Modelling Bedriaga's rock lizard distribution in Sardinia: An ensemble approach

Pierluigi Bombi[*], Daniele Salvi, Leonardo Vignoli, Marco A. Bologna

Abstract. Many techniques for predicting species potential distribution were recently developed. Despite the international interest for these procedures, applications of predictive approaches to the study of Italian fauna distribution are exceptionally rare. This paper aimed at: (a) detecting climatic exigencies of *A. bedriagae* in Sardinia; (b) predicting the *Archaeolacerta bedriagae* Sardinian potential distribution; (c) identifying the most vulnerable Italian populations of the species. Literature and field data were utilized as presence records. Six modelling procedures (BIOCLIM, DOMAIN, ENFA, GAM, GLM, and MAXENT) were adopted. The species climatic requirements were defined using the WorldClim databank for deriving the environmental predictors. AUC and Kappa values were calculated for models validation. AUC values were compared by using Anova Monte Carlo. The best four models were combined through the weighted average consensus method for producing a univocal output. GAM and MAXENT had the best performances (respectively: AUC $= 0.93 \pm 0.03$, Kappa $= 0.77 \pm 0.08$; AUC $= 0.93 \pm 0.03$, Kappa $= 0.78 \pm 0.07$). Good results were also obtained by GLM and DOMAIN (respectively: AUC $= 0.89 \pm 0.04$, Kappa $= 0.72 \pm 0.05$; AUC $= 0.88 \pm 0.04$, Kappa $= 0.69 \pm 0.07$). BIOCLIM and ENFA gained relatively low performances (respectively: AUC $= 0.78 \pm 0.07$, Kappa $= 0.57 \pm 0.14$; AUC $= 0.75 \pm 0.06$; Kappa $= 0.49 \pm 0.10$). In Sardinia *A. bedriagae* is mainly influenced by seasonality, which causes the evidenced range fragmentation. Moreover, the general importance of multi-methods approaches and consensus techniques in predicting species distribution was highlighted.

Keywords: BIOCLIM, DOMAIN, ENFA, GAM, GLM, Lacertidae, MAXENT, Sardinia.

Introduction

During the last decades, the development of geostatistical analyses, supported by improved Geographic Information System (GIS) technology, provided new instruments to conservation biology. In particular, the possibility of applying predictive modelling techniques for studying species distribution revealed its great usefulness, being ever more commonly used (e.g. Guisan and Zimmermann, 2000; Graham et al., 2004; Sanchez-Cordero, Munguia and Townsend-Peterson, 2004; Elith et al., 2006). For this reason, many different approaches were developed for elaborating habitat suitability models on the basis of species-environment interactions (e.g. Ferrier, 1984; Verner, Morrison and Ralph, 1986; Margules and Austin, 1991; Franklin, 1995; Austin, 1998; Segurado and Araújo, 2004). Some authors utilized an a-

priori definition, based on expert knowledge, for defining the species ecological requirements (e.g. Donovan, Rabe and Olson, 1987; Breininger, Provancha and Smith, 1991; Congalton, Stenback and Barrett, 1993; Boitani et al., 2002). On the contrary, many other techniques applies specific algorithms for deriving species environmental needs from point data (e.g. Nix, 1986; Carpenter, Gillison and Winter, 1993; Manel, Dias and Ormerod, 1999; Guisan and Zimmermann, 2000; Hirzel et al., 2002; Brotons et al., 2004; Segurado and Araújo, 2004; Elith et al., 2006). A group of algorithms compares presence data with reliable absence data for defining the species ecological niche. This approach is utilized by several algorithms, such as classification and regression tree analyses (Breiman et al., 1984), generalized linear models (GLM; Pearce and Ferrier, 2000), generalized additive models (GAM; Pearce and Ferrier, 2000), and artificial neural networks (Thuiller, 2003). A second group of algorithms defines the species ecological requirements on

Department of Environmental Biology, University "Roma Tre", Viale G. Marconi 446, 00146 Rome, Italy
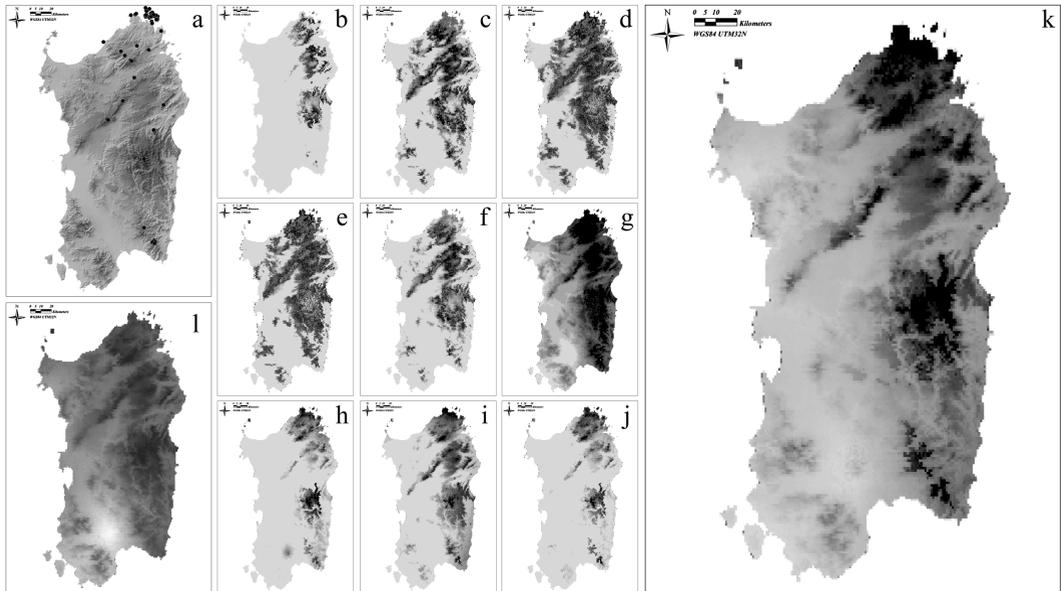[*]Corresponding author; e-mail: bombi@uniroma3.it

the basis of presence data only. BIOCLIM (Nix, 1986), DOMAIN (Carpenter, Gillison and Winter, 1993), Mahalanobis distance (Farber and Kadmon, 2003), and GARP (Stockwell and Peters, 1999) can be used as examples of this category. Another approach, adopted by ENFA (Hirzel et al., 2002) and MAXENT (Phillips, Dudik and Schapire, 2004) among others, consists of comparing presence data with background conditions. Generally, the first group of algorithms has very good predictive performances (Brotons et al., 2004) but, on the other hand, the required absence data are rarely available in faunistic datasets. This problem can be partially overcome by utilizing pseudo-absence data, generated through alternative procedures (e.g. randomly, or derived from preliminary habitat suitability maps), for representing real-absence data (Engler, Guisan and Rechsteiner, 2004; Chefaoui and Lobo, 2008).

