University of Naples “Federico II”
Department of Agriculture
Ph.D. in Agriculture and Food Sciences
XXX Cycle

A methodological assessment of Species Distribution Models as tools to plan species conservation and niche modelling

Supervisor
Prof. Danilo Russo

Ph.D. candidate
Sonia Smeraldo

Contents

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter I

1. General introduction and research objectives

1.1 . Species distribution models and ecological niche theory

The knowledge of a species distribution has always been essential to understand species' ecology and its conservation needs (Franklin, 2010). In the last decades, predictive modelling of species distribution has become a widely used technique and an increasingly important tool in many fields of natural and biological sciences to address various issues in ecology, biogeography, evolution, conservation biology and climate change research (Guisan & Thuiller, 2005; Elith & Leathwick, 2009; Russo et al., 2016). The quantification of species-environment relationships represented the core of predictive geographical modelling in ecology (Guisan & Zimmermann, 2000).

"Species distribution models" (SDMs), "ecological niche models" (ENMs) or "bioclimatic envelope models" are different names that have been used to describe similar mechanistic or correlative models which infer the ecological requirements of the species from field observations on the basis of statistically or theoretically derived response surfaces (Peterson, 2006; Kearney & Porter, 2009; Sillero, 2011).

SDMs is the term most frequently used, especially for correlative models (e.g. Austin, 2002; Guisan & Thuiller, 2005; Pearson, 2007; Lobo et al., 2010) which "characterize the multivariate environmental space delimiting species' distributions, and project this subset of environmental space back onto geography" (Raes, 2012). Many researchers consider the definitions SDMs, ENMs or bioclimatic envelope models mainly as synonyms. However, it is worth to note that some others researchers pointed on the fact that these terms (particularly SDM and ENM) are not fully interchangeable and their differences are not merely semantic (Peterson, 2006; Peterson & Soberón, 2012) but depend on the choice of the set of variables used to predict species' niche and distribution as well as to the focus of the study (i.e. niche quantification versus spatial predictions). As proposed by Araújo & Peterson (2012), bioclimatic envelope models estimate the "multivariate space of climatic variables best matching the observed species' distribution". Raes (2012) interpreted ENMs as restricting the bioclimatic envelope to variables that are meaningful to the ecological niche of the species, without inferring any geographical projection. ENMs "link the envelope to elements of ecological niche theory" instead of simply estimating the bioclimatic envelope. The niche concept has become a central theme in species distribution modelling (Kearney, 2006; Soberon, 2007; Pearman et al., 2008) and has been used and defined in many different ways (Chase & Leibold, 2003).

The Grinnellian theory embodies the idea of the niche as a subdivision of the habitat containing the environmental conditions that enable individuals of a species to survive and reproduce, based on broad-scale variables (climate) that are not affected by species density (Hirzel & Le Lay, 2008; Wiens et al., 2009).

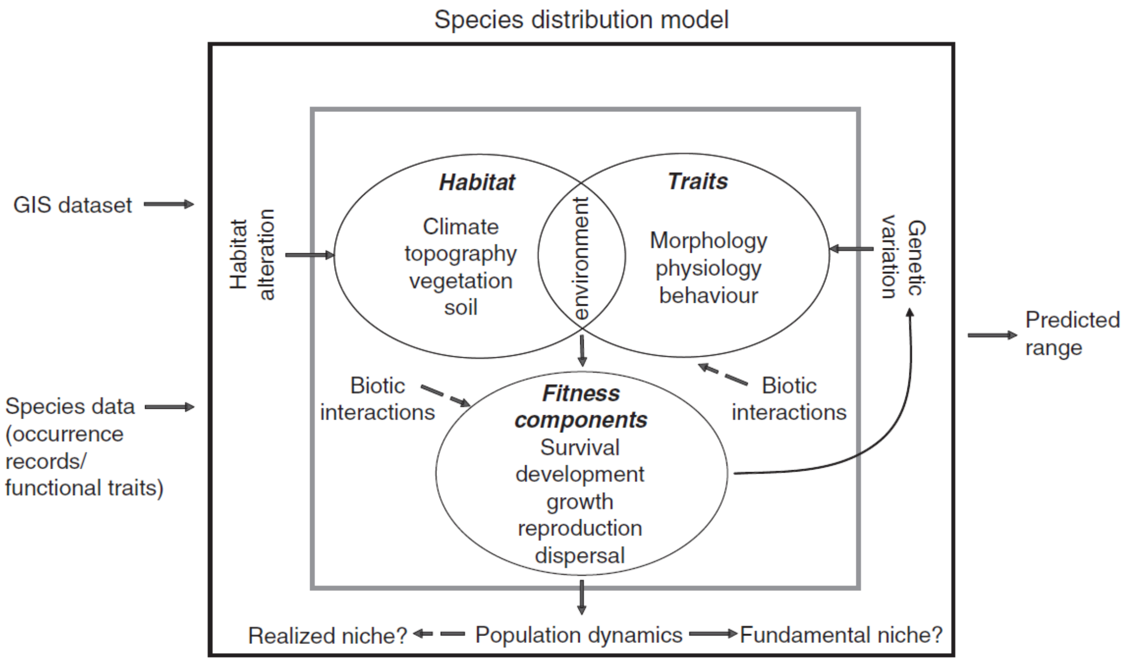


Figure1 Major ecological processes captured by a species distribution model (SDM). Modified picture by Kearney & Porter (2009).

In the context of SDMs it is considered in the Hutchinsonian manner. Hutchinson (1957) defined the species "fundamental niche" as the n-dimensional volume in the environmental space where a species can maintain a viable population and persist along time. When a species does not occupy its entire fundamental niche due to niche exclusion by competition, Hutchinson used the term "realized niche". However, the species can be absent from suitable habitats for historical reasons or due to limitations in its ability to disperse to those habitats (Pulliam, 2000; Holt, 2003). Here, we prefer to use the term SDM because this unifies the niche concept with its geographical projection (Guisan & Zimmermann, 2000; Kearney, 2006; Soberón & Nakamura, 2009; Wiens et al., 2009; Sillero, 2011). In fact, the development of a SDM can be conceived of as first constructing an ecological niche model, and then projecting it onto space to infer a species potential distribution (Raes, 2012). The aim of SDMs is to predict the probability of occurrence of a given taxon, given a set of environmental variables (climate, topography, land cover, etc.) that are assumed to be related to the distribution and habitat preferences of the taxon under study. A habitat suitability map (HSM) of a given species is then extrapolated on the basis of the conditions observed in known occurrence sites (Elith & Leathwick, 2009).

SDMs have been described as estimating the fundamental (potential) niche, realized (actual) niche, the multivariate species niche (Rotenberry et al., 2006), or, when conditioned only on climate variables, the "climatic niche" (Franklin, 2010). It should be noted, however, that any defined ecological niche space derived from the observed distribution of species in geographical space is, at best a realized niche. The full extent of a species' fundamental niche cannot be revealed by the environmental conditions at observed collection localities. Estimation of the fundamental niche can only be achieved by experimental studies and physiological models (Colwell & Rangel, 2009; Carretero & Sillero, 2016). This limitation should be kept in mind while interpreting any correlative model derived from observed collection localities and the abiotic conditions at those localities.

Ecological niche models are developed very differently for correlative and mechanistic SDMs (Dormann et al., 2012). In a correlative approach, we start from species occurrence data (geo-referenced locality records for a species). A presence (or high abundance) record suggests that, at some stage, individuals of that species were able to develop, survive to the adult stage and successfully reproduce in that location (Pulliam, 2000). Spatial conditions also geo-referenced to that site, such as climate or soil, are then inferred to be within that species tolerance range (Austin, 2002). This alters model parameters or coefficients such that other sites with similar conditions are weighted in favour of a prediction of presence, ultimately defining a multivariate space of suitable environmental conditions. Many of the ecological processes and interactions that lead to successful persistence of the species at that site are implicitly captured by statistical analyses (Kearney & Porter, 2009; Figure 1). In a mechanistic (or process-based) approach, functional trait data are linked to GIS (Geographic Information System) data through a model that explicitly captures the key processes by which traits and habitat features interact to determine the species' environment (Dormann et al., 2012). If a species' niche is to be modelled mechanistically to make inference on its potential range, the organism must not enter the model as a point on a map but rather as a set of behavioural, morphological and physiological traits. Both mechanistic and correlative modelling approach may embody advantages as well as critical issues (Kearney & Porter, 2009; Dormann et al., 2012) to the prediction of species distribution.

In this work we will focus on correlative models which have obtained major success in the development of SDMs because they require little knowledge of the mechanistic links between organisms and their environments, which often represent an advantage for rare and poorly studied species as well as for alien species, also characterised by scarce information on their distribution in the introduction range and poorly known adaptation to environments outside the native range.

As we will discuss afterwards, correlative models provide numerous algorithms and variable selection procedures (e.g. regressive or machine-learning based approaches) which can be more easily tailored to fit available data. They may exploit a wide range of proxy spatial data types and may use datasets at a great variety of scales, capturing processes at different scales within a single model. Furthermore, correlative models provide a simple output indirectly representing many different processes and corresponding to a dimensionless habitat suitability index or estimates of probability of occurrence or abundance.

Preparation of environmental and occurrence data usually take much longer than the actual modelling procedure itself. This fast deployment time is probably another advantage of correlative modelling approach. Mechanistic models commonly take a long time to develop, as they often simulate nonlinear dynamics and hence have to deal with issues such as numerical diffusion and time stepping (Dormann et al., 2012). In many circumstances, however, it is desirable to explicitly incorporate potentially range-limiting processes.

For instance, a mechanistic SDM may provide deep understanding of the constraints limiting the distribution and abundance, and may be more robust in the context of environmental and climate change (Davis et al., 1998; Dormann, 2007).

1.2. Framework for modelling species distribution

The following elements are required for modelling and predicting species distributions (Guisan & Zimmermann, 2000; Franklin, 2010):

- A theoretical or conceptual model of the abiotic and biotic factors controlling species distributions in space and time, and at different scales, and the expected form of the response functions;
- Data on species occurrence (location) in geographical space (a measure of presence, habitat use, abundance, or some other property) or expert knowledge about habitat requirements or preferences;
- Digital maps of environmental variables representing those factors (or their surrogates) determining habitat quality, or correlated with it. These are generally derived from remote sensing, from spatial models of environmental processes, or from some other source, and stored in a GIS;
- A model linking habitat requirements or habitat use (species occurrence) to the environmental variables;
- Data and criteria to validate the predictions and a way to interpret error or uncertainty in the analysis.

1.2.1. Conceptual model formulation

The theoretical framework of SDMs is based on the concept of ecological niche (see previous paragraph) and may be split into many sub- phases. At the very start of the study, researchers must define a conceptual model of the system to be simulated based on the existing knowledge on species ecologyand they must clearly define the objectives as well as the multiple hypotheses of the work (Guisan & Theurillat, 2000).

A central and recurrent problem in SDM building is identifying the appropriate scale for modelling (Wiens, 2002). Scale is usually best expressed independently as *resolution* (cell size) and *extent* of the study area, because modelling a large area does not necessarily imply considering a coarse resolution (Guisan & Thuiller, 2005). Understanding the theory and processes driving the observed distribution patterns is very essential to avoid a mismatch between the scale used for modelling and the one at which key processes occur. For instance, interspecific competition can only be detected at a resolution where organisms interact and compete for the same resources (Huston, 2002). Finer resolution usually provides better predictions for fixed or very locally mobile organisms, while for highly mobile species a larger cell size can accountfor larger portions of the landscape that need to be included in order to consider all habitat types likely to be occupied by the species (Jaberg & Guisan, 2001).

Another possible mismatch can occur between the “resolution” at which species data were sampled (e.g. plot size in field surveys, grid size in atlas surveys) and the one at which environmental predictors are available. Optimally, both should be the same, but such coherence is not always possible.

Choosing the geographic extent for the study area is a critical step in SDM building and an inappropriate selection may yield misleading results (Guisan & Thuiller, 2005). Van Horn (2002) illustrates how an overly constrained extent can lead to an incorrect interpretation if only part of an important environmental gradient is sampled, in particular when using political instead of physical boundaries. This restriction of the study area extent may lead to partial SDMs, which do not include the full environmental variation under which the species is known to occur (Raes, 2012). A multiscale hierarchical modelling framework often may provide the solution to this spatial scaling issue. For instance, Gallien et al. (2012) developed a hierarchical approach to improve regional SDM performance while simultaneously accounting for both global and regional information.

The final phases of the theoretical models consist of assessing available and missing data, identifying an appropriate sampling strategy for collecting new data (Hirzel & Guisan, 2002) or for complementing existing sets, as well as to analysing the relevance of environmental predictors for the focal species (Thuiller et al., 2004).

1.2.2. Occurrence data

Species data can be simple presence, presence–absence, or abundance observations based on random or stratified field sampling, or based on observations obtained opportunistically, such as those from natural history collections (Graham et al., 2004). Often, the only data available about occurrence of a species are presence records. Methods for studying species distributions can use these data by analysing information about the environmental conditions at those presence locations. In contrast, data sets that also include species absence records are informative about sampling effort, hence their inference is much more robust to biases in sampling and they can estimate species occurrence probabilities.

An additional complication is the fact that species are often detected imperfectly, even if sessile (Chen et al., 2013). Two common and important types of errors can arise in species occurrence data: false negatives and false positives. The first ones are the most prevalent in ecological surveys and occur when species are missed in searches of occupied sites. The type and strength of the bias induced depends on how detectable a species is (Barve et al., 2011; Guillera-Arroita, 2017). False-positive errors are likely to occur in situations where it is difficult to distinguish between individual species observed, such as species with similar calls (e.g. grey tree frogs in the Eastern US), or morphology (cryptic species) (Ruiz-Gutierrez et al., 2016).

There are also a number of serious potential traps for the data accuracy. First, the occurrence localities may be biased. For example, they are often highly correlated with the nearby presence of roads, rivers or other access pathways (Reddy & D´Avalos, 2003). The location of occurrence sites may also exhibit spatial autocorrelation (e.g., if a researcher collects specimens from several nearby localities in a restricted area). Similarly, sampling intensity and sampling methods often vary widely across the study area (Anderson, 2003). In addition, data accuracy may be affected by transcription errors or lack of sufficient geographic details (especially in older records). Frequently, the number of occurrence sites may be too low to estimate the parameters of the model reliably (Stockwell & Peterson, 2002). Many studies highlighted the negative effects of small sample size in SDMs performance (Loiselle et al., 2008; Wisz et al., 2008; Mateo et al., 2010; Tessarolo et al. 2014; Proosdij et al., 2016). Model performance is known to rapidly decrease for sample sizes smaller than 20 (Stockwell & Peterson, 2002) or 15 (Papeş & Gaubert, 2007), and is dramatically poor for samples sizes smaller than 5 records (Pearson et al., 2007). Contrary to this, some modelling techniques, like MaxEnt algorithm, lead to high model accuracy using small samples (Hernandez et al., 2006).

1.2.3. Environmental variables

The formulation of conceptual models is based on the ecological knowledge of the species and, as a consequence, leads to the selection of a set of conceptually (e.g. physiologically) meaningful explanatory variables for the species predictive model (Guisan & Zimmermann, 2000). Nowadays environmental data are easily obtained from databases found on the Internet or from institutional sources in the form of maps or, more precisely, as digital spatial data (Romero et al., 2016; Petitpierre et al., 2017).

Environmental predictors can exert direct or indirect effects on species, arranged along a gradient from proximal to distal predictors (Austin, 2002), and are optimally chosen to reflect the three main types of influences on the species (modified from Guisan & Zimmermann, 2000; Huston, 2002): *limiting factors*, defined as factors controlling species eco-physiology (e.g. temperature, water, soil composition); *disturbances*, defined as all types of perturbations affecting environmental systems (natural or human-induced); *resources*, defined as all compounds that can be assimilated by organisms (e.g. energy and water). These relationships between species and their overall environment can cause different spatial patterns to be observed at different scales, often in a hierarchical manner (Pearson et al., 2004).

Errors may be present in the variables, perhaps due to errors in data manipulation, or inaccuracies in the climatic models used to generate climatic variables, or interpolation of lower-resolution data.

It is important to consider three particular aspects when building a set of predictors to project SDMs in time or space: *proximality*, *multicollinearity* and *over-parameterization*. Proximality is the use of limiting factors in order to bring the model closer to the real requirements of the species, thus allowing more robust predictions (Austin, 2007; Kearney & Porter, 2009). However, without an priori knowledge on the ecology and physiology of species, the choice of the most proximal variables is not obvious as they may be confounded with other, highly correlated variables. Multicollinearity (i.e. when two or more variables are correlated) can significantly decrease the accuracy of SDM predictions if the correlation matrices of the variables differ between the calibration and projection ranges (Dormann et al., 2008; Braunisch et al., 2013). A common rule of thumb is to avoid correlations between variables where Pearson's correlation |r| is higher than a fixed threshold (e.g.>0.7; Dormann et al., 2013). When several variables are correlated, it should be chosen the variable most proximal to the species' ecology (Austin, 2007; Austin & Van Niel, 2011). Over-parameterization can be the result of fitting a model with too many predictors relative to the number of available observations. It may result in modelling spurious relationships between biological and environmental variables without any ecological and causal relationship, thus potentially reducing transferability (Warren & Seifert, 2011). A common solution is the empirical rule of '1 in 10' (Harrell et al., 1984), i.e. the use of a maximum of one predictor for ten (but preferably 15– 20) species occurrence records.

1.2.4. Methods for the development of SDMs

In the past three decades, scientists have developed numerous models to generate habitat suitability maps based on the relationship between species occurrences and associated environmental variables. Each occurrence locality is simply a georeferenced data denoting a site where the species has been observed, while the environmental variables in GIS format all pertain to the same geographic area, the study area, which has been partitioned into a grid of cells (Phillips et al., 2006).

A major difference between kinds of species distribution data is related to the quality and amount (Li & Wang, 2013). On this account, modelling techniques can be classified in two main groups: "*presence/absence*" and "*presence-only*" methods (Brotons et al., 2004; Barbet-Massin et al., 2012). The first group of methods includes generalised linear models (GLM), generalised additive models (GAM), generalized boosting models (GBM), multivariate adaptive regression splines (MARS), classification and regression tree analyses (CART), artificial neural networks (ANN) and random forest (RF).

