Concepts & Synthesis

Favourability: concept, distinctive characteristics and potential usefulness

Pelayo Acevedo[1,2,3,4*] and Raimundo Real[1]

1- Biogeography, Diversity, and Conservation Research Team, Department of Animal Biology, Faculty of Sciences, University of Malaga, E-29071 Malaga, Spain.

2- CIBIO/UP – Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, Campus Agrario de Vairão, 4485-661 Vairão, Portugal.

3- Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), E-13071 Ciudad Real, Spain.

4- Centre de Recerca en Sanitat Animal (CReSA), UAB-IRTA, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.

*Corresponding author: pacevedo@uma.es (P. Acevedo). Tel.: +34 952132383; fax: +34 952131668.

Abstract

The idea of analysing the general favourability for the occurrence of an event was presented in 2006 through a mathematical function. However, even when favourability has been used in species distribution modelling, the conceptual framework of this function is not yet well perceived among many researchers. The present paper is conceived for providing a wider and more in-depth presentation of the idea of favourability; concretely we aimed to clarify both the concept and the main distinctive characteristics of the favourability function, especially in relation to probability and suitability, the most common outputs in species distribution modelling. As the capabilities of the favourability function go beyond species distribution modelling, we also illustrate its usefulness for different research disciplines for which this function remains unknown. In particular, we stressed that the favourability function has potential to be applied in all the cases where the probability of occurrence of an event is analyzed, such as, for example, habitat-selection or epidemiological studies.

Keywords: epidemiology, favourability function, habitat selection, habitat suitability, probability of occurrence, species distribution modelling.

Brief introduction

The favourability function – defined in Real et al. (2006) – assesses the variation in the probability of occurrence of an event in certain conditions with respect to the overall prevalence of the event. Consequently, it has potential to be applied in the cases where the probability of occurrence of an event is analyzed, such as species distribution modelling (Franklin 2009) or, among others, habitat-selection and epidemiological studies (Manly et al. 2002; Pfeiffer et al. 2008). In addition, it can be applied to all methods able to produce probability; although favourability was usually calculated from probabilities yielded by logistic regressions (Hosmer and Lemeshow 2000), favourability values can be derived, for example, from probabilities obtained using additive or Bayesian models (Hastie and Tibshirani 1990; Bernardo and Smith 2000). So far the concept and the main distinctive characteristics of favourability are not well perceived among many researchers, especially for disciplines different from species distribution modelling. The main aim of this study was to carry out a broader presentation of the favourability concept and to illustrate the usefulness of the favourability function to the scientific community.

Defining the favourability idea and function

Pierre-Simon Laplace defined probability in his first general principle about probability calculation as the ratio of the number of favourable cases to the whole number of possible cases (Laplace 1825, page 12). In this way, the concept of favourability was implicit from the beginning in that of probability. If all cases are equally, and totally, favourable – or unfavourable – then this ratio depends on the prevalence of the event. In his second principle Laplace stated that different cases could differ in possibility, conferring gradualness to the denominator in the probability ratio. However, it can be argued that the concept of possibility is not appropriate to be given a continuous and gradual value, as an event is completely possible even when it is highly unlikely, i.e., the event is completely possible if it is not completely impossible. Laplace's second principle makes sense, however, if it is applied instead to the numerator of the probability ratio, so pointing to a quality of each case which may be appropriately called favourability and may take continuous values that can be constrained to range between 0 and 1. Thus, the probability of an event occurring in certain conditions combines the general prevalence of the event and the local favourability for that event occurring precisely in those conditions. Favourability may thus be obtained as a function of probability and prevalence.

The favourability function was conceptually conceived in this context to assess and remove the effect of prevalence on each probability value. With the favourability function, output values for different events are levelled in relation to each event’s prevalence in the dataset. That is, a favourability value of 0.5 for an event in certain locality or conditions indicates that the probability for the event’s occurrence in that locality or condition is the same as the overall prevalence of the event in the dataset, i.e., local conditions neither increase nor decrease the probability of occurrence with respect to what could be expected according to mere prevalence, thus denoting neutral local favourability. Consequently, local favourability values higher than 0.5 indicate characteristics that favour the event’s occurrence and values below 0.5 denote detrimental conditions for the event, regardless of the event prevalence.