The large number of techniques developed over a relatively short period of time is proof of the great interest in predictive approaches to the study of species distribution. The conservationist interest for these procedures was still increased by the introduction of consensus methods (e.g. Araújo and New, 2007; Marmion et al., 2008). They allow at considering the heterogeneous results obtained by multiple models as an ensemble, combining models through alternative methods (e.g. Weighted Average, Mean, Median). The consensus methods have the great advantage of allowing to overcome the uncertainty deriving from multiple results, providing a univocal and objective instrument for biodiversity conservation (Araújo and New, 2007).

Despite the amount of attention paid to animal distribution modelling worldwide, these techniques are almost neglected in Italy. One of the few examples of potential distribution prediction for the Italian fauna is the Italian Ecological Network (REN; Boitani et al., 2002), which took into account all vertebrate species. Unfortunately, the application of the REN approach to amphibians and reptiles showed some difficulties (Boitani et al., 2002). Such difficul-

ties are related to the extremely fine-grain environment perception of these animals, which can be difficultly considered by the REN procedure.

The usefulness of these geo-statistic procedures in decision-making process aimed at biodiversity protection is that they allow at focalizing the conservation initiatives toward the most critical areas and populations. It can improve our efficiency in planning effective measures for contrasting the current biodiversity loss. The allocation of the chronically limited resources available for the protection of nature can be optimized if based on scientifically sound strategies. At the light of these considerations we applied habitat suitability modelling techniques to the Bedriaga's rock lizard (*Archaeolacerta bedriagae* [Camerano, 1885]), a lacertid species endemic to Corsica and Sardinia, in order to fill the gap of knowledge about its distribution and to provide new instruments for its conservation. This species is a rock dwelling lizard widespread from the sea level up to the highest peaks of both islands (Schneider, 1984; Castilla et al., 1989; Bauwens et al., 1990; Vanhooydonck, Van Damme and Aerts, 2000; Bombi and Vignoli, 2004; Bombi et al., 2009). The conservation status of *A. bedriagae* is poorly known, as well as its biology. This species is listed in Annex II of the Bern Convention (Council of Europe, Bern 19.9.1979) and in the Annex IV of the Habitat Directive (92/43/EEC), and is considered Vulnerable in the Red Data Book of the Italian Vertebrates (Capula, 1998). Actually, the fragmentation of the Italian range (fig. 1a) of this strictly rupicolous lizard is locally emphasized by the rocky habitat discontinuity that greatly increases the conservation relevance of the species. Aims of this paper are: (i) comparing different techniques of habitat suitability modelling, (ii) detecting the climatic variables affecting the species distribution, and (iii) producing a univocal consensus model, in order to provide a helpful instrument for *A. bedriagae* conservation in Italy.

Figure 1. (a) Distribution of *Archaeolacerta bedriagae* (black points) in Sardinia (the darkness of the pixels is proportional to the elevation [pale grey = sea level; black = 1900 m a.s.l.]). Models of habitat suitability produced by BIOCLIM (b), ENFA distance geometric mean (c), ENFA distance harmonic mean (d), ENFA minimum distance (e), ENFA median (f), DOMAIN (g), GLM (h), MAXENT (i), GAM (j); the darkness of the pixels is proportional to the habitat suitability value (pale grey: not suitable; black: highly suitable). (k) Consensus model obtained through the weighted average method; the darkness of the pixels is proportional to the habitat suitability value (pale grey: not suitable; black: highly suitable). (l) Isothermality in Sardinia (pale gray = 26.7, black = 40.0).
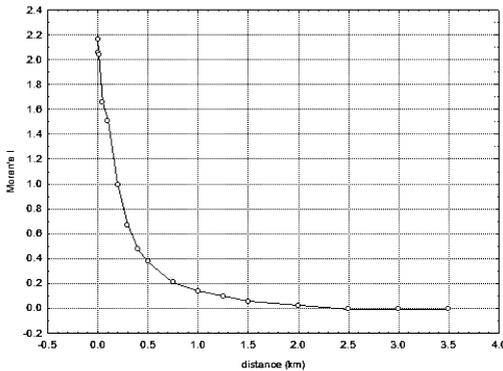
Materials and methods

Species distribution data

Data collection. The complete list of *A. bedriagae* localities in Sardinia was compiled by consulting all the available literature on the one hand (e.g., Peracca, 1905; Mertens, 1932; Stemmler, 1962; Cesaraccio and Lanza, 1984; Lanza, Cesaraccio and Malenotti, 1984; Bruno, 1986; Borri et al., 1988; Puddu, Viarengo and Erminio, 1988; In den Bosch, 1992; Poggesi et al., 1996; Michelot, 1997; Bombi and Vignoli, 2004) and by conducting intensive field research on the other. The fieldwork was carried out between June 2000 and April 2005 throughout Sardinia, including satellite islands, with the aim of confirming the species presence in cited locations, defining the exact coordinates of those localities, and identifying new sites. The presence of *A. bedriagae* was checked in all of the cited localities as well as in other rocky areas that appeared to be a-priori suitable for the particular autoecological requirements of the species (Bombi et al., 2009). Exact coordinates of presence sites were obtained by a GPS (Garmin, GPSmap 76S). The peculiar activity pattern and the generally low population density of this species (Bombi and Vignoli, 2004) did not allow at defining reliable absence data. Therefore, the apparently suitable sites visited but not confirmed, nor cited in literature, were excluded from the dataset, as well as the bibliographic localities apparently not suitable and derived from unreliable sources. Detailed coordinates of bibliographic sites

were assigned by individuating rocky outcrops in the area by means of detailed maps (IGM 1:25000), satellite images (http://earth.google.com), and field surveys. This process allowed us at obtaining 65 presence data throughout the entire Sardinian range of *A. bedriagae*, with a spatial resolution of few hundreds of meters or higher (fig. 1a). The study area was intentionally limited to Sardinia in order to focus the attention on the species conservation status in this island.