These methods require presence/absence data of good quality in order to generate statistical functions or discriminative rules that allow habitat suitability to be ranked according to distributions of presence/absence data of a species (Manel et al., 1999; Guisan & Zimmerman, 2000). To obtain a good data quality it is important to avoid any errors of identification or of georeferencing which may create a false positive record that would be not useful and informative (Peterson & Soberón, 2012). Presence records of a species should be collected in order to allow modellers to realistically capture a large part of species ecological niche (Lomba et al., 2010). Another consideration is that of spatial autocorrelation which may complicate niche modelling applications because the non–independence of occurrence points may cause problems in model calibration by over-emphasizing certain environmental conditions. Obviously, the main difficulty in a presence/absence model is to have accurate data on absences. In fact, these data are hard to obtain, especially for mobile or low-density species. Moreover, in all cases of presence/absence model, sites not easily accessible or practicable can distort predictions (Lahoz-Monfort et al., 2013; Syfert et al., 2013; Guillera-Arroita et al., 2016).

Anyway, such models are increasingly used when only presence data is available, by creating artificial "absence" data usually called "pseudo-absences" or "background" data (Barbet-Massin et al., 2012). Different strategies have been proposed to improve the generation of an appropriate pseudo-absence data seta and they lead to the conclusion that random selection yields the most reliable distribution models. The best results were also obtained by using a large number of pseudo-absences (e.g. 10,000) with presence and absences weighted equally (Phillips & Dudik, 2008; Barbet-Massin et al., 2012).

The second group of methods includes BIOCLIM (SRE), DOMAIN, Ecological Niche Factor Analysis (ENFA); Genetic Algorithm for Rule Set Production (GARP) and Maximum Entropy method (MaxEnt). These methods require presence data only and allow to the generation of models where the knowledge of absences is inadequate or unavailable. SDM such as MaxEnt or GARP indeed require the use of background data or pseudo-absence data (Barbet-Massin et al., 2012) , that often are automatically generated by the software implementing the method.

In this section, we briefly describe the characteristics of the SDM methods above listed and highlight their strengths and shortcomings.

- **Generalized linear models (GLMs)** (Nelder & Wedderburn, 1972; McCullagh & Nelder, 1989) are mathematical extensions of linear models based on an assumed relationship (called a link function) between the mean of the response variable (species presence or absence, but also abundance or population density) and the linear combination of the explanatory variables (Guisan et al., 2002). The common types of GLMs are linear regression, logistic regression and Poisson regression. The explanatory variables of GLMs can contain interaction terms and polynomial terms, so they are preferable for nonlinear yet simple relationship between species and environment variables. All model parameters of GLMs can be clearly interpreted with ecological meanings. Applying GLMs requires careful calibration: users must check the significance of each explanatory variable and remove the non-significant variables (model selection) (Li & Wang, 2013). A frequently-used GLM is the Gaussian logit model in which the logit of the predicted probability of occurrence (p) is

$$\alpha + \beta_1 f_1(x_1) + \gamma_1 f_1(x_1)^2 + \ldots + \beta_n f_n(x_n) + \gamma_n f_n(x_n)^2$$

  where the f_j are the environmental variables, α, β_j and γ_j are the fitted coefficients, x_n represents each grid cell of the study area and the logit function is defined by logit (p) = $\ln\left(\frac{p}{1-p}\right)$ (Phillips et al., 2006).

  GLM, in particular logistic regression, is one of the best established statistical frameworks for SDM. Case studies using GLM for predicting species distributions (and related biotic variables) were extensively reviewed in Franklin (2010), Guisan & Zimmermann (2000) and Guisan et al. (2002).

- **Generalized additive models (GAMs)** (Hastie & Tibshirani, 1987) are non-parametric extensions of GLMs, thus providing more flexibility and the potential for modelling complex shapes of ecological response than GLMs (Yee & Mitchell, 1991; Elith et al., 2006). GAMs use data-defined smoothing functions to fit nonlinear species–environment relationships. The smooth functions are computed independently for each explanatory variable and added to build the final model. The number of smoothing parameters should be reasonably small to avoid overfitting. GAMs are useful when the relationship between species and environmental variables has a more complex form not easily fitted by GLMs (Li &Wang, 2013). If probability of occurrence is modeled with a GAM using a logit link function, the logit of the predicted probability has the form

  $$p = g_1(f_1(x_1)) + \ldots + g_n(f_n(x_n))$$

  where the f_i are again environmental variables. The g_i are smooth functions fit by the model, with the amount of smoothing controlled by a width parameter (Phillips et al., 2006).

- **Generalized boosting models (GBMs)** (Friedman et al., 2000; Friedman, 2001) are designed to fit many simple models whose predictions are then combined to give more robust estimates of the relationship between species distribution and a set of environmental variables, whereas GLMs seek to fit the single model that best explains the relationship between species and environment (Friedman et al., 2000; Friedman, 2001). It can compute a sequence of simple classification and/or regression trees, where each successive tree is built for the prediction residuals of the preceding tree. GBMs can eventually produce a good fit even if the specific nature of the relationships between the predictor variables and the dependent variable is complex (e.g. nonlinear). GBM can be understood as a method for developing a model in a forward stage-wise fashion, at each step adding small modifications in parts of the model space to fit better the data (Friedman et al., 2000). GBMs can be used for regression as well as classification problems, with continuous and/or categorical predictors.

- **Multivariate adaptive regression splines (MARS)** (Friedman, 1991) are an extension of linear regression models that automatically model nonlinearities and interactions (Friedman, 1991; Li & Wang, 2013). A major assumption of linear models is that the coefficients are stable across all levels of the explanatory variables. In contrast, MARS allow changes in coefficients, and are suitable when it is suspected that the model' coefficients have different optimal values across different levels of the explanatory variables. MARS are particularly powerful when there are large numbers of explanatory variables and low-order interaction effects (Thuiller et al., 2009). While MARS are, like GAMs, able to model complex relationships between response and predictor variables, they are computationally fast, while GAMs can be slow or even impossible to fit for very large datasets (Franklin, 2010). In contrast with previous models, MARS consider interactions between variables, not globally (over the entire range of the predictors), but locally – between sub-regions of every basis function.

- **Classification and regression tree analyses (CART)** (Breiman et al., 1984) consist of recursive partitions of the dimensional space defined by the predictors into groups that are as homogeneous as possible in terms of response. A classification tree is built by repeatedly splitting the data (in or case, species presence or absence data) into two exclusive groups, defined by a simple rule based on a single explanatory variable at each step (Segurado & Araújo, 2004). CART is able to uncover complex interactions between predictors that may be difficult or impossible to model using traditional multivariate techniques (Li & Wang, 2013).

- **Artificial neural networks (ANNs)** (McCulloch & Pitts, 1943) are a machine learning approach, involving a network of simple processing elements (artificial neurons) that can exhibit complex global behavior (e.g. select a site as habitat based on numerous environmental variables), determined by the links between the neurons and associated functions (Li & Wang, 2013). The key feature of an ANN is that it contains at least one hidden layer. Each neuron in the hidden layer receives information from each input (predictors variables), sums the inputs, adds a constant (the bias), then transforms the result using a fixed function. The response variable, or "output" in neural network terminology, is a weighted combination of the derived features in the hidden layer. Weights are defined in an iterative way, during the so called "training phase" of a neural network, so that all the input data presented to the network generate at the network output the desired result, within a predefined accuracy threshold. ANNs can operate like multiple regressions when the outputs are continuous variables, or like classifications when the outputs are categorical. ANN usually requires long computation time, especially in the network training phase and it doesn't necessarily perform better than other statistical and machine learning methods.

- **Random forest (RF)** (Breiman, 2001), an extension of CART, is a classifier that consists of many decision trees, implementing Breiman's CART algorithm for classification and regression (Breiman, 2001). To classify a new object from an input vector, random forest puts the input vector down each of the trees in the forest. Each tree gives a classification (which is usually called the tree 'votes' for that class) and the forest chooses the classification with the most votes (over all the trees in the forest).
  RF is not sensitive to the problem of multicollinearity and runs efficiently with large databases. It may process thousands of input variables without variable deletion and provides estimates of what variables are important in the classification. RF is one of the most accurate machine-learning algorithms with high performance in predicting species distribution.

- **BIOCLIM (SRE)** (Busby, 1986; Nix, 1986) predicts suitable conditions in a “bioclimatic envelope”, consisting of a rectilinear region in environmental space representing the range (or some percentage) of observed presence values in each environmental dimension. Although it generally does not perform as good as some other modeling methods (Elith et al., 2006), particularly in the context of climate change (Hijmans & Graham, 2006), it is still used, among other reasons because the algorithm is easy to understand. It computes the similarity of a location by comparing the values of environmental variables at any location to a percentile distribution of the values at known locations of occurrence (’training sites’). The closer to the 50th percentile (the median), the more suitable the location is.

- **DOMAIN** (Carpenter et al., 1993) predicts a suitability index by computing the minimum distance in environmental space between environmental variables at any location and those at any of the known locations of occurrence (’training sites’). The Gower distance is then calculated as the mean of these distances over all environmental variables. The algorithm assigns to a place the distance to the closest known occurrence (in environmental space).

- **Ecological Niche Factor Analysis (ENFA)** (Hirzel et al., 2002) compares the eco-geographical predictor distribution for a presence data set consisting of locations where the species has been detected with the predictor distribution of the whole area. Like Principal Component Analysis, ENFA summarises all predictors into a few uncorrelated factors retaining most of the information (Hirzel et al., 2002).

  The advantage of ENFA is that it takes presence records together with environmental data for the entire study area, without requiring a sample of background points to be treated like absences. Nevertheless, ENFA tended to produce overly optimistic predictions of the extent of suitable habitat in a comparison with GAMs and GLMs (high commission error or low specificity) (Franklin, 2010).

- **The Genetic Algorithm for Rule-Set Prediction (GARP)** (Stockwell & Noble, 1992; Stockwell, 1999) uses an artificial intelligence framework called genetic algorithms. It produces a set of positive and negative rules that together give a binary prediction; rules are favoured in the algorithm according to their significance (compared with random prediction) based on a sample of background cells and presence cells.

- **MaxEnt** (Phillips *et al.*, 2006) is a machine learning model which estimates the likelihood of the species by finding the distribution of “maximum entropy” (i.e. closest to a multivariate uniform distribution) given the constraint that the expected value of each environmental variable under this estimated distribution matches its empirical mean. The cells of the study area represent the space on which the MaxEnt probability distribution is defined (*X*), whereas cells with known species occurrence records constitute the sample points, and the features are climatic variables, elevation, soil category, vegetation type or other environmental variables, and functions thereof. The unknown probability distribution π assigns a non-negative probability $\pi(x)$ to each point x, and these probabilities sum to 1. The approximation of π is also a probability distribution, and we denote it π^*. The entropy of π^* is defined as

$$H(\pi^*) = -\sum_{x \in X} \pi^*(x) \ln \pi^*(x)$$

where ln is the natural logarithm.

In many comparative studies using a large number of SDM methods, MaxEnt was usually among the top-performing methods in terms of prediction accuracy (Elith et al., 2006). Phillips et al. (2006) outlined some advantages and disadvantages of MaxEnt for SDM compared to other methods. It only requires presence data plus environmental information for the whole study area. The results are amenable to interpretation of the form of the environmental response functions. MaxEnt is a generative approach, rather than discriminative, which makes it very robust to limited amounts of training data (small samples) and it can utilize both continuous and categorical data also incorporating interactions between different variables. The output is continuous, allowing fine distinctions between the modelled suitability of different areas. If binary predictions are desired, there is great flexibility in the choice of threshold. If the application is conservation planning, the fine distinctions in predicted relative environmental suitability can be valuable to reserve planning algorithms.

Furthermore, MaxEnt uses an exponential model for probabilities, which is not inherently bounded above and can give very large predicted values for environmental conditions outside the range present in the study area. Extra care is therefore needed when extrapolating to another study area or to future or past climatic conditions.

1.2.5. Ensemble forecasting of species distribution

The number of techniques available for SDMs is large and is increasing steadily, making it difficult to select the most appropriate methodology for their needs (Elith et al., 2006; Heikkinen et al., 2006). Although many SDM are based on a correlative approach, they use different assumptions, mathematical algorithms and parameterizations. They may vary in how they model the shape, nature, and complexity of species' response, select predictor variables, weight variable contributions, or allow for interactions (Buisson et al., 2010).

Sometimes different models provide large discrepancies and then diverse predictions (Pearson et al., 2006; Randin et al., 2006).

A growing concern has recently emerged for “ensemble forecasting” approaches, which fit a number of alternative models (i.e., various initial conditions combined with several statistical methods) and explore the range of resulting projections (Araujo & New, 2007). Ensemble modelling experiments can be used to distinguish regions of consensus among models (i.e. where most models agree that the environment is suitable/unsuitable) from regions where there is disagreement. A number of studies have found that ensembles outperform individual models providing more reliable and robust predictions of the potential distribution of species (Araujo & New, 2007; Marmion et al., 2009; Grenouillet et al., 2011; Crossman et al., 2012; Guo et al., 2015; Beaumont et al., 2016).

The idea of ensemble forecasting dates back to 1969, when J.M. Bates and Granger observed that combined forecasts would yield lower mean error than any of the constituent individual forecasts (Araujo & New, 2007). For example, BIOMOD (Thuiller et al., 2009), including its recent version “biomod2”, is a package for the open-source R software (R Development Core Team 2011) which can automatically compare 10 different techniques to calculate an SDM and is able to suggest the “best” model (i.e., the model with the best predictive performance) among extensive model comparisons, as well as to use a consensus approach that computes a weighted sum of the variability within the ensemble predictions. The output of BIOMOD is an “ensemble model” graphically represented as a map where each cell indicates the probability of species presence ranging between 0 and 1. This final map can also be transformed into presence–absence values using a set of different thresholds provided by the modelling package.

1.2.6. Methods for evaluating models

Models generating presence–absence predictions are usually evaluated by comparing the predictions with a set of validation sites and constructing a confusion matrix that records the number of true positive (a), false positive (b), false negative (c) and true negative (d) cases predicted by the model (Allouche et al., 2006).

SDMs are best validated using new or independent data, i.e. data not used to estimate the parameters or fit the model. However, often it is not feasible to collect new data. In this situation, one approach very commonly used to validate SDMs (for review, see Guisan & Zimmerman, 2000) is to divide or partition the data into one portion used to calibrate the model, called the training data, and one portion used to validate the predictions, called the testing data (Smith, 1994; Franklin, 2010).

One simple measure of accuracy that can be derived from the confusion matrix is the proportion of correctly predicted sites, called overall accuracy. Two alternative measures that are often derived from the confusion matrix are *sensitivity* and *specificity*. Sensitivity is the proportion of correctly predicted observed presences, and therefore quantifies omission errors. Specificity is the proportion of observed absences that are predicted as such, and therefore quantifies commission errors.

Other methods for assessing the accuracy of ordinal score models derive from these two indexes.

- The **receiver operating characteristic (ROC)** curve (Fielding & Bell, 1997) is constructed by using all possible thresholds to classify the scores into confusion matrices, obtaining sensitivity and specificity for each matrix, and then plotting sensitivity against the corresponding proportion of false positives (equal to 1 − specificity). The use of all possible thresholds avoids the need for a selection of a single threshold, which is often arbitrary (Manel et al., 1999; 2001; Liu et al., 2005), and allows appreciation of the trade-off between sensitivity and specificity (Pearce & Ferrier, 2000).

- The **area under the ROC curve (AUC)** is often used as a single threshold-independent measure of model performance, therefore not affected by the arbitrary selection of a model threshold.
  This index ranges between 0 and 1 and a model will be considered to discriminate better than chance if the AUC is higher than 0.5.

- **Minimum difference between training and test data** (**AUC_{diff}**) is based on the intuitive notion that overfit models should generally perform well on training data but poorly on test data (Warren & Seifert, 2011). By minimizing the difference between training and test data, we minimize the risk that our model is over-parameterized in such a way as to be overlay specific to the training data. Excellent model performance is expressed by AUC_{diff} close to 0.

- **Cohen's kappa** (Cohen, 1960) corrects the overall accuracy of model predictions by the accuracy expected to occur by chance. The kappa statistic ranges from − 1 to 1, where 1 indicates perfect agreement and values of zero or less indicate a performance no better than random (Segurado & Araujo, 2004).

- The **true skill statistic (TSS)**, also known as the Hanssen–Kuipers discriminant, compares the number of correct forecasts, minus those attributable to random guessing, to that of a hypothetical set of perfect forecasts. It is defined as:

$$\text{TSS} = \text{Sensitivity} + \text{Specificity} - 1$$

where sensitivity and specificity are calculated based on the probability threshold for which their sum is maximized. Like kappa, TSS takes into account both omission and commission errors, and has the same range of values as mentioned above (Allouche et al., 2006).

1.3 . Applications of SDMs: their role in species' conservation

SDMs can be used to tackle many issues in applied ecology and conservation biology, the most basic of which is to understand the relationships between a species and its abiotic and biotic environment and to identify areas where a given species is likely to occur.

This application allows to detect new areas where the target species presence is unknown and is a useful tool for estimating the potential for species to occur in areas not previously surveyed (Hernandez et al., 2008). Such models are highly relevant for species that are elusive and difficult to observe (Razgour et al., 2016) and, in particular, for rare and threatened species which are usually a conservation priority and for which the knowledge gap is generally wider than for more common species (Bosso et al., 2016b; McCune, 2016; Proosdij et al., 2016; Shaffer-Smith et al., 2016). Detecting new potential population of rare species could be highly productive when planning field surveys on areas where a target taxon is more probably present, making field operation more cost-effective (Rebelo & Jones, 2010).

Furthermore, SDMs offer a very useful tool to assess the effectiveness of networks of protected areas (Wilson et al., 2005; Bosso et al., 2016b). Indeed, habitat suitability maps provided by SDMs may be used to carry out a gap analysis which, very often, highlights an inappropriate degree of protection offered by the overall reserve network to the potential geographic range of the species (Rodríguez et al. 2007; Li et al., 2016). This application results even more useful when SDMs predict areas of high-priority for conservation (Domígues-Vega et al., 2012; Smith et al., 2016; Di Minin et al., 2017), such as hotspots of species richness (Vale G. et al., 2016; Cooper-Bohannon et al., 2016) which need of special protection or of appropriate restoration (Razgour et al., 2011).