The mathematical rationale for the favourability function is presented in Real et al. (2006). Basically, the favourability function may take a form similar to the logistic probability in which the effect of the event’s prevalence is mathematically eliminated in the *logit* of a logistic regression equation. Among other forms, favourabilities (*F*) may be directly derived from probabilities (P) yielded by any mathematical method in the following way:

$$F = \frac{\frac{P}{(1-P)}}{\frac{n_1}{n_0} + \frac{P}{(1-P)}}$$

being n_1 and n_0 the respective number of positive and negative samples for a given event in the dataset, and, when using logistic regression, $P = \frac{e^y}{1+e^y}$; where e is the basis of the natural logarithm, and y is a regression equation of the form: $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$; where α is a constant and $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients of the n predictor variables $x_1, x_2, \ldots, x_n$ (Tabachnick and Fidell 1996, page 127).

It must be stressed that the favourability function does not provide a probability output independent of the sample prevalence, but a measure of the degree to which local conditions lead to a local probability higher or lower than that expected at random, being this random probability defined by the overall prevalence of the event, which in turn is what must be expected if maximum entropy is assumed (Real et al. 2006). Local probability depends both on the response of the dependent variable to the predictors and on the overall prevalence of the event (e.g. Cramer 1999), whereas favourability values depend only on the response of the dependent variable to the predictors in the study area (see below). Thus, favourability is not aimed at replacing probability but at complementing it, by providing, for example, a comparable measure of the response of each event to the predictors for events differing in prevalence. In this way, favourability may be used to detect, for example, conditions that favour in the same degree the occurrence of a rare disease and a common seasonal flu, even when the probability of suffering them differs due to their different prevalence. However, this concept was recently misunderstood as a way to obtain the probability of occurrence when event prevalence differs from 50% (Albert and Thuiller 2008).

Sample prevalence dependence: a statistical assessment for probability, favourability and suitability outputs

To bring to light the sample prevalence dependence in the probability, favourability and suitability outputs, we built a virtual species with a prevalence of 20% which was designed to logistically respond to an environment defined by only one environmental variable on a virtual landscape composed of 1000 units (i.e. 200 presences). From the species distribution, two samples of 125 territorial units with contrasted prevalences – one with 20% and another with 80% (see Figure 1) – were randomly extracted. Each sample was modelled using different procedures.

We compared the output of the favourability function (Real et al. 2006) with those resulting from probability and suitability obtained with Ecological Niche Factor Analysis (ENFA; Hirzel et al. 2002) and Maximum Entropy approach (MaxEnt; Phillips et al. 2006). Probabilities were obtained using logistic regression (Hosmer and Lemeshow 2000), and they were included as inputs into the favourability function (Real et al. 2006). ENFA was run in Biomapper 4.0 (freely available at http://www.unil.ch/biomapper/) with the median algorithm (Hirzel et al. 2008). MaxEnt version 3.1 (freely available at http://www.cs.princeton.edu/~schapire/maxent/) was run with default parameter values and the logistic output format (Elith et al. 2011).

Results of all models were projected to the whole landscape (Figure 2) and outputs obtained from samples with a prevalence of 20% and those with a prevalence of 80% were graphically compared (Figure 3), so that outputs independent from prevalence should yield a line close to the identity line. Figures 2 and 3 illustrate how the favourability function was the method most independent of prevalence, since quite similar results were obtained from samples with contrasted prevalences. But this did not occur for probability or the suitability outputs obtained from ENFA and MaxEnt. Slight mismatches observed with respect to the diagonal for the favourability function in Figure 3 are due to slightly different detected responses to the variable in each randomly selected sample. Our results contradicted those reported by Albert and Thuiller (2008) in which favourability was suggested to be biased by sample prevalence, but they are consistent with previous studies and with the conceptual framework behind favourability (see Real et al. 2006).