Data preparation. Spatial autocorrelation in presence data was assessed by producing the Moran's *I* correlogram (fig. 2), which allowed to identify the threshold of 1 km as minimum distance between points for minimizing spatial autocorrelation effect. On the basis of this preliminary analysis seven points were removed to avoid bias introduced by clustered distribution of data, and one single point per cell was used in the analyses. The remaining dataset was split in ten random repetitions of two subsets. The first group of subsets, containing between 26 and 37 records (approximately 55% of data), was used to elaborate the models (train subset). The second group of subsets, containing between 21 and 32 records (approximately 45% of data), was used for validation (test subset). A number of pseudo-absence points, twice the number of original data, were added to each subset, both train and test. These pseudo-absence data were generated randomly (Engler, Guisan and Rechsteiner, 2004) within the study area and used for validating all of the obtained models, as well as for performing the two regression procedures.

Figure 2. Moran's *I* correlogram for assessing spatial auto-correlation effect.

Environmental predictors

Variables acquisition. We used as environmental predictors the climatic data provided by WorldClim (version 1.4) (Hijmans et al., 2004). The WorldClim databanks consist of climate surfaces with global coverage for 19 climatic variables (table 1) interpolated using data for the period 1950-2000. In particular, we adopted the climate surfaces with pixels of 30″ of geographic degree, corresponding to a resolution of approximately 1 km (see Hijmans et al., 2005 for details). The spatial resolution of 30″ was maintained in all of the following analyses as well as in the output. We utilized climatic variables only as predictors because of the extremely fine-grain environment perception of reptiles, which is difficultly captured in land-cover maps (Boitani et al., 2002).

Variables selection. Climatic predictors were included in the models according to the specific properties of each algorithm, by adopting different variable selection approaches.

Models production

Six widely used procedures (BIOCLIM, DOMAIN, ENFA, GLM, GAM, and MAXENT) were applied for defining the species climatic exigencies on the basis of climatic data. Each algorithm was calibrated by using all of the ten independent train repetitions, thus ten different models were obtained from each different procedure. The ten models were independently validated and summarized in one mean model for each algorithm.

BIOCLIM/DOMAIN. BIOCLIM (Nix, 1986) defines the bounding box enclosing the species records in climatic hyper-space (e.g. Busby, 1991; Elith and Burgman, 2003; Farber and Kadmon, 2003; Parra, Graham and Freile, 2004 for details). DOMAIN (Carpenter, Gillison and Winter, 1993) measures the pixel's Gower distance from the most similar presence point in the environmental hyper-space (e.g. Carpenter, Gillison and Winter, 1993; Elith and Burgman, 2003 for details). BIOCLIM and DOMAIN were implemented by using DIVA-GIS 5.2 (www.diva-gis.org). With the aim of avoiding over-parameterization (Kriticos

Table 1. Climatic variables used to elaborate the models.

var1 = Annual mean temperature
var2 = Mean diurnal range (mean of monthly×
   (max temp − min temp))
var3 = Isothermality ((var2/var7) ∗ 100)
var4 = Temperature seasonality (standard deviation ∗ 100)
var5 = Maximum temperature of warmest month
var6 = Minimum temperature of coldest month
var7 = Temperature annual range (var5 − var6)
var8 = Mean temperature of wettest quarter
var9 = Mean temperature of driest quarter
var10 = Mean temperature of warmest quarter
var11 = Mean temperature of coldest quarter
var12 = Annual precipitation
var13 = Precipitation of wettest month
var14 = Precipitation of driest month
var15 = Precipitation seasonality
   (standard deviation/mean)
var16 = Precipitation of wettest quarter
var17 = Precipitation of driest quarter
var18 = Precipitation of warmest quarter
var19 = Precipitation of coldest quarter

and Randall, 2001; Williams, Bolitho and Fox, 2003) applying BIOCLIM and DOMAIN, the variables in the final models were selected by comparing the validation scores of preliminary univariate models and including only those parameters with scores higher than a specific threshold value (i.e. AUC > 0.7; Kappa > 0.4) (Fielding and Bell, 1997; Tape, 2006).

ENFA. The Ecological Niche Factor Analysis (ENFA) (Hirzel et al., 2002) compares the distributions of the environmental variables between the occurrence locations and the whole study area. This analysis extracts factors that explain the relationship between species distribution and ecological variables. ENFA does not produce values of habitat suitability directly, but they can be derived combining the extracted factors. Four different algorithms can be used to produce, on the basis of the computed factors, maps of habitat suitability: Median, Distance Geometric Mean, Distance Harmonic Mean, Minimum Distance (e.g. Hirzel et al., 2002; Hirzel, Hausser and Perrin, 2004 for details). Bio-Mapper 3.1 (Hirzel, Hausser and Perrin, 2004) was adopted to perform ENFA and the relative mapping algorithms. IDRISI 15.0 Andes (Clark Labs, 1987-2006) was used for importing and exporting data to and from BioMapper. The number of variables considered by ENFA does not affect the resulting models, because this procedure computes a sort of "factor loading" of each predictor that weights relative importance of variables (Hirzel et al., 2002; Hirzel, Hausser and Perrin, 2004). Hence all the variables were considered together, after the exclusion of those correlated (variables n°1, 9, 11, 13, 14, 16, 17, and 19 were excluded). The factors obtained by ENFA were used to compute the final models of habitat suitability by using all of the four algorithms (Median, Distance Geometric Mean, Distance Harmonic Mean, Minimum Distance), obtaining 10 models for each of the four mapping algorithms.

GLM/GAM. Generalized Linear Models (GLMs) (McCullagh and Nelder, 1989) are mathematical extensions of linear models that allow for non-linearity and non-constant variance structures in the data (e.g. Guisan, Edwards and Hastie, 2002; Brotons et al., 2004 for details). Generalized Additive Models (GAMs) (Hastie and Tibshirani, 1986) are semi-parametric extensions of GLMs (e.g. Guisan, Edwards and Hastie, 2002; Brotons et al., 2004 for details): in GAMs some predictors can be modelled non-parametrically in addition to linear and polynomial terms for other predictors. R 2.6.2 (R Development Core Team, 2008) was used to perform GLMs and GAMs. The two regression methods (GLM and GAM), with a binomial probability distribution and a logit link, were fitted for *A. bedriagae* based on climatic variables. The climatic predictors were previously tested for correlation and related variables were excluded (i.e. $r^2 > 0.8$; variables n°1, 6, 9, 10, 11, 14, 16, 18, and 19 were excluded). GLMs were performed including predictors, chosen after correlation analysis, in the linear form only. Variables in the models were selected through a stepwise procedure based on Akaike's Information Criterion (AIC; Akaike, 1974) that operates, both backward and forward, by excluding and adding variables one by one from an initial full model, and comparing the resulting AIC values. Similarly, an AIC-based selection of predictors was applied to calibrate our GAMs: this stepwise procedure excludes and adds predictors from the initial model to minimize the AIC score.

