

Modelling suitable areas for *Hyla meridionalis* under current and future hypothetical expansion scenarios

Neftalí Sillero

Abstract. New records of the frog *Hyla meridionalis* have been reported on the Iberian Central mountain range, although it is not clear if those records correspond to an expansion process or to a lack of sampling. The species' global distribution was modelled to identify possible areas of expansion in the current environment and in a near future environment. To verify the effects of modelling a dynamic situation, the hypothetical expansion process from south to north was simulated in the current scenario. The species records were partitioned in four different datasets and modelled iteratively over a study area enclosing all records: records only from North Africa, from North Africa and South Iberian Peninsula, from North Africa and all the Iberian Peninsula, and from the entire distribution. The complete dataset was also used to extrapolate the model to a current and future global environmental scenarios. The models were compared by subtraction per pairs and by Cohen's Kappa. No spatio-temporal trends were detected in species records; therefore, the hypothesis of a recent expansion is not supported. The species spread only to the Mediterranean part of the Iberian Peninsula when modelling only with African records, and to Europe when modelling also with Iberian records. The extrapolation models were moderately similar: large suitable areas were predicted in all the continents. Temperature variables contributed more to models. The species can expand to new suitable areas, although it is not possible to know if it would expand outside the biogeographical regions where the species is present.

Keywords: biogeography, ecological modelling, expansion, *Hyla meridionalis*, Iberia, North Africa.

Introduction

Hyla meridionalis (Anura, Hylidae) occurs in several countries of the Palearctic region (*sensu* Borkin, 1999; fig. 1): Morocco, Algeria and Tunisia in north-western Africa; Portugal, Spain, south and south-eastern France and north-western Italy in Europe, as well as Madeira, Canary Islands and Menorca (Balearic Islands) (Sillero, 2008 and references therein). In the Iberian Peninsula, its range is divided in three isolated nuclei (García-París, 1997; Tejedo and Reques, 2002): one in the central and south-western parts of the Peninsula, another in the north-eastern extreme of Spain (Catalonia region) and a third in the Basque Country (Spain). It is introduced in Madeira, Canary Islands and Menorca (Sillero, 2008).

During the last two decades, several works presented new records of *H. meridionalis* in the south-eastern part of the Salamanca province (Bueno, 1991; Merchán et al., 2004; García, 2007; Sillero and Carretero, 2007), and in the south-western part of the Madrid province (Aceituno, 2001; Martínez-Solano and Aceituno, 2001). This fact led Martínez-Solano and Fernández-González (2003) and Merchán et al. (2004) to suggest that the species might be in expansion in both provinces. The Salamanca records would be highly relevant under such a hypothesis, as they constitute the only population located northern to the Central mountain range, which is considered an effective geographic barrier (Merchán et al., 2004). This population is apparently isolated (Sillero and Carretero, 2007) and could have remained undiscovered due to the lack of intense sampling in the province of Salamanca. However, Martínez-Solano and Fernández-González (2003) rejected this possibility for Madrid, as it is considered to be well-sampled (García-París et al., 1989), and argued that a recent expan-

Centro de Investigação em Ciências Geo-Espaciais (CICGE), Universidade do Porto, Departamento de Matemática Aplicada, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal
e-mail: neftali.pablos@fc.up.pt

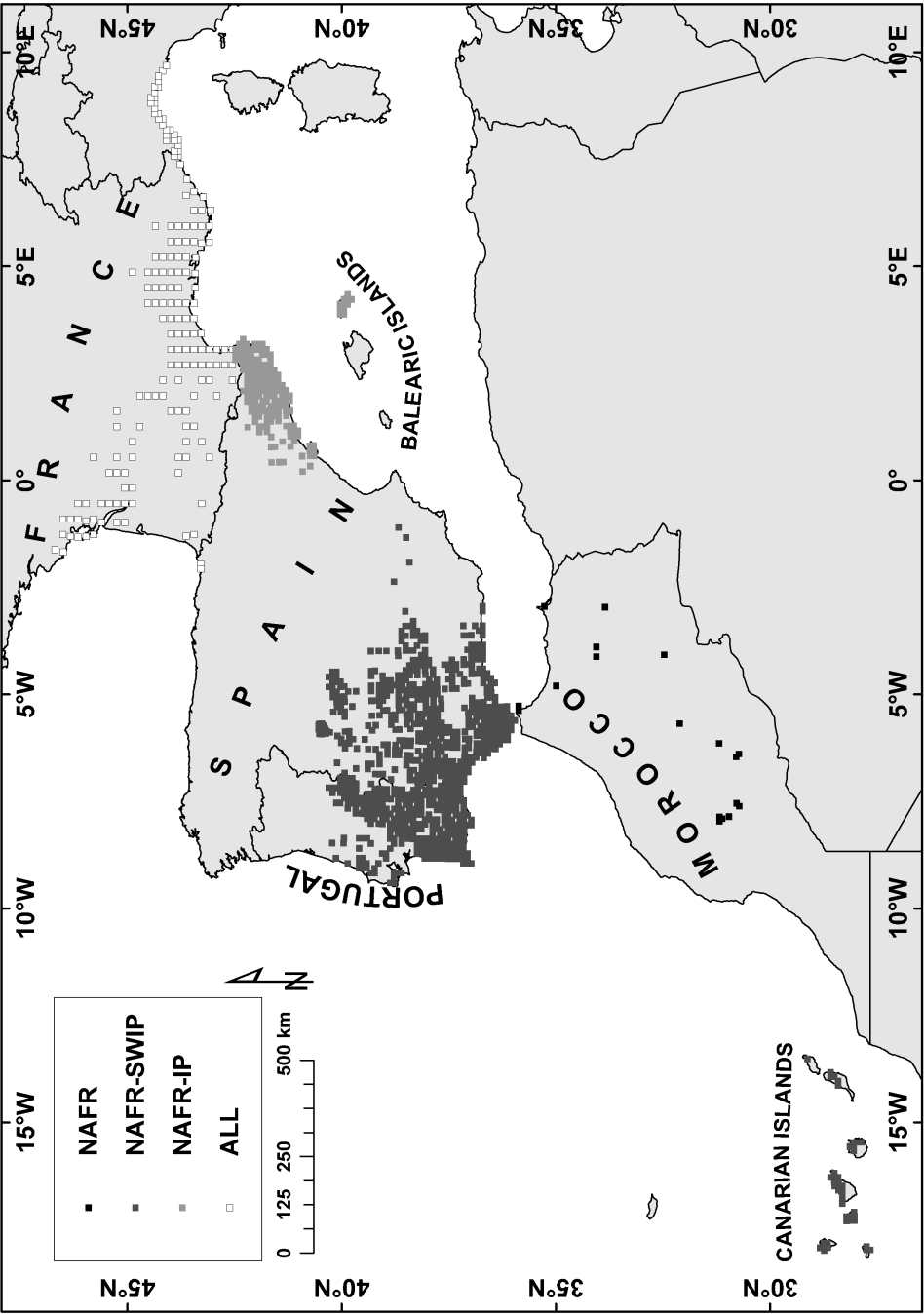


Figure 1. The global distribution of *Hyla meridionalis*. The NAFR model was calculated with records only from North Africa; the NAFR-SWIP with records from North Africa plus south-western Iberian Peninsula; the NAFR-IP with records from North Africa and all the Iberian Peninsula; and the ALL with records from the whole distribution. See more details in the Material and methods section.

