

# Using crisp and fuzzy modelling to identify favourability hotspots useful to perform gap analysis

Alba Estrada · Raimundo Real · J. Mario Vargas

Received: 24 April 2007 / Accepted: 9 January 2008 / Published online: 25 January 2008  
© Springer Science+Business Media B.V. 2008

**Abstract** In this study, we propose the use of a favourability function to perform Gap Analysis. To exemplify this, we modelled the distribution of terrestrial mammal species in Andalusia (South of Spain) on the basis of their presence/absence on a grid of 10 km × 10 km UTM cells ( $n = 961$ ). Using logistic regression and 30 variables related with the environment, space and human influence, we obtained probabilities of occurrence for each species in each cell. We computed a crisp favourability index considering the areas as favourable or unfavourable for a species if the probability of occurrence was higher or lower than the species prevalence, respectively. We also used a favourability function and fuzzy logic to level all species to the same threshold of favourability, which allowed to compare and to combine species distributions. Adding up the fuzzy favourability values for each species in each cell we obtained a fuzzy favourability index that we compared with species richness (sum of species in each cell) and with the crisp favourability index. We performed Gap Analysis by overlapping these results with the current reserve network of Andalusia. Gaps were grouped in fewer and bigger zones after applying the favourability indices. Considerations and recommendations for the use of the favourability function to select areas of conservation interest are discussed.

**Keywords** Andalusia · Conservation planning · Fuzzy logic · Mammals · Natural reserve network · Spain · Spatial modelling

## Abbreviations

RENPA	Red de Espacios Naturales Protegidos de Andalucía (Natural Reserve Network of Andalusia)
UTM	Universal Transverse Mercator
TIN	Triangulated Irregular Network
TINSURF	Triangulated Irregular Network Surface
GIS	Geographic Information System

---

A. Estrada (✉) · R. Real · J. M. Vargas  
Biogeography, Diversity, and Conservation Research Team, Department of Animal Biology,  
Faculty of Sciences, University of Málaga, 29071 Málaga, Spain  
e-mail: aestrada@uma.es

FDR	False Discovery Rate
P	Logistic Probability
F	Favourability
AUC	Area Under the Curve
FEM	Fuzzy Envelope Model
ENFA	Ecological Niche Factor Analysis

## Introduction

The establishment of protected areas began in the 19th century, with the declaration of Yosemite as a state-protected natural reserve in 1864. Protected areas became increasingly relevant in nature conservation with the declaration of Yellowstone in 1872 as the first National Park in the world, being Yosemite upgraded to this category in 1890 (Greene 1987). In Spain, the first National Park was declared in 1918, and it was in 1929 when the first natural reserve was declared in Andalusia (Consejería de Medio Ambiente 2004). The number of protected areas increased sharply during the second half of the 20th century. This caused the need for evaluation of the conservation value of different areas to prioritize their preservation in the context of the whole network of reserves.

The assessment and design of a protected area network are considered as important applications of Conservation Biogeography, which was defined by Whittaker et al. (2005) as: “*the application of biogeographical principles, theories, and analyses, being those concerned with the distributional dynamics of taxa individually and collectively, to problems concerning the conservation of biodiversity*”.

Gap Analysis is a methodology designed to compare the distribution of biodiversity with the network of protected areas in a territory. Scott et al. (1989, 1993) began to use this term in the early 1990s and applied Gap Analysis to know which areas were more important for conservation and to prioritize which must be protected next. To assess the conservation value of an area it is necessary to have surrogates of biodiversity such as species richness (see, for instance, Araújo 1999; Yip 2004), rarity (Rey Benayas and de la Montaña 2003; Real et al. 2006b), vulnerability (de la Montaña and Rey Benayas 2002), maintenance of patterns and processes (Rouget et al. 2003), ecosystem representativeness (Sierra et al. 2002), or minimum area to support viable populations (Allen et al. 2001). Species richness is the most frequent measure of biodiversity (Brose et al. 2003; Grenyer et al. 2006) and one of the most important biological properties of a territory when evaluating its conservation status (Real 1992). However, species richness does not guarantee the correct protection of all species in a region, and the mere aggregation of species in an area does not imply that this area is important for these species.

Normally, the databases of these studies are atlases of species distribution that can be biased depending on the survey effort. As natural reserves sometimes are the most surveyed areas, they often appear in the atlases as supporting more species than the surrounding zones. Spatial modelling can attenuate this problem by providing potential distribution of each species, which do not depend so much on the survey effort (Real et al. 2006b). However, different species tend to have different prevalences in a territory, so causing different bias in their spatial modelling and precluding the joint use of the output models.

The favourability function proposed by Real et al. (2006a) levels all species at the same threshold of favourability, independently of their proportion of presences, and this allows

direct comparison of the distribution of different species. In this approach, an area is not absolutely favourable or absolutely unfavourable for a species, but it has a degree of favourability. Fuzzy logic is applicable in this situation, as the process of environmental modelling can be understood as the identification of the grade of membership of each area to the fuzzy set of favourable areas for each species.

The term favourability, however, differs from habitat suitability or ecological niche. The favourability for the occurrence of a species may not represent the fundamental or the realized niche, and is not always related to the suitability of the habitat. According to the source-sink model (Pulliam 1988), for example, the geographical distribution of a species includes sink areas where the habitat is unsuitable but populations are maintained by dispersal from source areas. Hence distribution models may be describing an amalgam of realized niche (or suitable habitats) and sink areas close to sources (Austin 2002). Conversely, an area with high habitat suitability may be unfavourable for the occurrence of a species due to historic causes (past events that prevented the species from inhabiting the area).