MAXENT. In order to estimate the target probability distribution, MAXENT (Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006) finds the probability distribution of maximum entropy that is constrained by considered ecological parameters (e.g. Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006; Peterson, Papes and Eaton, 2007 for details). MaxEnt 3.1 (Phillips, Anderson and Schapire, 2006) was utilized to apply MAXENT approach. Because MAXENT can combine predictors to manage over-fitting by regularizing factors (Phillips, Anderson and Schapire, 2006; Phillips and Dudík, 2008), we performed this algorithm by using all of the climatic variables.

Models validation

For validation purposes, we adopted two of the most widely used procedures (e.g. Pearce and Ferrier, 2000; Elith, Burgman and Regan, 2002; Lehmann, Overton and Leathwick, 2002; Zaniewski, Lehmann and Overton, 2002; Segurado and Araújo, 2004; Elith et al., 2006). The first method compares the predicted values of habitat suitability assigned to presence and pseudo-absence data in the test subset by producing the "Receiver Operating Characteristic" (ROC) plots (Fielding and Bell, 1997) and deriving the relative "Area Under Curve" (AUC) value (Faraggi and Reiser, 2002). The main positive feature of this procedure consists of being a single threshold-independent measure for model performance (Manel, Williams and Ormerod, 2001; Lehmann, Overton and Leathwick, 2002; Thuiller, 2003; Brotons et al., 2004; McPherson, Jetz and Rogers,

2004; Thuiller, Lavorel and Araújo, 2005; Allouche, Tsoar and Kadmon, 2006; Peterson, Papes and Eaton, 2007). Secondly, the maximum Kappa (Cohen, 1960) was calculated. Cohen's kappa (Shao and Halpin, 1995; Manel, Williams and Ormerod, 2001; Loiselle et al., 2003; Petit et al., 2003; Berg, Gardenfors and Von Proschwitz, 2004; Parra, Graham and Freile, 2004; Pearson, Dawson and Liu, 2004; Rouget et al., 2004; Segurado and Araújo, 2004) corrects the overall accuracy of model predictions by the accuracy expected to occur by chance. Other advantages of Kappa are its simplicity, the fact that both commission and omission errors are accounted for in one parameter, and its relative tolerance to zero values in the confusion matrix (Manel, Williams and Ormerod, 2001; Allouche, Tsoar and Kadmon, 2006). Generally, the validation score are considered excellent if $0.9 < \text{AUC} < 1$ and $0.75 < \text{Kappa} < 1$, good if $0.80 < \text{AUC} < 0.90$ and $0.4 < \text{Kappa} < 0.75$, fair if $0.70 < \text{AUC} < 0.80$, poor if $0.60 < \text{AUC} < 0.70$ and Kappa < 0.4, and fail if $0.50 < \text{AUC} < 0.60$ (e.g. Landis and Koch, 1977; Fielding and Bell, 1997). All of the models were validated in DIVA-GIS by using the test subsets, obtaining ten AUC and Kappa values for each modelling approach.

Models comparison

In order to evaluate the predictive performances, we compared the ten AUC values of each modelling procedure for finding non-random differences and, therefore, identifying homogeneous groups of algorithms. We adopted a Null Models approach for performing such comparison, contrasting the observed Anova *F* index with those simulated by 3×10^4 random Monte Carlo permutations in EcoSim 7.0 (Gotelli and Entsminger, 2001). This number of permutations ensures that algorithm biases are avoided (Lehsten and Harmand, 2006). Random simulations were generated by shuffling the repetition algorithms and retaining the original AUC values of single repetitions (Gotelli and Entsminger, 2001). *F* index was calculated for the original data as well as for the simulated matrices and results were compared, calculating the probability (P) of the null hypothesis that the observed *F* index (F_{obs}) was drawn at random from the distribution of the simulated *F* indexes (F_{exp}) (Edgington, 1987; Gotelli, 2000). Non-random differences were assumed when $P_{F_{obs} \geqslant F_{exp}} \leqslant 0.05$ (Gotelli and Graves, 1996).

Ensemble

The mean models obtained by summarizing the ten repetitions of the best four techniques (according to their AUC scores) were combined by means of the weighted average consensus method (Araújo and New, 2007; Marmion et al., 2008). The habitat suitability of the *i*th grid cell (HS_i) in the consensus model was calculated as:

$$HS_i = \frac{\sum_j (AUC_{mj} \times mj_i)}{\sum_j AUC_{mj}},$$

where mj_i are the probability of presence of *A. bedriagae* in the *i*th grid cell as predicted by the model elaborated with

the *j*th algorithm. ArcGIS 9.2 (ESRI Inc., Redlands, CA, USA) was utilized for summarizing the ten repetitions of each procedure and for producing the consensus model.
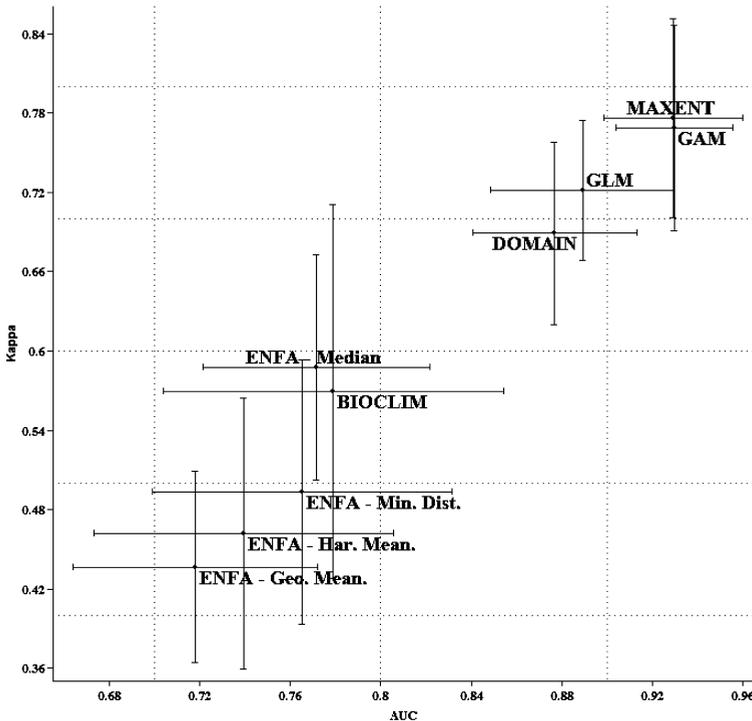
Results