SDMs may provide the degree of fragmentation of species' habitat due to natural barriers or when species may not be able to migrate across human-modified landscapes which represents an obstacle to gene flow between populations occurring in suitable patches. Connectivity analyses applied to SDMs may support conservation actions identifying dispersal corridors for species movement across fragmented habitat (Puddu and Maiorano, 2016) or patches of reserve networks.

The ability of SDMs to project the potential effects of anthropogenic global warming on species distribution and ecosystem properties has received an increasing attention (Thuiller, 2003; Beaumont et al., 2007; Bagchi et al., 2017; Casajus et al., 2016; Benedetti et al., 2017; Bosso et al., 2017a). The Intergovernmental Panel on Climate Change (IPCC) has developed a set of climate change scenarios that describe future greenhouse gas (GHG) emissions (Nakicenovic et al., 2000) and include a wide variety of possible socio-economic projections. 'Worst case' scenarios generate more severe projections, with annual average temperatures rising up to 5.8 °C by the end of the century in some regions, while other scenarios are less severe. One of the major impacts of climate change may be the movement of populations from their original locations to new and unoccupied areas. In this process, depending on the dispersal ability of the species, populations may become highly fragmented and local extinctions may occur (Guisan & Thuiler, 2005; Rebelo et al., 2010; Aguair et al., 2016). If these situations persist over long periods of time, it is likely that other conservation issues may arise (Hughes et al., 2012). SDMs have an important role in highlighting likely shifts of suitable habitat for a species due to climate change (Araujo et al., 2011; Pio et al., 2014; Brown et al., 2015; Beaumont et al., 2016; Bagchi et al., 2017; Ofori et al., 2017), detecting a potential range contraction and identifying the areas likely to be used as corridors by the species to move across a future fragmented habitat and which need major conservation efforts in order to guarantee species survival (Thuiller, 2003; Beaumont et al., 2007; Rebelo et al., 2010; Aguair et al., 2016).

Many studies have used SDMs to project species suitable habitat in the past (most frequently to the Last Glacial Maximum) in order to test the existence of a niche conservatism in the present (Rebelo et al., 2012; Carotenuto et al., 2016; De Castro et al., 2016). These models permit to determinate the areas where species could have persisted, namely glacial refugia. A very recent study applied SDMs in order to test whether extinct megafauna responded in a similar manner to extant species while coping with the changing environmental conditions that characterized the end of the Pleistocene (Di Febbraro et al., 2017).

Given the ever growing concern caused by biological invasions and the threat posed by pest species (also pathogens; see also Baxter & Possingham, 2011; Bosso et al., 2016a and Schatz et al., 2017), recent studies have applied SDMs to predict the potential distribution of alien species, identifying current potential hotspots of invasion and assessing climate change influences on the probability of such biological invaders to colonize and spread in new areas (Thuiller et al., 2005; Beaumont et al., 2009; Bellard et al., 2013; Di Febbraro et al., 2016; Bosso et al., 2016a; Bosso et al., 2017a,b).

SDMs are applied to explore niche segregation in sympatric species (Peers et al., 2013; Russo et al., 2014). In fact, interspecific competition might involve changes in the environmental niche of a species where the latter is sympatric to competitors. The importance of biotic factors, like competition, in affecting the species distribution has also been highlighted by other recent studies (Razgour et al., 2011; Razgour et al., 2016; Jones et al., 2016).

Habitat suitability mapping is frequently used to support appropriate management plans for species recovery and to identify suitable sites for species reintroduction (Olsson & Rogers, 2009; Wilson et al., 2011; Ardestani et al. 2015).

Finally, an important application of SDMs involves the prediction of the impacts of anthropogenic activities on the landscape connectivity. The latter is a critical factor which may greatly influence the distribution of animals when it alters their movements and their ability to reach foraging grounds (Henry et al., 2007). In particular, some recent studies have used SDMs to identify the main animal connectivity routes, such as those for seasonal migration, and have analysed the potential interference of anthropogenic infrastructures on them (Roscioni et al., 2014). Such an approach makes SDMs a very useful tool to achieve a sustainable development of renewable energy infrastructures, like wind farms, and proper surveys for monitoring the wildlife fatalities.

1.4 . Objectives of the study

After assessing the many different applications of SDMs supporting species' conservation actions, our study focused on other important but less handled issues which have been analysed through three study cases.
We sought to answer the following questions:

- Does the spatial scale of the analyses influence the SDMs' ability to detect niche differences for two sympatric species which apparently share the same habitat and ecological requirements?
  Are the management practices adopted for one species also effective for the other one?

- May SDMs be applied to predict the distribution of reintroduced species and support appropriate management plans to identify and eventually restore suitable sites which might be colonized in a near future by the species?

- Are presence data, which do not take into account the phenology of species, effective to develop SDMs representing the full ecological niche of a species?

1.5. References

Aguiar, L. M., Bernard, E., Ribeiro, V., Machado, R. B., & Jones, G. (2016). Should I stay or should I go? Climate change effects on the future of Neotropical savannah bats. *Global Ecology and Conservation*, *5*, 22-33.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, *43*(6), 1223-1232.

Anderson, R. P., Lew, D., & Peterson, A. T. (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological modelling*, *162*(3), 211-232.

Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., J Hijmans, R., Huettmann, F., ... & A Loiselle, B. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129-151.

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution*, *22*(1), 42-47.

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, *93*(7), 1527-1539.

Ardestani, E. G., Tarkesh, M., Bassiri, M., & Vahabi, M. R. (2015). Potential habitat modeling for reintroduction of three native plant species in central Iran. *Journal of arid land*, *7*(3), 381-390.

Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling*, *157*(2), 101-118.

Austin, M. (2007). Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological modelling*, *200*(1), 1-19.

Austin, M. P., & Van Niel, K. P. (2011). Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, *38*(1), 1-8.

Bagchi, R., Hole, D. G., Butchart, S. H., Collingham, Y. C., Fishpool, L. D., Plumptre, A. J., ... & Willis, S. G. (2017). Forecasting potential routes for movement of endemic birds among important sites for biodiversity in the Albertine Rift under projected climate change. *Ecography*.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in Ecology and Evolution*, *3*(2), 327-338.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., ... & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*(11), 1810-1819.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Or*, 451-468.

Baxter, P. W., & Possingham, H. P. (2011). Optimizing search strategies for invasive pests: learn before you leap. *Journal of Applied Ecology*, *48*(1), 86-95.

Beaumont, L. J., Pitman, A. J., Poulsen, M., & Hughes, L. (2007). Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, *13*(7), 1368-1385.

Beaumont, L. J., Gallagher, R. V., Thuiller, W., Downey, P. O., Leishman, M. R., & Hughes, L. (2009). Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, *15*(3), 409-420.

Beaumont LJ, Graham E, Duursma DE, et al. (2016). Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecological Modelling*, *342*, 135-146.

Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M., & Courchamp, F. (2013). Will climate change promote future invasions?. *Global Change Biology*, *19*(12), 3740-3748.

Benedetti, F., Guilhaumon, F., Adloff, F., & Ayata, S. D. (2017). Investigating uncertainties in zooplankton composition shifts under climate change scenarios in the Mediterranean Sea. *Ecography*.

Bosso, L., Febbraro, M., Cristinzio, G., Zoina, A., & Russo, D. (2016)a. Shedding light on the effects of climate change on the potential distribution of *Xylella fastidiosa*. *Biological invasions*, *18*(6), 1759-1768.

Bosso, L., Mucedda, M., Fichera, G., Kiefer, A., & Russo, D. (2016)b. A gap analysis for threatened bat populations on Sardinia. *Hystrix, the Italian Journal of Mammalogy*, *27*(2).

Bosso, L., Luchi, N., Maresi, G., Cristinzio, G., Smeraldo, S., & Russo, D. (2017)a. Predicting current and future disease outbreaks of Diplodia sapinea shoot blight in Italy: species distribution models as a tool for forest management planning. *Forest Ecology and Management*, *400*, 655-664.

Bosso, L., De Conno, C., & Russo, D. (2017)b. Modelling the Risk Posed by the Zebra Mussel Dreissena polymorpha: Italy as a Case Study. *Environmental Management*, 1-10.

Braunisch, V., Coppes, J., Arlettaz, R., Suchant, R., Schmid, H., & Bollmann, K. (2013). Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography*, *36*(9), 971-983.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brotons, L., Thuiller, W., Araújo, M. B., & Hirzel, A. H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, *27*(4), 437-448.

Brown, J. L., & Yoder, A. D. (2015). Shifting ranges and conservation challenges for lemurs in the face of climate change. *Ecology and evolution*, *5*(6), 1131-1142.

Buisson, L., Thuiller, W., Casajus, N., Lek, S., & Grenouillet, G. (2010). Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, *16*(4), 1145-1157.

Busby, J. R. (1986). Bioclimate prediction system (BIOCLIM). *User's manual version*, *2*.

Carotenuto, F., Di Febbraro, M., Melchionna, M., Castiglione, S., Saggese, F., Serio, C., ... & Raia, P. (2016). The influence of climate on species distribution over time and space during the late Quaternary. *Quaternary Science Reviews*, *149*, 188-199.

Carpenter, G., Gillison, A. N., & Winter, J. (1993). DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and conservation*, *2*(6), 667-680.

Carretero, M. A., & Sillero, N. (2016). Evaluating how species niche modelling is affected by partial distributions with an empirical case. *Acta Oecologica*, *77*, 207-216.

Casajus, N., Périé, C., Logan, T., Lambert, M. C., de Blois, S., & Berteaux, D. (2016). An objective approach to select climate scenarios when projecting species distribution under climate change. *PloS one*, *11*(3).

Chase, J. M., & Leibold, M. A. (2003). *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press.

Chen, G., Kéry, M., Plattner, M., Ma, K., & Gardner, B. (2013). Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology*, *101*(1), 183-191.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), 37-46.

Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19651-19658.

Cooper-Bohannon, R., Rebelo, H., Jones, G., Monadiem, A., Schoeman, M. C., Taylor, P., & Park, K. (2016). Predicting bat distributions and diversity hotspots in southern Africa. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Crossman, N. D., Bryan, B. A., & Summers, D. M. (2012). Identifying priority areas for reducing species vulnerability to climate change. *Diversity and Distributions*, *18*(1), 60-72.

Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B., & Wood, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature*, *391*(6669), 783.

De Castro, O., Di Maio, A., Di Febbraro, M., Imparato, G., Innangi, M., Véla, E., & Menale, B. (2016). A multi-faceted approach to analyse the effects of environmental variables on geographic range and genetic structure of a perennial psammophilous geophyte: the case of the sea daffodil *Pancratium maritimum* L. in the Mediterranean basin. *PloS one*, *11*(10), e0164816.

Di Febbraro, M., Martinoli, A., Russo, D., Preatoni, D., & Bertolino, S. (2016). Modelling the effects of climate change on the risk of invasion by alien squirrels. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Di Febbraro, M., Carotenuto, F., Castiglione, S., Russo, D., Loy, A., Maiorano, L., & Raia, P. (2017) Does the jack of all trades fare best? Survival and niche width in Late Pleistocene megafauna. *Journal of Biogeography*.

Di Minin, E, Soutullo, A., Bartesaghi, L., Rios, M., Szephegyi, M.N. , Moilanen, A. (2017). Integrating biodiversity, ecosystem services and socio-economic data to identify priority areas and landowners for conservation actions at the national scale. *Biological Conservation*. *206*, 56-64

Domíguez-Vega, H., Monroy-Vilchis, O., Balderas-Valdivia, C. J., Gienger, C. M., & Ariano-Sánchez, D. (2012). Predicting the potential distribution of the beaded lizard and identification of priority areas for conservation. *Journal for Nature Conservation*, *20*(4), 247-253.

Dormann, C. F. (2007). Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global ecology and biogeography*, *16*(2), 129-138.

Dormann, C. F., Purschke, O., Márquez, J. R. G., Lautenbach, S., & Schröder, B. (2008). Components of uncertainty in species distribution analysis: a case study of the great grey shrike. *Ecology*, *89*(12), 3371-3386.

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., ... & Singer, A. (2012). Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, *39*(12), 2119-2131.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Münkemüller, T. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27-46.

Elith, J., & Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, *32*(1), 66-77.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics*, *40*, 677-697.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, *24*(1), 38-49.

Franklin, J. (2010). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Fridedman, J. H. (1991). Multivariate adaptive regression splines (with discussion). *Ann. Statist*, *19*(1), 79-141.

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, *28*(2), 337-407.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E., & Thuiller, W. (2012). Invasive species distribution models–how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, *21*(11), 1126-1136.

Graham, C. H., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. T. (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in ecology & evolution*, *19*(9), 497-503.

Grenouillet, G., Buisson, L., Casajus, N., & Lek, S. (2011). Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, *34*(1), 9-17.

Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, *40*(2).

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2), 147-186.

Guisan, A., & Theurillat, J. P. (2000). Equilibrium modeling of alpine plant distribution: how far can we go?. *Phytocoenologia*, *30*(3/4), 353-384.

Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, *157*(2), 89-100.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, *8*(9), 993-1009.

Guo, C., Lek, S., Ye, S., Li, W., Liu, J., Li, Z. (2015). Uncertainty in ensemble modelling of large-scale species distribution: effects from species characteristics and model techniques. *Ecological modelling*, 306, 67-75.

Harrell, F. E., Lee, K. L., Califf, R. M., Pryor, D. B., & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in medicine*, *3*(2), 143-152.

Hastie, T., & Tibshirani, R. (1987). Generalized additive models: some applications. *Journal of the American Statistical Association*, *82*(398), 371-386.

Heikkinen, R. K., Luoto, M., Araújo, M. B., Virkkala, R., Thuiller, W., & Sykes, M. T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, *30*(6), 751-777.

Henry, M., PONS, J., & COSSON, J. F. (2007). Foraging behaviour of a frugivorous bat helps bridge landscape connectivity and ecological processes in a fragmented rainforest. *Journal of Animal Ecology*, *76*(4), 801-813.

Hernandez, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, *29*(5), 773-785.

Hernandez, P. A., Franke, I., Herzog, S. K., Pacheco, V., Paniagua, L., Quintana, H. L., ... & Vargas, J. (2008). Predicting species distributions in poorly-studied landscapes. *Biodiversity and conservation*, *17*(6), 1353-1366.

Hijmans, R. J., & Graham, C. H. (2006). The ability of climate envelope models to predict the effect of climate change on species distributions. *Global change biology*, *12*(12), 2272-2281.

Hirzel, A., & Guisan, A. (2002). Which is the optimal sampling strategy for habitat suitability modelling. *Ecological modelling*, *157*(2), 331-341.

Hirzel, A. H., & Le Lay, G. (2008). Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, *45*(5), 1372-1381.

Holt, R. D. (2003). On the evolutionary ecology of species' ranges. *Evolutionary ecology research*, *5*(2), 159-178.

Huston, M. A. (2002). Introductory essay: critical issues for improving predictions. *Predicting species occurrences: issues of accuracy and scale*, 7-21.

Jaberg, C., & Guisan, A. (2001). Modelling the distribution of bats in relation to landscape structure in a temperate mountain environment. *Journal of Applied Ecology*, *38*(6), 1169-1181.

Jones, M., Bertola, L. D., & Razgour, O. (2016). Predicting the effect of interspecific competition on habitat suitability for the endangered African wild dog under future climate and land cover changes. *Hystrix, the Italian Journal of Mammalogy*, *27*, 8.

Kearney, M. (2006). Habitat, environment and niche: what are we modelling?. *Oikos*, *115*(1), 186-191.

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*, *12*(4), 334-350.

Lahoz-Monfort, J. J., Guillera-Arroita, G. & Wintle, B. A. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, *23, 504–515.* doi:10.1111/geb.12138

Li, X., & Wang, Y. (2013). Applying various algorithms for species distribution modelling. *Integrative Zoology*, *8*(2), 124-135.

Li, Y.; B. Cui; X. Qiu; C. Ding; I. Batool (2016). Management reference for nature reserve networks based on MaxEnt modeling and gap analysis: a case study of the brown–eared pheasant in China. *Animal Biodiversity and Conservation*. *39* (2) , 241-252.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), 385-393.

Lobo, J. M., Jiménez‐Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103-114.

Loiselle, B. A., Jørgensen, P. M., Consiglio, T., Jiménez, I., Blake, J. G., Lohmann, L. G., & Montiel, O. M. (2008). Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes?. *Journal of Biogeography*, *35*(1), 105-116.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, *143*(11), 2647-2657.

Manel, S., Dias, J. M., & Ormerod, S. J. (1999). Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological modelling*, *120*(2), 337-347.

Manel, S., Williams, H. C., & Ormerod, S. J. (2001). Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of applied Ecology*, *38*(5), 921-931.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and distributions*, *15*(1), 59-69.

Mateo, R. G., Croat, T. B., Felicísimo, Á. M., & Munoz, J. (2010). Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions*, *16*(1), 84-94.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115-133.

McCullagh, P., & Nelder, J. A. (1989). Generalized Linear Models, no. 37 in Monograph on Statistics and Applied Probability.

McCune, J. L. (2016), Species distribution models predict rare species occurrences despite significant effects of landscape context. *Journal of Applied Ecology*, 53: 1871–1879.

Miller, J., & Franklin, J. (2002). Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modelling*, *157*(2), 227-247.

Naimi, B., & Araújo, M. B. (2016). sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*(4), 368-375.

Nakicenovic, N., Alcamo, J., Grubler, A., Riahi, K., Roehrl, R. A., Rogner, H. H., & Victor, N. (2000). *Special Report on Emissions Scenarios (SRES), A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Nelder, J. A., & Baker, R. J. (1972). *Generalized linear models*. John Wiley & Sons, Inc..

Nix, H. A., & Busby, J. (1986). BIOCLIM, a bioclimatic analysis and prediction system. *Annual report CSIRO. CSIRO Division of Water and Land Resources, Canberra*.

Ofori, B. Y., Stow, A. J., Baumgartner, J. B., & Beaumont, L. J. (2017). Combining dispersal, landscape connectivity and habitat suitability to assess climate-induced changes in the distribution of Cunningham's skink, Egernia cunninghami. *PloS one*, *12*(9).

Olsson, O., & Rogers, D. J. (2009). Predicting the distribution of a suitable habitat for the white stork in Southern Sweden: identifying priority areas for reintroduction and habitat restoration. *Animal Conservation*, *12*(1), 62-70.

Papeş, M., & Gaubert, P. (2007). Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity and distributions*, *13*(6), 890-902.

Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological modelling*, *133*(3), 225-245.