The modelled response of the virtual species to the variable was the same (and correct) for the probability and favourability functions, being the differences in the results only due to the effect of sample prevalence on the probability outputs. Two different responses of the species were obtained for ENFA and MaxEnt. ENFA was not able to detect the subjacent monotonic response of the species to the environment. With both samples, ENFA identified Gaussian responses (e.g. Acevedo et al. 2007) and the maximum response value was obtained in both cases because in this procedure suitability values are rescaled (Hirzel et al. 2002). For these reasons, two different relationships were established between suitability values derived from the different samples (one in each tail of the curve), but none of them was close to the identity line. The results obtained for MaxEnt show that quite different responses were modelled on each sample, which may be related to the fact that MaxEnt produces a number of indices that are not directly related to the probability of occurrence (Royle et al. 2012). Thus, with ENFA and MaxEnt the response of the species to the environment cannot be segregated from the effect of sample prevalence on the suitability output. The results here provided show that probability and suitability are biased in their outputs when working with samples – of the same species – differing in prevalence, which is not the case with favourability.

The concept of favourability for biogeographers

Many researchers working with species distribution models produce maps showing continuous gradients of how environmental characteristics are appropriate – in a broad sense – for a target species (Guisan and Thuiller 2005). Model's predictions can be either considered as gradients or used only to classify localities as appropriate on inappropriate, but the latter option limits the informative capacity of the model. Thus, when models are aimed to guide conservation strategies they are more useful as continuous gradients (Barbosa et al. 2010).

Nevertheless, the continuous model's predictions should be levelled in order to determine those characteristics in the study area which actually favour the species presence. That is what the favourability function does. So, using the favourability function those localities with environmental conditions that favour the presence of the species ($F>0.5$) can be easily distinguished from those with detrimental characteristics ($F<0.5$) for its presence. This makes the favourability function particularly useful in conservation biology, for example, to identify expansion routes of invasive species (Muñoz and Real 2006; Nielsen et al. 2008), or to identify areas where a species may be more vulnerable to habitat or climate changes (e.g. Guitiérrez-Illán et al. 2010).

The concept behind the favourability function was also raised by biogeographers working with probability (Liu et al. 2005; Jiménez-Valverde and Lobo 2007) and profile methods (see Hirzel et al. 2006). A rationale conceptually close to favourability was used to reclassify the suitability scores obtained with the ENFA (Hirzel et al. 2002). The suitability score over which the model predicts more presences than expected by chance can be used as a threshold to identify the localities that actually are favourable for the target species. Liu et al. (2005) and Jiménez-Valverde and Lobo (2007), for example, proposed several methods to obtain the best threshold to split the localities into two categories, which tend to locate the threshold near the point where probability equals prevalence. These categories could appropriately be called favourable and unfavourable, as they represent probabilities higher or lower than prevalence, respectively. So, the determination of those conditions enhancing the probability of species presence over the probability expected by chance – the concept behind the favourability function – is widely considered sound in biogeography. The favourability function not only provides the favourability threshold more easily (F=0.5) but also provides information about the degree to which every locality is favourable.

In addition, the favourability function has other distinctive characteristics that make it especially applicable in conservation biogeography and other research disciplines.

Main distinctive characteristics of favourability values

The main distinctive characteristics of the favourability function in relation to common outputs in other modelling techniques (probability and/or suitability) are summarised in the following five points:

1- Given the definition of favourability as the assessment (between 0 and 1) of the variation in the probability of occurrence of an event in certain conditions with respect to the overall prevalence of the event, there is only one way of obtaining favourability values from probabilities and prevalences. In this aspect favourability differs from suitability, as for each modelling method, suitability is an idiosyncratic way of ranking local sites according to their capacity to hold the species that is not directly related to probability (e.g. Guisan and Zimmermann 2000). This is why different modelling techniques produce differing suitability values with the same dataset, but all ways of obtaining favourability should yield the same favourability values from the same dataset.

2- Favourability values – like probability values and unlike suitability – are interpretable in absolute terms, as they indicate how local presence’s probability differs from that expected by chance in the whole sample. However, suitability values, such as, for example, those derived from ENFA, ensemble forecasting approaches (Araújo and New 2007) or some of the outputs from MaxEnt (Phillips et al. 2006), are only relative and therefore uninformative in absolute terms. For example, the suitability value assigned to each focal locality in ENFA for each factor axis is based on a count of all localities with species presence that lay as far or farther apart from the median than the focal locality (Hirzel et al. 2002). This count is normalized in such a way that the suitability index always ranges from zero to one (see Figure 3). In ensemble forecasting suitability values are the result of merging, in some occasions, methods generating probability with others that yield suitability scores (e.g. Thuiller et al. 2009). Consequently, the suitability values obtained by these kinds of methods cannot be easily interpreted, especially when comparing different models, even if they are calibrated against a dataset with equal species prevalence.