Single-models performances

The predictive performances of the adopted algorithms were non-homogeneous. BIOCLIM obtained medium AUC and Kappa values (AUC = 0.779 ± 0.075, Kappa = 0.569 ± 0.141, fig. 1b). DOMAIN got high validation scores (AUC = 0.876 ± 0.036, Kappa = 0.688 ± 0.069, fig. 1g). Models derived by ENFA reached medium values (Median: AUC = 0.771 ± 0.050, Kappa = 0.587 ± 0.085, fig. 1f; Minimum Distance: AUC = 0.765 ± 0.066, Kappa = 0.493 ± 0.100, fig. 1e; Distance Harmonic Mean: AUC = 0.739 ± 0.066, Kappa = 0.461 ± 0.103, fig. 1d; Distance Geometric Mean: AUC = 0.718 ± 0.054, Kappa = 0.436 ± 0.072, fig. 1c). GLM produced models

with high validation values (AUC = 0.889 ± 0.041, Kappa = 0.721 ± 0.053, fig. 1h). Models obtained through GAM procedure reached very high AUC and Kappa (AUC = 0.929 ± 0.026, Kappa = 0.768 ± 0.078, fig. 1j). Predictions generated by MAXENT got very high AUC and Kappa (AUC = 0.929 ± 0.031, Kappa = 0.776 ± 0.075, fig. 1i).

Models comparison

The Anova performed through Monte Carlo procedure for comparing predictive performances of competing methods revealed three significantly different groups of algorithms (fig. 3). The first group, homogeneous in terms of AUC values ($F_{obs} = 0.001$, $P_{Fobs \geqslant Fexp} = 0.977$), comprises GAM and MAXENT. This group obtained validation values significantly higher ($F_{obs} = 18.943$, $P_{Fobs \geqslant Fexp} < 0.001$) than values reached by GLM and DOMAIN, which have similar performances ($F_{obs} = $

Figure 3. Predictive performance of the competing algorithms. Mean and standard deviation of AUC and Cohen's maximum Kappa obtained by the ten repetitions of the applied procedures.

0.481, $P_{Fobs \geqslant Fexp} = 0.492$). GLM-DOMAIN obtained AUC values significantly higher ($F_{obs} = 59.764$, $P_{Fobs \geqslant Fexp} < 0.001$) than the third group, composed by BIOCLIM and the four mapping algorithms of ENFA. This group includes procedures with homogeneous AUC values ($F_{obs} = 1.550$, $P_{Fobs \geqslant Fexp} = 0.201$), but ranging between a relatively large interval (0.779 [BIOCLIM] −0.718 [ENFA − Distance Geometric Mean]). In addition, the algorithms belonging to this group have a standard deviation significantly higher ($F_{obs} = 24.395$, $P_{Fobs \geqslant Fexp} = 0.007$) than the standard deviation of the two best groups.

Variables importance

Three variables ("Maximum Temperature of Warmest Month", "Isothermality", and "Precipitation Seasonality") produced the best models using BIOCLIM (respectively: Mean $AUC_{var5} = 0.813$, Mean $AUC_{var3} = 0.759$, and Mean $AUC_{var15} = 0.720$) and four ("Precipitation Seasonality", "Isothermality", "Maximum Temperature of Warmest Month", and "Mean Diurnal Temperature Range") using DOMAIN (respectively: Mean $AUC_{var15} = 0.833$, Mean $AUC_{var3} = 0.815$, Mean $AUC_{var5} = 0.793$, and Mean $AUC_{var2} = 0.770$).

ENFA showed the importance of "Precipitation of Warmest Quarter", "Precipitation of Driest Month", "Precipitation of Driest Quarter", and "Annual Precipitation" that made the highest contribution to marginality (i.e. mean score > 0.15). "Precipitation of Warmest Quarter" (mean score = 0.40) had the highest score ten out of ten times. "Precipitation of Driest Month" (mean score = 0.36) got the second highest score eight out of ten times. "Precipitation of Driest Quarter" (mean score = 0.28) reached the third highest score six times. "Annual Precipitation" (mean score = 0.15) got the second highest score two times.

GLMs and GAMs evidenced the importance of "Mean Temperature of Wettest Quarter" (selected eight out of ten times by both GLM and GAM), "Temperature Seasonality" (selected six

out of ten times), and "Isothermality" (selected five out of ten times).

"Isothermality", "Mean Diurnal Temperature Range", "Maximum Temperature of Warmest Month", and "Precipitation Seasonality" made the most important contribution to MAXENT (i.e. mean contribution > 15%). In particular, "Isothermality" (mean contribution = 25.91%) was the most important variable five out of ten times, "Mean Diurnal Temperature Range" (mean contribution = 21.13%) four times, "Maximum Temperature of Warmest Month" (mean contribution = 17.25%) one time, and "Precipitation Seasonality" (mean contribution = 15.08%) appeared as the second most important parameter three out of ten times.

Habitat suitability pattern

The best models shared a common geographic pattern as highlighted by the consensus model (fig. 1k). The range fragmentation of Bedriaga's rock lizard in Sardinia was evidenced. Three main sub-ranges in eastern Sardinia were identified. In addition, one thin corridor of suitability in central Sardinia, corresponding with Marghine and Goceano mountain chains, and two smaller areas of potential presence in the South-East of the island, corresponding with Mount Genis and Mount Settefratelli, can be observed.

Discussion

Models performances

GAM regression analysis and MAXENT are the best procedures, among those we tested, in estimating climatic exigencies and modelling habitat suitability for *A. bedriagae* in Sardinia. This result is partially in agreement with previous comparative analyses of different modelling approaches (e.g. Elith et al., 2006), which highlighted MAXENT as one of the top-performing algorithms. In contrast, the extremely good models obtained by GAM – significantly better than results of GLM – were in some mea-

sure unexpected, and they are probably due to the particular relationship between the presence of this species and the climatic predictors that is difficultly described by the linear combination of explanatory variables utilized in GLM. This partially unexpected result puts in evidence the strong influence played by the target species ecology and distribution on the efficiency of different modelling procedures.

The regression approaches utilized in GLMs, including predictors in linear form, as well as the Gower metric utilized by DOMAIN still appear as good methods in modelling the potential distribution of our target species. The similar results obtained by these two algorithms are partially surprisingly, being shown in previous papers (Segurado and Araújo, 2004; Elith et al., 2006) the relatively scarce performances of DOMAIN. This unexpected good efficiency of DOMAIN is probably due to the complex influences of climatic predictors on Bedriaga's rock lizard distribution and to the relatively low number of records, which are particularly fitted to be described by the Gower distance analysis (Guisan and Zimmermann, 2000).

