis (CTA; Breiman *et al.*, 1984), generalized boosting model (GBM; Friedman, 2001), generalized linear models (GLM, McCullagh & Nelder, 1989), multiple adaptive regression splines (MARS; Friedman, 1991), flexible discriminant analysis (FDA; Hastie, Tibshirani & Buja, 1994) and random forest for classification and regression (RF; Breiman, 2001). We preliminary 10-fold, cross-validated each model by calculating the area under curve (AUC) (Fielding & Bell, 1997; Thuiller, Lavorel & Araújo, 2005). For each species, we used the models with AUC > 0.7 to produce a single consensus model by calculating the weighted average of the single models (see equation 1 in Marmion *et al.*, 2008). To this end, we utilized as model weight the scores from cross-validation, which are proportional to those from an independent test dataset ($r = 0.665$; $P < 0.005$).

We used an independent dataset to validate the consensus models by calculating the AUC values in the R package PresenceAbsence (Freeman, 2007). This independent dataset consists of 211 detailed locality records derived from field samples and spread throughout the study area (Speybroeck *et al.*, unpublished data). Most of these locali-

ties were intensively sampled for several years; thus, they can be considered valid presence/absence data for each species. In addition, we converted the continuous models of habitat suitability (HS) to binary maps, according to the minimal predicted area criterion (Engler, Guisan & Rechsteiner, 2004), by identifying the HS threshold to achieve a model sensitivity of 0.9. We chose this threshold by producing a set of thresholds from different criteria (i.e. maximum percentage of presence and absence correctly predicted, maximum kappa, maximum true skill statistic) and selecting the criterion that produced the highest threshold, in order to counterbalance the commission error of downscaled models (Bombi & D'Amen, 2011).

Measures of conservation priority

In order to compare conservation priorities at different spatial resolutions, we used atlas maps and downscaled models for deriving: (1) species richness, conservation value and complementarity of each cell; (2) the relative conservation importance of protected areas; (3) the list of underrepresented species. Species richness was derived by counting the number of species estimated to occur in each cell. Conservation value was estimated by calculating the irreplaceability of each cell, considering every cell as unprotected. Complementarity was measured as the irreplaceability value of each unprotected cell, taking into account the existence of reserves. We evaluated the relative conservation importance of each protected area by calculating the reserves' irreplaceability. Finally, we performed a gap analysis for quantifying the extent to which each species is represented in the reserve network. We obtained two independent results for each measure at the two resolutions, thus allowing a comparison of outcomes.

Boundaries of nationally designed protected areas and sites included in the European Natura 2000 network are available from the Italian Ministry for the Environment (<ftp://ftp.scn.minambiente.it/Cartografie>). In order to match reserve boundaries with species distribution maps, we defined 35% of protected surface per cell as cut-off value for determining whether reserves are present or absent in the cell. This threshold is critical because it could produce an under- or overestimation of the real protected surface (Araújo, 2004). For solving this issue, we selected the threshold that selected a number of cells with a total surface equal to the total surface of Sardinian protected areas (approximately 5000 km²) (D'Amen *et al.*, 2011).

We assumed that each species should be represented proportionally to its range size. Accordingly, we set the minimum and maximum conservation targets to 50% of the number of occupied cells that are currently protected for the species with the smallest and largest distribution in Sardinia. Targets for the other species were calculated by interpolating the extreme targets using a linear regression on the log-transformed number of occupied cells (Rodrigues *et al.*, 2004; Maiorano, Falcucci & Boitani, 2006). We measured the specific extent of occurrence as number of occupied cells in both atlas maps and downscaled models independently.

To define the landscape of conservation importance, we estimated the summed irreplaceability of each cell using C-Plan (Version 4; Pressey *et al.*, 2009). Irreplaceability is the degree to which one site is required in a reserve network to achieve established conservation targets (see Ferrier *et al.*, 2000 for further details). We calculated the summed irreplaceability in each cell by adding the irreplaceability values for all the species in the cell (Pressey *et al.*, 1994). We assigned to each species an arbitrary priority score (Table 1) according to the relative endemism and abundance in the study area. More specifically, we assigned the highest scores to endemic and uncommon taxa, and the lowest to species that are probably introduced in the area; common but native or endemic species received intermediate scores. The value of summed irreplaceability can range from zero to a large number. A high value indicates a crucial site for many species; values close to zero indicate not important sites for most of the species. We predicted the irreplaceability of each cell both considering all cells as unprotected (conservation value) and taking into account the existing reserves (complementarity).