The aim of this paper is to show how the favourability function can be used to perform fuzzy Gap Analysis. We detected important areas for the conservation of biodiversity in Andalusia, using mammal favourability hotspots as important conservation areas, and modelling the distribution of mammals with logistic regression and with the favourability function to obtain crisp and fuzzy favourability hotspots, respectively. We compared these results with those obtained using observed mammal richness.

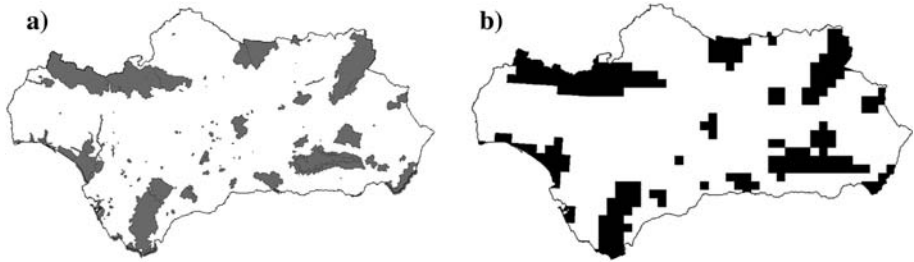
## Material and methods

### Study area

Andalusia is the southern region of mainland Spain, with almost 87,600 km<sup>2</sup> of extension. It is one of the 17 Spanish autonomous regions and is divided in eight administrative provinces (Fig. 1). The climate is Mediterranean with an important precipitation gradient, from 170 mm/year to more than 1800 mm/year. The elevation gradient is also important,

**Fig. 1** Situation of Andalusia in the Iberian Peninsula





**Fig. 2** (a) Natural reserve network of Andalusia (RENPA). (b) 10 km  $\times$  10 km UTM cells with at least 25% of their area covered by the RENPA

ranging from sea level to almost 3500 m.a.s.l. Climatic and orographic heterogeneity results in high habitat diversity in Andalusia.

We selected the Andalusian region to perform the study because conservation policy powers in Spain, including the power to declare protected natural areas, were devolved to autonomous regions in 1989. Andalusia is the Spanish autonomous region with the largest surface declared as protected natural areas and the third with the highest number of protected zones (Consejería de Medio Ambiente 2004). The Natural Reserve Network of Andalusia (RENPA: Red de Espacios Naturales Protegidos de Andalucía) covers 20% of the territory and comprises 150 natural areas with different status of protection (Fig. 2a).

The region supports 56 indigenous terrestrial mammal species, according with the Atlas of terrestrial mammals of Spain (Palomo and Gisbert 2002). Twenty-six of these species are threatened (Consejería de Medio Ambiente 2001a). The Cabrera vole (*Microtus cabreræ*) and the wolf (*Canis lupus*) are critically endangered in Andalusia (Consejería de Medio Ambiente 2001a) and the Iberian lynx (*Lynx pardinus*) is critically endangered in Spain and in the world (Palomo and Gisbert 2002; IUCN 2004).

The 10 km  $\times$  10 km UTM cell map ( $n = 961$ ) was obtained overlapping the digital outline of Andalusia on the digital UTM cell map of the Iberian Peninsula (resulting by fusion of the maps used by Barbosa et al. (2003) for Spain and Portugal), using Cartalinx software. The distribution of the 56 indigenous terrestrial mammals inhabiting Andalusia on the basis of their presence/absence data were taken from the Atlas of terrestrial mammals of Spain (Palomo and Gisbert 2002), except for the Iberian lynx (*Lynx pardinus*) which were taken from Guzmán et al. (2004) because these are more recent data reporting a sharp reduction in the distribution of the species. Appendix 1 shows the list of mammal species inhabiting Andalusia.

## Variables

We used 30 variables related to the environment, spatial situation, and human influence (Table 1) to model the distribution of terrestrial mammals on the basis of their presence/absence data, because our aim was to assess the RENPA according to the distribution of the environmental, spatial, and human-related favourability for each species. We digitized the variables to a vector format (except for *Alti* which was already available in digital version in a raster format) using Cartalinx, and processed them using Idrisi GIS software. Isoline variables (*HJan* through *Long*) were interpolated in a raster format, with resolution scale of 1 pixel  $\approx$  1 km<sup>2</sup>, from a triangulated irregular network with the Idrisi TIN and TINSURF modules performing Bridge and Tunnel (B/T) edge removal. *Con1* is a climatic index