The simple bounding box computed by BIO-CLIM, as well as the pseudo-PCA performed by ENFA, are acceptable methods when used in defining a general pattern of distribution and particularly helpful if applied in preliminary analyses (Engler, Guisan and Rechsteiner, 2004). In fact, from our results, it is clear that the pattern of suitability predicted by these two approaches (fig. 1b-f) is, with the obvious differences and uncertainties, similar to those produced by more performing methods (fig. 1g-j). The comparable efficiency here demonstrated for BIOCLIM and ENFA is in accordance with previous studies (Tsoar et al., 2007), which highlighted similar performances for these two algorithms. On the other hand, Tsoar et al. (2007) did not obtain differences between BIOCLIM-ENFA and DOMAIN as we found.

Regarding the four mapping algorithms applied to ENFA results, we observed some differences in terms of predictive performances. In particular, Median appears the best one, in accordance with the suggestions of Hirzel et al. (2002), and Distance Geometric Mean seems the worst. Nevertheless, it should be taken into account as Null Models analysis demonstrated that such differences could be occurred by chance.

Methods comparison and ensemble approaches

On the one hand, it should be noted that the only other published model of habitat suitability for *A. bedriagae* (Boitani et al., 2002) is not updated in terms of faunistic knowledge. In addition, such model was produced through a deductive approach, based on expert judgment, which is hard to be applied to amphibians and reptiles (Boitani et al., 2002) that are characterized by a scale of environment perception different from the scale of the considered ecogeographic variables. On the other hand, our results of model performances are partially different from those of previous papers (e.g. Segurado and Araújo, 2004; Elith et al., 2006; Tsoar et al., 2007). This evidence demonstrates the strong influence of the target species' ecology and distribution on the predictive performances of different modelling techniques and highlights the importance of multi-methods approaches in predicting species distribution. On its turn, if conservation issues should be addressed, this influence makes particularly important the adoption of consensus methods for deriving one univocal prediction from an ensemble of different models (Araújo and New, 2007; Marmion et al., 2008).

Variables importance

The Sardinian distribution of *A. bedriagae* is mainly influenced by seasonality in terms of both temperature and precipitation. In fact, "Isothermality" was evidenced as an important predictor by all of the modelling procedures but ENFA. Moreover, "Precipitation Seasonality" and "Temperature Seasonality" were important variables in three and two out of six

procedures respectively. These variables have a negative effect on *A. bedriagae* and this could be one of the causes of its absence from western Sardinia, which is characterized by marked seasonality and, more specifically, by high values of Isothermality (fig. 1l). Nevertheless, also extreme values in temperature influence the Bedriaga's rock lizard distribution. "Maximum Temperature of Warmest Month" was an important predictor in three out of six modelling procedures, causing that the species is distributed at low altitudes in the coastal and sub-coastal areas of north-eastern Sardinia only, which is characterized by relatively low maximum summer temperature. This result puts in evidence that our target species is relatively not thermophylous, being excluded from the areas characterized by extremely high summer temperatures. At the light of this scenario, it is possible to hypothesize that extreme climatic seasonality can affect lizard survivorship by influencing its seasonal activity pattern.

Obviously, many other factors may interact for influencing the actual species distribution. Among these adjunctive parameters, interspecific interactions, phylogeographic constrains, as well as suitability of microhabitat conditions, and topographic situations can certainly play a crucial role in shaping species range. Nevertheless, such variables can only act as limiting factors within a general climatic suitability, reducing the real probability of occurrence in a specific area despite of its suitable climatic conditions.

Conservationist conclusions

In terms of general conservation strategy, the evident differences between the unique available model for *A. bedriagae* distribution and our results strongly sustain the importance of overcoming the Italian Ecological Network approach to amphibians and reptiles. It should be addressed by supporting the future use of multiple modelling techniques in predicting the potential range of species belonging to the Italian herpetofauna.

More concretely, this paper represents a helpful instrument for optimizing the initiatives aimed at the conservation of *A. bedriagae*, as well as for focalizing the conservation measures toward the most critical populations. The distribution predicted by our models puts in evidence the strong isolation from the main subranges of certain areas, such as Mount Genis and Mount Settefratelli in south-eastern Sardinia as well as Marghine and Goceano Mountains in central Sardinia. These isolated areas require specific attention by conservationists for avoiding local extinction under dynamic environmental conditions. Indeed, such peripheral suitable nuclei are particularly exposed to threats from stochastic phenomena (e.g. diseases), as well as from changing climate that could drive these areas out of the species envelope.

References

Akaike, H. (1974): A new look at the statistical model identification. IEEE Trans. Aut. Contr. **19**: 716-723.

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.

Araújo, M.B., New, M. (2007): Ensemble forecasting of species distributions. Trends Ecol. Evol. **22**: 42-47.

Austin, M.P. (1998): An ecological perspective on biodiversity investigations: example from Australian eucalypt forests. Ann. Miss. Bot. Gard. **85**: 2-17.

Bauwens, D., Castilla, A.M., Van Damme, R., Verheyen, R.F. (1990): Field body temperatures and thermuregolatory behaviour of the high altitude lizard, *Lacerta bedriagae*. J. Herpetol. **24**: 88-91.

Berg, A., Gardenfors, U., Von Proschwitz, T. (2004): Logistic regression models for predicting occurrence of terrestrial molluscs in southern Sweden: importance of environmental data quality and model complexity. Ecography **27**: 83-93.

Boitani, L., Corsi, F., Falcucci, A., Maiorano, L., Marzetti, I., Masi, M., Montemaggiori, A., Ottaviani, D., Reggiani, G., Rondinini, C. (2002): Rete Ecologica Nazionale. Un approccio alla conservazione dei vertebrati italiani. Università di Roma "La Sapienza", Dipartimento di Biologia Animale e dell'Uomo; Ministero dell'Ambiente, Direzione per la Conservazione della Natura; Istituto di Ecologia Applicata. Available at http://www.gisbau.uniroma1.it/REN.

Bombi, P., Vignoli, L. (2004): Distribution, ecology and conservation of *Archaeolacerta bedriagae* in Sardinia (Reptilia, Lacertidae). In: The Conservation Status of Threatened Amphibian and Reptile Species of the Italian Fauna. Italian Journal of Zoology, Vol. 71 (Suppl. 1), p. 135-144. Bologna, M.A., La Posta, S., Eds, Modena, Mucchi Editore.

Bombi, P., Salvi, D., Luiselli, L., Bologna, M.A. (2009): Modelling correlates of microhabitat use of two sympatric lizards: a model selection approach. Anim. Biol. **59**: 109-126.