We also measured the summed irreplaceability for each protected area. In this calculation, we estimated how much each species is represented in the reserve network in terms of number of occupied parks. We considered as occupied all the parks that intersect the cells indicated by atlas maps and downscaled models. For each species, we chose a representation target of one occupied park. In the prediction of parks' summed irreplaceability, we weighed each species according to the priority score mentioned earlier (Table 1). Therefore, we classified all the protected areas according to their relative contribution to the representativeness of the entire network.

In the gap analysis, species distributions are compared with the distribution of conservation areas (Jennings, 2000). We determined the percentage of species distribution intersecting the reserve network. For each species, this percentage was compared with the relative conservation target. We considered those species not represented in any protected area as 'total gap' and species that met only a portion of their conservation target as 'partial gap' (Rodrigues *et al.*, 2004). We followed this procedure for atlas maps and downscaled models independently and obtained two lists of gap species.

Results

Downscaling performances

We observed large differences in predictive performances of downscaled models across species. The validation scores of downscaled models are proportional to those estimated for low-resolution models produced directly from atlas data for the same species ($r = 0.650$; $P < 0.01$). We obtained excellent downscaled models ($AUC \geq 0.9$) for four species (Table 1); good models for another four species ($AUC \geq 0.8$); and satisfactory models for a third group of

four species ($AUC \geq 0.7$). By contrast, for a fourth group of five species, we obtained insufficient results ($AUC < 0.7$). The latter group consists of very common and widespread species and their predictions indicated a potential occupancy of almost all the study area. Because both known and modelled distributions agree indicating these species as ubiquitous over the study area, we assumed that their predictions could be included in prioritization exercises, without compromising our outcomes.

Conservation priorities

Geographic patterns of species richness deriving from atlas data and downscaled models were dissimilar (Fig. 1a,d). Comparing cell values of species richness at high resolution with those of the corresponding cells at low resolution, by multiple sampling low-resolution cells, we observed a weak but significant correlations ($r = 0.394$; $P < 0.001$). The most obvious difference was the coarser texture of output maps from atlas data. The 5% most specious cell (N species > 8) calculated by atlas data corresponded to Asinara Island, La Maddalena Archipelago, Mount Limbara, Sassari, Supramonte massif, Cagliari, and scattered areas in Sulcis-Iglesiente and Sarrabus-Gerrei (see Supporting Information Fig. S1a for place names). These areas contain the 10% of the 5% most specious cells indicated by the downscaled models ($N > 11$), which correspond to the coast of Gallura and La Maddalena archipelago, individual mountains in north-eastern Sardinia, the coastal areas south of the Gulf of Orosei, and a large area in south-eastern Sardinia from Ogliastra to Sarrabus-Gerrei.

The geographic pattern of conservation value is similar to that of species richness at both high and low resolutions ($r = 0.738$ and $r = 0.747$, respectively; $P < 0.001$ in any case) (Fig. 1b,e). In addition, the pattern of conservation value is different when derived from atlas data or downscaled models ($r = 0.312$; $P < 0.001$). The 5% most irreplaceable cells (irreplaceability > 2.4), as calculated by atlas data, aggregated around the coast of Gallura and La Maddalena Archipelago as well as Mount Limbara, and contained the 16% of the 5% most irreplaceable cells (irreplaceability > 1.7) from downscaled maps. According to the downscaled maps, the high-value area along the coast of Gallura expanded with respect to the species richness map. At the same time, the mountains between Mount Limbara and Mount Albo become more prominent, and a new important area around Cagliari appeared.

The map of complementarity diverged from those of species richness ($r = 0.249$; $P < 0.001$) and conservation value ($r = 0.314$; $P < 0.001$) when calculated from downscaled models (Fig. 1f). The pattern of complementarity was more similar to those of species richness ($r = 0.415$; $P < 0.001$) and conservation value ($r = 0.437$; $P < 0.001$) when calculated from atlas data (Fig. 1c). The complementarity values of the high-resolution cells were correlated weakly with values of the corresponding cells at low resolution ($r = 0.196$; $P < 0.001$). Atlas data produced a map with the 5% most irreplaceable cells restricted, with few