sion of the species' range would be a more plausible explanation. Merchán et al. (2004) suggested that the global climatic warming is the most probable cause of the hypothetical expansion of *H. meridionalis*. Therefore, it is not clear if the species is currently in expansion or not. Additionally, Recuero et al. (2007) concluded using phylogenetic analyses that Iberian *H. meridionalis* have a double Moroccan origin and spread to the Iberian Peninsula from Morocco (see fig. 1 in their work) in two separate expansion events. The first, natural expansion took place from southern to northern Morocco and from there to the south of the Iberian Peninsula. Later, the population of northern Morocco became extinct and this area was then recolonized by a differentiated population from the south of Morocco. The second event begun with the human-mediated introduction of *H. meridionalis* from this differentiated north Moroccan population to the Mediterranean coast of France. From there, the species expanded eastward, southward and northward and was also introduced to the Canary Islands. Therefore, *Hyla meridionalis* may be expanding in three places of continental Europe: Salamanca, Madrid and north-eastern Iberia/south-eastern France.

Ecological modelling (Guisan and Zimmermann, 2000) is a useful tool for predicting and monitoring the expansion of indigenous or introduced species (Peterson and Vieglais, 2001), as well as the risk of invasion in specific areas (Thuiller et al., 2005), the effects of climatic change on species' ranges (Roura-Pascual et al., 2004), and the biogeographical relationships between species (Thuiller et al., 2006). Furthermore, the introduction of exotic species is a global conservation problem (Ricciardi, 2004), being considered one of the main causes for the amphibian worldwide decline (Cruz, Rebelo and Crespo, 2006).

The main objective of this study was to verify if there are new suitable habitats for the expansion of *Hyla meridionalis* outside its known range, over two study areas: an area enclosing all species records, and the whole world. There-

fore, the species' fundamental environmental niche (*sensu* Pearson, 2007) was modelled to verify if the species can expand or not, and which factors are driving the expansion. Specifically, two questions were addressed: (1) which suitable areas are not currently occupied by the species (and thus available for a hypothetical expansion) considering both study areas, and (2) which areas could be occupied in the future, considering a global climatic change scenario. Additionally, the usefulness of ecological modelling for predicting the distribution of a species in expansion was measured, as habitat suitability models are static and suppose that the species is in equilibrium with the environment (Guisan and Zimmermann, 2000). New records of an expanding species could strongly modify previous potential models and can be located in places classified by the model as unsuitable (Welk, 2004). Therefore, if the species was in expansion and not in equilibrium, the whole possible ecological range would not be represented in the records. To verify the effects of this problem on predicting models, the *H. meridionalis*' records were split in four datasets, following the groups described by García-París (1997), in order to simulate the hypothetical expansion process from south to north.

Material and methods

Study area

Models were calculated over a rectangle enclosing all species records (the Sahara area was excluded), and then, extrapolated to the whole world in a current and future environmental scenario. This strategy was used to avoid artificial high values in model outputs (such as AUC) and changes in the probability values towards upper and lower environmental boundaries, that can modify strongly the response curves describing the suitability of variables.

Species data

The distribution data of *H. meridionalis* were obtained from several distribution atlases (fig. 1): Morocco (15 records; Bons and Geniez, 1996), Portugal (977 records; Loureiro et al., 2008), Spain (3151 records; Pleguezuelos, Márquez and Lizana, 2002), France (136 records; Castanet and Guyétant, 1989) and Italy (29 records; Sindaco et al., 2006). There

are no records available for Algeria and Tunisia (Schleich, Kästle and Kabisch, 1996). All the records were located on a 10×10 km UTM grid.

Spatio-temporal trends

The data of collection was only available for records from Spain (Pleguezuelos, Márquez and Lizana, 2002) and Portugal (Loureiro et al., 2008). A map of records classified by year of collection was drawn in order to examine two possible spatio-temporal trends in the Iberian Peninsula: one expansion from south to north, and another from the Mediterranean coast of France south-west into the Iberian Peninsula.

Environmental data for current conditions

Ecogeographical variables (hereafter EGVs, table 1) for the current scenario included climatic and topographical data, obtained respectively from WorldClim series (Hijmans et al., 2005; <http://www.worldclim.org/>) and from the second version of the Shuttle Radar Topography Mission (SRTM; <http://www2.jpl.nasa.gov/srtm/>). These variables were clipped using both study areas (rectangular and whole world).

Environmental data for future scenario

EGVs for the future global scenario (table 1) were obtained from the WorldClim series (Hijmans et al., 2005; <http://www.worldclim.org/>). This set of EGVs was the same as used in the current scenario. The coupled atmosphere-ocean general circulation model HadCM3 (Hadley Centre Coupled Model, version 3) and the socio-economic scenario 'a2a' were used for a near future (2050), developed at the Hadley Centre in the UK (<http://www.metoffice.gov.uk/research/hadleycentre/>). A scenario is a coherent, internally consistent and plausible description of a possible future state of the world (IPCC-TGICA, 2007). It is not a forecast; rather, each scenario is one alternative image of how the future can unfold. A projection may serve as the raw material for a scenario, but scenarios often require additional information (e.g., about baseline conditions). A set of scenarios

Table 1. Description and units of the ecogeographical variables used for calculating models. The size of the pixel was 10 km^2 (see Material and methods section for details). WorldClim: <http://www.worldclim.org/>; SRTM: <http://www2.jpl.nasa.gov/srtm/>.

Variable	Origin	Units
Mean precipitation	WorldClim	mm
Maximum precipitation	WorldClim	mm
Minimum precipitation	WorldClim	mm
Mean maximum temperature	WorldClim	°C
Maximals of maximum temperature	WorldClim	°C
Mean minimum temperature	WorldClim	°C
Minimals of minimum temperature	WorldClim	°C
Altitude	SRTM	m

is often adopted to reflect, as well as possible, the range of uncertainty in projections. The scenario 'a2a' consists in a very heterogeneous world with continuously increasing global population and regionally oriented economic growth that is more fragmented and slower than in other story lines, without a posterior stabilization (IPCC-TGICA, 2007). This scenario has been frequently used in ecological modelling studies (e.g., Roura-Pascual et al., 2004; Araújo, Thuiller and Pearson, 2006). The same altitude data from the current global models were also used for the future global models, as topography is not expected to change in a near future.

Modelling procedure

To verify the capacity of predicting the distribution of a species hypothetically in expansion, i.e., in a dynamic and disequilibrium situation, the *H. meridionalis* records were split into four datasets, following the groups described by García-París (1997), in order to simulate the hypothetical expansion process (fig. 1): (1) records only from North Africa (NAFR); (2) records from North Africa and South Iberian Peninsula (NAFR-SWIP); (3) records from North Africa and all the Iberian Peninsula (NAFR-IP); and (4) records from the whole distribution (ALL). Thus, a model was produced for each sample, using always the rectangular study area and the current environmental conditions. Models calculated with ALL dataset were extrapolated over the whole world in both current and future global scenarios.