3- Favourability values – like suitability values and unlike well calibrated probability – are dependent on the extent of the study area if modifying the extent entails a modification of the species prevalence. Conceptually, a locality where the probability of finding a species is intermediate should be considered unfavourable for the species in the context of the core of the species range, but highly favourable in the context of a huge area where the species range represents a small portion. The favourability function quantifies this difference of consideration of a same probability value according to the differing prevalence of the species in – and normally due to the different extent of – the background area. This implies that favourability (and suitability) values obtained from models built in different study areas should be compared with these characteristics in mind, as each favourability is relative to its own study area (Barbosa et al. 2009).

4- The inherent quality of the favourability function of being expressed in relation to the event’s prevalence in the study area enables direct comparison and combination when several species are involved in the analytical design. For example, this is needed when using models for multiple species as a basis for defining relevant areas for conservation (Estrada et al. 2008), which cannot be built based on probability values because these are higher in common than in rare species, so the values for the former would prevail over those for the latter.

5- In addition, but closely related to point 4, favourability values – unlike probability or suitability values – can be regarded as the degree of membership of the localities to the fuzzy set of sites with conditions that are favourable for the species, which enables the easy application of fuzzy logic operations to distribution modelling (e.g. Robertson et al. 2004). Fuzzy logic operations expand the potential of the favourability function for comparison between models. For example, this function and the fuzzy indices derived from it were successfully used to study the biogeographical relationships in predator-prey systems (Real et al. 2009) and also between native and exotic sympatric species (Acevedo et al. 2010). Similarly, the transferability of models to other times, for example in climate change scenarios (Real et al. 2010; Acevedo et al. 2012) or land use changes (Acevedo et al. 2011), or to different resolution scales (Barbosa et al. 2010), can be better assessed with the combined use of the favourability function and fuzzy logic. For instance, an overall assessment of expected modification in species’ distribution in climate change scenarios can be obtained using fuzzy logic, since the favourability forecasted for a given species in the future can be deconstructed into the percentage that is expected to increase, overlap, be maintained and shift in relation to its favourability in the present (Real et al. 2010). On this point, it is worth mentioning that the spatial-temporal transference of models is risky and invites caution and careful considerations (e.g. Jiménez-Valverde et al. 2011). There have been increasing concerns about the use of correlative models for projecting species distribution into novel situations such as new territories or future climate change scenarios (e.g. Sutherst and Bourne 2009; Webber et al. 2011). Nevertheless, it should be noted that the concepts of favourability, probability and suitability are equally applicable to mechanistic and correlative modelling approaches, as they refer to the output which is produced by the models, and not to the inference method used to obtain these outputs.