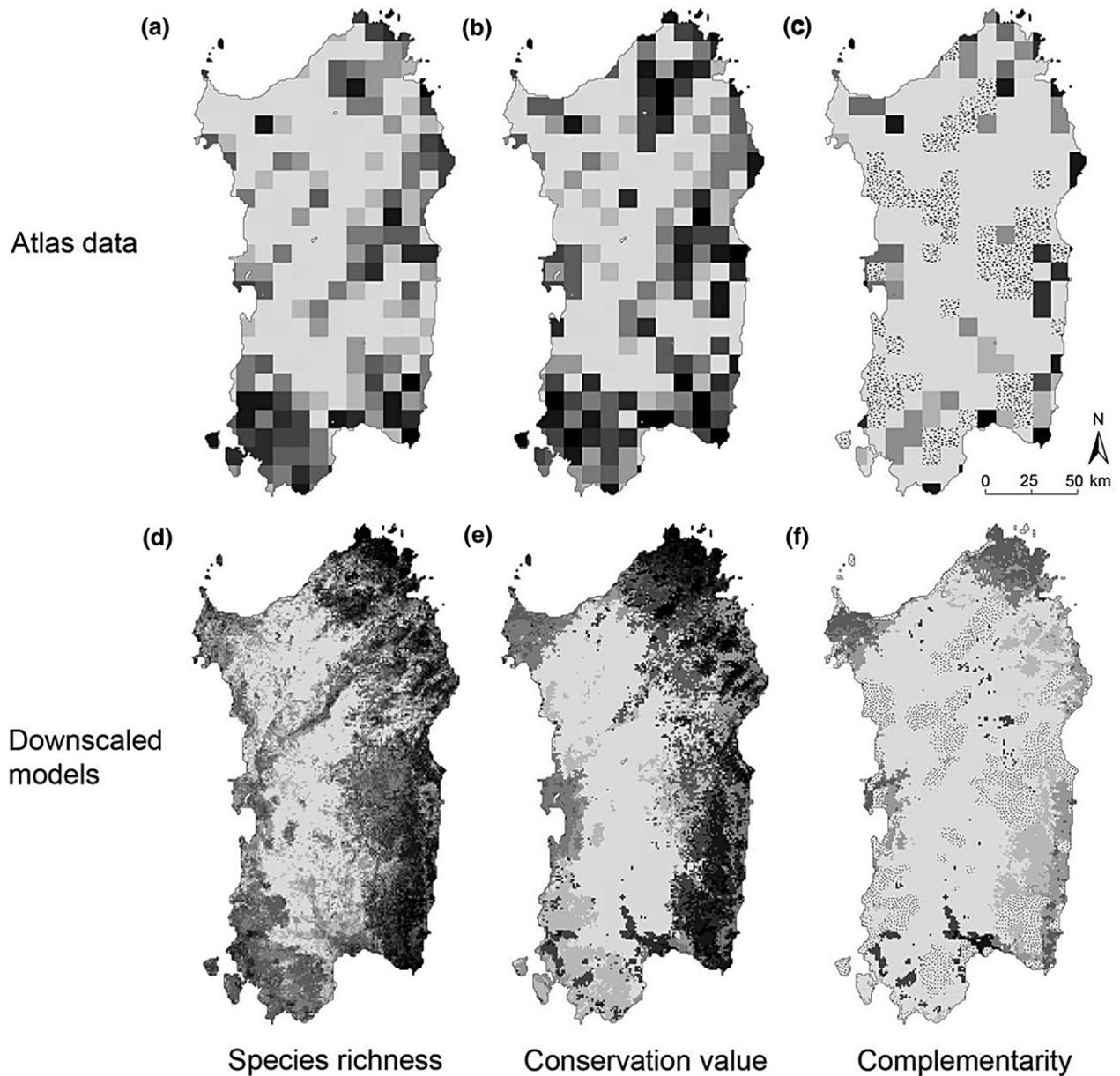


Figure 1 Priority maps for Sardinian reptiles calculated from atlas data (first row) and downscaled models (second row). The darkness of each areal unit is proportional to the relative value of: species richness [first column; pale grey = 1 species, black = 12 and 16 species in (a) and (d), respectively], conservation value (irreplaceability calculated considering none unit as protected) [second column; pale grey = 0, black = 3.6 and 2.4 in (b) and (e) respectively] and complementarity (irreplaceability calculated considering protected units) [second column; pale grey = 0, black = 2.8 and 1.8 in (c) and (f) respectively]. In (c) and (f) dotted areas evidence protected units.

exceptions, along the coasts of the island. These areas contain the 6.9% of the 5% most irreplaceable cells indicated by the downscaled models. The complementarity map produced by downscaled models showed the area of Cagliari and some areas in Sulcis-Iglesiente as the most important for conservation targeting. Further areas at high complementarity were identified in Gallura and Nurra.