**Table 1** Variables used to model mammals distribution in Andalusia

Code	Variable
<i>Area</i>	Cell area (m <sup>2</sup> )
<i>Alti</i>	Mean altitude (m) <sup>a</sup>
<i>Slop</i>	Slope (°) (calculated from <i>Alti</i> )
<i>HuJan</i>	Mean relative air humidity in January at 07:00 (%) <sup>b</sup>
<i>HuJul</i>	Mean relative air humidity in July at 07:00 (%) <sup>b</sup>
<i>HRan</i>	Annual relative air humidity range (%) (=   <i>HuJan</i> - <i>HuJul</i>  )
<i>PET</i>	Mean annual potential evapotranspiration (mm) <sup>b</sup>
<i>AET</i>	Mean annual actual evapotranspiration (mm) (= min [ <i>Prec</i> , <i>PET</i> ])
<i>Inso</i>	Mean annual insolation (h/year) <sup>b</sup>
<i>SRad</i>	Mean annual solar radiation (kW h/m <sup>2</sup> /day) <sup>b</sup>
<i>TJan</i>	Mean temperature in January (°C) <sup>b</sup>
<i>TJul</i>	Mean temperature in July (°C) <sup>b</sup>
<i>Temp</i>	Mean annual temperature (°C) <sup>b</sup>
<i>TRan</i>	Annual temperature range (°C) (= <i>TJul</i> - <i>TJan</i> )
<i>DFro</i>	Mean annual number of frost days (minimum temperature ≤0°C) <sup>b</sup>
<i>DPre</i>	Mean annual number of days with precipitation ≥0.1 mm <sup>b</sup>
<i>Prec</i>	Mean annual precipitation (mm) <sup>b</sup>
<i>MP24</i>	Maximum precipitation in 24 h (mm) <sup>b</sup>
<i>RMP</i>	Relative maximum precipitation (= <i>MP24</i> / <i>Prec</i> )
<i>DSno</i>	Mean annual number of snow days <sup>c</sup>
<i>ConI</i>	Continental index <sup>c</sup>
<i>HumI</i>	Humidity index <sup>c</sup>
<i>Pirr</i>	Pluviometric irregularity <sup>d</sup>
<i>ROff</i>	Mean annual runoff (mm) <sup>e</sup>
<i>Lati</i>	Latitude (°N) <sup>f</sup>
<i>Long</i>	Longitude (°E) <sup>f</sup>
<i>DHi</i>	Distance to the nearest highway (km) <sup>f</sup>
<i>U100</i>	Distance to the nearest urban centre with more than 100,000 inhabitants (km) <sup>f</sup>
<i>U500</i>	Distance to the nearest urban centre with more than 500,000 inhabitants (km) <sup>f</sup>
<i>Perm</i>	Soil permeability <sup>e</sup>

Sources: <sup>a</sup> US Geological Survey (1996). <sup>b</sup> Font (1983). <sup>c</sup> Font (2000). <sup>d</sup> Montero de Burgos and González-Rebollar (1974). <sup>e</sup> IGME (1979). <sup>f</sup> IGN (1999). Data on the number of inhabitants of the urban centres were taken from the Instituto Nacional de Estadística, Spain (<http://www.ine.es>)

depending on the annual temperature range and latitude, while *HumI* depends on the precipitation and evaporation. Secondary variables, defined in Table 1 by an algebraic operation in parentheses, were calculated from primary variables using the Idrisi Image Calculator. Distance variables (*DHi*, *U100* and *U500*) were calculated from the digitized highways and major towns using the Idrisi DISTANCE module. *Perm* was obtained from a map of synthesis of groundwater aquifers, a categorical map with four classes of aquifers (IGME 1979); we reclassified them in three classes of increasing permeability and the final value of *Perm* was determined by calculating the mean of the values assigned to the pixels within each 10 km × 10 km UTM cell. The RENPA vector map (Consejería de Medio Ambiente 2001b) was transformed to a raster map with a resolution scale of 1 pixel.

Afterwards, we obtained the mean value of each variable in the 961 (10 km × 10 km) UTM cells of Andalusia using the Idrisi EXTRACT module.

### Modelling method

By performing logistic regression of each species presence/absence on each variable separately we selected a subset of variables significantly related to each species distribution. To control the increase in type I error due to multiple tests (Benjamini and Hochberg 1995; García 2003), we only accepted the variables that were significant under a False Discovery Rate (FDR) of  $q < 0.05$ , using the procedure for all forms of dependency among statistics proposed by Benjamini and Yekutieli (2001). We then performed forward–backward stepwise logistic regression of each species on the subset of significant predictor variables to obtain a multivariate logistic model. In forward steps the variable most significantly related to the residuals not explained by the previous variables is selected, but the significance of all the variables in the model is tested after each step of forward inclusion, and non-significant variables are excluded before the next forward selection step (Legendre and Legendre 1998).

We performed a crisp logic model considering logistic probabilities ( $P$ ) of each species as favourable when  $P$  was higher than the prevalence (i.e. proportion of presences) and unfavourable when  $P$  was lower than the prevalence. Since probabilities are dependent on the prevalence, of each species, a probability value of 0.5 can indicate a favourable zone if the species has a restricted range or an unfavourable zone if the species is common. The output of this crisp favourability model has two possible values: 1 for favourable areas and 0 for unfavourable areas.

We transformed all this procedure into a fuzzy model by applying the favourability function proposed by Real et al. (2006a) on the logistic regression output, so converting logistic probabilities ( $P$ ) into favourability ( $F$ ) values.  $F$  values are independent of prevalence, since a value of  $F = 0.5$  is assigned to the predictor conditions for which  $P = \text{prevalence}$  (Real et al. 2006a). Favourability values higher than 0.5 correspond to areas where the probability of presence is higher than that expected according to prevalence, and the opposite occurs in areas with favourability below 0.5. Whereas  $P$ -values for different species are not comparable because of the prevalence bias,  $F$ -values are directly equivalent, and the models for all mammals are then levelled to the same threshold of favourability and can be compared and combined directly. So, a favourability value of 0.5 is a neutral value for all species. The favourability value for a species in a cell can be interpreted as the grade of membership of the cell to the fuzzy set of cells that are favourable for the species, and then the favourability function is the membership function, so allowing the use of concepts and applications of fuzzy logic to the resulting spatial analysis of the species. This approach has the advantage over other fuzzy modelling approaches (Robertson et al. 2004; Levinsky et al. 2007) that it is related to logistic regression, which most modellers are familiar with, and provides an explicit equation based on multivariate statistics, and thus allows for selecting among the predictor variables and assigns different importance to them.

The discrimination power of the models was assessed calculating their correct classification rate (CCR), sensitivity, and specificity, using the favourability value of  $F = 0.5$  as classification threshold, and the Area Under the Curve (AUC) of the Receiver Operating Characteristic, which is independent of any favourability threshold (Hosmer and Lemeshow 2000).