MaxEnt model

The behaviour and activity patterns of *H. meridionalis* (Sillero, 2008) complicate the accurate determination of absences in a given square. Also, as the species could be in expansion, the identification of new suitable areas could be hampered by using absence records (that is, if its realized environmental niche was calculated; Pearson, 2007). Therefore, a Maximum Entropy model was used to identify the environmental factors related to the distribution and to locate the areas of probable occurrence. This is a general-purpose machine learning method that uses presence-only occurrence data (Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006; Phillips and Dudik, 2008), and thus calculates the fundamental environmental species niche. Models were developed with MaxEnt 3.2.19 software (<http://www.cs.princeton.edu/~schapire/maxent/>), which is particularly well suited to noisy or sparse information, as is typical of species occurrence data, and capable of dealing with continuous and categorical EGVs at the same time (Phillips, Anderson and Schapire, 2006). MaxEnt models consistently outperform more established methods, either presence-only methods (Bioclim, Domain) or presence-absence methods (GAM, GLM), especially when samples sizes are low (Elith et al., 2006; Hernandez et al., 2006).

Essentially, Maximum Entropy models are based on a simple reasoning: when characterizing some unknown events with a statistical model, the one that has maximum entropy should be chosen, i.e., the one that produces the most uniform distribution but still infers as accurately as

possible the observed data (e.g., maximize entropy for a given chi-squared value). MaxEnt estimates the range of a species with the constraint that the expected value of each EGV (or its transform and/or interactions) should match its empirical average, i.e., the average value for a set of sample points taken from the species-target distribution (Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006). MaxEnt randomly selects uniformly distributed absence data from the background squares. Hence, a set of ten models was calculated over the rectangular study area through an iterative process (Araújo and New, 2007; Martínez-Freiría et al., 2008; Sillero, in press) for each of the datasets. MaxEnt weights each EGV by a constant and the estimated probability distribution is exponential in the sum of the weighed features, divided by a scaling constant to ensure that the probability values range from 0.0 to 1.0 and sum to 1.0. The program starts with a uniform probability distribution (gain = 0) and iteratively alters one weight at a time to maximize the likelihood of the occurrence data set. The gain is a measure of the likelihood of the samples. The algorithm converges to the optimum probable distribution, thus, the gain can be interpreted as representing how much better the distribution fits the sample points than the uniform distribution does (Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006; Phillips and Dudik, 2008).

MaxEnt was run randomly selecting 75% of the presence records as training data and 25% as test data. The ten iterative models were run with autofeatures and tested with receiver operated characteristics (ROC) plots. The area under the curve (AUC) of the ROC plot was taken as a measure of the overall fit of the current and future models (Liu et al., 2005). AUC was selected because it is independent of prevalence (the proportion of presence in relation with the total dataset size) as assessed by its mathematical definition (Bradley, 1997; Forman and Cohen, 2005; Fawcett, 2006). ROC values are calculated using proportions (sensitivity vs. 1-specificity), corresponding each of them to a different column in the confusion matrix, i.e., both rates of presence and absence records are included. The importance of each EGV for explaining the species distribution was determined by: (1) jackknife analysis of the average AUC with training and test data; and (2) average percentage contribution of each EGV to the models. For that, EGVs were excluded in turn and a model created with the remaining variables; then a model was created using each individual variable. When extrapolating the models to other areas, MaxEnt identifies which areas in the environment space are outside the range represented by the training data, and represents them in the clamping map.

Ensemble forecasting

The Maximum Entropy model classifies squares with a continuous value of probability of occurrence between 0 and 1. Traditionally, a threshold is estimated to reclassify probability of occurrence maps into areas of probable presence/absence (e.g., Liu et al., 2005). However, the issue of selecting an optimized threshold with MaxEnt and selecting the 'best' model from a model ensemble is rather

subjective (Phillips, Dudik and Schapire, 2004; Phillips, Anderson and Schapire, 2006; Phillips and Dudik, 2008). Therefore, instead of picking the 'best' model from an ensemble, a more promising approach is to explore the resulting range of projections, a procedure known as ensemble forecasting (Araújo and New, 2007; Martínez-Freiría et al., 2008; Sillero, in press). With this method, the arithmetic average and the standard deviation of the ten iterative models were calculated (and the ALL and Future global extrapolation models), as well as the mean of the 'Balance training omission, predicted area and threshold value logistic threshold' (Phillips and Dudik, 2008).

Model comparisons

The ensemble models were compared in two ways:

- (1) The habitat suitability models (HSM) produced were subtracted per pairs. The pixels of the models were considered as similar when the absolute value of the subtraction was equal to 0, and dissimilar when the absolute value of the subtraction was different than 0.
- (2) The HSM were transformed into presence/absence models using the mean of the 'Balance training omission, predicted area and threshold value logistic' as threshold value. These potential models were compared per pairs using the Cohen's Kappa statistical measure of association (Cohen, 1960; Liu et al., 2005). The results of Kappa are usually classified as *Very poor* (0.0-0.2), *Poor* (0.2-0.4), *Regular* (0.4-0.6), *Good* (0.6-0.8) and *Very good* (0.8-1.0). The number of equal pixels in the suitable areas among current presence/absence models, i.e. those pixels with equal values (0 or 1) in both models, were also calculated.

Results

Spatio-temporal trends

No spatio-temporal trends were detected in *H. meridionalis*' records (fig. 2). The observed patterns did not support a hypothetical expansion from the South of the Iberian Peninsula or from the Mediterranean coast.

Current environmental scenario models

Four sets with 10 iterative models were calculated for a current environmental scenario over the rectangular study area using the four different presence records' datasets (NAFR, NAFR-SWIP, NAFR-IP, and ALL). All current global models obtained an AUC higher than 0.90:

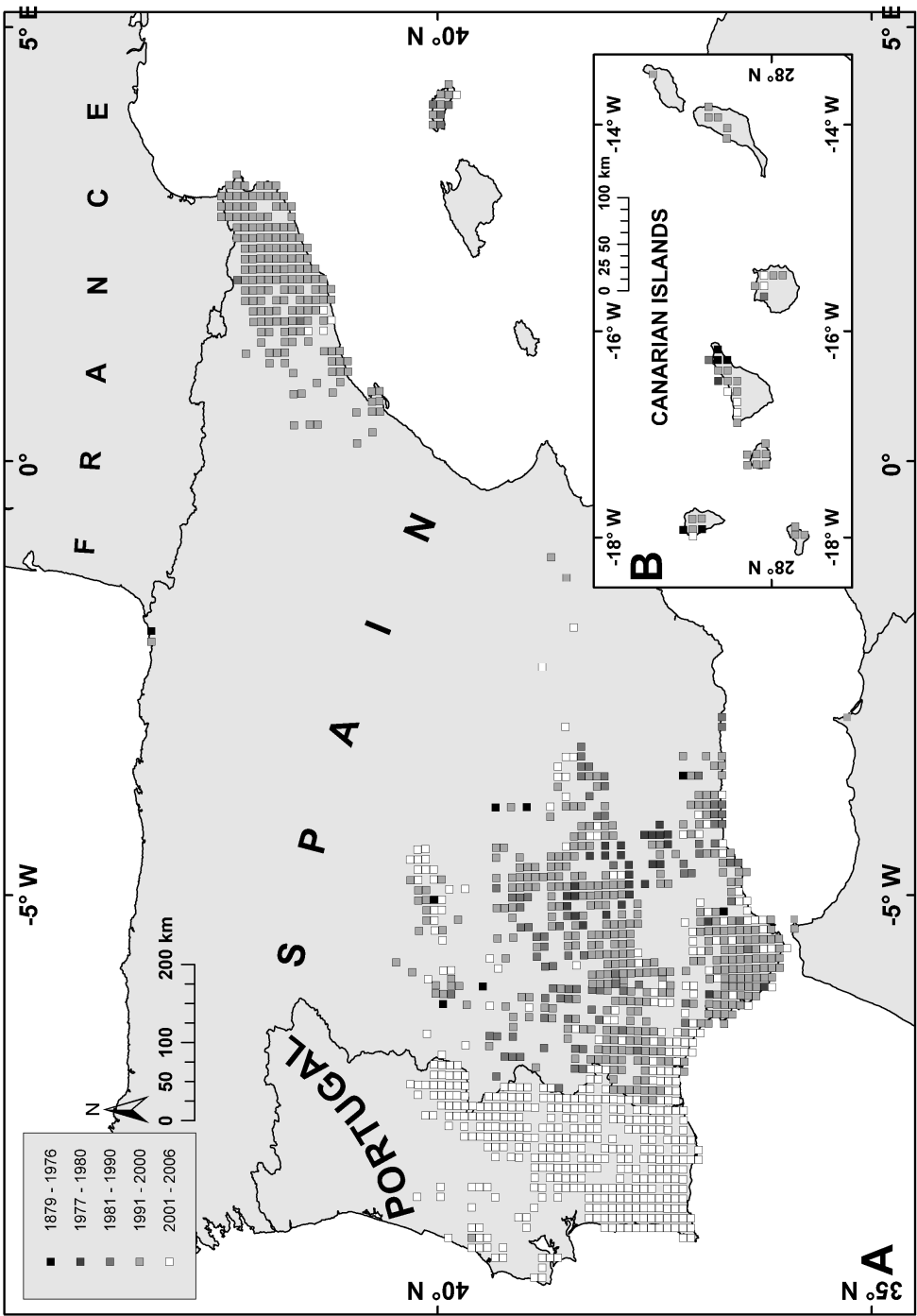


Figure 2. Spatio-temporal trends in the records of *H. meridionalis*. Records with collection date are only available for Spain and Portugal. The trends did not show a hypothetical expansion from the south of the Iberian Peninsula or from the Mediterranean coast.

Table 2. Suitable and unsuitable areas for models (measured as km²) calculated over the rectangular study area. The habitat suitability models were transformed into presence (suitable areas)/absence (unsuitable areas) models using the mean of the 'Balance training omission, predicted area and threshold value logistic' as threshold value, and compared per pairs. For abbreviations see fig. 1.

Model	Threshold	Suitable	Unsuitable
NAFR	0.0598 ± 0.0158	27 510	14 022
NAFR-SWIP	0.0396 ± 0.0102	31 515	10 017
NAFR-IP	0.0653 ± 0.0055	28 054	13 478
ALL	0.0747 ± 0.0055	24 604	16 928

NAFR, 0.958 ± 0.012 ; NAFR-SWIP, 0.966 ± 0.001 ; NAFR-IP, 0.951 ± 0.001 ; and ALL, 0.934 ± 0.002 . The largest suitable (presence) area (in number of pixels) corresponded to the ALL model; the NAFR and NAFR-IP models had similar suitable areas, meanwhile the NAFR-SWIP model had the smallest suitable area (see table 2). The suitable areas of the NAFR model (fig. 3A) occupied the mountainous regions of North Africa and the south-western part of the Iberian Peninsula. NAFR-SWIP model (fig. 3B) had similar suitable areas, but more restricted in North Africa. These areas maintained in NAFR-IP model (fig. 3C), but it included also some areas in the north of the Iberian Peninsula, in the western coast of France and in the Mediterranean coast of Spain, France and Italy. The ALL model (fig. 3D) presented more suitable areas, mainly in the south-western part of France and northern part of Spain. The islands of Corsica and Sardinia appeared as suitable in all four models, although the degree of suitability grew towards ALL model. This trend was apparent also in the Iberian Peninsula and France. In the opposite, mountainous regions of North Africa lost suitability in the ALL model.

The EGVs with a larger percentage contribution to the models were not similar in all models (fig. 4). In NAFR model, minimum precipitation and mean minimum temperature were the EGV with the largest contribution. In NAFR-SWIP model, these variables were the minimum and mean precipitation as well as the mean maximum temperature. The NAFR-

IP and ALL models presented similar results: the largest contribution corresponded to mean precipitation, followed by mean and minimum of minimum temperature, mean maximum temperature, and minimum and maximum precipitation; altitude and maximum of maximum temperature had the lowest contribution. The mean percentage contribution of models without one variable were quite similar among variables, i.e., variable removal did not strongly affect the models, and the shape of this trend is maintained in all four datasets (fig. 5). However, the importance of the contribution to models decreased from NAFR to ALL models, with the exception of NAFR dataset in AUC values (fig. 5C). When the models were calculated only with one variable, the contribution varied strongly among EGVs (the mean maximum temperature and the maximum precipitation were the less important variables), but again trends were similar in all datasets with the exception of NAFR dataset (fig. 5).

Model comparisons

The number of equal pixels among pairs of HSM and presence/absence (suitable/unsuitable areas) models decreased from the comparisons with the NAFR model to the comparisons with the ALL models, i.e., the most different models were the NAFR and ALL, and the most similar were the NAFR-IP and ALL (table 3). The values of Kappa decreased in a similar way: i.e., the most different models were the NAFR and ALL, and the most similar were the NAFR-IP and ALL (table 3).

Extrapolation models

The extrapolation models under the current and future global environmental scenarios were moderately similar (Kappa = 0.4240). The number of equal pixels between the ALL and future models was 568 354 (24.95%), and the differences in the suitable areas between the ALL and future models were out of 638 891 pixels. The current global extrapolation model