Differences between atlas- and model-based results emerged for the relative irreplaceability of protected areas

($r = 0.474$; $P < 0.001$) (Fig. 2). Atlas data produced a map with great heterogeneities between neighbouring reserves (Fig. 2a). Indeed, the 5% most irreplaceable sites are scattered all across the island. Differently, downscaled models generated well-defined clusters of protected areas (Fig. 2b). In this case, the 5% most irreplaceable sites were all located in the North of the island and a second group of important reserves were in Sulcis-Iglesiente. Intermediate levels of irreplaceability were assigned to mountain and coastal

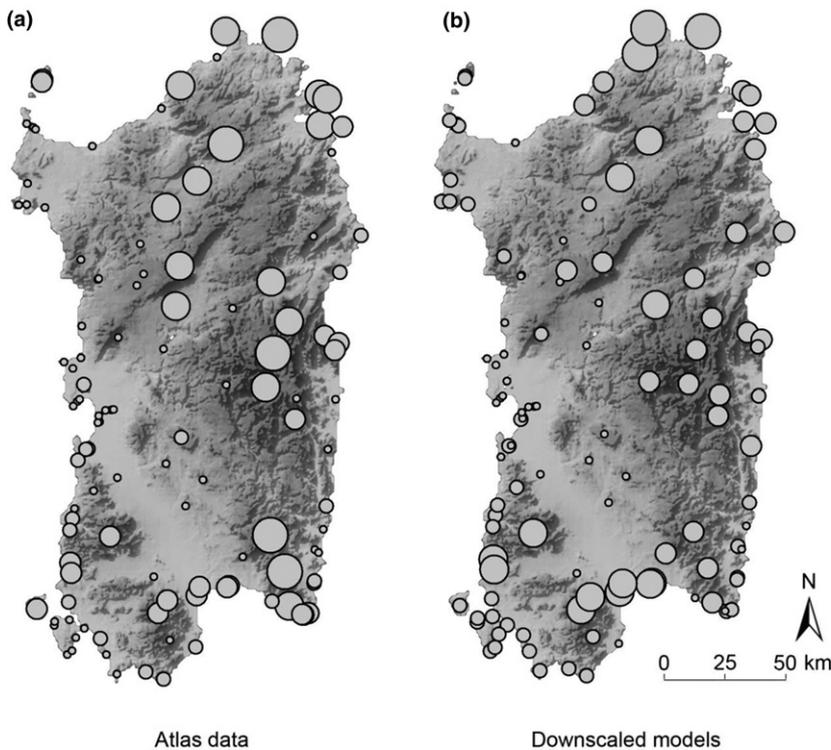


Figure 2 Irreplaceability of Sardinian protected areas. Each circle represents one protected area (nationally designed or Natura 2000) and the circles' dimension is proportional to the relative irreplaceability [min = 0, max = 0.9 and 1.1 in (a) and (b) respectively].

sites in eastern Sardinia, Sulcis-Iglesiente, and Nurra (see Supporting Information Fig. S1b and Table S2 for details).

The gap analysis carried out on atlas data and downscaled models produced similar outcomes (Table 1). The two approaches shared most of the species indicated as underrepresented. The percentages of target met calculated by the two approaches diverged but five species appeared always as partial gap species. *Testudo hermanni* was evidenced as a total gap species by atlas maps and a partial gap species by downscaled models. In addition to these shared species, atlas maps indicated *Testudo marginata* as a total gap species, whereas *Archaeolacerta bedriagae* was a partial gap species according to downscaled models. All of the other species met their respective conservation targets in both the analyses.

Discussion

Downscaling effects

The increase of spatial resolution in species' distribution maps has two main effects on prioritization process. First, the additional details allow directly the definition of specific sites for practical conservation measures. For instance, our downscaled models show individual mountains as hotspots of conservation value in north-eastern Sardinia rather than a unique, large, important area (Fig. 1). Second, describing species distribution at high resolution reduces the uncertainty in the representativeness of protected areas (Araújo,

2004). Indeed, species and reserves generally occupy only portions of the cells where they occur and these portions may not coincide. A reduction in the cell surface increases the probability that the portions occupied by species and reserves overlap.