We calculated three indices. The first one is species richness, i.e. number of mammal species in each cell  $j$ :

$$\sum_{i=1}^{56} (c_{ij}) \quad (1)$$

High values of index 1 represent species richness hotspots.

The second index is the result of adding up, for each species  $i$  in each cell  $j$ , the areas favourable for the presence of the species:

$$\sum_{i=1}^{56} (p_{ij}) \quad (2)$$

where  $p$  is the output (1 or 0) of the crisp favourability modelling. High values of index 2 represent crisp favourability hotspots.

The third index is the result of adding up the fuzzy favourability value for each species  $i$  in each cell  $j$ :

$$\sum_{i=1}^{56} (F_{ij}) \quad (3)$$

When applying index 3 we summed directly the output of the favourability values without transforming them to crisp values. High values of this index represent fuzzy favourability hotspots.

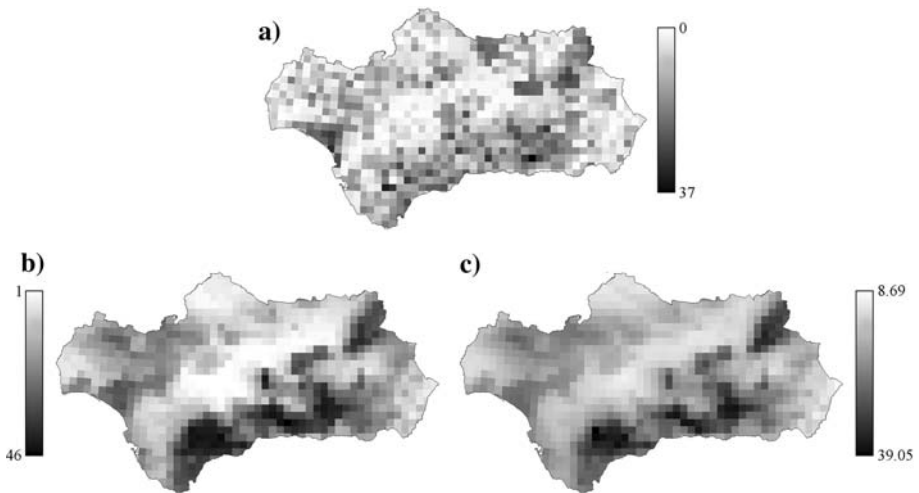
Only cells with at least 25% of their area covered by the RENPA were considered as protected (Fig. 2b). We performed Gap Analysis by overlapping those cells to the 20% of cells with highest richness values after applying the indices. We established the threshold in 20% of the territory because the RENPA covers approximately this proportion of Andalusia. Those cells above this threshold, which were not protected were considered gaps in the protection of mammals.

## Results

We did not obtain significant models for three species, which are marked with an asterisk in Appendix 1. The mean values of the discrimination power of the 53 remaining significant models were: 74.48% for the CCR, 74.36% for the sensitivity, 75.13% for the specificity and 0.81 for the AUC. According to the general rule proposed by Hosmer and Lemeshow (2000), this mean AUC value indicates excellent mean discrimination. All the modelled species were represented in the RENPA. At least 22% of the overall favourability in Andalusia for every species corresponded to areas covered by the RENPA.

Figure 3a shows mammal richness distribution in Andalusia after applying index 1. Several cells with a high number of species border with cells without species. On the contrary, when we applied index 2 (Fig. 3b) and index 3 (Fig. 3c) differences in the scores among neighbouring cells were less marked. Maps obtained after applying indices 2 and 3 were more similar with each other than with the map reporting number of species (index 1).

Figure 4 (a, c and e) shows the 20% of cells with the highest values of the three indices. After overlapping the RENPA (Fig. 2b) to these results, we obtained the gaps in the protection of biodiversity, considering mammal richness, crisp favourability hotspots, and fuzzy favourability hotspots as indicative of conservation importance, respectively. Gaps



**Fig. 3** Value obtained for each cell after applying: (a) index 1 based on species richness, (b) index 2 based on crisp modelling, and (c) index 3 based on fuzzy modelling

differed in number of cells (108, 103 and 91, respectively) and in location (Fig. 4b, d and f). Besides, considering index 1 gaps formed 32 groups of cells, while after applying indices 2 and 3 only 7 groups were formed.