Pearman, P. B., Guisan, A., Broennimann, O., & Randin, C. F. (2008). Niche dynamics in space and time. *Trends in Ecology & Evolution*, *23*(3), 149-158.

Pearson, R. G., Dawson, T. P., & Liu, C. (2004). Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, *27*(3), 285-298.

Pearson, R. G. (2007). Species' distribution modeling for conservation educators and practitioners. *Synthesis. American Museum of Natural History*, *50*.

Peers, M. J., Thornton, D. H., & Murray, D. L. (2013). Evidence for large-scale effects of competition: niche displacement in Canada lynx and bobcat. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*(1773), 20132495.

Peterson, A.T. (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, **3**, 59–72.

Peterson, A. T., & Soberón, J. (2012). Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, *10*(2), 102-107.

Peterson, A. T., & Soberón, J. (2012). Integrating fundamental concepts of ecology, biogeography, and sampling into effective ecological niche modeling and species distribution modeling. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, *146*(4), 789-796.

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, *26*(3), 275-287.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3), 231-259.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161-175.

Proosdij, A. S., Sosef, M. S., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542-552.

Puddu, G., & Maiorano, L. (2016). Combining multiple tools to provide realistic potential distributions for the mouflon in Sardinia: species distribution models, spatial pattern analysis and circuit theory. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology letters*, *3*(4), 349-361.

Raes, N. (2012). Partial versus full species distribution models. *Natureza & Conservação*, *10*(2), 127-138.

Randin, C. F., Dirnböck, T., Dullinger, S., Zimmermann, N. E., Zappa, M., & Guisan, A. (2006). Are niche-based species distribution models transferable in space?. *Journal of biogeography*, *33*(10), 1689-1703.

Razgour, O., Hanmer, J., & Jones, G. (2011). Using multi-scale modelling to predict habitat suitability for species of conservation concern: the grey long-eared bat as a case study. *Biological Conservation*, *144*(12), 2922-2930.

Razgour, O., Rebelo, H., Di Febbraro, M., & Russo, D. (2016). Painting maps with bats: species distribution modelling in bat research and conservation. *Hystrix*.

Rebelo, H., & Jones, G. (2010). Ground validation of presence-only modelling with rare species: a case study on barbastelles Barbastella barbastellus (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, *47*(2), 410-420.

Rebelo, H., Tarroso, P., & Jones, G. (2010). Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology*, *16*(2), 561-576.

Rebelo, H., Froufe, E., Brito, J. C., Russo, D., Cistrone, L., Ferrand, N., & Jones, G. (2012). Postglacial colonization of Europe by the barbastelle bat: agreement between molecular data and past predictive modelling. *Molecular Ecology*, *21*(11), 2761-2774.

Reddy, S., & Dávalos, L. M. (2003). Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, *30*(11), 1719-1727.

Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, *13*(3), 243-251.

Romero, D., Olivero, J., Brito, J. C., & Real, R. (2016). Comparison of approaches to combine species distribution models based on different sets of predictors. *Ecography*, *39*(6), 561-571.

Roscioni, F., Rebelo, H., Russo, D., Carranza, M. L., Di Febbraro, M., & Loy, A. (2014). A modelling approach to infer the effects of wind farms on landscape connectivity for bats. *Landscape ecology*, *29*(5), 891-903.

Rotenberry, J. T., Preston, K. L., & Knick, S. T. (2006). GIS-BASED NICHE MODELING FOR MAPPING SPECIES'HABITAT. *Ecology*, *87*(6), 1458-1464.

Ruiz-Gutierrez, V., Hooten, M. B., & Campbell Grant, E. H. (2016). Uncertainty in biological monitoring: a framework for data collection and analysis to account for multiple sources of sampling bias. *Methods in Ecology and Evolution*, *7*(8), 900-909.

Russo, D., Di Febbraro, M., Rebelo, H., Mucedda, M., Cistrone, L., Agnelli, P., ... & Bosso, L. (2014). What story does geographic separation of insular bats tell? A case study on Sardinian Rhinolophids. *PLoS One*, *9*(10), e110894.

Russo, D., Maiorano, L., Rebelo, H., & Preatoni, D. G. (2016). Distribution Models: exploring patterns and processes in mammal ecology. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Schaffer-Smith, D., Swenson, J., & Boveda-Penalba, A. (2016). Rapid conservation assessment for endangered species using habitat connectivity models. *Environmental Conservation, 43*(3), 221-230

Schatz, A. M., Kramer, A. M., & Drake, J. M. (2017). Accuracy of climate-based forecasts of pathogen spread. *Royal Society Open Science*, *4*(3), 160975.

Segurado, P., & Araujo, M. B. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography*, *31*(10), 1555-1568.

Sillero, N. (2011). What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, *222*(8), 1343-1346.

Smith, P. A. (1994). Autocorrelation in logistic regression modelling of species' distributions. *Global ecology and biogeography letters*, 47-61.

Smith, A., Schoeman, M.C., Keith, M., Erasmus, B.F.N., Monadjem, A., Moilanen, A., Di Minin, E. (2016). Synergistic effects of climate and land-use change on representation of African bats in priority conservation areas. Ecological Indicators. *69,* 276-283.

Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology letters*, *10*(12), 1115-1123.

Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19644-19650.

Stockwell, D. R., & Noble, I. R. (1992). Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and computers in simulation*, *33*(5-6), 385-390.

Stockwell, D. (1999). The GARP modelling system: problems and solutions to automated spatial prediction. *International journal of geographical information science*, *13*(2), 143-158.

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological modelling*, *148*(1), 1-13.

Syfert, M. M., Smith, M. J. & Coomes, D. A. (2013). The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLOS ONE*, *8*(7), 10.1371

Tessarolo, G., Rangel, T. F., Araújo, M. B., & Hortal, J. (2014). Uncertainty associated with survey design in Species Distribution Models. *Diversity and Distributions*, *20*(11), 1258-1269.

Thuiller, W. (2003). BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. *Global change biology*, *9*(10), 1353-1362.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, *27*(2), 165-172.

Thuiller, W., Richardson, D. M., PYŠEK, P., Midgley, G. F., Hughes, G. O., & Rouget, M. (2005). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, *11*(12), 2234-2250.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369-373.

Vale, C. G., Campos, J. C., Silva, T. L., Gonçalves, D. V., Sow, A. S., Martínez-Freiría, F., ... & Brito, J. C. (2016). Biogeography and conservation of mammals from the West Sahara-Sahel: an application of ecological niche-based models and GIS. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Van Horn, B. (2002). Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. Á In: Scott, JM et al.(eds), Predicting species occurrences: issues of accuracy and scale.

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*(2), 335-342.

Wiens, J. A. (2002). Predicting species occurrences: progress, problems, and prospects. *Predicting species occurrences: issues of accuracy and scale*, 739-749.

Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19729-19736.

Wilson, K. A., Westphal, M. I., Possingham, H. P., & Elith, J. (2005). Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, *122*(1), 99-112.

Wilson, C. D., Roberts, D., & Reid, N. (2011). Applying species distribution modelling to identify areas of high conservation value for endangered species: A case study using Margaritifera margaritifera (L.). *Biological Conservation*, *144*(2), 821-829.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and distributions*, *14*(5), 763-773.

Yee, T. W., & Mitchell, N. D. (1991). Generalized additive models in plant ecology. *Journal of vegetation science*, *2*(5), 587-602.

Drinking bat *Barbastella barbastellus* (Jens Rydell)

Rosalia alpina (Luca Cistrone)

Chapter II

2. Protecting one, protecting both? Scale-dependent ecological differences in two species using dead trees, the rosalia longicorn beetle and the barbastelle bat

2.1. Abstract

Organisms sharing the same habitats may differ in small-scale microhabitat requirements or benefit from different management. In this study, set in Italy, we focused on two species of high conservation value, the cerambycid beetle *Rosalia alpina* and the bat *Barbastella barbastellus*, which often share the same forest areas and in several cases the same individual trees. We compared the potential distribution and, at two spatial scales, the niches between such species. The predicted distributions largely overlapped between the beetle and the bat. The niches proved to be similar on a broad scale, yet not at a finer one. Compared with *B. barbastellus*, *R. alpina* tends to occur at lower altitude in more irradiated sites with lower canopy closure and uses shorter trees with wider diameters. *B. barbastellus* trees occurred more often within forest or along its edges, whereas *R. alpina* lays eggs in trees found in clearings. *B. barbastellus* plots were more frequent in forest, *R. alpina* ones in forested pasture and open-shredded forest. Overall, solar radiation influenced more critically site and tree selection by *R. alpina,* as a warm microclimate is essential for larval development. Although *B. barbastellus* reproduction may be favoured by warmer roosting conditions, bats may also find such conditions in dense forest, in strongly-irradiated cavities high up in tall trees that project above the canopy. We emphasize that subtle differences in the ecological requirements of syntopic taxa could be missed at broad scales, so multiple-scale assessment is always advisable.

2.2. Introduction

Species conservation that is undertaken by adopting management practices developed for organisms other than the target one that apparently share the same habitats and ecological requirements may be tantalizing yet extremely risky. In fact, syntopic species may often differ markedly in characteristics only revealed at different spatial and/or temporal scales: this may make management practices useful for some species, yet ineffective or even harmful to others (Martin and Lopez, 2002; Russo *et al.,* 2005). This is one of the reasons why the so-called "umbrella species" strategy (managing one species sharing its home range or habitat preferences with several others to benefit a large set of organisms besides the target one) has raised significant concern (Simberloff, 1998; Martin & Lopez, 2002; Roberge & Angelstam, 2004).

Only ecological analyses conducted at multiple spatial scales may reveal spatial or temporal niche differences that would otherwise go unnoticed, but such studies in conservation biology are still infrequent (Razgour et al., 2011). In this study we test whether a two-scale analysis may reveal differences in two management-dependent forest species that are phylogenetically distant yet apparently share similar habitat preferences: a tree-roosting bat (the barbastelle *Barbastella barbastellus* Schreber) and a saproxylic cerambycid beetle (the rosalia longicorn *Rosalia alpina* L.) whose life cycle largely depends on the availability of dead wood.

R. alpina is often mentioned as a flagship species due to its distinctive colouring (Duelli & Wermelinger, 2005, Caci et al., 2013) and is listed as a Priority Species under Annex II of 357/97/EC Habitats Directive. The strong habitat specialization of saproxylic beetles makes them especially vulnerable to human impact (e.g. Berg et al., 1994, Vié et al*.,* 2009) and exposes them to adverse effects of intensive forestry (Jonsson et al., 2005; 2006). In summer, adults lay eggs in dead wood, so the species only occurs in patches of mature forest rich with this substrate, as dead wood is necessary for larval development (Duelli & Wermelinger, 2005; Russo et al., 2011).

B. barbastellus is a medium-sized vespertilionid bat featuring in Annexes II, IV of the Habitats Directive. Its diet consists largely of moths (Sierro & Arlettaz, 1997), which are hunted in a range of habitats (Hillen et al., 2011; Zeale et al., 2012; Ancillotto et al., 2014). Although in some cases *B. barbastellus* maternity groups also roost in live trees (Russo et al., 2010) or even rock crevices (Ancillotto *et al.,* 2014), typically they are found beneath loose bark on standing dead trees (Russo *et al.,* 2004), where small groups of females gestate, give birth and rear young (Russo et al., 2004).

Despite the very different life cycles, mobility and scales at which saproxylic beetles and snag-roosting bats perceive their habitat and landscapes, both species are threatened by the removal of dead or defective trees. This is one of the typical consequences of commercial forestry in temperate forests (Dudley et al., 2004) driven by concerns over forest health, productivity, fire risk, as well as the spread of pests. Short harvesting rotations characteristic of commercial forestry result in trees being logged before they complete their natural cycle, so that forest turnover will not lead to significant dead wood production (Hunter, 1999). Forestry is a prime cause for the disappearance of reproductive substrate for cerambycids (e.g. Russo et al., 2011a) as well as for the loss of bat roosting habitat (Hayes & Loeb, 2007). Typically, bats roost in small groups spread over large forest areas and switch frequently between roosts (Lewis, 1995), so that the persistence of even a small number of bats requires many trees (Russo et al., 2005). Since snags are a rare resource even in unmanaged forests (Russo et al., 2004), the reduction of dead tree density due to forestry has adverse consequences for bats such as *B. barbastellus*, that specialize on roosting in snags, especially if they use snags as nurseries (Kurta et al., 2002; Russo et al., 2004).

In the course of previous studies carried out in the Italian Peninsula (Russo et al., 2004; 2007; 2010; 2011a) both the bat and the beetle were found to share the same forest areas and in several cases even the same individual trees (Russo et al., 2011a). An apparently obvious assumption would be that since the two organisms may share trees, conservation practices favouring the bat would also benefit the beetle and vice versa. This could be achieved by leaving patches of forest unmanaged and promoting the presence of snags. In our case, two potentially important factors might discourage this approach, 1) the considerable phylogenetic distance between the two species, which may hide subtle yet crucial differences in ecological requirements; and 2) the differing mobility and sensitivity to multiple spatial scales linked with their life histories (Russo et al., 2004; 2011a; Drag et al., 2011; Hillen et al., 2011; Ancillotto et al., 2014). For these reasons, variables other than tree type might influence selection patterns in different ways and only reveal themselves at specific spatial scales.

Our study was based in Italy, where both species occur sympatrically and sufficient data were available for our modelling exercise. We carried out a niche comparison between *R. alpina* and *B. barbastellus* and hypothesized that differences in ecological requirements, if any, could emerge at a small spatial scale (i.e. the tree used and its immediate surroundings), best reflecting microhabitat needs. If the ecological needs of the two species differ, conserving the bat and the beetle would require a multiple spatial scale approach within areas of sympatry and ad-hoc management practices. Although previous studies had analyzed tree and habitat selection in *B. barbastellus* (Russo et al., 2004) and *R. alpina* (Russo et al., 2011a) and developed a spatially explicit model for the latter (Bosso et al., 2013) to offer a basis for gap analysis, ours is the first that compares the two species' niches at two markedly different spatial scales to explicitly test whether common management strategies may be revealed at either scale and adopted to improve the status of two taxa of conservation concern.

Within this framework, we generated the following predictions:

1. The co-occurrence of the two species in areas characterized by old-growth stands will lead to significant overlap in their potential distribution and similarity in ecological niches estimated using factors selected at a broad geographical scale.
2. At a smaller scale, *R. alpina* will prefer trees in semi-open habitat whose immediate surroundings favour sun irradiation, a factor deemed of vital importance for the development of larvae (Duelli & Wermelinger, 2005; Russo et al., 2001; Bosso et al., 2013). On the other hand, although reproductive females of *B. barbastellus* may be favoured by roosting in sun-exposed cavities, they may do so in dense forest too by roosting higher up in trees (Russo et al., 2004; 2007). Roosting in more forested habitats would shelter emerging bats from aerial predators and allow them to start foraging earlier (Russo et al. 2007). Therefore, unlike *R. alpina* the bat will tend to occur in more densely forested habitats and will select taller trees.

2.3. Materials and methods

2.3.1. Study area

We considered the entire Italian territory between latitudes 45° N–36° N and longitudes 6°E–18°E (corresponding to ca. 301.000 km^2, elevation range = 0–4810 m a.s.l.) including the Alpine, Continental and Mediterranean biogeographical regions according to the EC/92/43 Habitats Directive classification.

2.3.2. Model development and validation

Presence records for *R. alpina* (n = 350) and *B. barbastellus* (n = 223) originated from the authors' personal databases and previous publications (Bosso et al., 2013; Russo et al., 2004; 2005; 2010). Records were screened in ArcGIS (version 9.2) using average nearest neighbour analyses and Moran's I measure of spatial autocorrelation to remove spatially correlated data points and guarantee independence and equivalence to random sampling. After this selection, 37 and 25 data respectively for *R. alpina* and *B. barbastellus* were used to generate species distribution models and niche analysis (Fig. 1).

All *B. barbastellus* data used to generate the models were collected between 2000 and 2013, while 82% of *R. alpine* records were collected between 2000 and 2010 and the remaining 18% between 1990 and 2000. Overall, this ensures a good temporal matching with the time period of the land cover map production (2006).

In the recent ecological literature, a diversity of meanings is attributed to the term "niche" so that to avoid confusion constant qualification is advisable (McInerny & Etienne, 2012). In this paper, we follow Phillips et al. (2006) and assume that environmental conditions at the occurrence localities represent samples from realized niche. Therefore, given a certain study area and corresponding environmental dimensions, a niche – based model provides an approximation of the species' realized niche.

To generate models and carry out niche analysis for *R. alpina* and *B. barbastellus*, we used a set of 21 environmental predictor variables. We included one topographical and 19 bioclimatic variables obtained from the WorldClim database (www.worldclim.org/current) (Hijmans et al., 2005). Land cover was obtained from Global Land Cover 2000 (http://bioval.jrc.ec.europa.eu/products/glc2000/products.php). All variable formats were raster files (in ESRI format) with a resolution of 30 arc seconds (0.93×0.93 km = 0.86 km^2 at the equator) and 1,307,195 grid cells. In order to remove the highly correlated variables for the final distribution models, we calculated a correlation matrix using the Pearson's correlation coefficient and selected only variables which showed a weak correlation with the others ($r<0.5$), as done in Russo et al.

(2014). We came up with 11 final environmental predictor variables and used them to model habitat suitability and niche analysis of both species in ASCII files. Bosso et al. (2013) also developed a species distribution model for *R. alpina*, but in the present study we generated a new model as niche comparison required use of the same environmental predictor variables for both species.

We used the most applied method with scarce presence-only data to develop species distribution models for *R. alpina* and *B. barbastellus*, i.e. MaxEnt ver. 3.3.3k (http://www.cs.princeton.edu/~schapire/maxent). This method usually results in good predictive models compared to other presence-only models (e.g. Elith et al., 2006). It uses a generative approach, rather than a discriminative one, which can be an important advantage when the amount of training data is limited (Phillips et al., 2006). Furthermore, it has a good ability to predict new localities for poorly known species (Rebelo & Jones, 2010; Bosso et al., 2013; Russo et al., 2014).

To build the models, we used the presence records for *R. alpina* and *B. barbastellus* selected as described above, and the following environmental predictor variables: Altitude, Land cover, Mean Diurnal Range, Isothermality, Temperature Seasonality, Temperature Annual Range, Mean Temperature of Wettest Quarter, Mean Temperature of Driest Quarter, Precipitation Seasonality, Precipitation of Wettest Quarter and Precipitation of Coldest Quarter. Further details on environmental predictor variables are given in Table 1.

