Comparison of approaches to combine species distribution models based on different sets of predictors

David Romero[1,*], Jesús Olivero[1], José Carlos Brito[2], Raimundo Real[1]

[1]Grupo de Biogeografía, Diversidad y Conservación, Departamento de Biología Animal, Universidad de Málaga, Campus de Teatinos, 29071 Málaga, Spain.

[2]CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, Instituto de Ciências Agrárias de Vairão, R. Padre Armando Quintas, 4485-661 Vairão, Portugal.

[*]Author for correspondence, e-mail: davidrp@uma.es

Keywords: Akaike Weight Averaging, Bayesian Integration, Ensemble Forecasting, Favourability Function, Fuzzy Logic, Model Updating, Stepwise Selection.

Distribution models should take into account the different limiting factors that simultaneously influence species ranges. Species distribution models built with different explanatory variables can be combined into more comprehensive ones, but the resulting models should maximize complementarity and avoid redundancy. Our aim was to compare the different methods available for combining species distribution models. We modelled 19 threatened vertebrate species in mainland Spain, producing models according to three individual explanatory factors: spatial constraints, topography and climate, and human influence. We used five approaches for model combination: Bayesian inference, Akaike weight averaging, stepwise variable selection, updating, and fuzzy logic. We compared the performance of these approaches by assessing different aspects of their classification and discrimination capacity. We demonstrated that different approaches to model combination give rise to disparities in the model outputs. Bayesian integration was systematically affected by an error in the equations that are habitually used in distribution modelling. Akaike weights produced models that were driven by the best single factor and therefore failed at combining the models effectively. The updating and the stepwise approaches shared recalibration as the basic concept for model combination, were very similar in their performance, and showed the highest sensitivity and discrimination capacity. The fuzzy-logic approach yielded models with the highest classification capacity according to Cohen’s kappa. In conclusion: i) Bayesian integration, employing the currently used equation, and the Akaike weight procedure should be avoided; ii) the updating and stepwise approaches can be considered minor variants of the same recalibrating approach; and iii) there is a trade-off between this recalibrating approach, which has the highest sensitivity, and fuzzy logic, which has the highest overall classification capacity. Recalibration is better if

unfavourable conditions in one environmental factor may be counterbalanced with favourable conditions in a different factor, otherwise fuzzy logic is better.

INTRODUCTION

Species distributions are affected by limiting factors such as climate, topography, human activities, history, ecological interactions, and population dynamics, among others (Lomolino et al. 2005 pp. 77-96), which should thus be taken into account by species distribution models (Real et al. 2003). In fact, biogeographers are as interested in the individual role of each factor in driving species distribution as in the combined effect of several of them. This is particularly important in the context of the current biodiversity crisis, because, as Lomolino (2004) pointed out, the distribution of a species must be preserved in order to conserve it, and this cannot be achieved without knowing the processes, both individually and in combination, that affect the species range. This approach needs to be applied to threatened species in particular because their spatial ranges are often limited and their populations are subject to negative factors. As the risk of species extinction could increase due to the effects of multiple factors acting alone or in synergy (Brook et al. 2008), it is important to determine the maximum number of these factors.

Currently, there is a high and increasing number of environmental variables that are available, and each factor affecting a species distribution can be analysed in detail. These driving factors may be used together to perform a model (Carvalho et al. 2011), or used individually to build different models that could be subsequently combined (Márquez et al. 2011). However, methodological questions continue to arise regarding the way in which to combine different driving factors into comprehensive explanatory distribution models (Márquez et al. 2011, Beans et al. 2012).

Bates and Granger (1969) addressed a similar problem and proposed a way to combine forecasts that would yield lower mean errors than any of the constituent individual forecasts, provided that one of the forecasts was based on variables or information that the other forecast had not considered and that both forecasts were unbiased. The combined forecast would be obtained by a linear combination of the two sets of forecasts, giving a weight k to the first set of forecasts and a weight (1 - k) to the second set, k being obtained from a combination of the errors of the individual forecasts. Combinations of forecasts have been used in economics (Gregory et al. 2001), biomedicine (Nilsson et al. 2000), meteorology (Sanders 1963), and climatology (Benestad 2004), among others. However, rather than always being predictive, species distribution models are also descriptive or explanatory, and errors in forecasting are rarely provided when they are used to forecast the occurrence of the species. In addition, a model may correctly consider a locality as favourable according to a driving factor even if, due the lack of a different limiting factor, the species does not occur there. Araújo and New (2007) proposed the use of ensemble forecasting to predict species distributions by exploring the range of forecasts across more than one set of initial conditions, model classes, parameters, and boundary conditions. Their approach differed from ours, in that our aim was to combine models based on different sets of variables, using the best aspect of each model while avoiding redundancy.

This article addresses the question of how to combine variables representative of different biotic and abiotic factors in the modelling process (Márquez et al. 2011, Beans et al. 2012). We compared five methods for model combination derived from different approaches: Bayesian inference, information theory, stepwise variable selection, updating procedures, and fuzzy logic. The use of Bayesian inference for species distribution modelling is increasing (see Pereira and Itami 1991, Osborne et al. 2001), as it may be used to incorporate new knowledge, such as the modelled effect of a factor on a species distribution, into prior existing knowledge, such as the modelled effect of a different factor (Niamir et al. 2011). Information theory is widely used by the scientific community across multiple disciplines, including biogeography, to determine the most economic, rapid, and reliable (i.e. the most parsimonious) method to obtain robust results using different criteria (Burnham and Anderson 2002, Márquez et al. 2011). In the first steps, stepwise methods tend to select variables acting on a larger scale (Muñoz et al. 2005, Muñoz and Real 2013) and in subsequent steps to only add variables significantly related to the residuals not accounted for by previously incorporated variables (Barbosa et al. 2009). Updating methods have been used to adjust previously developed models to different circumstances (Steyerberg et al. 2004, Gaston and Viñas 2010) using shrinkage rules that can highlight the effect of the most relevant factors (Tibshirani 1996, Hastie et al. 2001, Guisan et al. 2002). Fuzzy logic uses the concept of favourability, which reflects the degree to which, in a continuous gradient form 0 to 1, the conditions are good and auspicious for the species. Fuzzy logic treats local environmental favourability as the degree of membership of every locality in the fuzzy set of localities whose environmental conditions are favourable to the species (Estrada et al. 2008). This approach allows different mathematical operations of fuzzy logic to be applied (Zadeh 1965) — intersection, union, and inclusion — which are useful for model combination. We compared the performance of these approaches by assessing different aspects of their classification and discrimination capacity.