Downscaled models reduce, at the same time, commission error because of the coarse resolution of atlas maps and omission error because of the gaps of knowledge in atlas data. The low resolution of atlas maps obscures the fact that species are often not present everywhere in the occupied cells. This problem, which is due to the scale component of MAUP, artificially inflates the species' extent of occurrence and introduces a bias when species' distributions are intersected with reserve boundaries. Such uncertainty may cause an overestimation of species' representation in protected areas. On the other hand, atlas data are often an incomplete representation of true species distributions. This underestimation of the real extent of occurrence artificially reduces the observed species' representation (Bombi, Luiselli & D'Amen, 2011). In any case, these two opposite effects introduce a variable amount of uncertainty, which can jeopardize the entire prioritization process.

The incomplete knowledge in atlas data can explain the differences in the spatial patterns of low- and high-resolution conservation priorities. The atlas-based maps evidence the importance of several cells that correspond to large towns (e.g. Sassari and Cagliari) and well-sampled areas (e.g. Sulcis-Iglesiente, Mount Limbara, Supramonte massif). Whereas, areas ignored by sampling activities (e.g.

Ogliastro and Baronie) were identified as priorities only from downscaled models. In addition, atlas maps produce large differences of relative importance between renowned and well-sampled parks compared with neighbouring and neglected reserves (see Supporting Information Fig. S1b and Table S2). Differently, downscaled models define eco-geographic clusters of protected areas with similar irreplaceability.

Downscaled models remove some biases of atlas data but also introduce a certain degree of uncertainty in the prioritization. In particular, Bombi & D'Amen (2011) showed that, although the models remain useful up to a 12-fold grain-size reduction, the predictive performance decrease proportionally to the intensity of downscaling. Moreover, they demonstrated that cross-scale predictions are prone to commission error. Because model errors propagate through the process of overlaying species distributions (Araújo *et al.*, 2005), it is crucial minimizing the uncertainty in species models. In this light, it is a primary exigency reducing as far as possible the overestimation of species ranges and stabilizing the model sensitivity.

Several approaches for downscaling coarse presence data to high-resolution distribution maps have been proposed in the last years (e.g. Pearson *et al.*, 2004; Araújo *et al.*, 2005; McPherson *et al.*, 2006; Mitikka *et al.*, 2007; Trivedi *et al.*, 2008; Anderson *et al.*, 2009). Each approach is based on different assumptions, has specific strengths, and is affected by peculiar weaknesses. Recently, Bombi & D'Amen (2011) compared the efficacy of different approaches and analysed the concept underpinning each of them. Our downscaling approach assumes that suitable environmental conditions are the features that the occupied cells have in common, so that sufficient random sample points could capture such features. This approach, which is based on a fixed number of points, allows controlling data prevalence, increasing model specificity (e.g. McPherson, Jetz & Rogers, 2004; Allouche *et al.*, 2006), and overcoming the reduction of occupied cells because of the aggregation of data in coarse grids. At the same time, the random placement of data points across the occupied atlas cells stabilizes the rate of omission error across spatial resolutions (see Bombi & D'Amen, 2011 for a detailed discussion).

The differences evidenced in spatial patterns do not reflect similar incongruences in the degree of species representation. With few exceptions, both atlas- and model-based gap analyses evidence the same set of species as underrepresented. Indeed, six of the seven gap species identified by the two approaches are the same. This result reinforces our threshold-based criterion for selecting protected cells, evidencing that the uncertainty associated with this process does not change across scales. A certain degree of error is necessarily caused by matching reserve boundaries and species distributions at different resolutions (Araújo, 2004). In particular, small protected areas can be overlooked so that the real level of protection can be under- or overestimated. The cross-scale robustness of our criterion allows comparing the degrees of representation from alternative sources.

Deficiencies in atlas data lead to poor parameterization of SDMs, both at coarse and fine resolutions. Our models have heterogeneous performances at both high and low resolution, as shown by the correlation between AUCs of downscaled and coarse-grain models. The species that produce the worst models share common distribution patterns and ecological strategy. All of them are largely widespread across Sardinia and extremely adaptable to different conditions. These features have negative influences on SDMs (Brotons *et al.*, 2004; McPherson & Jetz, 2007). Such inadequacy of distribution data for species utilizing a wide variety of habitats can explain the failure of some attempts to downscale species distribution from low-resolution data (Collingham *et al.*, 2000; McPherson *et al.*, 2006; Mitikka *et al.*, 2007).