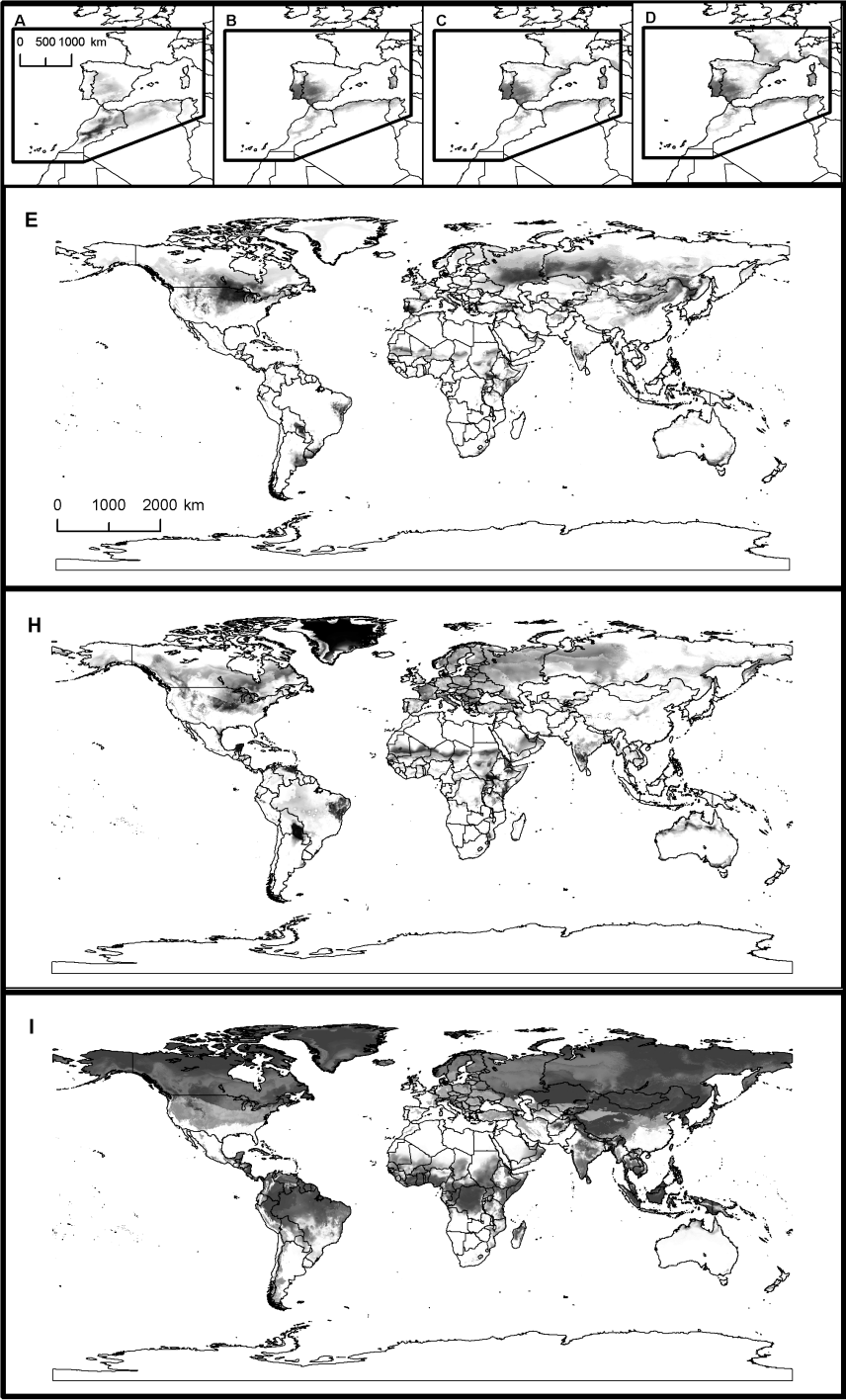


Figure 3. MaxEnt models of *H. meridionalis* calculated with records from NAFR (Map A), NAFR-SWIP (Map B), NAFR-IP (Map C) and ALL (Map D) datasets. Maps E and H represent the extrapolations of ALL to the current and future global scenario. Map I represent clamping values for future model. Clamping identifies those areas in the environment space with values outside the range represented by the training data. The suitability index increases from white (0) to grey colours (1). Clamping values increases from white (0) to grey colours (1). See abbreviations in fig. 1.

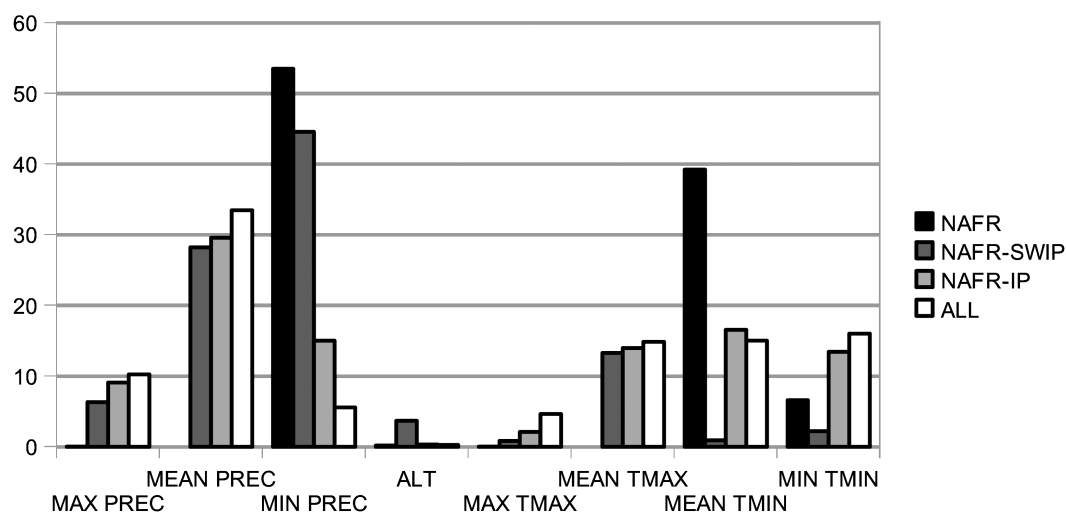


Figure 4. Variable mean percentage contribution for explaining the variance of each model. MEAN PREC: Mean precipitation; MAX PREC: Maximum precipitation; MIN PREC: Minimum precipitation; MEAN TMAX: Mean maximum temperature; MAX TMAX: Maximals of maximum temperature; MEAN TMIN: Mean minimum temperature; MIN TMIN: Minimals of minimum temperature; ALT: Altitude. See abbreviations for model names in fig. 1.

(fig. 3E) presented as suitable areas large extensions on central Europe and Asia, as well as sub-Saharan Africa (Sahel region) and central North America. Some regions of South America are also suitable: Paraguay/Uruguay, and the eastern extreme of Brazil. The suitable areas in the future global extrapolation model (fig. 3H) for the year 2050 were moderately similar to the current model, but those areas had a larger extension, specially in North and South America, Africa and Oceania. However, clamping values (fig. 3I), either for current and future models, were high in these regions, suggesting that their prediction reliability were low.

Discussion

The current and future global extrapolation models identified larger extensions of suitable habitats for *Hyla meridionalis* in Europe, Asia, Africa and America. Apparently, the species had many potential areas for expansion. Notwithstanding, MaxEnt's clamping results showed that most of the predicted areas are not reliable. In fact, clamping values were particularly high in areas such as Greenland, South

Argentina and Chile, which were surprisingly predicted as highly suitable. Clamping identified those values outside the values represented in the species records. This could be due to the difficulties of the extrapolation process, that applies the resulting model to the whole area instead of modelling again. The model should identify areas with a Mediterranean climate, the biogeographical region where the species is present (Sillero et al., 2009). Instead, the model identified areas moderately arid. However, these results may confirm the hypothesis by Merchán et al. (2004), where climatic change is responsible for the supposed expansion of *H. meridionalis*. In fact, some variables with a higher influence in the species distribution were temperature related. The altitude was in almost all cases the variable with the lowest contribution. The species lives in Mediterranean areas, independently of the altitude (there are records from sea level to mountain tops), avoiding regions too humid and cold (Sillero, 2008).