## Discussion

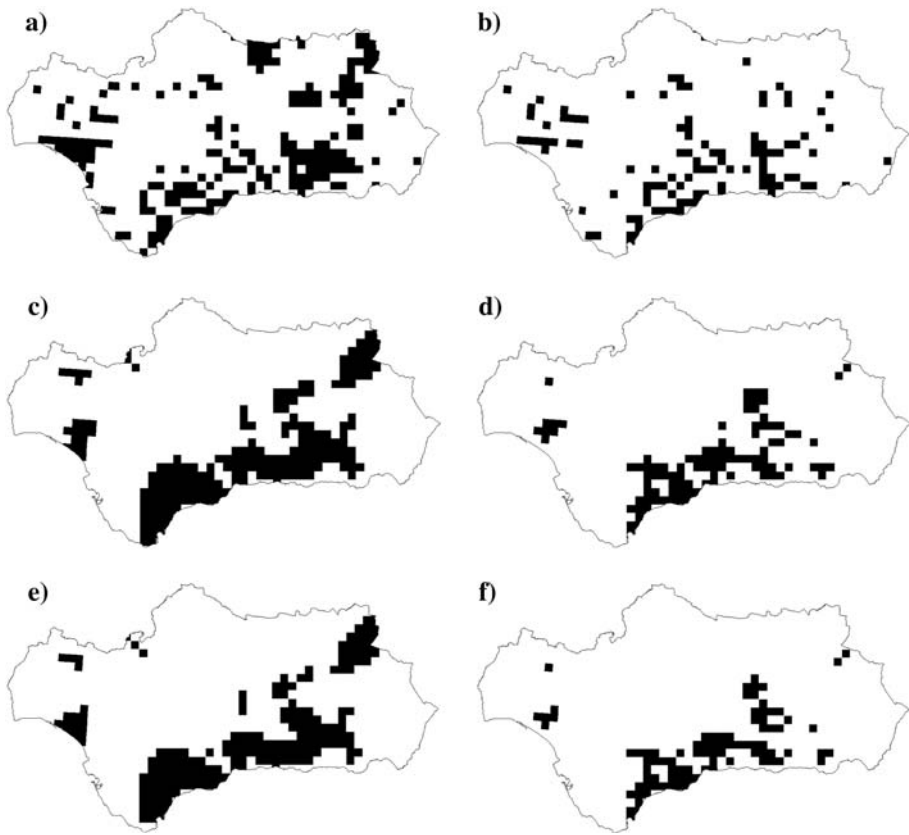
### Spatial models in Gap Analysis

Our results do not reflect the habitat suitability of the species. Habitat suitability models must be applied at a scale that is normally smaller than that used in our favourability models. A 100 km<sup>2</sup> cell is a good resolution scale when the variation scale of the study is large, but it has a considerable variety of habitats inside, which advises against its use in habitat suitability models. However, the term habitat suitability has been applied elsewhere without taking into account these considerations (see, for instance, Thuiller et al. 2005) and using variables that are not characteristic of the habitat, such as, for example, the geographical location (Brotons et al. 2004).

Some species have scattered distributions in the Atlas of terrestrial mammals of Spain, which probably do not represent the status of the species but results from an insufficient survey effort. Spatial models solve this problem because areas with high scores tend to be together and predicted values decrease gradually to other areas with lower values (Barbosa et al. 2005). This is why different authors have applied spatial modelling to perform Gap Analysis (Allen et al. 2001; Pearlstine et al. 2002; Maiorano et al. 2006) or to select areas for species conservation (Araújo and Williams 2000; Araújo et al. 2002; van Teeffelen et al. 2006).

However, most spatial modelling techniques yield scores that are dependent on the prevalence of the species, which advises against their use in procedures such as Gap Analysis where every species should be levelled to the same threshold (Jiménez-Valverde and Lobo 2006). Species that have a restricted distribution due to ecological, geographical,





**Fig. 4** Twenty per cent of Andalusian UTM grid cells with the highest conservation values considering index 1 based on species richness (a), and biodiversity gaps obtained by overlapping the natural reserve network of Andalusia to those cells (b). Idem considering index 2 based on crisp modelling (c and d) and index 3 based on fuzzy modelling (e and f)

or historical factors, have a low probability value even in the favourable areas where they are present. Conversely, common species have high probability of presence even in unfavourable areas. The use of the favourability function allows overcoming this drawback, because  $F$  values are high where the conditions are favourable for the species independently of the prevalence.

#### Surrogates of biodiversity

Normally, the conservation value of a territory is established using different criteria, such as richness, rarity or vulnerability of the species inhabiting the area (Rey Benayas and de la Montaña 2003; Real et al. 2006b; Grenyer et al. 2006; Estrada et al. in press). In this study we compared richness hotspots with favourability hotspots, the latter meaning zones favourable for a high number of mammal species.

Although species richness has been considered a good surrogate of biodiversity in many studies (Real et al. 1993; Araújo 1999; Maes et al. 2005), there are some authors who have

shown that species richness is more related with the distribution of generalist species than to the distribution of rare ones (Vázquez and Gaston 2004). However, the majority of the fuzzy favourability hotspots for mammals obtained in this study coincide with those of fuzzy mammal rarity in Andalusia (Real et al. 2006b).

Reserve-selection techniques based on complementarity seek to maximize representation of biodiversity within the limitations of cost. Hotspots of complementarity have been shown to tend to be located in areas with high local species richness and in areas concentrating species of restricted range size (Araújo and Williams 2001). Complementarity is a first logical procedure when a country or region lacks a natural reserve network, because with a limited budget the new network would represent all species in few areas. This is not the case of Andalusia where more than 17,500 km<sup>2</sup> are protected, the reserve network is composed of 150 natural areas, and all modelled mammal species are represented in the RENPA.

In fact, there will never be a perfect surrogate or suite of surrogates of biodiversity (Groom et al. 2006). Results of our study should be analyzed together with those of fuzzy rarity (Real et al. 2006b), vulnerability, endemism or representation for mammals and other groups of animals and plants in the area, to identify important zones for biodiversity conservation in Andalusia. In this context, fuzzy logic could be applied using the intersection and union operations of the fuzzy indices, and we would obtain important areas for all groups and indices simultaneously, and important areas for at least one of them, respectively.

### Fuzzy favourability hotspots

The aggregation of the favourability for the whole set of species identifies the areas favourable for many species, whose conservation value is higher than that of areas supporting many species. It is important to protect zones with high favourability value because: (a) they can act as refuges for those mammal species that are rare in the whole territory, i.e., for species that generally address adverse environmental conditions; (b) they may represent source areas where birth rate is higher than mortality and which export individuals to sink areas (Pulliam 1988); and (c) even unoccupied favourable areas are important for conservation when metapopulation dynamics are involved (Levins 1969; Hanski and Simberloff 1997; Muñoz et al. 2005). These arguments are relevant for individual species, but we argue that areas that fulfil these conditions for a high number of species merit attention for conservation purposes.