MATERIAL AND METHODS

Species and study area

We modelled threatened vertebrate species whose European distributions are either restricted or almost completely restricted to the Iberian Peninsula. We chose these species because spatially restricted distributions usually show higher vulnerability to changes in environmental conditions (Carvalho et al. 2010), and thus determining the ability to integrate different driving factors into distribution models for these species is relevant for conservation purposes. We analysed five amphibian, four reptile, four mammal, and six bird species that are threatened or near threatened in mainland Spain according to the Red Lists in which the IUCN criteria have been adapted to this country (Palomo et al. 2007, Pleguezuelos et al. 2004, Madroño et al. 2004) (see Table 1).

We applied Grinnell's (1917) niche concept, according to which a niche is a land property and which defines the environmental features that make an area suitable to a given species. Thus, our models were built with the specific purpose of distinguishing which areas in Spain are favourable to the presence of the target species. Using 10x10-km UTM squares, presences were obtained from Palomo et al. (2007), Pleguezuelos et al. (2004), and Martí and Del Moral (2003) (see examples in Fig. 1). Only breeding distributions were considered in the case of birds. Absence data for the modelled taxa corresponded to the set of grid cells in which presences were not recorded in the atlases. We assumed equivalence between "not recorded" and "absent", because additional sampling efforts were made by the authors of the atlases regarding grid cells for which data were scarce in the previous literature.

The Iberian Peninsula has a very high diversity of environmental conditions due to its geographic position between two continents — Europe and Africa — and between two large water masses — the Mediterranean Sea and the Atlantic Ocean, and because of its great physiographical heterogeneity (Fig. 1). An uneven rainfall gradient increases toward the northwest, with annual rainfall ranging from around 200 mm to around 2000 mm. The temperature gradient has annual averages ranging from around 6ºC to 18ºC, and increases toward the south (Font 2000). Two macro-bioclimatic regions have been defined: temperate in the northernmost areas and Mediterranean in the rest of the peninsula (López et al. 2008).

Predictor factors

Three explanatory factors were taken into account to build the distribution models: spatial constraints, topography and climate, and human influence (Table 2). These predictor factors involve variables affecting species distribution at different geographic scales. Climate and topography are suitable predictors for the regional scale (200-2000 km) and land cover is suitable for the landscape scale (10-200 km) (Pearson and Dawson 2003). We used a resolution of 10x10-km, which can detect landscape changes and, when applied to the extent of our study area, can detect regional variations in climate and topography. Taking human activity in general into account is important because it may have a serious impact on the availability and quality of habitats of many species in Spain (Delibes-Mateos et al. 2009). Additionally, variables describing spatial effects (i.e. pure spatial structure, autocorrelation) allow the possible roles of population dynamics, dispersal capacities, and historical events on species distributions to be inferred (Legendre 1993, Real et al. 2003).

Human activities were described using different variables: water bodies, population, distance to highways and to urban centres, and land-cover (see sources of variables in Table 2). Land cover layers were processed so that the surface area of each class (see Delibes-Mateos et al. 2009) was used in preference to either the presence of each class or the distance from each class (see Romero et al. 2012). Latitude and longitude were used as the spatial variables and were treated according to a specific methodology based on trend surface analysis (Legendre 1993, see below). Topography was represented by three variables: slope, degree of southward exposure, and degree of westward exposure, which were derived from digital elevation models. Climate variables (temperature and precipitation) were obtained from the datasets supplied by the Spanish Institute of Meteorology (Agencia Estatal de Meteorología, AEMET 2007). These data derive from the regionalization for Spain of climate change models generated by the Intergovernmental Panel on Climate Change (IPCC). The distribution models based on the climatic factor were built using variables for the period 1960-1990 according to the ocean-atmosphere general circulation model (ECHAM4) provided by the Max Planck Institut für Meteorologie. The original resolution adopted for the variables was 1 km^2 per pixel. We then extracted the mean values of the variables for each 10x10-km UTM square of mainland Spain to match the resolution of the atlas data.

Spearman correlation coefficients were calculated to control multicollinearity between variables. When a set of variables belonging to the same factor were correlated with $r > 0.75$, we selected just one variable to enter in the models. The main selection criterion was to choose the variable that showed the lowest correlation with the other ones in the factor. In addition, we chose climatic variables describing the most extreme intra-annual conditions, e.g. July mean temperature was used instead of the annual mean temperature. Regarding human activity, we used variables describing well-defined land uses, e.g., oak woods rather than mixed-tree woods, and variables describing natural and well-represented environments, such as woods rather than mining areas. The variance inflation factor (VIF) of the set of candidate variables for the models was controlled using the range proposed by Montgomery and Peck (1992), according to whom a model is acceptable up to a VIF value of 10. The variables defining space were not submitted to multicollinearity analysis because they are mere coordinates of location.

Distribution models and factor integration

Three models were performed for each species using logistic regression (hereafter referred to as "partial probability models"), in which the dependent variable was species presence/absence in the UTM squares:

1. A spatial model was built using trend surface analysis (Legendre 1993); thus, a series of variables based on the average longitude (LO) and the average latitude (LA) — LO, LA, LO^2, LA^2, LO^3, LA^3, LA×LO, $LA \times LO^2$, $LA^2 \times LO$ — were entered in a logistic regression, so that the trend surface was considered to be the resulting logit (that is, the 'y_{sp}' lineal combination of spatial variables).
2. A topo-climatic model was built using the topographic and climate variables (Table 2), from which the logit 'y_{tpcl}' (that is, the 'y_{tpcl}' lineal combination of topo-climatic variables) was extracted.
3. A model for human influence was built using human influence variables, from which the logit 'y_h' (that is, the 'y_h' lineal combination of human variables) was extracted.

Benjamini and Hochberg's (1995) procedure was used to control the false discovery rate (FDR) in the topo-climatic and human influence partial probability models, in order to avoid type I errors arising from the large number of variables used (Benjamini and Hochberg 1995). Only the variables that were significantly ($p < 0.05$) related to the species distribution under an FDR of $q < 0.05$ were accepted, where q is the false discovery rate.

The challenge at this point was to integrate the three partial probability models in a comprehensive model. As the fuzzy intersection approach demands the use of fuzzy set membership degrees, such as favourability values (see below), we expressed all the outputs of the combining approaches in terms of favourability values in order to make them commensurate. Thus, probabilities were transformed into favourability values using the following formula (Real et al. 2006, Acevedo and Real 2012):

$$F = [P/(1-P)]/[(n_1/n_0)+(P/[1-P])],$$

where P is the probability value obtained according to each model, and n_1 and n_0 are the number of squares with presences or absences, respectively. The favourability function reflects the degree (between 0 and 1) to which the probability values obtained in each model differ from that expected according to the species' prevalence, where 0.5 indicates no difference between both probability values. Probability depends both on the response of the species to the predictors and on the overall prevalence of the species (Cramer 1999), whereas favourability values only reflect the response of the species to the predictors (Acevedo and Real 2012). When using favourability values, the responses of all the species are also expressed in the same "currency" and may be commensurate.