The potential of the favourability function

To date, applications of the favourability function are nearly restricted to species distribution modelling , which is likely because the main research discipline of the developers was biogeography. Taking into account the concept behind this function and the distinctive characteristics of favourability values previously described, and similarly to other logistic models (e.g. Keating and Cherry 2004 and references therein), the potential of the favourability function in other research disciplines is high. The concept of favourability is quite relevant, for instance, in habitat-selection studies for determining the sampling units in which the process under study, e.g., nesting success, is favoured, i.e., those sampling units with a higher probability of event occurrence than expected by chance. For processes differing in prevalence favourability values provide comparable measures of the response of each process to the predictors; for example, with the favourability function it is possible to quantify in the same terms the degree to which the local environmental characteristic are favouring bird nesting occurrence and nesting success for each sampling unit (see Amici et al. 2009). In another example, Real et al. (2009) used the favourability function to identify areas autoecologically favourable for the rare Iberian lynx (*Lynx pardinus*) but autoecologically unfavourable for its common staple prey the wild rabbit (*Oryctolagus cuniculus*), so highlighting the lack of trophic resources in parts of the potential range for a critically endangered species. This would be unattainable with probabilities, as the very common, and prevalent, rabbit tend to yield higher values of probability of occurrence than the scarce lynx, even in localities where rabbit densities are unable to support lynx populations. The concept of favourability and its distinctive characteristics are also promising in epidemiology. Epidemiological studies in wildlife try to identify risk factors that increase the frequency of pathogens (e.g. Vicente et al. 2007) and to create risk maps in which the probability of their transmission is shown (e.g. Rochlin et al. 2011). Including the concept of favourability in these studies entails two main advantages. First, those populations (or individuals, it depends on the sampling unit used in the study) in which the probability of presence of the pathogen is higher than expected by chance (F>0.5) in the study area can be identified. These are key populations for disease control and monitoring (Mörner et al. 2002). Similarly, those values of a given risk factor over which the probability for the presence of a pathologic condition is higher than expected by chance can also be identified. For example, Fernández et al. (2000) studied the relationships between coronary artery anomalies and aortic valve morphology obtaining that the probability of occurrence of anomalous coronary artery patterns increases continuously according to the degree of deviation of the aortic valve from its normal (tricuspid) design according, for example, to the following logit expression: y= -2.0976 + 0.3136*group (where group referred to six groups of valve conditions into which the continuous spectrum of aortic valve morphology was divided, from 0=tricuspide to 5=bicuspide). By including the favourability concept in this study, the authors could have determined over which aortic valve morphotype (from 0 to 5) the probability of occurrence of the anomalous coronary pattern was higher than expected by chance, and therefore, the anomalies were being promoted. So, given the expression previously reported and considering that 220 out of 968 of the coronary artery patterns were anomalous, a favourability value higher than 0.5 is obtained for valve morphotype value higher than 2.7, so these are the values that actually favour the anomalous coronary pattern.

Secondly, favourability is also a promising function for biogeography of diseases where interactions among – hosts and vectors – species differing in prevalence are relevant (Peterson 2008) and where time series are usually available (e.g. Boadella et al. in press). As previously stated , the use of the favourability function and fuzzy logic allows direct comparisons and/or combinations between more than one model (host, vector and pathogen), which enables a more complete assessment of the distribution of the disease transmission risks (see Estrada-Peña et al. 2008) by obtaining reliable multi-host, multi-pathogen and/or multi-scenario risk maps. In this context, Boadella et al. (in press) analyzed the factors associated to the detection of a group of parasites – *Trichinella* spp. – infecting wild boar (*Sus scrofa*). The inclusion of the idea of favourability in this study (first time in spatial epidemiology) was needed to combine the risks obtained for each of the 12 years included in the study in order to obtain two proxies of the risk for *Trichinella* spp. infection for the study period. One index was defined to identify areas where the conditions for *Trichinella* spp. infection were favourable during the study period (endemic areas for the parasites), and another was designed to determine the global distribution of these parasites during the study period.

So, the combined used of the favourability function and fuzzy logic operations enabled a more-in-depth assessment of the risks for a given parasite group in a multi-scenario context.

Concluding remarks

The main aim of this study was to carry out a broad presentation of the favourability concept and the favourability function to the scientific community. In addition to the studies in conservation biogeography, here we highlighted the usefulness of this function in two other disciplines (habitat-selection and epidemiology). We think that its capabilities go beyond these examples, and that the examination of the concept and the exploration of its usefulness for other disciplines will prove to be helpful in all cases where the probability of occurrence of an event is analyzed.

Acknowledgements

The authors acknowledge funding from Spanish Plan Nacional de Investigación and FEDER CGL2009-11316/BOS. P. A. was supported by the Vicerrectorado de Investigación of the University of Malaga and currently by a Beatriu de Pinós fellowship funded by Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa, of the Generalitat de Catalunya, and the COFUND Programme - Marie Curie Actions under 7th Marc Programme of the European Community.

References

Acevedo P, Cassinello J, Hortal J, Gortázar C (2007) Invasive exotic aoudad (*Ammotragus lervia*) as a major threat to native Iberian ibex (*Capra pyrenaica*): a habitat suitability model approach. Divers Distrib 13:587-597.

Acevedo P, Farfán MA, Márquez AL, Delibes-Mateos M, Real R, Vargas JM (2011) Past, present and future of wild ungulates in relation to changes in land use. Landscape Ecol 26:19-31.