Borri, M., Agnelli, P., Cesaraccio, G., Corti, C., Finotello, P., Lanza, B., Tosini, G. (1988): Preliminary notes on the herpetofauna of the satellite islands of Sardinia. Boll. Soc. Sarda Sc. Nat. **26**: 149-165.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984): Classification and Regression Trees. Belmont, CA, Wadsworth International Group.

Breininger, D.R., Provancha, M.J., Smith, R.B. (1991): Mapping Florida scrub jay habitat for purposes of land-use management. Photogramm. Eng. Rem. S. **57**: 1467-1474.

Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H. (2004): Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography **27**: 437-448.

Bruno, S. (1986): Tartarughe e Sauri d'Italia. A. Martello, Giunti Firenze.

Busby, J.R. (1991): BIOCLIM – A bioclimate analysis and prediction system. Nature conservation: cost effective biological surveys and data analysis. In: Nature Conservation, p. 64-68. Margules, C.R., Austin, M.P., Eds, Melbourne, CSIRO.

Capula, M. (1998): Lucertole minacciate. In: Libro Rosso degli Animali d'Italia, Vertebrati, p. 45. Bulgarini, F. et al., Eds, Rome, WWF Italia.

Carpenter, G., Gillison, A.N., Winter, J. (1993): DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodiv. Conserv. **2**: 667-680.

Castilla, A.M., Bauwens, D., Van Damme, R., Verheyen, R. (1989): Notes on the biology of the high altitude lizard *Lacerta bedriagae*. Herpetol. J. **1**: 400-403.

Cesaraccio, G., Lanza, B. (1984): Nuovi dati sull'erpetofauna dell'arcipelago della Maddalena (Sardegna NE). Boll. Soc. Sarda Sci. Nat. **23**: 137-143.

Chefaoui, R.M., Lobo, J.M. (2008): Assessing the effects of pseudo-absences on predictive distribution model performance. Ecol. Model. **210**: 478-486.

Cohen, J. (1960): A coefficient of agreement for nominal scales. Edu. Psychol. Meas. **20**: 37-46.

Congalton, R.G., Stenback, J.M., Barrett, R.H. (1993): Mapping deer habitat suitability using remote sensing and geographic information systems. Geocarto International **3**: 23-33.

Clark Labs (1987-2006): Clark University, available at http://www.clarklabs.org.

Donovan, M.L., Rabe, D.L., Olson, C.E. (1987): Use of geographic information systems to develop habitat suitability models. Wildlife Soc. B **15**: 574-579.

Edgington, E.S. (1987): Randomization Tests. New York, Marcel Dekker.

Elith, J., Burgman, M.A. (2003): Habitat models for PVA. In: Population Viability in Plants. Conservation, Management and Modelling of Rare Plants, p. 203-235. Brigham, C.A., Schwartz, M.W., Eds, New York, Springer-Verlag.

Elith, J., Burgman, M.A., Regan, H.M. (2002): Mapping epistemic uncertainty and vague concepts in predictions of species distribution. Ecol. Model. **157**: 313-329.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E. (2006): Novel methods improve predictions of species distribution from occurrence data. Ecography **29**: 129-151.

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.

Faraggi, D., Reiser, B. (2002): Estimation of the area under the ROC curve. Stat. Med. **21**: 3093-3106.

Farber, O., Kadmon, R. (2003): Assessment of alternative approaches for bioclimatic modelling with special emphasis on the Mahalanobis distance. Ecol. Model. **160**: 115-130.

Ferrier, S. (1984): The status of the Rufous Scrub-Bird *Atrichornis rufescens*: habitat, geographical variation and abundance. Ph.D. Thesis, University of New England, Armidale, Australia.

Fielding, A.H., Bell, J.F. (1997): A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Cons. **24**: 38-49.

Franklin, J. (1995): Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geog. **19**: 474-499.

Gotelli, N.J. (2000): Null model analysis of species co-occurrence patterns. Ecology **81**: 2606-2621.

Gotelli, N.J., Entsminger, G.L. (2001): EcoSim: Null Models Software for Ecology. Version 7. Jericho, VT, Acquired Intelligence Inc. & Kesey-Bear. Available at http://garyentsminger.com/ecosim/index.htm.

Gotelli, N.J., Graves, G.C. (1996): Null Models in Ecology. Washington, DC, Smithsonian Institution Press.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C., Peterson, A.T. (2004): New developments in museum-based informatics and application in biodiversity analysis. Trends Ecol. Evol. **19**: 497-503.

Guisan, A., Zimmermann, N.E. (2000): Predictive habitat distribution models in ecology. Ecol. Model. **135**: 147-186.

Guisan, A., Edwards, T.C., Hastie, J.T. (2002): Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol. Model. **157**: 89-100.

Hastie, T., Tibshirani, R. (1986): Generalized additive models. Stat. Sci. **1**: 297-318.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A. (2004): The WorldClim Interpolated Global Terrestrial Climate Surfaces. Version 1.3. Available at http://biogeo.berkeley.edu/.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A. (2005): Very high resolution interpolated climate surfaces for global land areas. Int. J. Clim. **25**: 1965-1978.

Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N. (2002): Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? Ecology **83**: 2027-2036.

Hirzel, A.H., Hausser, J., Perrin, N. (2004): Biomapper 3.1. Lab. of Conservation Biology, Department of Ecology and Evolution, University of Lausanne. Available at http://www.unil.ch/biomapper.

In den Bosch, H.A.J. (1992): A new locality for *Lacerta bedriagae* (Camerano) in northern Sardinia with data on reproduction. Amphibia-Reptilia **13**: 197-201.

Kriticos, D.J., Randall, R.P. (2001): A comparison of systems to analyse potential weed distributions, In: Weed Risk Assessment, p. 61-79, Groves, R.H., Panetta, F.D., Virtue, J.G., Eds, Collingwood, CSIRO Publishing.

Lanza, B., Cesaraccio, G., Malenotti, P. (1984): Note su *Archaeolacerta bedriagae* (Camerano) (Reptilia Lacertidae). Boll. Soc. Sarda Sci. Nat. **23**: 145-153.

Lehmann, A., Overton, J.McC., Leathwick, J.R. (2002): GRASP: generalized regression analysis and spatial prediction. Ecol. Model. **157**: 189-207.

Lehsten, V., Harmand, P. (2006): Null models for species cooccurrence patterns: assessing bias and minimum iteration number for the sequential swap. Ecography **29**: 786-792.

Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G., Williams, P.H. (2003): Avoiding pitfalls of using species distribution models in conservation planning. Conserv. Biol. **17**: 1591-1600.