Five alternative combining methodologies were applied:

1. A Bayesian integration approach: We used the mathematical equation presented in Maynard (1981) as applied by Pereira and Itami (1991) and Osborne et al. (2001) for model combination:

$$P_{post}=1/1+\exp[\log((1-P_{rev})/P_{rev})-\log(P_{prior}/(1-P_{prior}))],$$

where P_{post} is the new probability obtained by combining a prior probability (P_{prior}) and a new probability used for revision (P_{rev}). Maynard's (1981) equation only uses two models to build the new probability; we combined the topo-climatic and the human partial probability models, and subsequently combined the obtained output with the spatial model.

2. An Akaike weight (AICw) averaging integration approach, based on information theory: A relative Akaike weight was given to each of the partial probability models and these weights were used to produce a weighted averaged model (Burnham and Anderson 2002, p. 75). Thus, this combined model was expressed as the probability of finding a species within each square and the probability was transformed into a favourability function.
3. A stepwise-procedure integration approach: We obtained a combined model by performing a forward-backward stepwise logistic regression with the variables that were entered in the three partial probability models (see Márquez et al. 2011). Probabilities were again transformed into favourability values.
4. An updating approach: A combined model was obtained by performing a logistic regression using the equations ‘y’ of the partial models as variables:

$$y_{new}=\alpha+ß_1y_{sp}+ ß_2y_{tpcl}+ ß_3y_h,$$

where y_{new} is the new logit equation from which the new updated combined probabilities were obtained; α is a new constant; $ß_1$, $ß_2$ and $ß_3$ are new coefficients; and ‘y_{sp}’, ‘y_{tpcl}’ and ‘y_h’ are the logits according to each partial model. The new probabilities were also transformed into favourability values.

5. A fuzzy intersection (fuzzy logic approach): Probabilities resulting from the three partial probability models were transformed into favourability values, which can be interpreted as degrees of membership of every UTM square in the fuzzy set of squares whose environmental conditions are favourable for the species (Estrada et al. 2008). We assumed that every factor with which a significant partial model can be obtained acts as a limiting factor for the species distribution. We then calculated the intersection between the partial favourability models as the largest fuzzy set which was contained in the three models; that is, for every square, the minimum favourability value derived from any of the three partial models was chosen (Zadeh 1965).

Statistical analyses were performed using IBM SPSS statistics 19. Favourability models were geographically represented using ArgGIS 9.3.

Comparative assessment

The performance of the five model integration approaches was compared according to their discrimination and classification capacity. Measures of classification accuracy were based on the 0.5 favourability threshold, because when favourability is 0.5 probability is equal to overall prevalence (Acevedo and Real 2011). The following classification capacity indices were used: sensitivity; specificity; correct classification rate (CCR); Cohen's Kappa (Fielding and Bell 1997); and under- and over-prediction rates (Barbosa et al. 2013). Under-prediction refers to the proportion of observed presences in the predicted absence area, and over-prediction refers to the proportion of absences in the predicted presence area (see formulas in Barbosa et al. 2013).

Good classification is indicated by high sensitivity, specificity, CCR, and Kappa values, and low under-prediction and over-prediction rates. We used the Area Under the Receiver Operating Characteristic Curve (AUC, Lobo et al. 2008) as an index of discrimination capacity. The average value obtained for the 19 species analysed was calculated for each index and methodological approach.

RESULTS

Thirty-eight variables were finally used as candidates for model training: nine spatial, nine topo-climatic (four topographic and five climatic), and twenty human variables (Table 2). One representative species from each vertebrate group is shown as an example: *Chioglosa lusitanica* represents amphibians, *Iberolacerta cyreni* represents reptiles, *Pterocles orientalis* represents birds, and *Galemys pyrenaicus* represents mammals (see Fig. 2 for single-factor models and Fig. 3 for comprehensive combined models). The results obtained for all the species are shown in Supplementary material Appendix 1 for single-factor models and in Supplementary material Appendix 2 for combined models.

The combination approach using the Akaike criterion produced models mainly based on only one factor: a weight higher than 0.99 was given to the main factor for all but one species (Table 3). Using this approach, the favourability values of the combined models were practically the same as those of the respective partial model with the highest weight, and thus the final model was not effectively refined by the effect of other factors (see the maps of the four species used as examples in Fig. 2 and Fig. 3). Therefore, this approach produced models based on topo-climatic factors alone for 37% of the species, on spatial factors alone for 53%, and on human influence factors alone for 11% of them.

In contrast, the other approaches produced models based on more than one factor. The models combined using the Bayesian method, and to a certain extent those using fuzzy logic, showed favourability values that were often lower than the models performed using the other approaches (see Fig. 3). For every species, the models obtained by using the stepwise and the updating combination approaches were qualitatively and quantitatively very similar (see Fig. 3).

There were significant differences in classification capacity indices between partial and combined models, and between partial models (see Fig. 4, and Table 4 for significance tests). The Bayesian approach produced models with the best values for specificity, CCR, and the over-prediction rate, but their sensitivity was very low and they had the worst under-prediction rate. The Akaike weight approach had the worst values for most of the classification assessment criteria (specificity, CCR, Kappa, and the over-prediction rate). The performance of the stepwise and updating procedures did not significantly differ according to any of the classification indices, although the stepwise-procedure always performed slightly better than the updating procedure. These two approaches generated models that scored well according to all the assessment criteria, and had the highest values for sensitivity. The fuzzy logic approach had the best values according to the Kappa index and a good over-prediction rate, but had low sensitivity and a high under-prediction rate.

The stepwise combined model alone had a significantly higher discrimination capacity (AUC value) than the partial spatial model; all the combined models had a significantly higher discrimination capacity than the partial human-influence model; and none of the combined models had a discrimination capacity that was significantly different from that of the topo-climatic model. There were no significant differences between combined models according to AUC values.

DISCUSSION

Our results indicate that the descriptive capacity of distribution models can be substantially improved when several factors are combined compared to models based on a single factor. Combining procedures generated models with a better overall classification rate, higher specificity, and a lower over-prediction rate. However, the discrimination capacity (AUC) and sensitivity of the partial models did not significantly improve when applying the combining procedures. Lobo et al. (2008) have already warned against using AUC as an overall index of model performance, because it only measures the discrimination capacity of the models. Furthermore, species with a restricted range, which is the usual situation of threatened species, typically occupy territories that are easy to discriminate from unoccupied localities, even with clearly deficient models. On the other hand, given that sensitivity measures the ability of the model to correctly classify actual presences as presences, sensitivity has been customarily considered a more valuable index of model performance than specificity (Brito et al. 1999), partly because presences are more certain than absences, and thus a model that misclassifies actual presences of species is unacceptable. However, sensitivity may be easily increased by predicting presences in larger areas of the environment, and thus a partial model with a high sensitivity value obtained in this way is difficult to improve by using a better combined model. In this aspect, sensitivity must be balanced by the over-prediction rate (Barbosa et al. 2013), and it is this index which is clearly improved with the combined models.