The spatial patterns of conservation priority produced by downscaled models are different from those by atlas data. Nevertheless, further studies are required for quantifying the contribution of cross-scale predictions to the practical conservation of the species. Such studies should compare spatial patterns of priority derived from coarse-grain observations and their downscaled projections with priorities directly obtained from non-downscaled high-resolution models (i.e. from fine-scale distribution data). The gap of variations between the tested patterns (from atlas maps and downscaled models) and the reference pattern (from high-resolution data) would measure the enhancement in priority selection allowed by cross-scale predictions. This kind of assessment would need fine-scale field observations directly obtained at high resolution, which are seldom available in sufficient number in the real world. The high-resolution records should be used for simulating coarse-scale atlas data; then high-resolution records, simulated atlas data, and downscaled models from simulated atlas data should be used for defining alternative priority patterns for comparison.

Conservation priorities

We generated high-resolution maps of priority areas for local conservation initiatives. The most important areas, in terms of both species richness and irreplaceability, are mainly located in eastern Sardinia. More specifically, coastal and sub-coastal areas of Gallura and individual mountains from Mount Limbara to Mount Albo, as well as hilly areas of Ogliastra and Sarrabus-Gerrei (Fig. 1) represent the main pinnacles of conservation value. In addition, plain areas around Cagliari could contribute at increasing the overall representativeness of the regional reserves' network. Differently, Maiorano *et al.* (2006) identified highly irreplaceable areas for vertebrates in western Sardinia. Such incongruence evidences the specificity of prioritization exercises for the considered taxa. The calculation of parks' irreplaceability allows identifying a specific set of key sites. The coastal and insular reserves of Gallura have the highest priority. A second group of important reserves were in Sulcis-Iglesiente. Intermediate levels of irreplaceability were assigned to mountain and coastal sites in eastern Sardinia, Sulcis-Iglesiente and Nurra. Future management

strategies should consider this set of key sites as conservation priorities for Sardinian reptiles.

Our gap analysis evidences a set of species that are under-represented in protected areas. Three of these species (*Archaeolacerta bedriagae*, *Euleptes europaea* and *Tarentola mauritanica*) often co-occur in rocky sites of eastern Sardinia (Bombi *et al.*, 2009a; 2009b; Salvi & Bombi, 2010). Thus, future conservation strategies could be facilitated as they should focus on co-occurrence sites. Because of the threats to Sardinian populations of *Hemorrhois hippocrepis* (Filippi & Luiselli, 2000; Bombi *et al.*, 2011), increasing the representation of this species in reserves' network is a priority. The highly complementary areas around Cagliari could contribute to meet the conservation target for this species. Finally, new reserves in Gallura could improve the protection of *Emys orbicularis*, *Testudo graeca* and *T. hermanni* at the same time.

Conclusion

Our results suggest that species distribution atlases could represent a priceless source of data for identifying conservation priorities, even at a local scale. In this light, we clarified that appropriate downscaling procedures allow to define the geographic patterns of priority and to quantify the relative importance of protected areas at a higher resolution, suitable for local conservation planning. At the same time, downscaling does not influence the identification of under-represented species. In addition, this paper is the first attempt to suggest the most urgent targets for conservation initiatives on Sardinian reptiles in terms of species and areas. Further studies are needed for improving the quality of available data and the assessing of the actual protection regimes of each reserve, but here we furnish a first detailed priority scheme that could represent a practical basis for setting regional strategies.

Acknowledgements

We wish to thank Jeroen Speybroeck for his courtesy in providing field data for validation, Manuela D'Amen for her competent comments to the manuscript, Antonello Loddo and Nicola Cabitta (Oasi WWF Monte Arcosu) for logistic support, Martin Bennett for the improvement of the English style, and two anonymous reviewers for their helpful comments. PB has been partially supported by funds of WWF Oasi, DS by a FCT grant (Grant No. SFRH/BPD/66592/2009), and MAB by a fund of the Italian Ministry for Instruction, University and Research (Grant No. PRIN 20085YJMTTC).

References

Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232.