The geographic sequence observed in the predicted suitable/unsuitable areas partially support a hypothetical expansion process. When modelling only with African records, the species was predicted to be present exclusively in the

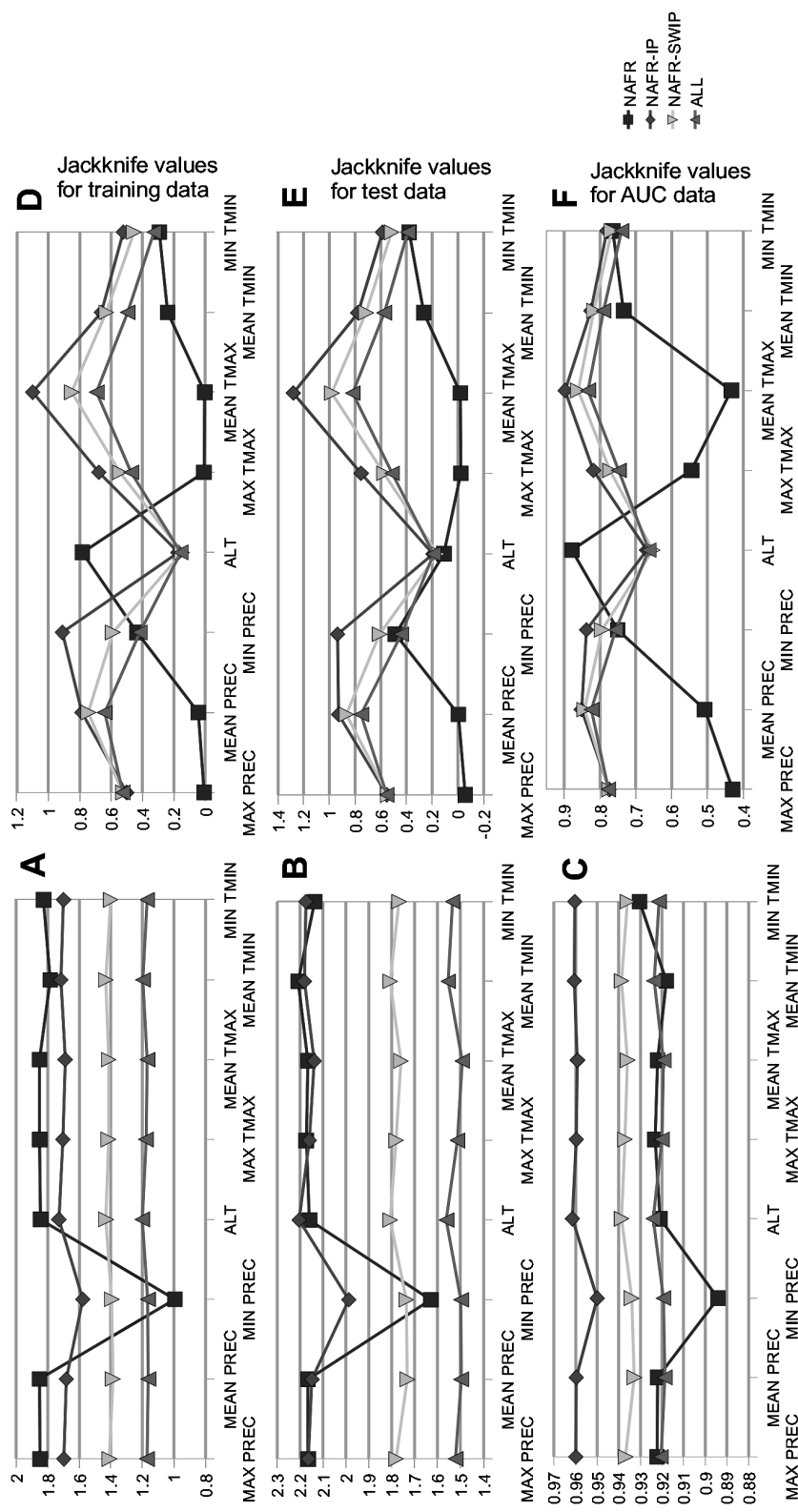


Figure 5. Jackknife results of the contribution of variables to each model for the training (A and D) and test (B and E) records, as well as AUC values (C and F). The first column corresponds to models calculated without one variable. The second column corresponds to models calculated only with one variable. See more details in Material and methods. See abbreviations for model names in fig. 1 and for variables in fig. 4.

Table 3. In each row, the *first lines* represent number of equal pixels among current models (measured as km²), calculated as the absolute difference equal to zero between each pair of models; the *second lines* represent equal pixels in the suitable areas (measured as km²) among current presence/absence models, i.e., those pixels with equal values (0 or 1) in both models; and the *third lines* represent Cohen's Kappa statistical measure of association results of the comparisons of current models. The results of Kappa are usually classified as *Very poor* (0.0-0.2), *Poor* (0.2-0.4), *Regular* (0.4-0.6), *Good* (0.6-0.8) and *Very good* (0.8-1.0). Values decreased from top to bottom, and from right to left. For abbreviations see fig. 1.

Comparison	NAFR	NAFR-SWIP	NAFR-IP
NAFR-SWIP	30 894		
	33 663		
	0.5445		
NAFR-IP	27 221	35 937	
	27 424	35 179	
	0.2604	0.6262	
ALL	24 038	32 440	36 903
	25 292	32 027	37 540
	0.1680	0.4939	0.7944

Mediterranean part of the Iberian Peninsula (*sensu* Sillero et al., 2009). The species spread through the Atlantic part of the Iberian Peninsula and Europe when the records from North Iberia were included in the models (NAFR-IP and ALL models). Even with the records from south-western Iberia, the species spread only to the Mediterranean region of Europe, outside the Iberian Peninsula (*sensu* Rivas-Martínez, Penas and Díaz, 2004). Furthermore, the NAFR-IP and ALL models were very similar (Kappa = 0.79); in fact, they were the most similar models: the number of similar pixels, as well as the Kappa values (tables 3 and 4), decreased from the comparisons with the ALL model to the comparisons with the NAFR model. However, the analysis of species records does not identify any spatio-temporal trends and do not support the hypothesis of a recent expansion.

The model results may indicate that a species can be modelled, independently of the sample size, whenever the species records are inside the same biogeographical area. This is a logical result as models suppose that the species is in equilibrium with the environment, and

the records represent all the ecological range (Guisan and Zimmermann, 2000). However, this could be a problem when modelling a species in expansion. Only if the species expands inside a particular biogeographical region, new potential areas would be identified in that region, and thus, the prediction would be considered valid. On the other hand, new potential areas outside that biogeographical region would be identified exclusively when the species expands there. In other words, it is only possible to identify new potential areas when those habitats were included in the species' distribution range (Beaumont et al., 2009).