Although altitude is an indirect variable usually correlated with other direct variables such as temperature, we used it because it was easily recorded for the plot-level analyses and applied to the two-scale analyses for consistency.

In the MaxEnt panel, we selected the following options: random seed; remove duplicate presence records; write plot data; regularization multiplier (fixed at 1); 10,000 maximum number of background points; 1000 maximum iterations; and, finally, 20 replicate effects with cross-validate replicated run type. For the latter procedure, 80% of records were randomly extracted for training and 20% for testing the model and the procedure was repeated 20 times. The average final map obtained had a logistic output format with suitability values from 0 (unsuitable habitat) to 1 (suitable habitat).

We used Jackknife analysis to estimate the actual contribution that each variable provided to the geographic distribution models. During this process, Maxent generated three models: first, each environmental predictor variable was excluded in turn and a model created with the remaining variables to check which of the latter was most informative. Second, a model was created using each environmental predictor variable individually to detect which of them had the most information not featuring in the others. Third, a final model was generated based on all variables. Response curves derived from univariate models were plotted to determine how each environmental predictor variable influences the probability of presence.

For the placement of pseudo-absence points following Barve et al. (2011) as background area, we used all ecoregions found in the Italian Peninsula including presence points.

We tested the predictive performance of the models with different methods: the receiver operating characteristics (ROC), analysing the area under curve (AUC) (Fielding & Bell, 1997); the true skill statistic (TSS) (Allouche et al., 2007); and the minimum difference between training and testing AUC data (AUC_{diff})

(Warren & Seifert, 2011). Such statistics were averaged across the 20 replicates run on the 80% (training) vs. 20% (testing) dataset split. These model evaluation statistics range between 0 and 1: excellent models performances are expressed respectively by AUC and TSS value close to 1 and AUC_{diff} close to 0.

Table 1 List of ecogeographical variables used for this study, type and measurement unit.

Type	Ecogeographical Variable	Unit	Source
Topographical	Elevation	m a.s.l.	WorldClim database
Habitat	Land cover	-	GLC 2000
Climatic	Mean Diurnal Range	°C	WorldClim database
	Isothermality	%	WorldClim database
	Temperature Seasonality	°C	WorldClim database
	Temperature Annual Range	°C	WorldClim database
	Mean Temperature of Wettest Quarter	°C	WorldClim database
	Mean Temperature of Driest Quarter	°C	WorldClim database
	Precipitation Seasonality	%	WorldClim database
	Precipitation of Wettest Quarter	mm	WorldClim database
	Precipitation of Coldest Quarter	mm	WorldClim database

Figure 1. Presence records of *Rosalia alpina* (black dots) and *Barbastella barbastellus* (grey dots) considered for the development of a maximum entropy model.

2.3.3 Protecting one, protecting both? Scale-dependent niche differences

We performed niche overlap analyses using the analytical framework proposed by Broennimann et al. (2012) and recently adopted by Di Febbraro et al. (2013) and Russo et al. (2014). The procedure follows three steps: data pre-processing, calculation of the niche overlap measure and testing niche similarity. For niche comparison, we used the Schoener's D index, ranging from 0 (no overlap) to 1 (niche identical). This index was calculated using the "ecospat" package (ran.rproject.org/web/packages/ecospat/index.html; Broenniman et al., 2014).

Niche analysis was conducted both at broad and plot scales. At a broad scale, corresponding to the whole Italian territory, we compared the differences in environmental predictor variable values used to build the species distribution models between species. To quantify niche overlap we used the ordination method of principal component analysis calibrated on the whole environmental space including the presence records where the species occur (for details see Broennimann et al., 2012 and Di Febbraro et al., 2013).

We also compared the niches of *R. alpina* and *B. barbastellus* at the plot scale by exploring differences in tree and site descriptors for 42 trees used by the former for reproduction and 66 nursery roosts of the latter occurring in the same study area (Abruzzo, Lazio and Molise National Park). Our plots are represented by the focal tree and the characteristics found in the space corresponding to its canopy. Trees were found in earlier studies (Russo et al., 2004; 2005; 2007; 2010; 2011a) in which adults of *R. alpina* were observed inspecting suitable trees and *B. barbastellus* roosts were found by capturing bats at drinking sites, fitting them with radiotags and tracking them in daytime to their roost-trees (see Russo et al., 2004; 2011a for further details on methodology). Following Russo *et al.*, (2011a) we considered the following features: (1) tree diameter at breast height (in meter); (2) tree height (in meter, obtained with a clinometer); (3) percentage canopy closure (assessed visually as the percentage of the tree's branches and foliage which was in contact with those of surrounding trees); (4) sun index (estimated on a 0–12 scale: 0, fully shaded tree; 12, sun from all directions); (5) tree condition (live; dead and fallen tree); (6) crown condition (full foliage, foliage or branches partly missing, crown absent); (7) tree position (in forest, in clearing, on forest edge); (8) undergrowth height (0, between 0 and 1 m, >1 m); (9) altitude (in m a.s.l.); (10) habitat type (forest, forested pasture and open shredded forest). "Forest" corresponded to stands mainly unlogged since 1956 or subject to occasional very limited and highly selective logging; "forested pasture" was characterized by pastures interspersed with (or surrounded by) old trees present at low densities and shrubs; and "open-shredded forest", so called because trees used to be pruned by 'shredding' (removing all side branches), a traditional form of management aimed to keep open space for pasture and use removed branches and foliage as animal fodder (Russo et al., 2011a).

To quantify niche overlap at plot scale, we used the Factor Analysis for Mixed Data (FAMD) calibrated on the above records. FAMD is a principal component method to explore datasets including both continuous and categorical variables. The continuous variables are scaled to unit variance and the categorical variables are transformed into a disjunctive data table (crisp coding) in order to balance the influence of both continuous and categorical variables in the analysis.

To explain niche differences between species at both scales we compared environmental predictor variables recorded at presence points and plot descriptors measured in the field on trees used with an Analysis of Variance (ANOVA) followed by a Tukey's test post hoc and χ^2-test at p-value<0.05 using XLSTAT version 2013.1 (Addinsoft, Paris, France).

2.4. Results

2.4.1. Broad geographical scale - model development and validation

As predicted, at the broad geographical scale the potential distribution generated by our models for the two species largely overlapped. The model for *R. alpina* identified substantially uninterrupted areas of occurrence in Italy from the Alps to southern Apennines (Fig. 2). In the remaining area, only a few scattered sites were predicted as suitable. The model for *B. barbastellus* predicted a higher probability of presence in the central Apennines, but also in the southern Apennines and in the western Alps.

The environmental predictor variable sets contributing the most to the models for the two species were similar. In the *R. alpina* model (limiting our analysis to environmental predictor variable with percent contribution > 2%) land cover (45%) and altitude (36%) were the main factors influencing model performance. A slightly lower contribution was provided by isothermality (9%), precipitation seasonality (5%), mean temperature of driest quarter (2.4%) and precipitation of coldest quarter (2.4%). Based on the model's predictions, *R. alpina* has a higher probability of occurring in broadleaved deciduous forest found at relatively high altitudes (1000–1500 m a.s.l.). The environmental predictor variables that were most important to explain the potential distribution of *B. barbastellus* were: land cover (47%), altitude (27%), mean temperature of driest quarter (10%), isothermality (8%), precipitation seasonality (4%) and precipitation of coldest quarter (2.1%). We observed that in the areas predicted as suitable for *B. barbastellus* (Fig. 2), land cover was mainly characterized by broadleaved deciduous forest and deciduous shrubs. Furthermore, in those areas altitude ranged between 1000-1500 m a.s.l., and the mean temperature of the driest quarter between 15-20 ° C.

Species distribution models achieved excellent levels of predictive performance for *R. alpina* and *B. barbastellus* as seen from the AUC, TSS and AUC_{diff} values (Table 2).

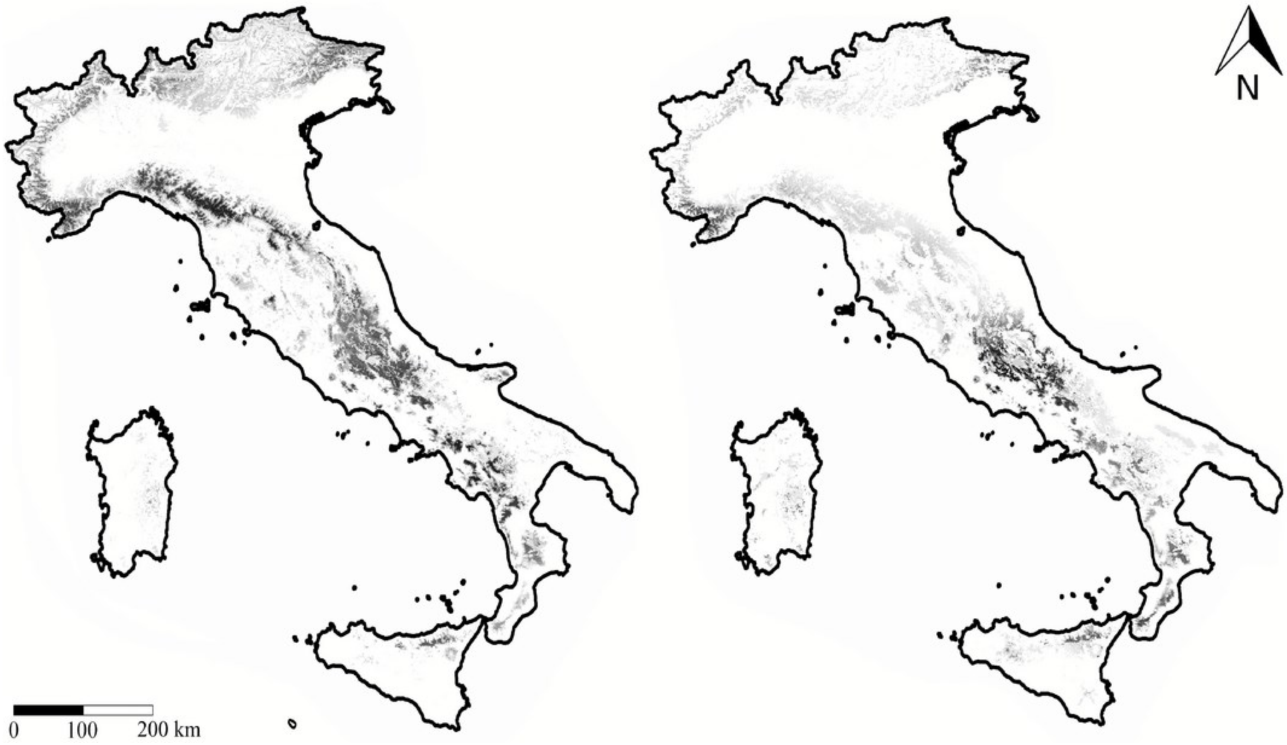


Figure 2. Maximum entropy model developed for *Rosalia alpina* (left) and *Barbastella barbastellus* (right) in Italy. Likelihood of species occurrence is expressed as shades of grey (white = 0, black = 1).

Table 2 Predictive performances achieved for SMDs developed for *R. alpina* and *B. barbastellus*.
AUC, area under curve; AUCdiff, difference between training and testing AUC data; SD, standard deviation, TSS, true skill statistic.

Species	AUC Training	SD	AUC Test	SD	AUC_{diff}	SD	TSS	SD
B. barbastellus	0.953	0.004	0.918	0.029	0.035	0.003	0.701	0.112
R. alpina	0.948	0.011	0.914	0.014	0.034	0.002	0.721	0.155

2.4.2. Protecting one, protecting both? Scale-dependent niche differences

The outcome of niche similarity tests greatly depended on the scale considered. In agreement with our first prediction, the two species' niches were similar at a broad geographical scale, yet they were rather different at the plot scale (Table 3). The graphical representation of the niches of the two species in environmental space at both scales is shown in Fig. 3.

At a broad spatial scale, relatively few differences were spotted between the potential areas of occurrence for the two species, their large-scale ecological requirements being mostly similar (Fig. 3; Table 4). *B. barbastellus* was found to occur in areas whose mean temperature of driest quarter, and to some extent isothermality, were significantly greater than those occupied by *R. alpina* (Fig. 3). Although both species mostly occurred in similar land use types, *B. barbastellus* was predicted to occur in cultivated and managed areas more often than *R. alpina*, whereas the latter was more often predicted to occur in mosaics of tree cover interspersed with other natural vegetation and in semi-open areas of shrubland and deciduous forest.

At a plot scale, in accordance with our second prediction we identified marked differences in the type of habitat used by the two species (Fig. 3; Table 4). Although both species may occur in the same plot and share the same tree, *R. alpina* tends to occur at sites characterized by a lower altitude, lower canopy closure, higher sun index (i.e. sites are more irradiated) and to select shorter trees with a larger diameter than *B. barbastellus* (Figs 3 and 4). The trees used by *B. barbastellus* (mostly snags) occurred significantly more often within forests or along forest edges, whereas *R. alpina* tends to lay eggs in trees found in clearings.

B. barbastellus plots were significantly more frequent in forest, while those of *R. alpina* were more frequent in forested pasture and open-shredded forest (Table 4).

Table 3 Outcomes of niche equivalence and niche similarity tests for *R. alpina* and *B. barbastellus* at broad (Italy) and plot (tree and its immediate surroundings) scales. (n.s.) = not significant

FAMD, factor analysis for mixed data; NS, not significant; PCA-ENV, principal component analysis calibrated on the whole environmental space including the presence records where the species occur.

Scale	Method	Schoener's D	Niche equivalency	Niche similarity	
				R. alpina → *B. barbastellus*	*B. barbastellus* → *R. alpina*
Broad	PCA-ENV	0.487	0.01	0.01	0.01
Plot	FAMD	0.319	0.003	(n.s.)	NS

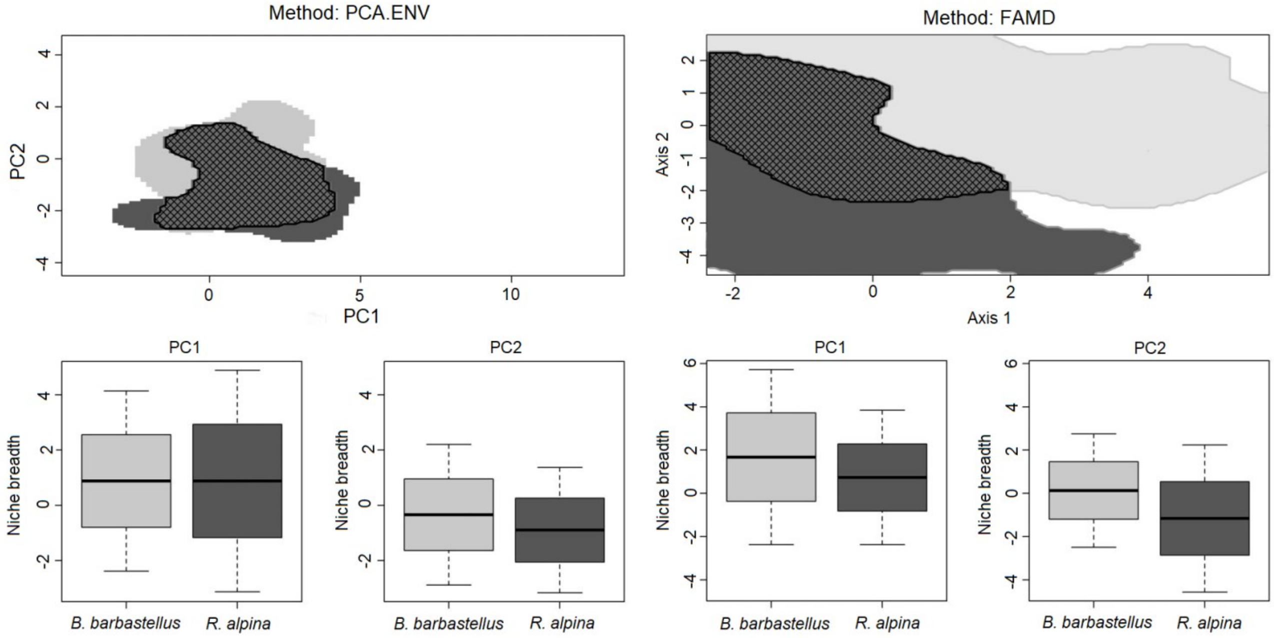


Figure 3. Graphical representation of the environmental niches of *Rosalia alpina* (dark grey) and *Barbastella barbastellus* (light grey). The upper right panel represents the niche of the two species along the two-first axis (PC1 and PC2) of the principal component analysis (PCA-ENV) calculated at broad scale.

Increasing values of PC1 mainly correspond to increasing precipitation of coldest quarter and decreasing mean diurnal range [mean of monthly (maximum temperature – minimum temperature]) and mean temperature of wettest quarter. Increasing values of PC2 mainly correspond to increasing values of isothermality, mean temperature of driest quarter and precipitation seasonality (coefficient of variation). The upper left panel shows the species' niches along the two-first axis (Axis1 and Axis2) of the factor analysis for mixed data (FAMD) calculated at plot scale. Increasing values of PC1 mainly correspond to increasing percentage of canopy closure and tree height and decreasing sun index and tree diameter, whereas PC2 mainly increases with altitude and tree height. For each scale, the corresponding box-and-whisker plots illustrated below represent the median, quartiles, and extremes of the two component values.

Table 4 Statistical comparisons (analyses of variance or chi-square test) respectively of environmental predictor variables values used to build the species distribution models on a broad geographical scale (Italy) and plot characteristics measured in an area of sympatry (in central Italy) between *Rosalia alpina* and *Barbastella barbastellus*. Sample size for Italy: n = 25 (*B. barbastellus*), n = 37 (*R. alpina*), broad scale; n = 66 (*B. barbastellus*); n = 42 (*R. alpina*), plot scale.

[a] = F; [b] = χ^2; $p < 0.001$ = ***; $p < 0.01$ = **; $p < 0.05$ = *; NS = not significant.