- Anderson, B.J., Arroyo, B., Collingham, Y., Etheridge, B., Fernandez-De-Simon, J., Gillings, S., Gregory, R.D., Leckie, F., Sim, I. & Thomas, C.D. (2009). Using distribution models to test alternative hypotheses about a species' environmental limits and recovery prospects. *Biol. Conserv.* **142**, 488–499.
- Araújo, M.B. (2004). Matching species with reserves – uncertainties from using data at different resolutions. *Biol. Conserv.* **118**, 533–538.
- Araújo, M.B., Densham, P.J. & Williams, P.H. (2004). Representing species in reserves from patterns of assemblage diversity. *J. Biogeogr.* **31**, 1037–1050.
- Araújo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *TREE* **22**, 42–47.
- Araújo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005). Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Glob. Ecol. Biogeogr.* **14**, 17–30.
- Belsley, D.A., Kuh, E. & Welsch, R.E. (1980). *Regression diagnostics*. New York: John Wiley.
- Blanchet, F.G., Legendre, P. & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology* **89**, 2623–2632.
- Blondel, J., Aronson, J., Bodiou, J.Y. & Boeuf, G. (2010). *The Mediterranean region: biological diversity in space and time*. 2nd edn. New York, NY: Oxford University Press.
- Bombi, P., Capula, M., D'Amen, M. & Luiselli, L. (2011). Climate change threatens the survival of highly endangered Sardinian populations of the snake *Hemorrhois hippocrepis*. *Anim. Biol.* **61**, 239–248.
- Bombi, P. & D'Amen, M. (2011). Scaling down distribution maps from atlas data: a test of different approaches with virtual species. *J. Biogeogr.* doi:10.1111/j.1365-2699.2011.02627.x.
- Bombi, P., Luiselli, L. & D'Amen, M. (2011). When the method for mapping species matters: defining priority areas for conservation of African freshwater turtles. *Divers. Distrib.* **17**, 581–592.
- Bombi, P., Salvi, D., Vignoli, L. & Bologna, M.A. (2009a). Modelling Bedriaga's rock lizard distribution: an ensemble approach. *Amphibia-Reptilia* **30**, 413–424.
- Bombi, P., Salvi, D., Luiselli, L. & Bologna, M.A. (2009b). Modelling correlates of microhabitat use of two sympatric lizards: a model selection approach. *Anim. Biol.* **59**, 109–126.
- Breiman, L. (2001). Random forests. *Mach. Learn.* **45**, 5–32.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984). *Classification and regression trees*. New York: Chapman and Hall.
- Brotos, L., Thuiller, W., Araújo, M.B. & Hirzel, H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* **27**, 437–448.

- Carroll, C. (2010). Role of climatic niche models in focal-species-based conservation planning: assessing potential effects of climate change on Northern Spotted Owl in the Pacific Northwest, USA. *Biol. Conserv.* **143**, 1432–1437.
- Collingham, Y.C., Wadsworth, R.A., Huntley, B. & Hulme, P.E. (2000). Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *J. Appl. Ecol.* **37**, 13–27.
- Costa, G.C., Nogueira, C., Machado, R.B. & Colli, G.R. (2010). Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiv. Conserv.* **19**, 883–899.
- D'Amen, M., Bombi, P., Pearman, P.B., Schmatz, D.R., Zimmermann, N.E. & Bologna, M.A. (2011). Will climate change reduce the efficacy of protected areas for amphibian conservation in Italy? *Biol. Conserv.* **144**, 989–999.
- Dennis, R. (2003). Gains and losses of French butterflies: tests of predictions, under-recording and regional extinction from data in a new atlas. *Biol. Conserv.* **110**, 131–139.
- Engler, R., Guisan, A. & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J. Appl. Ecol.* **41**, 263–274.
- Ferrier, S., Pressey, R.L. & Barrett, T.W. (2000). A new predictor of the irreplaceability of areas for achieving a conservation goal, its application to real-world planning, and a research agenda for further refinement. *Biol. Conserv.* **93**, 303–325.
- Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **24**, 38–49.
- Filippi, E. & Luiselli, L. (2000). Status of the Italian snake-fauna and assessment of conservation threats. *Biol. Conserv.* **93**, 219–225.
- Freeman, E. (2007). *PresenceAbsence: an R package for presence-absence model evaluation*. cran.r-project.org
- Friedman, J. (1991). Multivariate adaptive regression splines. *Ann. Stat.* **19**, 1–67.
- Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **29**, 1189–1232.
- Gehlke, C. & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *J. Am. Stat. Assoc.* **29**, 169–170.
- Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009.
- Guisan, A. & Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecol. Modell.* **135**, 147–186.
- Hastie, T., Tibshirani, R. & Buja, A. (1994). Flexible discriminant analysis by optimal scoring. *JASA* **89**, 1255–1270.
- IUCN (2010). *IUCN red list of threatened species*. Version 2010.4. <http://www.iucnredlist.org>
- Jennings, M.D. (2000). Gap analysis: concepts, methods, and recent results. *Landsc. Ecol.* **15**, 5–20.
- Loiselle, B.A., Graham, C.H., Goerck, J.M. & Smith, K.G. (2003). Avoiding pitfalls of using species distribution models in conservation planning. *Conserv. Biol.* **17**, 1591–1600.
- Maiorano, L., Falcucci, A. & Boitani, L. (2006). Gap analysis of terrestrial vertebrates in Italy: priorities for conservation planning in human dominated landscape. *Biol. Conserv.* **133**, 455–473.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2008). Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* **15**, 59–69.
- Marx, B.D. & Smith, E.P. (1990). Weighted multicollinearity in logistic regression: diagnostics and biased estimation techniques with an example from lake acidification. *Can. J. Fish. Aquat. Sci.* **47**, 1128–1135.
- McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models*. London: Chapman and Hall.
- McPherson, J.M. & Jetz, W. (2007). Effects of species ecology on the accuracy of distribution models. *Ecography* **30**, 135–151.
- McPherson, J.M., Jetz, W. & Rogers, D.J. (2004). The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* **41**, 811–823.
- McPherson, J.M., Jetz, W. & Rogers, D.J. (2006). Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions – possibilities and limitations. *Ecol. Modell.* **192**, 499–522.
- Mitikka, V., Heikkinen, R.K., Luoto, M., Araújo, M.B., Saarinen, K., Pöyry, K. & Fronzek, S. (2007). Predicting range expansion of the map butterfly in Northern Europe using bioclimatic models. *Biodiv. Conserv.* **17**, 623–641.
- Mittermeier, R.A., Gil, R.P., Hoffmann, M., Pilgrim, J. & Brooks, T. (2004). *Hotspots revisited: earth's biologically richest and most threatened terrestrial ecoregions*. Mexico: CEMEX.
- Niamir, A., Skidmore, A.K., Toxopeus, A.G., Muñoz, A.R. & Real, R. (2011). Finessing atlas data for species distribution models. *Divers. Distrib.* **17**, 1173–1185. doi: 10.1111/j.1472-4642.2011.00793.x.
- Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich, UK: Geo Abstracts.
- Pearson, R.G., Dawson, T.P. & Liu, C. (2004). Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* **27**, 285–298.
- Pressey, R.L., Johnson, I.R. & Wilson, P.D. (1994). Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. *Biodivers. Conserv.* **3**, 242–262.