The potential areas outside the Mediterranean basin can only be considered as potential areas of accidental introduction, as there is no geographical connection among them and the Mediterranean basin. This is a potential problem that would have not been detected if models were not extrapolated to the whole world. Many species present in a region without signs of expansion, may become invasive if introduced outside of their natural range (Richardson and Thuiller, 2007). In some extreme cases, some species can perform better in the invaded range than in the autochthonous ones (Ebeling, Hensen and Auge, 2008). Since the expansion of exotic, invasive species constitutes one of the main threats to global amphibian biodiversity (Ricciardi, 2004), studies predicting potential occurrence areas outside the species' natural range may alarm against human-mediated introductions to such areas and enhance amphibian conservation.

All the models obtained an AUC value higher than 0.90. Recently, a debate is ongoing about the reliability of the use of ROC curves in ecological modelling. Some authors maintain that ROC is not independent of prevalence (Wiley et al., 2003; Jiménez-Valverde et al., 2008; Lobo, Jiménez-Valverde and Real, 2008). However, these papers proved in fact that specialized species are easier to model (Seoane et al., 2005; Hernandez et al., 2006; Sillero et al., 2009) as they modify the species' sample size

instead the extension of the study area. The specialist condition is a relative value: it depends on the environmental width of the study area, which is related with the study area size. Furthermore, several studies proved that prevalence and sample size can affect model results, but not AUC (McPherson, Jetz and Rogers, 2004; Raes and ter Steege, 2007; Albert and Thuiller, 2008). In the case of *H. meridionalis*, AUC values remained similar whenever the sample size changed: for instance, ALL model presented the lowest value.

Recuero et al. (2007) considered the northern Iberia and southern France populations as introduced by humans from northern Morocco. The models showed that the expansion is possible from these areas. However, it is not possible to visualize the hypothetical expansion process in the records, as no spatio-temporal trends were detected (Welk, 2004). This result is a consequence of the recent age of the herpetological Iberian atlases (Pleguezuelos, Márquez and Lizana, 2002; Loureiro et al., 2008). Recuero et al. (2007) cannot determine neither the date nor the point of origin, natural or human-mediated, of northern Iberia-southern France populations with their data. They suggest as probable starting point some locality in the Mediterranean coast of France (in coincidence with the starting point of the recent introduction of *Discoglossus pictus* in western Europe; Llorente et al., 2002). From there, it has spread eastward, westward and southward. Further northward and eastward expansion may have been prevented by the absence of ecologically favourable habitats and/or the presence of *H. arborea* and *H. intermedia*, parapatric to *H. meridionalis* in the zone. However, *H. meridionalis* is syntopic with *H. arborea* in a large part of its range, and no competition problems have been described until now (Sillero, 2008). The origin of the introduced populations could be any coastal point in Morocco (Recuero et al., 2007). A similar introduction situation has been recently described for *Tarentola mauritanica* (Harris et al., 2004): the single haplotype that characterizes popu-

lations across Spain, Portugal, Italy, Menorca, Crete, and Tunisia suggests introduction by humans followed by rapid expansion throughout southern Europe. A detailed population genetics' study, combined with landscape genetic tools, may be able to shed light to the starting points and timing of the introduction and expansion processes hypothesized for *H. meridionalis* in the Iberian Peninsula.

Conclusions

Ecological modelling could be a useful tool for identifying new suitable areas for a species in expansion, but only inside those biogeographical regions where the species is present, and independently of the record number. This is very important when determining the study area for modelling a species, which should include all the biogeographical regions with species presences. Further studies are needed to monitor the *H. meridionalis* populations in Salamanca and Madrid, as well as north-eastern Iberia/south-eastern France, in order to verify the hypothetical expansion process; and to study the influence and competence of *H. arborea* in the distribution range of *H. meridionalis*.

Acknowledgements. Thanks to A. Kaliontzopoulou and F. Martínez-Freiría for improving an earlier version of the manuscript, and to M.A. Carretero for ideas and help. I am supported by a post-doctoral grant (SFRH/BPD/26666/2006) from FCT. This work is dedicated to the memory of Daniel Omolo.

References

- Aceituno, J. (2001): *Hyla meridionalis* (Ranita meridional), presencia en suroeste de Madrid y nuevas citas en Toledo. Bol. AHE **12**: 9-10.
- Albert, C.H., Thuiller, W. (2008): Favourability functions versus probability of presence: advantages and misuses. *Ecography* **31**: 417-422.
- Araújo, M.B., New, M. (2007): Ensemble forecasting of species distributions. *TREE* **22**: 42-47.
- Araújo, M.B., Thuiller, W., Pearson, R.G. (2006): Climate warming and the decline of amphibians and reptiles in Europe. *J. Biog.* **33**: 1712-1728.

- Beaumont, L., Gallagher, R., Thuiller, W., Downey, P., Leishman, M., Hughes, L. (2009): Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions* **15**: 409-420.
- Bons, J., Geniez, P. (1996): *Anfibios y Reptiles de Marruecos (Incluido Sáhara Occidental)*. Atlas biogeográfico, Monografías Herpetológicas. Valencia, Asociación Herpetológica Española.
- Borkin, L.J. (1999): Distribution of amphibians in North Africa, Europe, Western Asia, and the former Soviet Union. In: *Patterns of Distribution of Amphibians: a Global Perspective*, p. 329-420. Duellman, W.E., Ed., Maryland, The Johns Hopkins University Press.
- Bradley, A.P. (1997): The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Patt. Recog.* **30**: 1145-1159.
- Bueno, R. (1991): Hallazgo de Ranita Meridional (*Hyla meridionalis* Boettger, 1874) en la provincia de Salamanca. *Bol. AHE* **2**: 15.
- Castanet, J., Guyétant, R. (1989): *Atlas de répartition des Amphibiens et Reptiles de France*. Paris, Société Herpétologique de France.
- Cohen, J.A. (1960): A coefficient of agreement for nominal scales. *Edu. Psychol. Measur.* **20**: 37-46.
- Cruz, M.J., Rebelo, R., Crespo, E.G. (2006): Effects of an introduced crayfish, *Procambarus clarkii*, on the distribution of south-western Iberian amphibians in their breeding habitats. *Ecography* **29**: 329-338.
- Ebeling, S.K., Hensen, I., Auge, H. (2008): The invasive shrub *Buddleja davidii* performs better in its introduced range. *Div. Distr.* **14**: 225-233.
- Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Townsend Peterson, A., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R., Soberon, J., Williams, S., Wisz, M., Zimmermann, N.E. (2006): Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129-151.
- Fawcett, T. (2006): An introduction to ROC analysis. *Patt. Recog. Lett.* **27**: 861-874.
- Forman, G., Cohen, I. (2005): Beware the Null Hypothesis: Critical Value Tables for Evaluating Classifiers. *ECML* **2005**: 133-145.
- García, P. (2007): Nueva población de *Hyla meridionalis* Boettger, 1874 en la provincia de Salamanca. *Bol. AHE* **18**: 37-39.
- García-Paris, M. (1997): *Hyla meridionalis*. In: *Atlas of Amphibians and Reptiles in Europe*, p. 126-127. Gasc J.P. et al., Eds, Paris, Societas Europaea Herpetologica & Muséum National d'Histoire Naturelle (IEGB/SPN).
- García-Paris, M., Martín, C., Dorda, J., Esteban, M. (1989): Atlas provisional de los anfibios y reptiles de Madrid. *Rev. Esp. Herp.* **3** (2): 237-257.
- Guisan, A., Zimmermann, N.E. (2000): Predictive habitat distribution models in ecology. *Ecol. Mod.* **135**: 147-186.
- Harris, J.D., Batista, V., Lymberakis, P., Carretero, M.A. (2004): Complex estimates of evolutionary relationships in *Tarentola mauritanica* (Reptilia: Gekkonidae) derived from mitochondrial DNA sequences. *Mol. Phylog. Evol.* **30**: 855-859.
- Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L. (2006): The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* **29**: 773-785.
- Hijmans, R.J., Cameron, E., Parra, J.L., Jones, P.G., Jarvis, A. (2005): Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climat.* **25**: 1965-1978.
- IPCC-TGICA (2007): General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment. Version 2. Prepared by T.R. Carter on behalf of the Intergovernmental Panel on Climate Change. Task Group on Data and Scenario Support for Impact and Climate Assessment, 66 p.
- Jiménez-Valverde, A., Jorge, M., Lobo, J.M., Hortal, J. (2008): Not as good as they seem: the importance of concepts in species distribution modelling. *Div. Distr.* **14**: 885-890.
- Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G. (2005): Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**: 385-393.
- Llorente, G.A., Montori, A., Santos, X., Carretero, M.A. (2002): *Discoglossus pictus*. In: *Atlas de distribución y Libro Rojo de los Anfibios y Reptiles de España* (2ª impresión), p. 91-93. Pleguezuelos, J.M., Márquez, R., Lizana, M., Eds, Madrid, Dirección General de Conservación de la Naturaleza-Asociación Herpetológica Española.
- Lobo, J.M., Jiménez-Valverde, A., Real, R. (2008): AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biog.* **17**: 145-151.
- Loureiro, A., Ferrand, N., Carretero, M.A., Paulo, O. (2008): *Atlas dos Anfíbios e Répteis de Portugal continental*, Lisboa, Instituto de Conservação da Natureza.
- Martínez-Freiria, F., Sillero, N., Lizana, M., Brito, J.C. (2008): GIS-based niche models identify environmental correlates sustaining a contact zone between three species of European vipers. *Div. Distr.* **14**: 452-461.
- Martínez-Solano, I., Aceituno, J. (2001): Primeras citas de ranita meridional en la Comunidad de Madrid. *Quercus* **184**: 36.
- Martínez-Solano, I., Fernández-González, J.E. (2003): La colección de anfibios de Madrid del Museo Nacional de Historia Natural y su utilidad en conservación. *Graellsia* **59**: 108-159.
- 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.
- Merchán, T., Sillero, N., Lizana, M., Fontana, F. (2004): Nuevos hallazgos de la ranita meridional (*Hyla meridionalis* Boettger, 1874) en la provincia de Salamanca. *Bol. AHE* **15**: 81-85.