Manel, S., Dias, J., Ormerod, S.J. (1999): Comparing discrimant analysis, neural networks and logistic regression for predicting species distributions: a case of study with Himalayan river bird. Ecol. Model. **120**: 337-347.

Manel, S., Williams, H.C., Ormerod, S.J. (2001): Evaluating presence-absence models in ecology: the need to account for prevalence. J. Appl. Ecol. **38**: 921-931.

Margules, C.R., Austin, M.P. (1991): Nature Conservation: Cost Effective Biological Survey and Data Analysis. Canberra, Australia, CSIRO.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W. (2008): Evaluation of consensus methods in predictive species distribution modelling. Divers. Distr. doi: 10.1111/j.1472-4642.2008.00491.x.

McCullagh, P., Nelder, J.A. (1989): Generalized Linear Models, 2nd Edition. London, Chapman and Hall.

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.

Mertens, R. (1932): Zur Verbreitung und Systematik einiger *Lacerta*-Formen der Appenninischen Halbinsel und der Tyrrhenischen Inselwet. Senckenbergiana **14**: 235-239.

Michelot, M. (1997): *Lacerta bedriagae*. In: Atlas of Amphibians and Reptiles in Europe, p. 234-235. Gasc, J.-P. et al., Eds, Paris, Societas Europaea Herpetologica and Muséum National d'Histoire Naturelle.

Nix, H.A. (1986): A biogeographic analysis of Australian Elapid snakes. In: Atlas of Australian Elapid Snakes, Aust. Flora Fauna Ser. 8, p. 4-15. Longmore, R., Ed.

Parra, J.L., Graham, C.C., Freile, J.F. (2004): Evaluating alternative data sets for ecological niche models of birds in the Andes. Ecography **27**: 350-360.

Pearce, J., Ferrier, S. (2000): An evaluation of alternative algorithms for fitting species distribution models developed using logistic regression. Ecol. Model. **128**: 127-147.

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.

Peracca, M.G. (1905): Nuove osservazioni intorno alla *Lacerta sardoa* Peracca della Sardegna. Boll. Mus. Zool. Anat. Comp. Regia Univ. Torino **519**: 1-10.

Peterson, A.T., Papes, M., Eaton, M. (2007): Transferability and model evaluation in ecological niche modelling: a comparison of GARP and Maxent. Ecography **30**: 550-560.

Petit, S., Chamberlain, D., Haysom, K., Pywell, R., Vickery, J., Warman, L., Allen, D., Firbank, L. (2003): Knowledgebased models for predicting species occurrence in arable conditions. Ecography **26**: 626-640.

Phillips, S.J., Dudík, M. (2008): Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography **31**: 161-175.

Phillips, S.J., Dudik, M., Schapire, R.E. (2004): A maximum entropy approach to species distribution modelling. In: Abstracts of the Twenty-First Century International Conference on Machine Learning.

Phillips, S.J., Anderson, R.P., Schapire, R.E. (2006): Maximum entropy modelling of species geographic distributions. Ecol. Model. **190**: 231-259.

Poggesi, M., Agnelli, P., Borri, M., Corti, C., Finotello, P.L., Lanza, B., Tosini, G. (1996): Erpetologia delle isole circumsarde. Biogeographia (N.S.) **18**: 583-618.

Puddu, E., Viarengo, M., Erminio, C. (1988): Animali di Sardegna – Gli Anfibi e i Rettili. Cagliari, Ed. Della Torre.

R Development Core Team (2008): A Language and En-
vironment for Statistical Computing. R Foundation
for Statistical Computing, available at http://www.R-
project.org.

Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C.,
Egoh, B., Mgidi, T. (2004): Mapping the potential
ranges of major plant invaders in South Africa, Lesotho
and Swaziland using climatic suitability. Divers. Distrib.
10: 475-484.

Sanchez-Cordero, V., Munguia, M., Townsend-Peterson,
A. (2004): GIS-based predictive biogeography in the
context of conservation. In: Frontiers of Biogeography:
New Directions in the Geography of Nature, p. 311-323.
Lomolino, M.V., Heaney, L.R., Eds, Sunderland, MA,
Sinauer Associates.

Schneider, B. (1984): *Lacerta bedriagae* Camerano, 1885 –
Tyrrhenische Gebirseidechse. In: Handbuch der Rep-
tilien und Amphibien Europas, p. 211-224. Böhme, W.,
Ed., Aula-Verlag 2.

Segurado, P., Araújo, M.B. (2004): An evaluation of meth-
ods for modelling species distributions. J. Biogeogr. **31**:
1555-1568.

Shao, G., Halpin, P.N. (1995): Climatic controls of eastern
North American coastal tree and shrub distributions. J.
Biogeogr. **22**: 1083-1089.

Stemmler, O. (1962): Ein beitrag zur biologie und systema-
tik von *Lacerta bedriagae*. Aquaria **9**: 5-12.

Stockwell, D., Peters, D. (1999): The GARP modelling
system: problems and solutions to automated spatial
prediction. Int. J. Geogr. Inf. Sci. **13**: 143-158.

Tape, T.G. (2006): Interpreting Diagnostic Tests. Med-
ical Center, University of Nebraska. Available at
http://gim.unmc.edu/dxtests.

Thuiller, W. (2003): BIOMOD – optimizing predictions
of species distributions and projecting potential future
shifts under global change. Global Change Biol. **9**: 1353-
1362.

Thuiller, W., Lavorel, S., Araújo, M.B. (2005): Niche prop-
erties and geographical extent as predictors of species
sensitivity to climate change. Global Ecol. Biogeogr. **14**:
347-357.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., Kadmon,
R. (2007): A comparative evaluation of presence only
methods for modelling species distribution. Divers. Dis-
trib. **13**: 397-405.

Vanhooydonck, B., Van Damme, R., Aerts, P. (2000): Eco-
morphological correlates of habitat partitioning in Cor-
sican lacertid lizards. Funct. Ecol. **14**: 358-368.

Verner, J., Morrison, M.L., Ralph, C.J. (1986): Wildlife
2000: Modelling Habitat Relationships of Terrestrial
Vertebrates. Madison, WI, University of Wisconsin
Press.

Williams, S.E., Bolitho, E.E., Fox, S. (2003): Climate
change in Australian tropical rainforests: an impending
environmental catastrophe. Proc. Roy. Soc. B **270**: 1887-
1892.

Zaniewski, A.E., Lehmann, A., Overton, J.M.C. (2002):
Predicting species spatial distributions using presence-
only data: a case study of native New Zealand ferns.
Ecol. Model **157**: 261-280.