Species richness should not be always identified with favourability hotspots. An area supporting relatively few species may be highly favourable for them, so obtaining a high favourability score, and if an area favourable for a species is unoccupied, then it computes for favourability hotspot but not for species richness. On the other hand, at least three processes may lead to high species richness in generally unfavourable areas: (a) generalist species tend to be present even in areas with low favourability values; (b) sink areas for many species may overlap; and (c) some areas not particularly favourable for any species may be more thoroughly surveyed than more favourable areas.

The Atlas of terrestrial mammals of Spain is probably the best source available to know the distribution of mammals in Andalusia, but 36 cells have no species reported in the atlas, whereas numerous scattered individual cells stand out for having many species reported. This reflects the general fact that our knowledge of the distribution of species or other individual biodiversity entities is far from complete (Ferrier 2002). However, in our

analysis, the number of important cells covered by the reserve network was higher after applying the favourability function, although this function was performed partly to correct a possible sampling bias in favour of protected areas. This seems to indicate that the survey effort was not biased to protected areas but to outside them. Although the criteria to establish protected areas in Andalusia were others than the richness or the favourability for the species, this analysis shows that the natural reserve network was generally well distributed in relation to the areas favourable for many species or, at least, better located than the mere visualization of rich areas in the atlas would suggest.

Robertson et al. (2004) applied a fuzzy classification technique (fuzzy envelope model, FEM) for predicting species' distributions by using presence-only locality records. Areas with higher values after applying the FEM technique were interpreted as holding more favourable conditions for the organism. The favourability function proposed by Real et al. (2006a) also obtains favourable areas for the species but using presence–absence data, in a similar way as logistic regressions do. Consequently, the favourability function is to logistic regression approximately the same as FEM is to ENFA, favourability function and FEM to be used in the scope of fuzzy logic and logistic regression and ENFA in the scope of crisp logic. Robertson et al. (2004), as well as other authors (Brotons et al. 2004), pointed out that, in cases where absence records are available, models built using presence–absence data may perform better than presence-only data models. We used the favourability function because absence data may provide useful information of the conditions unfavourable for the species.

To protect the 20% of the territory with the highest richness value considering index 1, it would be necessary to establish many and dispersed protected areas in Andalusia, while fewer, bigger, and more concentrated reserves would be needed to cover the gaps detected with the crisp and fuzzy favourability modelling. Similar results were obtained after applying Gap Analysis to the rarity and fuzzy rarity of mammals (Real et al. 2006b), and to the richness and fuzzy favourability hotspots of amphibians (Estrada et al. in press). Fuzzy modelling is then more appropriate to implement the idea of Ferrier (2002), who affirmed that conservation areas not only should be selected to represent as many elements of biodiversity as possible, but also should be sufficiently large and well connected to promote long-term persistence of this diversity.

The establishment of important areas for mammal richness in Andalusia should not be based on a rigid methodology, but on a flexible procedure, which should be able to change in space and time. Maes et al. (2005) recommended that policy makers should make more use of modelling techniques as a proactive conservation tool, because it allows to better target sites predisposed to be included to a natural reserve network. In the same way, Ferrier (2002) proposed integrating biological and environmental data through predictive modelling as a strategy that may help alleviate some of the problems associated with using remotely mapped surrogates in conservation planning. As the favourability function produces an output value for each species in each cell in relation to environmental conditions, it is sensitive to changes of these conditions in space and even in time according to climate or land use change scenarios (Audsley et al. 2006; Thuiller et al. 2006). Consequently, this modelling method may be an important tool to assess both the present and future conservation value of a territory and to foresee the future design of reserve networks.

**Acknowledgements** This work was conducted as part of the project CGL2006-09567/BOS (I+D project) funded by the Ministerio de Educación y Ciencia (Spain), financed jointly by the FEDER; and the project P05-RNM-00935, funded by the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain). A. Estrada is a PhD student with a grant of this administration.

**Appendix 1** List of indigenous mammal species inhabiting Andalusia

Scientific name	Common name
<i>Erinaceus europaeus</i>	Western hedgehog
<i>Atelerix algirus</i>	Algerian hedgehog
<i>Talpa occidentalis</i>	Iberian mole
<i>Neomys anomalus</i>	Miller's water shrew
<i>Crocidura suaveolens</i>	Lesser white-toothed shrew
<i>Crocidura russula</i>	Greater white-toothed shrew
<i>Suncus etruscus</i>	Pygmy white-toothed shrew
<i>Rhinolophus ferrumequinum</i>	Greater horseshoe bat
<i>Rhinolophus hipposideros</i>	Lesser horseshoe bat
<i>Rhinolophus euryale</i>	Mediterranean horseshoe bat
<i>Rhinolophus mehelyi</i>	Mehely's horseshoe bat
<i>Myotis bechsteinii</i>	Bechstein's bat
<i>Myotis myotis</i>	Greater mouse-eared bat
<i>Myotis blythii</i>	Lesser mouse-eared bat
<i>Myotis nattereri</i>	Natterer's bat
<i>Myotis emarginata</i>	Geoffroy's bat
<i>Myotis daubentonii</i>	Daubenton's bat
<i>Myotis capaccinii</i>	Long-fingered bat*
<i>Pipistrellus pipistrellus</i>	Common pipistrelle
<i>Pipistrellus mediterraneus</i>	–
<i>Pipistrellus kuhlii</i>	Kuhl's pipistrelle
<i>Hypsugo savii</i>	Savi's pipistrelle
<i>Nyctalus leisleri</i>	Leisler's bat
<i>Nyctalus lasiopterus</i>	Greater noctule
<i>Eptesicus serotinus</i>	Serotine
<i>Barbastrella barbastrellus</i>	Barbastelle*
<i>Plecotus austriacus</i>	Grey long-eared bat
<i>Miniopterus schreibersii</i>	Schreibers' bat
<i>Tadarida teniotis</i>	European free-tailed bat*
<i>Canis lupus</i>	Wolf
<i>Vulpes vulpes</i>	Red fox
<i>Mustela nivalis</i>	Weasel
<i>Mustela putorius</i>	Polecat
<i>Martes foina</i>	Beech marten
<i>Meles meles</i>	Eurasian badger
<i>Lutra lutra</i>	Otter
<i>Herpestes ichneumon</i>	Egyptian mongoose
<i>Genetta genetta</i>	Small-spotted genet
<i>Felis silvestris</i>	European wildcat
<i>Lynx pardinus</i>	Iberian lynx
<i>Sus scrofa</i>	Wild boar
<i>Cervus elaphus</i>	Red deer
<i>Capreolus capreolus</i>	Roe deer