Acevedo P, Jiménez-Valverde A, Melo-Ferreira J, Real R, Alves PC (2012) Parapatric species and the implications for climate change studies: a case study on hares in Europe. Global Change Biol doi: 10.1111/j.1365-2486.2012.02655.x

Acevedo P, Ward AI, Real R, Smith GC (2010) Assessing the biogeographical relationships of ecologically related species using favourability functions: a case study on British deer. Divers Distrib 16:515-528.

Albert CH, Thuiller W (2008) Favourability functions versus probability of presence: advantages and misuses. Ecography 31:417-422.

Amici A, Pelorosso R, Serrani F, Boccia L (2009) A nesting site suitability model for rock partridge (*Alectoris graeca*) in the Apennine Mountains using logistic regression. Ital J Anim Sci 8:751-753.

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

Barbosa AM, Real R, Vargas JM (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 AM, Real R, Vargas JM (2010) Use of Coarse-Resolution Models of Species' Distributions to Guide Local Conservation Inferences. Conserv Biol 24: 1378-1387.

Bernardo JM, Smith AFM (2000) Bayesian Theory. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken.

Boadella M, Barasona JA, Pozio E, Montoro V, Vicente J, Gortázar C, Acevedo P (in press) Spatial-temporal trend and risk factors for Trichinella sp. infection in wild boar (*Sus scrofa*) populations of central Spain: a long-term study. Int J Parasitol

Cramer JS (1999) Predictive performance of binary logit model in unbalanced samples. J R Stat Soc D 48:85-94.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologist. Divers Distrib 17:43-57.

Estrada A, Real R, Vargas JM (2008) Using crisp and fuzzy modelling to identify favourability hotspots useful to perform gap analysis. Biodivers Conserv 17:857-871.

Estrada-Peña A, Acevedo P, Ruiz-Fons F, Gortázar C, de la Fuente J (2008) Evidence of the Importance of Host Habitat Use in Predicting the Dilution Effect of Wild Boar for Deer Exposure to *Anaplasma* spp. PLoS ONE **3**(8): e2999. doi:10.1371/journal.pone.0002999.

Fernández MC, Durán AC, Real R, López D, Fernández B, de Andrés AV, Arqué JM, Gallego A, Sans-Coma V (2000) Coronary artery anomalies and aortic valve morphology in the Syrian hamster. Lab Ani-UK 34:145-154.

Franklin J (2009) Mapping species distributions. Spatial inference and prediction. Cambridge University Press, Cambridge.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993-1009.

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

Gutiérrez-Illán J, Gutiérrez D, Wilson RW (2010) The contributions of topoclimate and land cover to species distributions and abundance: fine-resolution tests for a mountain butterfly fauna. Global Ecol Biogeogr 19:15-173.

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. Chapman & Hall/CRC

Hirzel AH, Braunisch V, Le Lay G. Hausser J, Perrin N (2008) BIOMAPPER 4.0. Laboratoy of Conservation Biology, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. Available from http://www.unil.ch/biomapper.

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

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecol Model 199:142-152.

Hosmer DW, Lemeshow S (2000) Applied Logistic Regression. Wiley Interscience, New York.

Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either- or presence-absence. Acta Oecol 31:361-369.

Jiménez-Valverde A, Peterson AT, Soberón J, Overton J, Aragón P, Lobo J M (2011) Use of niche models in invasive species risk assessments. Biol Invasions 13:2785-2797.

Keating KA, Cherry S (2004) Use and interpretations of logistic regression in habitat-selection studies. J Wildlife Manage 68:774-789.

Laplace PS (1825) Essai philosophique sur les probabilités. Bachelier, Paris.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385-393.

Manly BFJ, McDonald LL, Thomas DL, McDonald TL, Erickson W (2002) Resource selection by animals: statistical design and analysis for field studies. Second edition. Kluwer Press, New York.

Mörner T, Obendorf D, Artois M, Woodford M (2002) Surveillance and monitoring of wildlife diseases. Rev Sci Tech Off Int Epiz 21:67-76.

Muñoz AR, Real R (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. Divers Distrib12:656-665.

Nielsen C, Hartvig P, Kollmann J (2008) Predicting the distribution of the invasive alien *Heracleum mantegazzianum* at two different spatial scales. Divers Distrib 14:307-317.