The multifactor approach is frequently used in distribution modelling (e.g. Márquez et al. 2011, Beans et al. 2012, Real et al. 2013, Romero et al. 2013). However, the most frequent practice is to only take into account climatic variables for forecasting distribution changes under climate change scenarios (e.g., Araújo et al. 2006, Carvalho et al. 2010, Pearman et al. 2010, Araújo et al. 2011). In our combined models, climate was the most important factor for birds (Table 3), probably because of their higher vagility (Lemoine et al. 2007), and hence their lower dependence on both historical events and geographical barriers (Cumming et al. 2012). In contrast, space was the most important factor for mammals, reptiles, and amphibians (Table 3). Our study supports the proposals of Pearson and Dawson (2003) and Márquez et al. (2011) that the quality of distribution forecasts improves when derived from models in which climate is combined with other influential factors. Therefore, it is of relevance to assess the different methodological options available for factor integration in a single model.

The Bayesian approach produced models with the worst under-prediction rate and sensitivity value, the latter being particularly low. This is because models combined using the Bayesian approach yielded favourability values that were lower than those of the partial models or the other combined models.

Herzog and Ostwald (2013) have encouraged the implementation of Bayesian statistics in experimental biology and other scientific disciplines. Bayes theorem can be used to modify a prior probability (hypothesis) with more reliable (less probabilistic) information, such as expert knowledge or a different hypothesis. However, during the course of our investigations, we found that the Bayesian assumptions were not met by the Bayesian integration approach proposed by Maynard in an unpublished PhD Thesis (1981) and first published by Pereira and Itami (1991):

$$P_{post}=1/1+\exp[\log((1-P_{rev})/P_{rev})-\log(P_{prior}/(1-P_{prior}))],$$

where P_{post} is the new probability obtained by combining a prior probability (P_{prior}) and a new probability used for revision (P_{rev}).

In fact, the combination of the two probabilities consists in adding the logit functions of the prior and revision logistic models to obtain the logit function of the posterior probability (see Supplementary material Appendix 3 for mathematical demonstration). Thus, the differential concepts of prior probability and probability used for revision are void of meaning because they are interchangeable. Besides, the variables defining the two probabilities should be independent, which has been overlooked in the literature. Finally, given the foregoing, this procedure ignores the fact that posterior probabilities are affected by the effect of prevalence twice, because this effect is taken into account in each of the logit functions. Given that the species studied in the present article always have low prevalence, this effect produces models with probability values that are consistently lower than expected (note that the favourability function corrects the effect of prevalence only once). This does not mean that the Bayesian approach should be discarded as a conceptual framework to combine distribution models. For example, Hoeting et al. (1999) proposed the use of Bayesian model averaging to account for model uncertainty. However, this approach has not been applied to the combination of species distribution models, possibly due to the difficulty of specifying the prior distributions on both the parameters and model space.

The Akaike weight approach obtained models based on a single factor. Márquez et al. (2011) obtained similar results when they used this integration method with four species. The main reason for this result is that AIC weights are based on absolute AIC differences irrespective of the total number of localities analysed. Thus, differences between two models with AIC values of 10 and 14, for example, are the same as those between models with AIC values of 5110 and 5114, and are weighted accordingly, although the difference is sizeable in the former case and nearly meaningless in the latter case. Another weighting procedure should be proposed in which the differences are relative rather than absolute, because with large N even small differences in model performance yield large differences in model weight. This and other limitations of the AIC are related to the fact that the penalty term is constant, which deserve greater attention (e.g., see Arnold 2010).

Of the methodological options proposed, our results suggest that only the stepwise, updating, and fuzzy logic approaches are good methods for combining species distribution models.

The stepwise procedure and the updating approach generated similar models, probably because both are re-parameterization procedures. The stepwise method re-parameterizes the variables and deletes those not significant, whereas the updating procedures only re-parameterize the variables. The advantages of the stepwise procedure were that it yielded more parsimonious models and that the models generated less multicollinearity, which is probably the reason for the stepwise-procedure performing slightly better than the updating procedure.

On the other hand, although the models based on fuzzy intersection were slightly inaccurate regarding presences, they obtained good values for the other criteria. Despite the fact that fuzzy intersection allows different factors to participate in a single model, when this method was used each grid was assigned a favourability value derived from only one factor. This method does not allow for recalibration or counterbalance between different environmental factors. This is essential when the different factors or variables cannot compensate for each other, i.e., when combining several limiting factors. For example, Real et al. (2009) found that the absence of rabbits could not be compensated for by other environmental factors favourable for the Iberian lynx, and Romo et al. (2014) found that the lack of larval food plant could not be compensated for by other factors favourable for monophagous butterflies.

These differences in performance could be of interest to studies with different aims (Guisan et al. 2013). For example, re-parametrization approaches minimise under-prediction and would therefore be a better choice when the objective is to ensure that all areas that can be invaded by an invasive species are monitored. Fuzzy intersection minimises over-prediction and would therefore be a better choice when the aim is to define management priorities in a setting of limited resources.

CONCLUSIONS

There are substantial differences in model outputs when different approaches are used to combine single-factor models. Such disparities can affect the capacity of a model to describe favourable areas as well as its capacity to predict trends, and thus they can have a direct effect on the usefulness of a model for conservation purposes.

The Akaike weight approach and the Bayesian integration equation should not be used to combine distribution models unless their mathematical foundation is revised.

Stepwise and updating approaches produce similar models, as both are recalibration methods; of these, the stepwise approach is more parsimonious.

Finally, the choice between recalibration and fuzzy intersection depends on the researcher’s objectives: recalibration procedures are the best if counterbalance between variables or factors is permitted, whereas fuzzy intersection is better when combining limiting factors that cannot be counterbalanced by more favourable factors.

ACKNOWLEDGEMENTS

D Romero was supported by a grant from the Ministerio de Educación: AP2007-03633. JC Brito was supported by FCT Compete-Programa Operacional Regional do Norte ON2. This study was partially financed by project 1098/2014 (Organismo Autónomo

Parques Nacionales, Spain). We thank S. Coxon for his help in the English revision of the manuscript. We are grateful to B. Leroy and two anonymous reviewers for their constructive comments on a previous version of the manuscript.