**Appendix 1** continued

Scientific name	Common name
<i>Capra pyrenaica</i>	Iberian wild goat
<i>Sciurus vulgaris</i>	Red squirrel
<i>Arvicola sapidus</i>	Southern water vole
<i>Chionomys nivalis</i>	Snow vole
<i>Microtus duodecimcostatus</i>	Mediterranean pine vole
<i>Microtus cabrerae</i>	Cabrera vole
<i>Apodemus sylvaticus</i>	Wood mouse
<i>Rattus rattus</i>	Black rat
<i>Mus domesticus</i>	House mouse
<i>Mus spretus</i>	Algerian mouse
<i>Eliomys quercinus</i>	Garden dormouse
<i>Lepus granatensis</i>	Iberian hare
<i>Oryctolagus cuniculus</i>	Rabbit

\* Species for which we did not obtain significant models

**References**

- Allen CR, Pearlstine LG, Kitchens WM (2001) Modeling viable mammal populations in Gap Analysis. *Biol Conserv* 99:135–144
- Araújo MB (1999) Distribution patterns of biodiversity and the design of a representative reserve network in Portugal. *Divers Distrib* 5:151–163
- Araújo MB, Williams P (2000) Selecting areas for species persistence using occurrence data. *Biol Conserv* 96:331–345
- Araújo MB, Williams PH (2001) The bias of complementarity hotspots toward marginal populations. *Conserv Biol* 15(6):1710–1720
- Araújo MB, Williams PH, Turner A (2002) A sequential approach to minimise threats within selected conservation areas. *Biodivers Conserv* 11:1011–1024
- Audsley E, Pearn KR, Simota C, Cojocaru G, Koutsidou E, Rounsevell MDA, Trnka M, Alexandrov V (2006) What can scenario modelling tell us about future European scale agricultural land use, and what not? *Environ Sci Policy* 9:148–162
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Modell* 157:101–118
- Barbosa AM, Real R, Olivero J, Vargas JM (2003) Otter (*Lutra lutra*) distribution modeling at two resolution scales suited to conservation planning in the Iberian Peninsula. *Biol Conserv* 114:377–387
- Barbosa AM, Segovia JM, Vargas JM, Torres J, Real R, Miquel J (2005) Predictors of Red Fox (*Vulpes vulpes*) Helminth parasite diversity in the provinces of Spain. *Wild Biol Pract* 1(1):3–14
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188
- Brose U, Martinez ND, Williams RJ (2003) Estimating species richness: sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology* 84(9):2364–2377
- Brotons L, Thuiller W, Araújo MB, Hirzel AH (2004) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448
- Consejería de Medio Ambiente (2001a) Libro rojo de los vertebrados amenazados de Andalucía. Junta de Andalucía, Sevilla, Spain
- Consejería de Medio Ambiente (2001b) Límites de EE.NN.PP. de Andalucía. Red de Información Ambiental de Andalucía (CD-Rom). Junta de Andalucía, Sevilla, Spain