- Pearson, R.G. (2007): Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. (<http://ncep.amnh.org/>: accessed December 18 2008).
- Peterson, A.T., Vieglais, D.A. (2001): Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience* **51**: 363-371.
- Phillips, S.J., Dudík, M. (2008): Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**: 161-175.
- Phillips, S.J., Dudík, M., Schapire, R.E. (2004): A maximum entropy approach to species distribution modeling. *Proc. XXI Int. Conf. Mach. Learn.*: 655-662.
- Phillips, S.J., Anderson, R.P., Schapire, R.E. (2006): Maximum entropy modeling of species geographic distributions. *Ecol. Mod.* **190**: 231-259.
- Pleguezuelos, J.M., Márquez, R., Lizana, M. (2002): Atlas de distribución y Libro Rojo de los Anfibios y Reptiles de España, 2ª impresión. Madrid, Dirección General de Conservación de la Naturaleza-Asociación Herpetológica Española.
- Raes, N., ter Steege, H. (2007): A null-model for significance testing of presence-only species distribution models. *Ecography* **30**: 727-736.
- Recuero, E., Iraola, A., Rubio, X., Machordom, A., Garcia-Paris, M. (2007): Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern. *J. Biog.* **34**: 1207-1219.
- Ricciardi, A. (2004): Assessing species invasions as a cause of extinction. *TREE* **19**: 619.
- Richardson, D.M., Thuiller, W. (2007): Home away from home-objective mapping of high-risk source areas for plant introductions. *Div. Distr.* **13**: 299-312.
- Rivas-Martínez, S., Penas, A., Dfaz, T.E. (2004): Biogeographic Map of Europe, Cartographic Service. University of León, Spain.
- Roura-Pascual, N., Suárez, A.V., Gómez, C., Pons, P., Touyama, Y., Wild, A.L., Peterson, A.T. (2004): Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proc. R. Soc. B: Biol. Sci.* **271**: 2527-2534.
- Schleich, H.H., Kästle, W., Kabisch, K. (1996): Amphibians and Reptiles of North Africa. Koeltz Scientific Books, 630 p.
- Seoane, J., Carrascal, L.M., Alonso, C.L., Palomino, D. (2005): Species-specific traits associated to prediction errors in bird habitat suitability modelling. *Ecol. Mod.* **185**: 299-308.
- Sillero, N. (2008): Ranita meridional – *Hyla meridionalis*. In: Enciclopedia Virtual de los Vertebrados Españoles, <http://www.vertebradosibericos.org/edn>. Carrascal, L.M., Salvador, A., Eds, Madrid, Museo Nacional de Ciencias Naturales.
- Sillero, N. (in press): Potential distribution of the new populations of *Hyla meridionalis* in Salamanca (Spain). *Acta Herpetologica*.
- Sillero, N., Carretero, M.A. (2007): A systematic survey on the extralimital populations of *Hyla meridionalis* in Salamanca (Spain). *Bol. AHE* **18**: 59-64.
- Sillero, N., Brito, J.C., Toxopeus, B., Skidmore, A.K. (2009): Biogeographical analysis of amphibians and reptiles from the Iberian Peninsula. *Amp. Rep.* **30**: 185-206.
- Sindaco, R., Doria, G., Razzeti, E., Bernini, F.E. (2006): Atlanti degli Anfibi e dei Rettili d'Italia/Atlas of Italian Amphibians and Reptiles. Firenze, Societas Herpetologica Italica, Editorial Polistampa.
- Tejedo, M., Reques, R. (2002): *Hyla meridionalis*. In: Atlas de distribución y Libro Rojo de los Anfibios y Reptiles de España (2ª impresión), p. 117-119. Pleguezuelos, J.M., Márquez, R., Lizana, M., Eds, Madrid, Dirección General de Conservación de la Naturaleza-Asociación Herpetológica Española.
- Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O., Rouget, M. (2005): Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Glob. Change Biol.* **11**: 2234-2250.
- Thuiller, W., Midgley, G.F., Rougeti, M., Cowling, R. (2006): Predicting patterns of plant species richness in megadiverse South Africa. *Ecography* **29**: 733-744.
- Welk, E. (2004): Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: Purple loosestrife (*Lythrum salicaria*) in North America. *Ecol. Mod.* **179**: 551-567.
- Wiley, E.O., McNyset, K.M., Peterson, A.T., Robins, C.R., Stewart, A.M. (2003): Niche modeling perspective on geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* **16**: 120-127.

Received: August 13, 2008. Accepted: June 30, 2009.