REFERENCES

Acevedo, P. and Real, R. 2011. Biogeographical differences between the two *Capra pyrenaica* subspecies, *C. p. victoriae* and *C. p. hispanica*, inhabiting the Iberian Peninsula: Implications for conservation. . Ecol. Model. 222: 814–823.

Acevedo, P. and Real, R. 2012. Favourability: concept, distinctive characteristics and potential usefulness. Naturwissenschaften 99: 515–522.

AEMET, 2007. Agencia Estatal de Meteorología of Spain. Ministerio de Medio Ambiente. Available at: http://www.aemet.es/es/elclima/cambio_climat/escenarios (accessed June 2007).

Araújo, M. B. et al. 2006. Climate warming and the decline of amphibians and reptiles in Europe. J. Biogeogr. 33: 1712–1728.

Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22: 42–47.

Araújo, M. B. et al. 2011. Impactos, vulnerabilidad y adaptación al cambio climático de la biodiversidad Española. 2 Fauna de vertebrados. Dirección general de medio Natural y Política Forestal. Ministerio de medio ambiente y medio rural y marino, Madrid.

Arnold, T. W. 2010. Uninformative parameters and model selection using Akaike's Information Criterion. J. Wildl. Manage. 74: 1175-1178

Barbosa, A. M. et al. 2009. Transferability of environmental favourability models in geographic space: the case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. Ecol. Model. 220: 747–754.

Barbosa, A. M. et al. 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. Divers. Distrib. 29: 1333–1338.

Bates J. M. and Granger, C. W. 1969. The combination of forecasts. Operat. Res. Quart. 20: 451–468.

Beans C. M. et al. 2012. Climate suitability and human influences combined explain the range expansion of an invasive horticultural plant. Biol. Invasions. 14: 2067–2078.

Benestad, R. E. 2004. Tentative probabilistic temperature scenarios for northern Europe. Tellus 56A: 89–101.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B. 57: 289–300.

Brito et al. 1999. Modelling wildlife distributions: logistic multiple regression vs overlap analysis. Ecography 22: 251–260.

Brook, B. et al. 2008. Synergies among extinction drivers under global change. Trends Ecol. Evol. 23: 453–460.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference. A practical information-theoretic approach, 2nd ed. Springer, New York.

Carvalho, S. B. et al. 2010. From climate change predictions to actions – conserving vulnerable animal groups in hotspots at a regional scale. Glob. Change Biol. 16: 3257–3270.

Carvalho, S. B. et al. 2011. Incorporating evolutionary processes into conservation planning using species distribution data: a case study with the western Mediterranean herpetofauna. Diversity Distrib. 17: 408–421.

Cramer, J. S. 1999. Predictive performance of binary logit model in unbalanced samples. J. R. Stat. Soc. D. 48: 85–94.

Cumming, G. S. et al. 2012. Towards a unification of movement ecology and biogeography: conceptual framework and a case study on Afrotropical ducks. J. Biogeogr. 39: 1401–1411.

Delibes-Mateos, M. et al. 2009. Long-Term changes in game species over a long period of transformation in the Iberian Mediterranean landscape. Environ. Manage. 43: 1256–1268.

Estrada, A. et al. 2008. Using crisp and fuzzy modelling to identify favourability hotspots useful to perform gap analysis. Biodivers. Conserv. 17: 857–871.

Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence-absence models. Environ. Conserv. 24: 38–49.

Font, I. 2000. Climatología de España y Portugal. Ediciones Universidad de Salamanca, Salamanca, Spain.

Gaston, A. and García-Viñas, J. I. 2010. Updating coarse-scale species distribution models using small fine-scale samples. Ecol. Model. 221: 2576–2581.

Guisan, A. et al. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol. Model. 157: 89–100.

Guisan, A. et al. 2013. Predicting species distributions for conservation decisions. Ecol. Lett. 16: 1424–1435.

Gregory, A. W. et al. 2001. Testing for forecast consensus. J. Bus. Econ. Stat. 19: 34–43.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. Auk 34: 427–433.

Hastie, T. J. et al. 2001. The elements of Statistical Learning: data mining, inference and prediction. Springer-Vergal, New York, pp. 739.

Herzog, S. and Ostwald, D. 2013. Experimental biology: sometimes Bayesian statistics are better. Nature 494: 35.

Hoeting, et al. 1999. Bayesian model averaging: a tutorial. Statistical Science 4: 382–417.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74: 1659–1673.

Lemoine, N. et al. 2007. Species richness of migratory birds is influenced by global climate change. Global Ecol. Biogeogr. 16: 55–64.

Lobo, J. et al. 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecol. Biogeogr. 17: 145–151.

Lomolino, M. V. 2004. Conservation biogeography. In: M. V. Lomolino and L. R. Heaney (eds.) Frontiers of biogeography. New directions in the geography of nature. Sinauer, Sunderland, pp. 293–296.

Lomolino, M. V. et al. 2005. Biogeography, 3rd ed. Sinauer, Sunderland.

López F. M. L. et al. 2008. Macrobioclimas, bioclimas y variantes bioclimáticas de la España peninsular y Balear, y su cartografía. Publicaciones de Biología, Universidad de Navarra, Serie Botánica 17: 229–236.

Madroño, A. et al. 2004. Libro rojo de las aves de España. Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornitología, Madrid, Spain.

Márquez, A. L. et al. 2011. Combining climate with other influential factors for modelling the impact of climate change on species distribution. Clim. Chan. 108: 135–157.

Martí, M. and Del Moral, J. C. 2003. Atlas de las aves reproductoras de España. Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornitología, Madrid, Spain.

Maynard, P. F. 1981. The logit classifier: a general maximum likelihood discriminant for remote sensing applications. Unpublished M.A. Thesis, department of geography, university of California, Santa Barbara, California.

Montgomery, D. C. and Peck, E. A. 1992. Introduction to Linear Regression Analysis. Wiley, New York.

Muñoz, A. R. et al. 2005. Modelling the distribution of Bonelli’s eagle in Spain: implications for conservation planning. Diversity Distrib. 11: 477–486.

Muñoz, A. R. and Real, R. 2013. Distribution of Bonelli's Eagle *Aquila fasciata* in southern Spain: scale may matter. Acta Ornithologica 48: 93–101.

Niamir, A. et al. 2011. Finessing atlas data for species distribution models. Diversity Distrib. 17: 1173–1185.

Nilsson, J. et al. 2000. Consensus predictions of membrane protein topology. FEBS Lett. 486: 267–269.

Osborne, P. E. et al. 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. J. Appl. Ecol. 38: 458–471.

Palomo, L. J. et al. 2007. Atlas y libro rojo de los mamíferos terrestres de España. Dirección General para la Biodiversidad-SECEM-SECEMU, Madrid.

Pearman, P. B. et al. 2010. Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. Ecography 33: 990–1003.

Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global Ecol. Biogeogr. 12: 361–371.

Pereira, J. M. C. and Itami, R. M. 1991. GIS-Based habitat modeling using logistic multiple regression: a study of the Mt. Graham Red Squirrel. Photogramm. Eng. Rem. S. 57: 1475–1486.

Pleguezuelos, J. M. et al. 2004. Atlas y libro rojo de los anfibios y reptiles de España. Dirección General de Conservación de la Naturaleza-Asociación Herpetológica Española, Madrid, Spain.

Real, R. et al. 2003. Relative importance of environment, human activity and spatial situation in determining the distribution of terrestrial mammal diversity in Argentina. J. Biogeogr. 30: 939–947.

Real, R. et al. 2006. Obtaining environmental favourability functions from logistic regression. Environ. Ecol. Stat. 12: 237–245.

Real, R. et al. 2008. Modelling chorotypes of invasive vertebrates in mainland Spain. Diversity Distrib. 14: 364–373.

Real, R. et al. 2009. Conservation biogeography of ecologically interacting species: the case of the Iberian lynx and the European rabbit. Diversity Distrib. 15: 390–400.

Real, R. et al. 2013. Estimating how inflated or obscured effects of climate affect forecasted species distribution. PLoS ONE 8: doi:10.1371

Romero, D. et al. 2012. Comparative assessment of different methods for using land-cover variables for distribution modelling of *Salamandra salamandra longirostris*. Environ. Conserv. 40: 48–59.

Romero, D. et al. 2013. Uncertainty in distribution forecasts caused by taxonomic ambiguity under climate change scenarios: a case study with two newt species in mainland Spain. J. Biogeogr. 41: 111–121.

Romo et al. 2014. Effects of climate change on the distribution of ecologically interacting species: butterflies and their main food plants in Spain. Ecography 37: 001-10.

Sanders, F. 1963. On subjective probability forecasting. J. Appl. Meteorol. 2: 191–201.

Steyerberg, E. W. et al. 2004. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Statis. Med. 23: 2567–2586.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso, J. Royal Statist. Soc. B 58: 267–288.

Wald, A. 1943. Tests of statistical hypotheses concerning several parameters with applications to problems of estimation. Trans. Amer. Math. Soc. 54: 426-482.

Zadeh, L. A. 1965. Fuzzy sets. Inform. Control 8: 338–353.

Supplementary material (Appendix ECOG-01477 at <www.oikosoffice.lu.se/appendix>). Appendix 1–3.

Table 1. Endangered vertebrates analysed by species. The number of observations refers to the study area (mainland Spain). Threat categories using IUCN criteria adapted to Spain are shown. NT: Near Threatened, VU: Vulnerable, EN: Endangered and CR: Critically Endangered.

Group	Species	Abbreviation	Number of observations	Category
Amphibians	*Chioglossa lusitanica*	*Chiolus*	166	VU
	Triturus pygmaeus	*Tripyg*	453	VU
	Alytes dickhilleni	*Alydic*	133	VU
	Rana pyrenaica	*Ranpyr*	25	VU
	Calotriton asper	*Calasp*	186	NT
Reptiles	*Testudo graeca*	*Tesgra*	38	EN
	Algyroides marchi	*Algmar*	30	VU
	Iberolacerta bonnali	*Ibebon*	23	VU
	Iberolacerta cyreni	*Ibecyr*	41	VU
Birds	*Aquila adalberti*	*Aquada*	163	EN
	Chersophilus duponti	*Chedup*	233	EN
	Emberiza schoeniclus	*Embsch*	124	EN
	Pterocles alchata	*Ptealc*	477	VU
	Pterocles orientalis	*Pteori*	864	VU
	Tetrax tetrax	*Tettet*	1339	VU
Mammals	*Galemys pyrenaicus*	*Galpyr*	425	VU
	Lynx pardinus	*Lynpar*	27	CR
	Microtus cabrerae	*Miccab*	219	VU
	Lepus castroviejoi	*Lepcas*	63	VU

Table 2. Environmental variables initially proposed for the distribution models defining the areas favourable to the endangered vertebrate species in Spain. Only variables with low correlations according to Spearman's index were retained to reduce the effect of multicollinearity. The variables finally considered for entering in the models are shown in bold.

Abbreviation	**Variable**	**Abbreviations**	**Variable**
Spatial situation			
La	**Latitude (ºN)(1)**	**Lo**	**Longitude (ºE)(1)**
Topography			
SE	**Southward exposure degree(2)**	**S**	**Slope (º) (3)**
WE	**Westward exposure degree(2)**		
Climate			
Temp	Annual mean temperature (ºC)(4)	**PSu**	**Summer precipitation (mm)(4)**
TJa	January mean temperature (ºC)(4)	PAu	Autumn precipitation (mm)(4)
TJul	**July mean temperature (ºC)(4)**	**PWi**	**Winter precipitation (mm)(4)**
TSp	Spring mean temperature (ºC)(4)	**TMin**	**Minimum temperature** (ºC)**(5)**
TSu	Summer mean temperature (ºC)(4)	**ContI**	**Continentality index** (ºC) **(6)**
TAu	Autumn mean temperature (ºC)(4)	GDD0	Growing degree days above 0ºC(5)
TWi	Winter mean temperature (ºC)(4)	GDD50	Growing degree days above 5ºC(5)
PAn	Annual precipitation (mm)(4)	IO0	Ombrothermal index above 0ºC(7)
PSp	Spring precipitation (mm)(4)	IO50	Ombrothermal index above 5ºC(7)
Human factor			
Land-cover/land-use variables			
PAST	**Pasture(8)**	**LW**	**Lush wood(9)**
SV	**Sparse vegetation(8)**	**RW**	**Riparian wood(9)**
SCLV	Sclerophyllous vegetation(8)	**MW**	**Mixed trees wood(9)**
OAKM	**Oak meadow(9)**	**HE**	**Heather(8)**
CM	**Conifer meadow(9)**	**NM**	**Natural meadows(8)**
OAKW	**Oak wood(9)**	BR	Bare rock(8)
CW	**Conifer wood(9)**	ARV	Artificial vegetation(8)
OLG	**Olive grove(8)**	RICE	Rice field(8)
AGR	Agroforestry(8)	**VIN**	**Vineyard(8)**
DHER	**Dry herbaceous crops(8)**	HETF	Heterogeneous farm(8)
FT	**Fruit tree crops(8)**	INDA	Industrial area(8)
IHER	**Irrigated herbaceous crops(8)**	MINING	Mining area(8)
SSOIL	Sandy soil(8)	TRAN	Transition area(8)
WR	**Water reservoirs(10)**	**LK**	**Lakes(10)**
SAL	Salt marsh (8)	**IMAR**	**Interior marshes(8)**

Other human activities

Dhi	**Distance to the nearest highway (km)(1)**	U500	Distance to the nearest urban centre with more than 500 000 inhabitants (km)(1)
U100	Distance to the nearest urban centre with more than 100 000 inhabitants (km)(1)	**HPd**	**Human population density in 2000 (number of inhabitants/km^2)(11)**

Sources:

(1) Spanish National Geography Institute (IGN). 1999. Mapa de Carreteras. Península Ibérica, Baleares y Canarias. Instituto Geográfico Nacional, Ministerio de Fomento. Madrid, Spain.