Italian territory

Variable	*B. barbastellus*	*R. alpina*	*Statistics*	*P*
Altitude (m)	984.91 ± 401.08	1100 ± 377.21	1.51 [a]	NS
Annual Mean Diurnal Range (°C)	7.64 ± 1.03	7.11 ± 1.43	3.04 [a]	NS
Annual Temperature Range (°C)	25.01 ± 1.82	24.58 ± 2.32	0.87 [a]	NS
Isothermality (%)	29.97 ± 2.24	28.25 ± 3.45	5.91 [a]	0.01*
Landcover			25.65 [b]	0.04*
Mean Temperature of Driest Quarter (°C)	16.45 ± 6.46	11.98 ± 8.70	5.75 [a]	0.01*
Mean Temperature of Wettest Quarter (°C)	8.83 ± 3.1	9.06 ± 3.22	0.09 [a]	NS
Precipitation of Coldest Quarter (mm)	213.02 ± 33.47	207.37 ± 35.29	0.46 [a]	NS
Precipitation of Wettest Quarter (mm)	283.88 ± 37.80	293.80 ± 48.43	0.89 [a]	NS
Precipitation Seasonality (%)	28.67 ± 9.57	27.51 ± 7.11	0.32 [a]	NS
Temperature Seasonality (%)	6125.23 ± 303.46	6146.85 ± 369.86	0.02 [a]	NS
Plot scale				
Altitude	1487.26 ± 115.78	1384.45 ± 137.54	7.86 [a]	0.01*
Undergrowth height			0.36 [b]	NS
Canopy closure	32.76 ± 37.71	13.80 ± 29.40	7.54 [a]	0.01*
Crown condition			0.48 [b]	NS
Habitat type			22.81 [b]	0.01*
Tree condition			21.89 [b]	0.02*
Sun index	5.01 ± 3.32	9.85 ± 2.90	5.70 [a]	0.01*
Tree diameter	0.92 ± 0.24	1.03 ± 0.31	4.14 [a]	0.03*
Tree height	20.36 ± 7.66	8.05 ± 4.81	8.55 [a]	0.01*
Tree position			34.80 [b]	0.01*

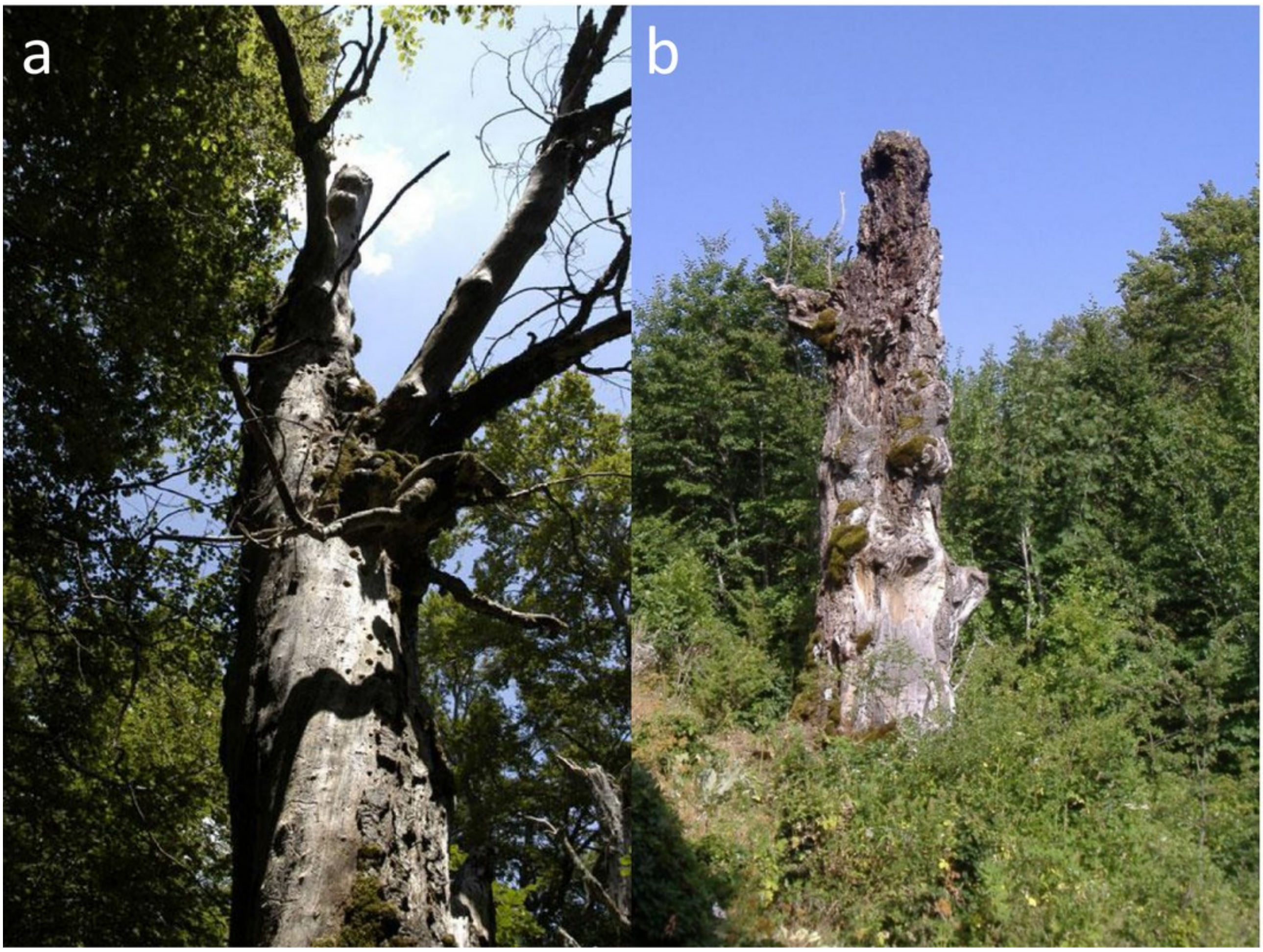


Figure 4. Typical trees selected by *Barbastella barbastellus* (a) and *Rosalia alpina* (b) in Central Italy.

2.5 Discussion

2.5.1 Scale-dependent niche differences between *Rosalia alpina* and *Barbastella barbastellus*

Our study showed that two organisms largely dependent on standing dead wood for a crucial part of their life cycle (reproduction), despite being found in the same areas and often in the same tree, exhibit different ecological requirements that mostly emerge when niches are compared at a small spatial scale. A broad geographical scale is too coarse-grained to reveal considerable differences. In accordance to our first hypothesis, the broad scale analysis essentially shows that the two species are largely sympatric and syntopic, occurring in the same areas of broadleaved forest, while failing to detect differences in microhabitats occupied.

At a large scale, only limited differences in preferred environmental predictor variables values were detected. Besides a small difference in isothermality (= mean diurnal range/temperature annual range, expressing temperature "evenness" over a year), *B. barbastellus* tends to occur in areas characterized by a higher temperature of the driest quarter than does *R. alpina*. Such mild conditions in the season when bats are active are important for reproduction. Pregnancy and lactation are energy-demanding because during these phases females need to be homeothermic and avoid using torpor to allow foetal development and to provide milk for their young, and because the energetic demands of lactation are high (Racey, 1973; Grinevitch et al., 1995). Higher roost temperatures (inevitably influenced by ambient temperature) reduce thermoregulatory costs and may thus be favourable for females during pregnancy and lactation (Grinevitch et al., 1995). Higher nocturnal temperatures also increase insect activity (Taylor, 1963) and thus bat foraging (Russo & Jones, 2003). For *R. alpina*, spring temperatures are more important than summer temperatures (Bosso et al., 2013): at least some warmth is needed daily in spring for larval development, when larvae build pupal cells beneath bark where they will eventually undergo metamorphosis in summer (Duelli & Wermelinger, 2005).

We found the potential distribution of the two species to largely overlap in terms of land use types, with both species predicted to occur in areas dominated by broadleaved forest. *B. barbastellus* was also predicted to occur in cultivated and managed areas more often than *R. alpina*. The beetle was instead predicted to occur more in mosaics of tree cover interspersed with other natural vegetation, and in semi-open areas of shrubland and deciduous forest. *B. barbastellus* forages in a broad range of habitats (Hillen et al., 2011; Zeale et al., 2012), including farmland, which may locally be of great importance (Ancillotto et al., 2014). In contrast, the more marked preference for semi-open landscapes detected for *R. alpina* may reflect its tendency to lay eggs in well-insulated trees (Duelli & Wermelinger, 2005; Russo et al., 2011a; Bosso et al., 2013) – which, as we will see, was revealed by plot-level comparisons.

Apart from these limited differences, the overlap of the species' potential distributions suggests that wide areas of syntopy can be protected, providing conservation benefits to both the beetle and the bat. In Italy, over 52% of habitat suitable for *R. alpina* is unprotected (Bosso et al., 2013). Given the distributional overlap, we may assume that the potential distribution of *B. barbastellus* also largely falls out of boundaries of parks and reserves. Broad scale analysis, however, reveals little of the detailed ecological differences characterizing the two species and gives no information on the type of management that the two species need.

In agreement with our second prediction, the niche comparison conducted at plot scale revealed marked differences in the ecological requirements of the two species, clearly rejecting the hypothesis that simply preserving snags in areas of syntopy will necessarily favour both taxa. The first obvious difference is that trees frequented by *R. alpina* had a much smaller canopy closure (on average almost one third of that recorded for *B. barbastellus* roosts). Consequently, these trunks are much more exposed to sun than those used by the bat, as can be seen by their significantly greater sun index. Such trees were also found to occur more often in clearings than those used by *B. barbastellus*, which were more frequent in forest interiors. Exposure of reproductive substrate to the sun is of utmost importance for many saproxylic beetles, as it provides the microenvironmental conditions that larvae require to develop (e.g. Lindhe et al., 2005; Buse et al., 2007; Vodka et al., 2009).

In the study area we focused on, clearings resulted from traditional shredding to open space for livestock grazing (Russo et al., 2011a). Shredded trees often attain large diameters and smaller heights, explaining the structural differences we found between trees favoured by *R. alpina* and *B. barbastellus*. For *B. barbastellus*, tall trees are important as bats tend to roost close to the canopy (Russo et al., 2004). Roosting at height is important for at least two reasons: first, bats are less accessible to predators; second, snag tops often protrude out of the canopy, allowing roosts to be irradiated even in dense forest (Russo et al., 2004; 2007). By selecting such roosts, reproductive females may achieve lower-cost thermoregulation while gaining some extra-foraging time through earlier roost emergence in a cluttered habitat, where predation risks are lower (Russo et al., 2007; 2011b). Female *B. barbastellus* roosting in dead trees may be favoured by the absence of living wood tissue, which reduces temperatures in roost trees (Law & Anderson, 2000).

B. barbastellus plots were more frequent within the forest (which is mostly unmanaged, or subject to very little management in the study area), whereas *R. alpina* plots occurred more often within open shredded forest and forested pasture. This result is in agreement with habitat selection studies carried out in the same area (Russo et al., 2005; 2011a) and matches well with tree structural characteristics preferred by the two species.

Despite the co-occurrence of *R. alpina* and *B. barbastellus* often in the same habitats and sometimes even in the same trees, the plot-level analysis revealed important differences in both habitat type and tree descriptors, indicating that different types of management may be required for the two species. This once more highlights that avoiding confusion between “used” and “preferred” resources or conditions is crucial, and emphasises the need to support management with sound scientific data instead of relying on merely qualitative observation-driven “expert judgement” which may be misleading.

Clearly, practices such as girdling may favour both *R. alpina* and *B. barbastellus* (Russo et al., 2010; 2011a) by increasing snag density but will not suffice to grant these species the conditions they require. In conclusion, fully reconciling the management requirements of the two species appears to be unviable on a tree/habitat scale: while the beetle would benefit from clearing some surrounding trees, the bat prefers its roosts to be protected by a denser canopy closure. Leaving unmanaged forest patches will increase the availability of dead wood (e.g. Ranius & Jansson, 2000), yet it will also lead to the expansion of closed forest avoided by *R. alpina*, at the expense of the beetle's preferred habitat.

In areas of syntopy, successful management may only be achieved through a landscape approach, i.e. by allowing the interspersion of forest patches with clearings and semi-open areas where snags are more sun-exposed. In the Italian Apennines, this landscape mosaic was once typical of traditionally exploited forest, when humans were more widespread in the mountains, and is now disappearing along with the loss of non-intensive management such as shredding and moderate livestock grazing. Today grazing intensity is in fact polarized, some areas being subjected to unsustainable livestock loads and consequently overgrazed, while others that were formerly semi-open are being invaded with re-growth of vegetation due to the absence of grazing by herbivores (Moreira & Russo, 2007).

2.5.2. Conservation implications

Our study has clear implications for conservation and management. First, it shows that important ecological differences may be unveiled only at one scale of analysis and missed at others (Levin, 1992). Predicting distribution ranges by modelling may offer important guidance in the development of conservation plans, for example by allowing more exhaustive gap analyses than when conducted only on ranges established on presence data (Bosso et al., 2013); by helping detect corridors, or low-suitability areas in need of restoration to improve connectivity (Doko et al., 2011); by making it possible to predict future distribution patterns under climate or land use change (Rebelo et al., 2010); or by assisting in the discovery of new areas of occurrence for rare species (Rebelo & Jones, 2010). However, subtle differences in ecological requirements could be missed at broad analysis scales, so it is always advisable to carry out analyses at multiple scales. We also once more highlight the value of a landscape approach to compromise the ecological requirements of species of conservation concern that may appear conflictual at habitat scale, and emphasise the importance of heterogeneity typical of traditional landscapes, shaped by hundreds of human generations (Di Salvo et al., 2009; Georgiakakis et al., 2010), to sustain diversity-rich communities.

2.6. References

Allouche, O., Tsoar, A. & Kadmon, R., (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology,* 43,1223–1232.

Ancillotto, L., Cistrone, L., Mosconi, F., Jones, G., Boitani, L. & Russo, D. (2014). The importance of non-forest landscapes for the conservation of forest bats: lessons from barbastelles (*Barbastella barbastellus*). *Biodiversity and Conservation*, DOI 10.1007/s10531-014-0802-7.

Berg, A., Ehnstrom, B., Gustafsson, L., Hallingbäck, T., Jonsell, M. & Weslien, J. (1994). Threatened plant, animal, and fungus species in Swedish forests: distribution and habitat associations. *Conservation Biology,* 8,718–731.

Bosso, L., Rebelo, H., Garonna, A.P. & Russo, D. (2013). Modelling geographic distribution and detecting conservation gaps in Italy for the threatened beetle *Rosalia alpina*. *Journal for Nature Conserv*ation., 21, 72–80.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* 21, 481–497.

Broennimann, O., Petitpierre, B., Randin, C., Engler, R., Breiner, F., D'Amen, M., ... & Di Cola, V. (2014). ecospat: Spatial ecology miscellaneous methods. *R package version*, *1.*

Buse, J., Schrőder, B. & Assmann, T. (2007). Modelling habitat and spatial distribution of an endangered longhorn beetle—a case study for saproxylic insect conservation. *Biology and Conservation,* 137:372–381

Caci, G., Biscaccianti, A., Cistrone, L., Bosso, L., Garonna, A.P. & Russo, D. (2013). Spotting the right spot: computer-aided individual identification of the threatened cerambycid beetle *Rosalia alpina*. *Journal of Insect Conservation,* 17, 787−795.

Di Febbraro, M., Lurz, P.W.W., Maiorano, L., Girardello, M. & Bertolino, S. (2013). The use of climatic niches in screening procedures for introduced species to evaluate risk of spread: the case of the American Eastern grey squirrel. *PLoS ONE,* 8: 1–10.

Doko, T., Fukui, H., Kooiman, A., Toxopeus, A.G., Ichinose, T., Chene, W., & Skidmore, A.K. (2011). Identifying habitat patches and potential ecological corridors for remnant Asiatic black bear (*Ursus thibetanus japonicus*) populations in Japan. *Ecological Modelling,* 222, 748–761.

Drag, L., Hauck, D., Pokluda, P., Zimmermann, K. & Cizek, L. (2011) Demography and dispersal ability of a threatened saproxylic beetle: a mark-recapture study of the rosalia longicorn (*Rosalia alpina*). *PLoS One,* 6:e21345.

Dudley, N., & Vallauri, D., France WWF. (2004). Dead wood—living forests. The importance of veteran trees and dead wood for biodiversity. WWF Report, World Wildlife Fund for Nature, Gland, Switzerland.

Duelli, P. & Wermelinger, B. (2005). Der Alpenbock (*Rosalia alpina*). Einseltener Bockkafer als Flaggschiff Art. *Merkblatt fur die Praxis,* 39,1–8.

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

Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence–absence models. *Environmental Conservation*, 24, 38–49.

Georgiakakis, P., Vasilakopoulos, P., Mylonas, M. & Russo, D. (2010). Bat species richness and activity over an elevation gradient in Mediterranean shrublands of Crete. *Hystrix Italian Journal of Mammalogy,* 21, 43–56.

Grinevitch, L., Holroyd, S.L. & Barclay, R.M.R. (1995). Sex differences in the use of daily torpor and foraging time by big brown bats (*Eptesicus fuscus*) during the reproductive season. *Journal of Zoology, 235,* 301–309.

Hayes, J.P. & Loeb, S.C. (2007). The influences of forest management on bats in North America. In: Lacki, M.J., Hayes, J.P., Kurta, A. (Eds.), Bats in forests: Conservation and management. Johns Hopkins University Press, Baltimore, Maryland, pp 207–235.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.

Hillen, J., Kaster, T., Pahle, J., Griebeler, E.M. & Veith, M. (2011). Sex-specific habitat selection in an edge habitat specialist, the western barbastelle bat. *Annales Zoologici Fennici,* 48,180–190.

Hunter, M.L.J. (1999). Maintaining Biodiversity in Forest Ecosystems. Cambridge University Press, Cambridge.

Jonsson, B.G., Kruys, N. & Ranius, T. (2005). Ecology of species living on dead wood–lessons for dead wood management. *Silva Fennica*, 39,289–309.

Jonsson, M., Ranius, T., Ekvall, H., Bostedt, G., Dahlberg, A., Ehnström, B., Nordén, B. & Stoklandf J.N. (2006). Cost-effectiveness of silvicultural measures to increase substrate availability for redlisted woodliving organisms in Norway spruce forests. *Biological Conservation,* 127, 443–462.

Kurta, A., Murray, S.W. & Miller, D.H. (2002). Roost selection and movements across the summer landscape. In: Kurta, A. and J. Kennedy, (Eds.), The Indiana bat: biology and management of an endangered species. Bat Conservation International, Austin, TX. pp. 118–129.