- Pressey, R.L., Watts, M.E., Barrett, T.W. & Ridges, M.J. (2009). The C-plan conservation planning system: origins, applications, and possible futures. In *Spatial models for conservation*: 211–234. Moilanen, A., Possingham, H.P. & Wilson, K.A. (Eds). Oxford: Oxford University Press. <http://www.uq.edu.au/~uqmwatts/cplan.html>
- R Development Core Team (2010). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ripley, B.D. (1996). *Pattern recognition and neural networks*. Cambridge University Press.
- Rodrigues, A.S.L., Akcakaya, H.R., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Chanson, J.S., Fishpool, L.D.C., Da Fonseca, G.A.B., Gaston, K.J., Hoffmann, M., Marquet, P.A., Pilgrim, J.D., Pressey, R.L., Schipper, J., Sechrest, W., Stuart, S.N., Underhill, L.G., Waller, R.W., Watts, M.E.J. & Yan, X. (2004). Global gap analysis: priority regions for expanding the global protected-area network. *BioScience* **54**, 1092–1100.
- Salvi, D. & Bombi, P. (2010). Reptiles of Sardinia: updating the knowledge on their distribution. *Acta Herpetol.* **5**, 161–177.
- Thuiller, W. (2003). BIOMOD-optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* **9**, 1353–1362.
- Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Glob. Change Biol.* **10**, 2020–2027.
- Thuiller, W., Lavorel, S. & Araújo, M.B. (2005). Niche properties and geographical extent as predictors of species sensitivity to climate change. *Glob. Ecol. Biogeogr.* **14**, 347–357.
- Trivedi, M.R., Berry, P.M., Morecroft, M.D. & Dawson, T.P. (2008). Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Glob. Change Biol.* **14**, 1089–1103.
- Turner, J.R.G., Lennon, J.J. & Lawrenson, J.A. (1988). British bird species distributions and the energy theory. *Nature* **335**, 539–541.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Geographic location of place names used in the text (a) and of protected areas considered in the analyses (b). Protected areas are represented with their centroids and the relative labels correspond to IDs in Table S2.

Table S1 Predictors considered in the analysis, relative variance inflation factor (VIF) of those included in the models, and data sources.

Table S2 Irreplaceability values of each protected area as calculated from the two approaches. The IDs correspond to the labels in Fig. S1b. In bold are indicated the ten most irreplaceable reserves according to the two approaches.

Please note: Wiley–Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.