(2) Farr, T.G. and Kobrick, M. 2000. Shuttle Radar Topography Mission produces a wealth of data. EOS Trans. Amer. Geophys. Union. 81: 583–585.

(3) United States Geological Survey (1996). GTOPO30. Land Processes Distributed Active Archive Center. EROS Data Center. https://lta.cr.usgs.gov/GTOPO30

(4) State Meteorology Agency (AEMET), Spanish Ministry of Agriculture, Food and Environment. http://www.aemet.es/es/elclima/cambio_climat/escenarios

(5) Pearson, R.G. et al. 2002. SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. Ecol. Model. 154: 289–300.

(6) Supan 1884. Mapa de oceanidad de la tierra. Índice simple de continentalidad.

(7) López Fernández, M.L. et al. 2008. Macrobioclimas, bioclimas y variantes bioclimáticas de la España peninsular y Balear, y su cartografía. Publicaciones de Biología, Universidad de Navarra, Serie Botánica, 17: 229–236.

(8) Corine Land Cover 2000 revisado para el 2006. CLC2000. Ministerio de Fomento. Gobierno de España.; http://www.ign.es/ign/layoutIn/corineLandCover.do/

(9) Mapa de Forestal de España (MFE200) 1997. Ministerio de Medio Ambiente. Gobierno de España.; www.magrama.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/mfe200_descargas.aspx

(10) Corine Land Cover 2000 revisado para el 2006. Ministerio de Fomento; y ampliado a su vez con la información del Sistema Integrado de Información del Agua (SIA) para España. Ministerio de Agricultura, Alimentación y Medio Ambiente. Gobierno de España.; www.magrama.gob.es/es/cartografia-y-sig/ide/descargas/agua/descargas_agua.aspx

(11) ORNL. 2001. LandScan 2000 Global Population Database. Oak Ridge National Laboratory. Oak Ridge. Tennessee.

Table 3. Akaike weight values (AICw) of each partial probability model in the favourability model combining the three factors. Each partial probability model measures the probability of finding a species according to a single predictor factor.

	Predictor factors		
	Spatial situation	**Topo-climatic**	**Human**
Amphibians			
Chioglossa lusitanica	0.0078	0.992	1.290^{E-44}
Triturus pygmaeus	1	2.857×10^{-12}	2.131^{E-109}
Alytes dickileni	1	2.896^{E-58}	2.766^{E-106}
Rana pyrenaica	1	1.641^{E-15}	2.812^{E-32}
Calotriton asper	1	3.052^{E-12}	2.353^{E-131}
Reptiles			
Testudo graeca	1	1.403^{E-18}	4.732^{E-32}
Algyroides marchi	1	9.193^{E-18}	4.487^{E-11}
Iberolacerta bonnali	1.274^{E-44}	1	3.267^{E-13}
Iberolacerta cyreni	1	4.458^{E-23}	2.973^{E-45}
Birds			
Aquila adalberti	0.0995	9.732^{E-05}	0.9004
Chersophilus duponti	5.100^{E-39}	1	1.306^{E-38}
Emberiza shoeniclus	2.546^{E-22}	0.994	0.00568
Pterocles alchata	4.334^{E-211}	1	8.188^{E-115}
Pterocles orientalis	3.337^{E-242}	1.52^{E-31}	1
Tetrax tetrax	1.203^{E-288}	0.999	4.859^{E-08}
Mammals			
Galemys pyrenaicus	2.759^{E-101}	1	3.762^{E-60}
Lynx pardinus	1	6.404^{E-18}	5.159^{E-45}
Microtus cabrerae	1	3.165^{E-74}	3.451^{E-45}
Lepus castroviejoi	1	2.553^{E-23}	1.173^{E-32}

Table 4. Significant differences (Wilcoxon's test) between models according to the mean values (N = 19 species) obtained for each classification capacity index. Sp: spatial model; Tpclim: topo-climatic model; Hum: human influence model. Letters indicate significant differences between models according to various indices: a = sensitivity; b = specificity; c = over-prediction; d = under-prediction; e = CCR; f = Kappa. Upright letters indicate p <0.001, and italic letters indicate p <0.05 .

	Partial models			**Combined models**				
	Sp	**Tpclim**	**Hum**	**Bayesian**	**Akaike**	**Stepwise**	**Updating**	**Fuzzy**
Bayesian	a,b,c,e	a,b,c,*d*,e	a,b,c,e,*f*		a,b,c,*d,e*	a,b,c,*d,e*	a,b,c,*d,e*	a,*b*,c
Akaike			*a,b,c,e,f*	a,b,c,*d,e*				*a,b,c,e,f*
Stepwise	*b,c,e,f*	*c,f*	*a*,b,*c*,e,f	a,b,c,*d,e*				a,*b,c*
Updating	*b,c,e,f*	*f*	*a*,b,*c*,e,f	a,b,c,*d,e*				a,*b,c*
Fuzzy	*a,b*,c,*e*,f	*a*,b,c,*e*,f	*a*,b,c,e,f	a,*b*,c	*a,b,c,e,f*	a,*b,c*	a,*b,c*	