Law, B.S. & Anderson, J. (2000). Roost preferences and foraging ranges of the eastern forest bat Vespadelus pumilus under two disturbance histories in northern New South Wales, Australia. *Austral Ecology,* 25, 352–367.

Levin, S.A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. *Ecology,* 73, 1943–1967.

Lewis, S.E. (1995). Roost fidelity of bats: a review. *Journal of Mammalogy,* 76, 481-496.

Lindhe, A., Lindelöw, Å. & Åsenblad, N. (2005). Saproxylic beetles in standing dead wood density in relation to substrate sun-exposure and diameter. *Biodiversity and Conservation*, 14, 3033–3053.

Martin, J. & Lopez, P. (2002). The effect of Mediterranean dehesa management on lizard distribution and conservation. *Biological Conservation,* 108,213–219.

Moreira, F. & Russo, D. (2007). Modelling the impact of agricultural abandonment and wildfires on vertebrate diversity in Mediterranean Europe. *Landscape Ecology,* 22,1461–1476.

Phillips, S.J., Anderson, R.P. & Schapire, R. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* 190, 231–259.

Racey, P.A. (1973). Environmental factors affecting the length of gestation in heterothermic bats. *Journal of reproduction and fertility. Supplement,* 19, 175–189.

Ranius, T. & Jansson, N. (2000). The influence of forest regrowth, original canopy cover and tree size on saproxylic beetles associated with old oaks. *Biological Conservation.*, 95, 85–94.

Razgour, O., Hanmer, J. & Jones, G. (2011). Using multi-scale modelling to predict habitat suitability for species of conservation concern: The grey long-eared bat as a case study. *Biological Conservation,* 144, 2922–2930.

Rebelo, H. & Jones, G. (2010). Ground validation of presence-only modelling with rare species: a case study on barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology,* 47, 410–420.

Rebelo, H., Tarroso, P. & Jones, G. (2010). Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology,* 16, 561–576.

Roberge, J.M. & Angelstam, P. (2004). Usefulness of the umbrella species concept as a conservation tool. *Conservation Biology,* 18, 76-85.

Russo, D. & Jones, G. (2003). Use of foraging habitats by bats (Mammalia: Chiroptera) in a Mediterranean area determined by acoustic surveys: conservation implications. *Ecography,* 26, 197-209.

Russo, D., Cistrone, L., Jones, G. & Mazzoleni S. (2004). Roost selection by barbastelle bats (*Barbastella barbastellus*, Chiroptera: Vespertilionidae) in beech woodlands of central Italy: consequences for conservation. *Biological Conservation,* 117, 73–81.

Russo, D., Cistrone, L. & Jones, G. (2005). Spatial and temporal patterns of roost use by tree-dwelling barbastelle bats *Barbastella barbastellus*. *Ecography,* 28, 769–776.

Russo, D., Cistrone, L. & Jones, G. (2007). Emergence time in forest bats: the influence of canopy closure. *Acta Oecologica,* 31,119–126.

Russo, D., Cistrone, L., Garonna, A.P. & Jones, G. (2010). Reconsidering the importance of harvested forests for the conservation of tree-dwelling bats. *Biodiversity and Conservation*, 19,2501–2515.

Russo, D., Cistrone, L. & Garonna, A.P. (2011a). Habitat selection in the highly endangered beetle *Rosalia alpina*: a multiple spatial scale assessment. *Journal of Insect Conservation*, 15, 685-693.

Russo, D., Cistrone, L., Garonna, A.P. & Jones, G. (2011b). The early bat catches the fly: daylight foraging in soprano pipistrelles. *Mammalian Biology*, 76, 87-89.

Russo, D., Di Febbraro, M., Rebelo, H., Mucedda, M., Cistrone, L., Agnelli, P., De Pasquale, P, Martinoli, A., Scaravelli, D., Spilinga, C. & Bosso, L. (2014). What story does geographic separation of insular bats tell? A case study on Sardinian rhinolophids. *PLoS ONE*, 9: e110894.

Simberloff, D. (1998). Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? *Biological Conservation*, 3, 247-257.

Sierro, A. & Arlettaz, R. (1997). Barbastelle bats (*Barbastella* spp.) specialize in the predation of moths: implications for foraging tactics and conservation. *Acta Oecologica,* 18,91–106.

Taylor, L.R. (1963). Analysis of the effects of temperature on insects in flight. *Journal of Animal Ecology,* 32, 99–117.

Vié, J.C., Hilton-Taylor, C. & Stuart, S.N. (2009). Wildlife in a changing world—an analysis of the 2008 IUCN red list of threatened species. IUCN, Gland.

Vodka, S., Konvicka, M. & Cìzek, L. (2009). Habitat preferences of oak feeding xylophagous beetles in a temperate woodland: implications for forest history and management. *Journal of Insect Conservation,* 13, 553–562.

Warren, D.L. & Seifert, S.N. (2011). Environmental niche modelling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications,* 21, 335–342.

Zeale, M.R.K., Davidson-Watts, I. & Jones, G. (2012). Home range use and habitat selection by barbastelle bats (*Barbastella barbastellus*): implications for conservation. *Journal of Mammalogy,* 93,1110–1118.

Castor fiber (www.geckoo-photography.com/european-beaver---castor-fiber.html)

Chapter III

3. Species distribution models as a tool to predict range expansion after reintroduction: a case study on Eurasian beavers (*Castor fiber*)

3.1. Abstract

Species Distribution Models (SDMs) may provide important information for the follow-up phase of reintroduction operations by identifying the main areas most likely to be colonized by the reintroduced species. We used SDMs to identify the potential distribution of Eurasian beavers (*Castor fiber*) reintroduced to Serbia and Bosnia and Herzegovina in 2004-2006 after being historically driven to extinction by overhunting. Models were also used to carry out a gap analysis to assess the degree of protection granted by the national reserve networks to the potentially expanding population. Distances from hydrographic network, broadleaved forest, main watercourses and farmland were the main factors influencing model performance. We estimated that suitable habitat covers 14.0 % (31,000 km^2) of the whole study area. In Serbia, in 2004-2013 beavers expanded their range at a mean colonization speed of 70.9 ± 12.8 km/year (mean ± SD). Only 2.89% of and 9.72% of beaver's suitable habitat lie within the national network of protected areas of Bosnia and Serbia respectively. We detected new potential areas where beavers will likely settle in the near future, advising on where further monitoring should be focused. We also identified low suitability areas to be targeted with appropriate management to improve their conditions as well as important regions falling outside reserve boundaries where protection should be granted.

3.2. Introduction

Reintroduction– i.e. the intentional translocation of species into parts of their historically known range from which they have been extirpated (IUCN/SSC, 2012) – is one of the key strategies adopted to restore biodiversity and ecosystem functioning. Forecasting the expansion of a reintroduced species over a given region as part of reintroduction's follow-up phase may be essential to anticipate events that might otherwise jeopardise the operation's success in order to secure appropriate protection in the newly colonised areas, prevent conflicts with humans, optimise monitoring strategies and apply adaptive management of habitat quality (e.g. Breitenmoser et al., 2001; Armstrong & Reynolds, 2012; McCarthy et al., 2012).

Habitat suitability mapping is frequently used to inform habitat restoration or preservation actions (Gibson et al., 2004) or to identify suitable sites for species reintroduction (e.g. Olsson & Rogers, 2009; Ardestani et al., 2015). This approach has become an important component of conservation planning in recent years, and a wide variety of modelling techniques have been developed for this purpose (Guisan & Thuiller, 2005; Elith et al., 2006; Elith & Leathwick, 2009). Although Species Distribution Models (hereafter SDMs) appear to be a promising tool to guide the planning of reintroduction operations by conservation biologists and landscape managers, so far their applications to this field have been scarce (see e.g. Wilson et al., 2011; Adhikari et al., 2012). Such models may predict the probability of species presence by relating current occurrences and environmental features at sample locations (Guisan & Zinnermann, 2000; Phillips et al., 2006). In this way the areas that are more likely to be re-colonised as well as the most probable routes followed in the process can be forecast, drawing valuable inferences on the establishment, expansion and persistence of a reintroduced species (Armstrong & Seddon, 2008).

A clear understanding of species-habitat relationships makes it possible to identify highly suitable release sites that offer the maximum chance of post-release survival or guide habitat restoration prior to reintroduction (Seddon et al., 2007).

Despite the outstanding potential of SDMs in predicting the expansion of a newly introduced species to a given region (Armstrong & Reynolds, 2012), their use to achieve this goal has been largely neglected. In our study we apply SDMs to assess the post-release expansion of Eurasian beavers (*Castor fiber*) in Serbia and Bosnia and Herzegovina and evaluate the potential role of nature reserves to assist this process. The species was once common across Eurasia from the British Isles to eastern Siberia, inhabiting freshwater habitats surrounded by forest but also reaching into the tundra and steppe zones (Nolet & Rosell, 1998; Macdonald & Barrett, 1993; Macdonald et al., 1995; Halley & Rosell, 2002). Prized for its fur, meat and castoreum (a urine-based fluid secreted from castor sacs for scent marking valued by humans for medic or cosmetic applications), beavers were wiped out by overhunting from most of their range by the middle 19th century (Djoshkin & Safonov, 1972). By the beginning of the 20th century, only 1200 individuals had persisted in eight isolated populations across the entire species' range (Nolet & Rosell, 1998). The remnant populations received legal protection, and since 1922 reintroductions started in many European regions to pursue species conservation and ecosystem restoration (Kollar & Seiter, 1990; Nolet & Rosell, 1994; Halley & Rosell, 2002). To date, the return of *C. fiber* has not yet taken place only in Portugal, Italy, southern Balkans (Greece, Albania, Bulgaria, Macedonia and Montenegro) and Ireland (Halley & Rosell, 2002; Halley et al., 2012), while two wild populations now occur in Scotland on a trial basis (Stringer & Gaywood, 2016). Currently, the species is strictly protected in the European Union under the Bern Convention (Appendix III) and the EU Habitats and Species Directive (Annex V for the Swedish and Finnish populations, Annex II and IV for all others).

The Serbian population of beavers was driven to extinction by overhunting by the second half of the 19th century, apart from an unconfirmed report of a beaver shot near Belgrade at the beginning of the last century (Ćirović et al., 2003,2007). By then, the species had also disappeared from the whole course of the Danube and its tributaries. In 1999, following the shooting of a beaver in Northern Serbia (Vojvodina, Bačka region) that had dispersed from the re-established population of Hungary, a reintroduction operation was started in Serbia and Bosnia and Herzegovina (Ćirović et al., 2001,2003). In 2004 and 2005, 75 subjects imported from Bavaria were reintroduced to the Obedska Bara and Zasavica Special Reserves in Serbia, and 40 were released in Semešnica and Sokočnica rivers in Bosnia and Herzegovina in 2005 and 2006 (Fig. 2).

Although many studies have assessed beaver habitat requirements (Zurowsky & Kasperczyk, 1990; Parker et al., 2001; Fustec et al., 2001; Vorel et al., 2008; Halley et al., 2012) only few applications of habitat suitability models have been carried out, e.g.in the Czech Republic (John et al., 2010) and Austria (Maringer & Slotta-Bachmayr, 2006). In our study, we quantify the potential distribution of *C. fiber* for Serbia and Bosnia and Herzegovina to predict the species' spatial pattern of expansion. We hypothesize that 1) given the overwhelming importance of riparian habitat and the tendency of beavers to feed on crops (Campbell-Palmer et al., 2016), hydrographical network, riparian broadleaved forest and farmland will be the main environmental variables influencing potential distribution; and that 2) since most reserves in the study region do not comprise rivers, the network of protected areas will offer little protection to the aforementioned distribution.

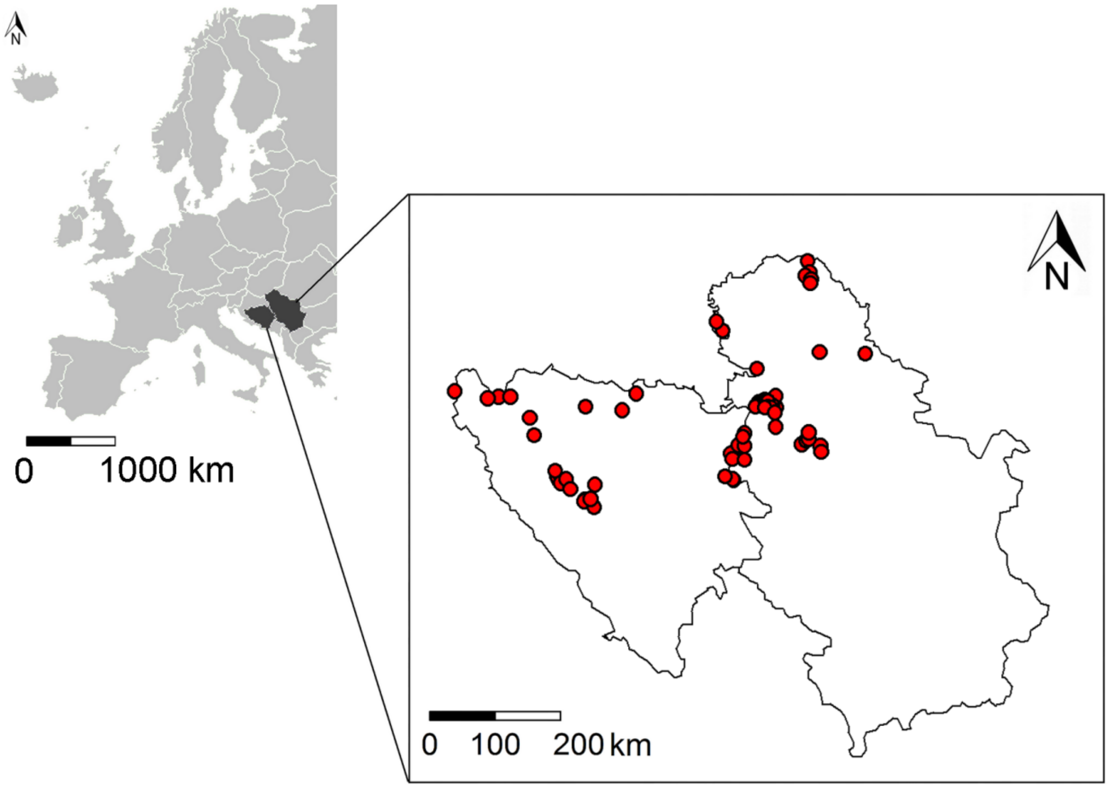
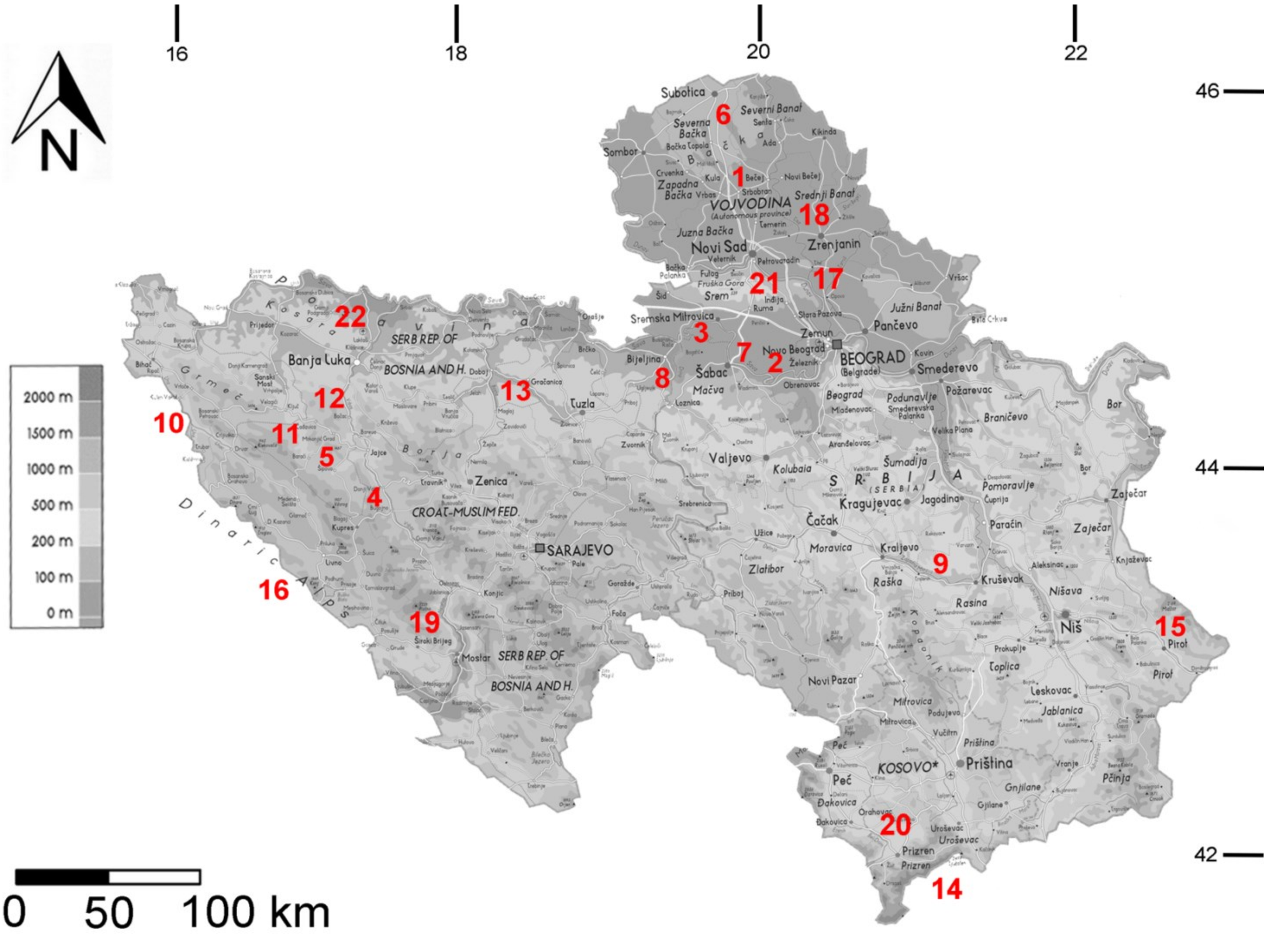


Figure 1. Geographic regions (Serbia and Bosnia and Herzegovina) and *Castor fiber* presence records considered for modelling.