- Consejería de Medio Ambiente (2004) La RENPA en cifras (CD-Rom). Junta de Andalucía. Sevilla
- De la Montaña E, Rey Benayas JM (2002) ¿Coinciden los espacios naturales protegidos con las áreas relevantes de diversidad de herpetofauna en España peninsular y Baleares? Ecosistemas 2002/2 (URL: <http://www.aect.org/ecosistemas/022/investigacion2.htm>)
- Estrada A, Márquez AL, Real R, Vargas JM (in press) Utilidad de los espacios naturales protegidos de Andalucía para preservar la riqueza de especies de anfibios. Munibe
- Ferrier S (2002) Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? Syst Biol 51(2):331–363
- Font I (1983) Atlas climático de España. Instituto Nacional de Meteorología, Madrid
- Font I (2000) Climatología de España y Portugal. Ediciones Universidad de Salamanca, Salamanca
- García LV (2003) Controlling the false discovery rate in ecological research. Tree 18:553–554
- Greene L (1987) Yosemite: the park and its resources. A history of the discovery, management, and physical development of Yosemite National Park. U.S. Department of the Interior/National Park Service, California
- Grenyer R, Orme CDL, Jackson SF, Thomas GH, Davies RG, Davies TJ, Jones KE, Olson VA, Ridgely RS, Rasmussen PC, Ding T-S, Bennett PM, Blackburn TM, Gaston KJ, Gittleman JL, Owens IPF (2006). Global distribution and conservation of rare and threatened vertebrates. Nature 444(2):93–96
- Groom MJ, Meffe GK, Carroll CR (2006) Principles of conservation biology, 3rd edn. Sinauer Associates, Inc. Sunderland. USA
- Guzmán JN, García FJ, Garrote G, Pérez R, Iglesias C (2004) El lince ibérico (*Lynx pardinus*) en España y Portugal. Censo-diagnóstico de sus poblaciones. Dirección General para la Biodiversidad, Madrid, 184 pp
- Hanski I, Simberloff D (1997) The metapopulation approach, its history, conceptual domain, and application to conservation. In: Hanski I, Gilpin M (eds) Metapopulation biology: ecology, genetics and evolution. Academic Press, San Diego, pp 5–26
- Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. John Wiley and Sons, Inc, New York
- IGME (1979) Mapa hidrogeológico nacional. Explicación de los mapas de lluvia útil, de reconocimiento hidrogeológico y de síntesis de los sistemas acuíferos, Vol. 81, 2nd edn. Instituto Geológico y Minero de España, Ministerio de Industria y Energía, Madrid
- IGN (1999) Mapa de carreteras. Península Ibérica, Baleares y Canarias. Instituto Geográfico Nacional, Ministerio de Fomento, Madrid
- IUCN (2004) Red list of threatened species (<http://www.redlist.org>)
- Jiménez-Valverde A, Lobo JM (2006) The ghost of unbalanced species distribution data in geographical model predictions. Divers Distrib 12:521–524
- Legendre P, Legendre L (1998) Numerical ecology, 2nd edn. Elsevier Science, Amsterdam
- Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. Bull Entomol Soc Am 15:237–240
- Levinsky I, Skov F, Svenning JC, Rahbek C (2007). Potential impacts of climate change on the distributions and diversity patterns of European mammals. Biodivers Conserv 16:3803–3816
- Maes D, Bauwens D, de Bruyn L, Anselin A, Vermeersch G, Van Landuyt W, de Knijf G, Gilbert M (2005) Species richness coincidence: conservation strategies based on predictive modelling. Biodivers Conserv 14:1345–1364
- Maiorano L, Faluccci A, Boitani L (2006) Gap analysis of terrestrial vertebrates in Italy: priorities for conservation planning in a human dominated landscape. Biol Conserv 133:455–473
- Montero de Burgos JL, González-Rebollar JL (1974) Diagramas bioclimáticos. ICONA, Madrid
- Muñoz AR, Real R, Barbosa AM, Vargas JM (2005) Modelling the distribution of Bonelli's eagle in Spain: implications for conservation planning. Divers Distrib 11:477–486
- Palomo LJ, Gisbert J (2002) Atlas de los mamíferos terrestres de España. Dirección General de Conservación de la Naturaleza-SECEM-SECEMU Madrid
- Pearlstine LG, Smith SE, Brandt LA, Allen CR, Kitchens WM, Stenberg J (2002) Assessing state-wide biodiversity in the Florida Gap Analysis project. J Environ Manage 66:127–144
- Pulliam HR (1988) Sources, sinks and population regulation. Am Nat 132:652–661
- Real R (1992) Las tendencias geográficas de la riqueza específica. Vargas JM, Real R, Antúnez A (eds) Monografías de Herpetología, 2:85–94. Asociación Herpetológica Española, Valencia
- Real R, Barbosa AM, Vargas JM (2006a) Obtaining environmental favourability functions from logistic regression. Environ Ecol Stat 13:237–245
- Real R, Estrada A, Barbosa AM, Vargas JM (2006b) Aplicación de la lógica difusa al concepto de rareza para su uso en *Gap Analysis*: el caso de los mamíferos terrestres en Andalucía. Serie Geográfica 13: 99–116

- Real R, Vargas JM, Antúnez A (1993) Environmental influences on local amphibian diversity: the role of floods on river basins. *Biodivers Conserv* 2:376–399
- Rey Benayas JM, de la Montaña E (2003) Identifying areas of high-value vertebrate diversity for strengthening conservation. *Biol Conserv* 114:357–370
- Robertson MP, Villet MH, Palmer AR (2004) A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. *Divers Distrib* 10: 461–474
- Rouget M, Richardson DM, Cowling RM (2003) The current configuration of protected areas in The Cape Floristic Region, South Africa—reservation bias and representation of biodiversity patterns and processes. *Biol Conserv* 112:129–145
- Scott JM, Csuti B, Estes JE, Anderson H (1989) Status assessment of biodiversity protection. *Conserv Biol* 3:85–87
- Scott JM, Davis F, Csuti B, Noss R, Butterfield B, Groves C, Anderson H, Caicco S, D'erchia F, Edwards Jr TC, Ulliman J, Wright RG (1993) *Gap Analysis: a geographic approach to protection of biological diversity*. *Wildlife Monogr* 123:1–41
- Sierra R, Campos F, Chamberlin J (2002) Assessing biodiversity conservation priorities: ecosystem risk and representativeness in continental Ecuador. *Landsc Urban Plan* 59:95–110
- Thuiller W, Lavorel S, Sykes MT, Araújo MB (2006) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Divers Distrib* 12:49–60
- Thuiller W, Richardson DM, Pysek P, Midgley GF, Hughes GO, Rouget M (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biol* 11: 2234–2250
- US Geological Survey (1996) *GTPO30 Land processes distributed active archive center* (<http://edcdaac.usgs.gov/gtopo30/gtopo30.asp>)
- Van Teeffelen AJA, Cabeza M, Moilanen A (2006) Connectivity, probabilities and persistence: comparing reserve selection strategies. *Biodivers Conserv* 15:899–919
- Vázquez LB, Gaston KJ (2004). Rarity, commonness, and patterns of species richness: the mammals of Mexico. *Global Ecol Biogeogr* 13:535–542
- Whittaker RJ, Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ (2005) Conservation biogeography: assessment and prospect. *Divers Distrib* 11:3–23
- Yip JY, Corlett RT, Dudgeon D (2004) A fine-scale Gap Analysis of the existing protected area system in Hong Kong, China. *Biodivers Conserv* 13:943–957