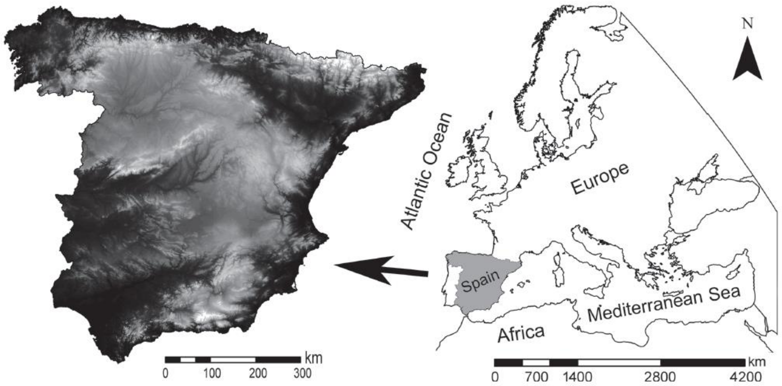
Figure captions

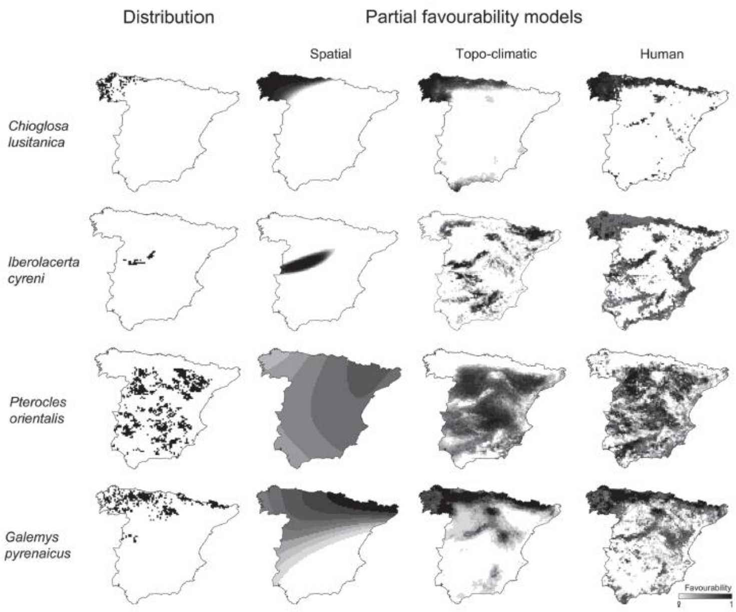
Figure 1. Study area in Europe. In the map on the left, grey shadows indicate elevation (black and white colours indicate elevation below 0 m.a.s.l. and above 3450 m.a.s.l., respectively).

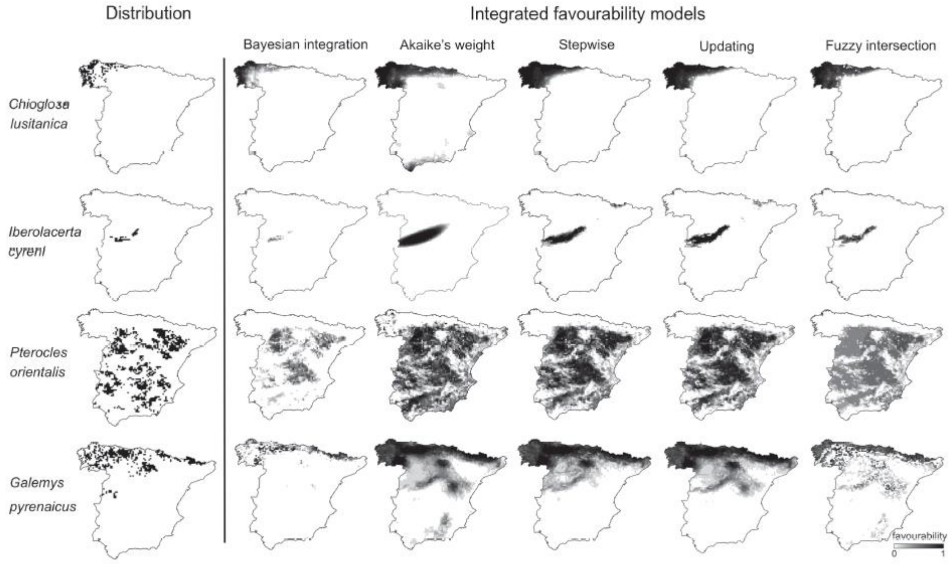
Figure 2. Current distributions in mainland Spain, and partial spatial, topo-climatic, and human-influence favourability models of four representative species: *Chioglossa lusitanica*, *Iberolacerta cyreni*, *Pterocles orientalis*, and *Galemys pyrenaicus* (favourability ranges from 0 to 1). The spatial resolution is 10x10-km.

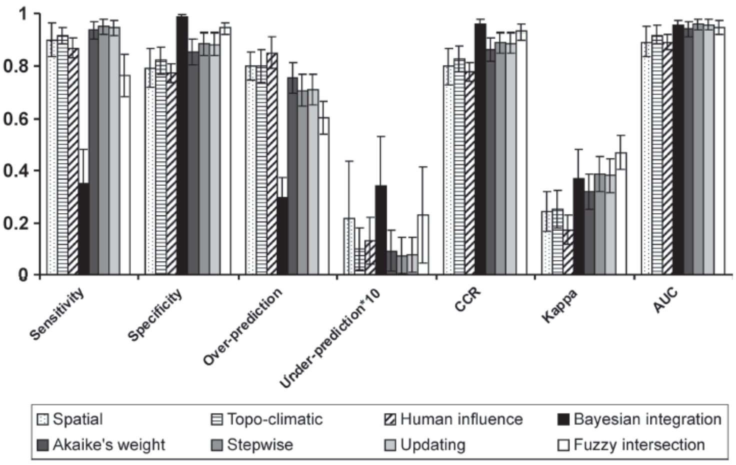
Figure 3. Favourability models combining different factors in mainland Spain of four representative species: *Chioglossa lusitanica*, *Iberolacerta cyreni*, *Pterocles orientalis*, and *Galemys pyrenaicus* (favourability ranges from 0 to 1). Five alternative methodological approaches for factor integration were applied: Bayesian, Akaike weight (AICw) averaging, stepwise-procedure, updating, and fuzzy intersection. The spatial resolution is 10x10-km.

Figure 4. Comparative assessments of five methodological integration approaches, and of three partial models according to the five indices used: sensitivity, specificity, correct classification rate (CCR), and Cohen’s kappa (all calculated using the favourability value 0.5 as the classification threshold), and the area under the curve of the receiver operating characteristic (AUC). Bars represent average values for the integrated models for the 19 vertebrate species (see the species list in Table 1) and 95% confidence intervals are represented above each bar. Under-predicted values were multiplied by ten for reasons of clarity.


N
Atlantic Ocean
Europe
Spain
Africa
Mediterranean Sea
km
0
100
200
300
km
0
700
1400
2800
4200


Distribution
Partial favourability models
Spatial
Topo-climatic
Human
Chioglosa lusitanica
Iberolacerta cyreni
Pterocles orientalis
Galemys pyrenaicus
Favourability
0
1


Distribution
Integrated favourability models
Bayesian integration
Akaike's weight
Stepwise
Updating
Fuzzy intersection
Chioglossa lusitanica
Iberolacerta cyreni
Pterocles orientalis
Galemys pyrenaicus
favourability
0
1


1
0.8
0.6
0.4
0.2
0
Sensitivity
Specificity
Over-prediction
Under-prediction*10
CCR
Kappa
AUC
Spatial
Topo-climatic
Human influence
Bayesian integration
Akaike's weight
Stepwise
Updating
Fuzzy intersection