Peterson AT (2008) Biogeography of diseases: a framework for analysis. *Naturwissenschaften* 95:483-491.

Pfeiffer D, Robinson T, Stevenson M, Stevens K, Rogers D, Clements A (2008) Spatial analysis in epidemiology. Oxford University Press, New York.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231-259.

Real R, Barbosa AM, Rodríguez A, García FJ, Vargas JM, Palomo LJ, Delibes M (2009) Conservation biogeography of ecologically interacting species: the case of the Iberian lynx and the European Rabbit. Divers Distrib 15:390-400.

Real R, Barbosa AM, Vargas JM (2006) Obtaining environmental favourability functions from logistic regression. Environ Ecol Stat 13:237-245.

Real R, Márquez AL, Olivero J, Estrada A (2010) Are species distribution models in climate warming scenarios useful for informing emission policy planning? An uncertainty assessment using fuzzy logic. Ecography 33:304-314.

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.

Rochlin I, Turbow D, Gomez F, Ninivaggi DV, Campbell SR (2011) Predictive Mapping of Human Risk for West Nile Virus (WNV) Based on Environmental and Socioeconomic Factors. PLoS ONE **6**(8): e23280.doi:10.1371/journal.pone. 0023280.

Royle JA, Chandler RB, Yackulic C, Nichols JD (2012) Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods in Ecology & Evolution doi: 10.1111/j.2041-210X.2011.00182.x

Sutherst RW, Bourne AS (2009) Modelling non-equilibrium distributions of invasive species: a tale of two modelling paradigms. Biol Invasions 11:1231-1237.

Tabachnick BG, Fidell LS (1996) Using multivariate analysis, 3rd edn. HarperCollins College Publishers, Northridge, California.

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. Ecography 32:369-373.

Vicente J, Hofle U, Garrido JM, Fernandez-de-Mera IG, Acevedo P, Juste R, Barral M, Gortázar C (2007) Risk factors associated with the prevalence of tuberculosis-like lesions in fenced wild boar and red deer in south central Spain. Vet Res 38:451-464.

Webber BL, Yates CJ, Le Maitre DC, Scott JK, Kriticos DJ, Ota N, McNeill A, Le Roux JJ, Midgley GF (2011) Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. Divers Distrib 17:978-1000.

Figure 1. Virtual landscape composed of 1000 units with (a) an environmental variable ranging from 0 (white) to 1000 (black); (b) a virtual species distribution (black circles show presences and white ones absences); and random samples of the species with prevalence of 20% (c) or 80% (d).

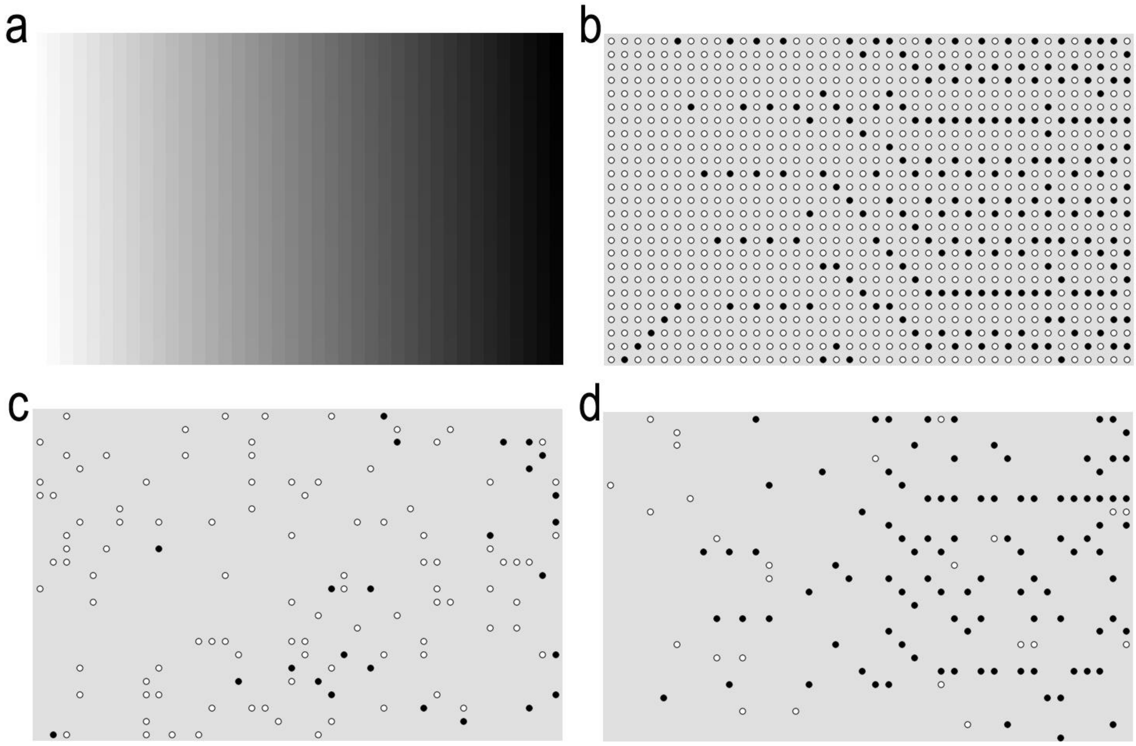

Figure 2. Predictions obtained for each sample (20% or 80%) and modelling procedure (probability, favourability, and suitability from ENFA and MaxEnt).

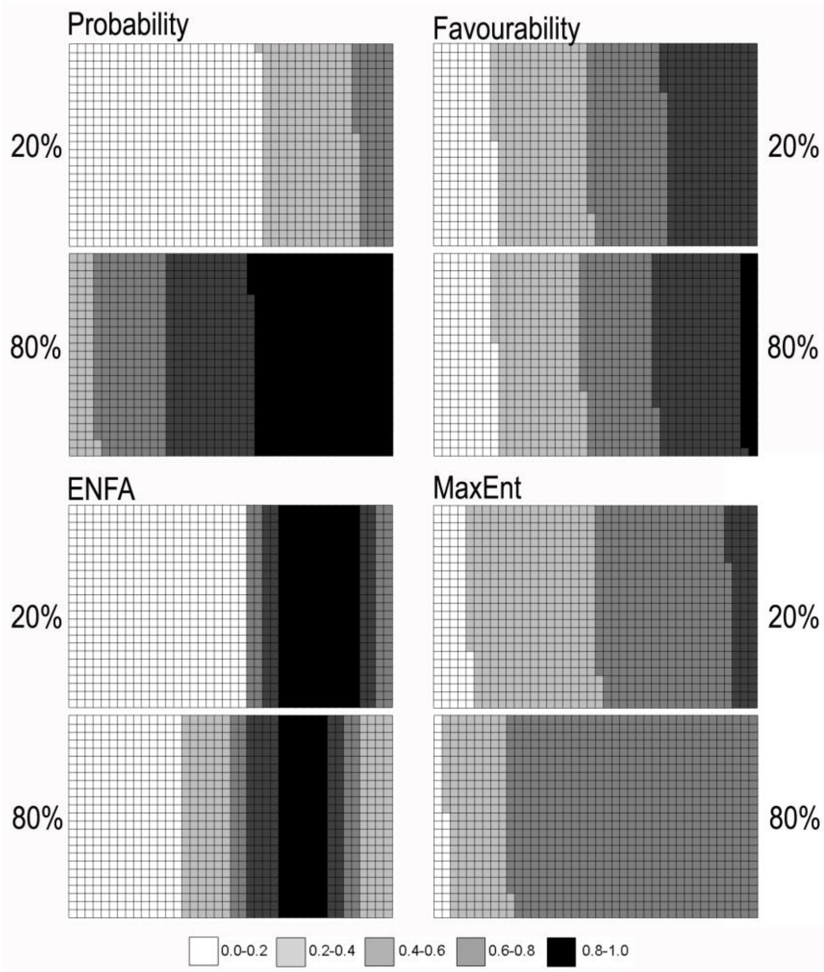

Figure 3. Comparison between outputs of models developed from a sample with prevalence of 20% against others from a sample with 80%. Lines are representing outputs of favourability (black-thick), probability (grey-thick), and MaxEnt (grey thin). Results from ENFA are represented with grey circles. The black-thin line represents de identity.