Legend

1 = Vojvodina
2 = Obedska Bara Special Nature Reserve
3 = Zasavica Special Nature Reserve
4 = Semešnica river
5 = Sokočnica river
6 = Pannonian Plain
7 = Sava (Serbia) river
8 = Drina (Serbia/Bosnia Herzegovina) river
9 = Morava river
10 = Una river
11 = Sana river
12 = Vrbas river
13 = Bosna river
14 = Šar mountain
15 = Stara planina
16 = Dinaric Alps mountain
17 = Danube river
18 = Tisa river
19 = Neretva river
20 = Beli Drim river
21 = Fruška Gora National Park
22 = Kozara National Park

Figure 2. Map of the study area showing the main locations and geographic features mentioned in the text.

3.3. Materials and methods

3.3.1. Study area

The study area included the whole territories of Serbia and Bosnia and Herzegovina, covering approximately 139,570 km² between latitudes 41°N-47°N and longitudes 15°E-23°E (Fig. 1). Elevation ranges from 0 up to 2500 m a.s.l. The area is largely mountainous and forested (c. 50% of Bosnia and 25% of Serbia). Arable land covers 53.2% of Serbia and 28.7% of Bosnia and Herzegovina (Corine land cover 2006, https://www.eea.europa.eu). Agricultural production is mostly prominent in the fertile Pannonian Plain situated in the northern part of Serbia (Vojvodina) and in the region of Serbia between the Sava, Drina and Great Morava rivers, while in Bosnia farmland is found in the valleys of Sava, Una, Sana, Vrbas, Bosna and Drina rivers.

3.3.2. Presence records of the Eurasian beaver

We used 71 presence records of Eurasian beavers from authors' personal databases obtained from the post-reintroduction monitoring in Serbia and Bosnia and Herzegovina (Grubešić et al., 2015) (Fig. 1). Records cover years 2004-2014, matching the year of production (2006) of the land cover map used for our study (Russo et al., 2014, 2015).

Each record represented the centroids of a beaver's territory, corresponding to a beaver's shelter (burrow or lodge) or, if this was not found, to the central part of the winter feeding territory. We checked for spatial autocorrelation in species occurrences by using Clark & Evans (1954)'s aggregation index – for further details, see also Ducci et al. (2015) and Di Febbraro et al. (2015).

We implemented a sampling procedure to obtain a representative set of the environmental conditions occurring throughout the area of each territory. Specifically, we defined the boundaries of each territory by intersecting two buffer areas:

1) a first 1.7 km radius circular buffer was drawn around each occurrence record, whose intersections with the watercourse's main axis were assumed to represent the territory's outer limits along the watercourse. The 1.7 km value corresponds to the maximum length known for a beaver's linear territory (Vorel et al., 2008);

2) a second 200 m buffer from each river bank was applied to encompass the territory portions alongside the river stretch, expressing the maximum territory width observed in the study area. This is approximately the longest distance covered by beavers from riverbanks to forage according to published observation and our own records (Allen 1983; D. Ćirović, unpublished data). Subsequently, a point was taken randomly from each of the areas comprised within the above buffers, repeating this sampling procedure 10 times and

obtaining 10 independent sets of 45 “occurrence” points, one for each territory. Each of the 10 replicated sets was used to train a separate SDM.

To fit SDMs with the global scale (Gallien et al., 2012 – see below), we used GBIF records.

3.3.3. Environmental variables

To generate SDMs we started from a set of 13 environmental predictors rasterized at a resolution of 100 m, including five topographical and seven habitat classification variables, in combination with the hydrographical network. The topographic predictors included a Digital Elevation Model (DEM) derived by Jarvis et al. (2008), from which altitude and slope were taken, and the following additional topographical indices (Wilson et al., 2007): Terrain Ruggedness Index (TRI), Topographic Position Index (TPI) and roughness index. Habitat predictors were calculated as Euclidean distances from the 2006 Corine Land Cover classes (European Environmental Agency; http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster) and from the hydrographic network of the Digital Chart of the World (DCW; http://www.diva-gis.org/gdata). The 13 predictors were checked for pairwise correlation and reduced to 11 considering a variance inflation factor less or equal to three (Zuur et al., 2010).

In order to choose the most appropriate set of environmental predictors, we developed a variable selection procedure taking into account model performance and overfitting.

The variables selection procedure was implemented for each of the 10 replicated occurrence datasets separately, then selecting the set of predictors that emerged as the best in most of the 10 replicates. First, we trained a full SDM (i.e. including all the 11 starting predictors), for which we calculated the training AUC and the relative variables importance through the functionalities provided in the biomod2 package. Subsequently, considering the rule of thumb proposed by Harrell et al. (1996) that prescribes the number of predictor variables in a model not to exceed one tenth of the number of occurrences, we selected the first seven most important variables from the full SDM. Starting from them, we trained four candidate SDMs including from the first four to all the seven predictors (see table below).

Predictors	SDM_I	SDM_II	SDM_III	SDM_IV
1. Euclidean distance from main watercourses	✓	✓	✓	✓
2. Euclidean distance from hydrographic network	✓	✓	✓	✓
3. DEM	✓	✓	✓	✓
4. Euclidean distance from farmlands	✓	✓	✓	✓
5. Euclidean distance from inland marshes		✓	✓	✓
6. Euclidean distance from broad-leaved forests			✓	✓
7. Euclidean distance from mixed forests				✓

For these four models, occurrences were randomly split into a 70% sample that was used to calibrate the model, and 30% to evaluate the model's predictive performance, repeating the procedure 10 times. Specifically, we calculated a training AUC on the 70% sample, a testing AUC on the 30% and the difference $AUC_{training} - AUC_{testing}$ (AUC_{diff}; Warren & Seifert, 2011). AUC_{diff} was interpreted as a measure of model overfitting (Radosavljevic & Andeson, 2014). The significance of the differences in $AUC_{testing}$ and AUC_{diff} among the candidate models was tested with the Wilcoxon signed–rank test. Finally, the group of predictors that produced an SDM with significant highest $AUC_{testing}$ and lowest AUC_{diff} out of all candidate models (i.e. the model with the highest predictive performance and the lowest overfitting) was selected as the final.

The variable selection procedure identified the following six predictors (Table 1) as those producing the best models: Elevation and Euclidean distance from hydrographic network, main watercourses, farmlands, inland marshes and broadleaved forests.

3.3.4. Modelling procedure

A growing amount of literature highlights that environmental truncation in niche estimation for areas encompassing only a small portion of a species' global range produces severely biased predictions (Barbet–Massin et al., 2010; Raes, 2012; Guisan et al., 2014). As *C. fiber* is distributed across the Palearctic (IUCN, 2012) and our study area represents a small portion of the entire range, SDMs were produced using a hierarchical structure from a global to regional scale (Pearson et al., 2004; Lomba et al., 2010; Gallien et al., 2012, Di Febbraro et al., 2015). Following this approach, models were first implemented to estimate the species' niche at their global range scale using bioclimatic variables (Global SDMs), then refined at a regional scale using land cover, hydrographical network and topographical variables as environmental predictors (Regional SDMs; Pearson et al., 2004; Lomba et al., 2010; Gallien et al., 2012, Di Febbraro et al., 2015).

Following Gallien et al. (2012) we fitted SDMs at the global scale with species occurrence data, pseudo-absences and bioclimatic predictors and then calculated a committee averaging as a main outcome. Instead of a 'traditional' probability of occurrence output, the committee averaging describes the percentage of agreement on the species presence between various model projections (for further details, see Thuiller et al., 2009) and was used as an additional input to train regional SDMs. Species occurrences were gathered from the Global Biodiversity Information Facility (GBIF) database. The accuracy of records was assessed by including only those whose positional values featured at least two decimal digits (0.01 decimal degrees, corresponding to 1.11 km at the equator; Strubbe et al., 2015). In addition, we filtered these data by removing duplicated records and those with unrealistic coordinates. A set of 10,000 pseudo-absences were randomly placed over a region identified by the WWF Terrestrial Ecoregions (Olson et al., 2001) that included beaver records (Barve et al., 2011). As initial set of environmental predictors, we considered the 19 bioclimatic variables derived from the WORLDCLIM database at a spatial resolution of 5 arc–minutes (ca. 10 km; Hijmans et al., 2005). To take into account the pairwise correlation between the predictors, the final set of variables was sub-selected considering a variance inflation factor (VIF) ≤ 3 (Zuur et al., 2010). The final set included Mean Diurnal Range (BIO2), Isothermality (BIO3), Mean Temperature of Wettest Quarter (BIO8), Precipitation Seasonality (BIO15), Precipitation of Warmest Quarter (BIO18) and Precipitation of Coldest Quarter (BIO19). All procedures were carried out with the packages “spatstat”, “maptools”, “rgeos” and “raster” in the R environment (R Development Core Team, 2015). The potential species distributions were predicted using an ensemble forecasting approach, as implemented in the R package “biomod2” (Thuiller et al., 2009). We considered the following seven modelling algorithms: generalized linear models (GLM), generalized additive models (GAM), generalized boosted models (GBM), random forests (RF), multivariate adaptive regression spline (MARS), maximum entropy models (MAXENT) and BIOCLIM (SRE) – for further details, see Thuiller et al. (2009). Each occurrence dataset was randomly split into a 70% sample, used for the calibration of the model, and a remaining 30%, used to evaluate model predictive performance, repeating the procedure 10 times and averaging the results. The predictive performance of each model was assessed by measuring the area under the receiver operating characteristic curve (AUC) (Hanley & McNeil, 1982) and the true skill statistic (TSS) (Allouche et al., 2006). Models were then projected over the study area using bioclimatic rasters at a resolution of 1 arc–seconds (ca. 1 km) following the direct downscaling approach (see Araujo et al. 2005; McPherson et al. 2006; Barbosa et al., 2010; Bombi & D’Amen 2012; Fernandes et al., 2014). To avoid using poorly calibrated models only projections from models with AUC ≥ 0.80 were considered in all subsequent analyses (Di Febbraro et al., 2015). As mentioned above, the committee averaging was chosen as the ensemble forecasting method (Gallien et al., 2012).

Regional SDMs were developed using an ensemble forecasting approach, as implemented in the package ”biomod2”in the R software (R Development Core Team, 2009) (Thuiller et al., 2009). Biomod2 is a modelling platform that makes it possible to train SDMs using different modelling techniques, as well as to evaluate them and perform different averaged outputs of the single-model predictions (see Thuiller et al., 2009). Using different statistical methods to model species distribution is highly recommended as prediction discrepancies between different techniques can be very large (Araujo et al., 2005; Thuiller et al., 2009). We considered the following seven modelling techniques (Thuiller et al., 2009; Jiguet et al., 2010; Ducci et al., 2015): (1) generalized linear models (GLM); (2) generalized additive models (GAM); (3) generalized boosted models (GBM); (4) random forests (RF); (5) multivariate adaptive regression spline (MARS), (6) maximum entropy models (MAXENT) and (7) BIOCLIM (Surface Range Envelope, SRE) – for further details, see Thuiller et al. (2009).

Following Pio et al. (2014), the modelling settings were tuned as follows. GLMs and GAMs were calibrated using a binomial distribution and a logistic link function. GBMs were calibrated with a maximum number of trees set to 5000, threefold cross-validation procedures to select the optimal numbers of trees to be kept and a value of seven as maximum depth of variable interactions. Random forest models were fitted by growing 750 trees with half the numbers of available predictors sampled for splitting at each node. MARS models were fitted with a maximum interaction degree equal to 2, MAXENT models were fitted with the default settings apart of a maximum value of 1000 iterations and default parameters were used to fit the SRE model.

Each occurrence dataset was randomly split into a 70 % sample, used for the calibration of the model, and the remaining 30 %, used to evaluate model performance. A set of 10,000 background points were randomly placed in the study area to characterize its environment and represent pseudo-absences. According to Gallien et al. (2012), these background points were not considered ‘true’ absences, i.e. we assumed that some absences probably reflect environmental conditions where the species cannot survive, while others reflect locations where the species has not been surveyed due to imperfect detection. These pseudo-absences were weighted by the committee averaging projections calculated with the Global SDMs (Appendix S2): where the Global SDMs showed a high level of agreement with an absence (i.e. a low habitat suitability) we attributed a high weight to that absence (i.e. a high probability of being a ‘true’ absence), and vice versa. The weight was calculated by an inverse logistic transformation (equation 1) to obtain a stronger discrimination between the predictions of absences and presences (Gallien et al., 2012):

$$\text{Weight}(x) = \frac{1}{1+\left(\frac{\text{projGlob}(x)}{\text{projGlob}(x)-1}\right)^2} \qquad (1)$$

where Weight(x) is the weight attributed to the pseudo–absence x, which depends on projGlob(x), the prediction of the Global SDMs at the location of x [if projGlob(x) = 1 then Weight (x) = 0].

The models’ predictive performances were assessed by measuring the area under the receiver operating characteristic curve (AUC) (Hanley & McNeil, 1982) and the true skill statistic (TSS) (Allouche et al., 2006). These validation methods have been widely used (e.g. Russo et al., 2014, 2015; Feuda et al., 2015; Bosso et al., 2016 a, b) and offer excellent performances. The data splitting procedure was repeated 10 times and the evaluation values averaged. For each of the 10 replicates of the species’ dataset, we ran a total of 70 Regional SDMs (seven algorithms x 10 splitting replicates for model evaluation). Model averaging (ensemble model) was performed by weighting the individual model projections by their AUC scores (only models with AUC ≥ 0.8) and averaging the result—a method shown to be particularly robust (Marmion et al., 2009). The relative importance of variables was also calculated from the ensemble model using the specifically devoted functionality available in the biomod2 package (Jiguet et al., 2010). The final projected distribution of the species was obtained by averaging the projections from our 10 replicated ensemble models. This final map was also transformed into presence–absence values using a threshold maximizing sensitivity (the percentage of correctly predicted presence) and specificity (the percentage of correctly predicted absence) (Fielding & Bell, 1997). Such threshold has been widely used (e.g. Algar et al., 2009;

Dubuis et al., 2011; Di Febbraro et al., 2015) and constitutes one of the most accurate approaches (Liu et al., 2005).

3.3.5. Assessment of potentially colonized suitable habitat

In order to discriminate predicted suitable areas already potentially colonized by beavers from those potentially colonizable in the future, we estimated the approximate distance covered by the beaver population since the release until the year of the most recent available record. For this analysis, we used only 39 records of beavers reintroduced to Serbia, because for Bosnia and Herzegovina we had gaps in the dataset for some years (2008 and from 2011 to 2013), and we also discarded some locations of subjects originated from Croatia and Hungary. First, we pooled the species' presence records according to their recording date in two-years groups from 2004 to 2013, obtaining 8 groups of points. Subsequently, we calculated the maximum distance between all the points within each group ("spDists" function in the *sp* package; Pebesma & Bivand, 2005). This represents the maximum distance covered by beavers in one year. Starting from the set of maximum distances calculated in this way, we computed the minimum, mean and maximum distance values, then generated three buffers around all occurrences points used for analysis with the "gBuffer" function of *rgeos* package (Bivand & Rundel, 2013). Suitable pixels falling inside the buffers refer to areas already potentially colonized by beavers up to 2014 (i.e. the year following that of the most recent available record). On the contrary, suitable areas placed beyond the maximum dispersal distance estimated up to 2014 (i.e. outside the buffers) represent areas that might be colonized by the species in the future.

3.3.6. Conservation gap analyses

To assess the degree of protection granted to Eurasian beavers by the reserve network of Serbia and Bosnia and Herzegovina, we carried out two conservation gap analyses, one based on the actual occurrence maps, the other based on the binarized potential distribution map (e.g. Bosso et al., 2013; Bosso et al., 2016c). We overlaid such maps with the shape files containing the boundaries of the nature reserves of Serbia and Bosnia and Herzegovina. The shape files of the protected areas for these two countries were downloaded from http://www.protectedplanet.net/ (UCN and UNEP-WCMC, 2016). We included national parks, Ramsar Network Areas (http://www.ramsar.org/activity/ramsar-culture-network) and Reserves (Natural, Managed Nature, Strict Nature and Special Reserves).

3.4. Results

3.4.1. Model performances and habitat suitability

Both Global and Regional SDMs showed good or excellent levels of predictive performance as indicated by the AUC and TSS values. AUC and TSS for Global SDMs had a mean value and a standard deviation respectively of 0.813 ± 0.019 and 0.537 ± 0.042, whereas evaluation scores for Regional SDMs were equal to 0.968 ± 0.006 and 0.824 ± 0.021. According to our hypothesis, the environmental predictors that were most important to explain the beaver's potential distribution were distances from hydrographic network, broadleaved forests and farmland. The species was also predicted to occur most likely at lower altitudes (Fig. 3), but might reach 1000 m a.s.l. in the suitable areas of Southern and Eastern Serbia (Šar mountain in the south and Stara planina mountain in the east) and Western Bosnia (Dinaric Alps) (Figs 2 and 4). In general, habitat suitability decreased for increasing distances from hydrographic network, broadleaved forests, agricultural lands, main watercourses and inland marshes. For main watercourses, the variable's response curve also showed a second peak at greater distances, likely corresponding to the places where secondary tributaries occurred (Fig. 3). Regional model predictions showed that the most suitable areas were concentrated in Northern Serbia and Bosnia and Herzegovina along the main rivers (i.e. Danube, Sava, Tisa and Drina rivers) and their tributaries, in Central Serbia and Bosnia and Herzegovina along the Great Morava, Bosna, Vrbas, Una and Sana Rivers and in Southern Serbia and Bosnia and Herzegovina along secondary river courses such as Neretva and Beli Drim Rivers, as well as many Great Morava's tributaries (Figs 2 and 4). The percentage of suitable habitat for the beaver accounts for ca. 14% (~31,000 km^2) of the whole study area.

When taken separately, outputs of the algorithms employed to implement the ensemble model all predicted a high amount of suitable habitat in Southern, South-eastern and Central regions of Serbia (Sar mountain and Beli Drim, Danube, Sava and Morava Rivers) and in Central-western regions of Bosnia (Sana and Sokočnica Rivers). Unsuitable areas were mostly detected for Northern Serbia and North-western Bosnia (Figs 2 and 4). The most restrictive predictions were provided by Random Forest (RF) and Generalised Boosting Models (GBM) whereas Generalised Linear Model (GLM) and Generalised Additive Model (GAM) detected suitable habitat mostly along the main rivers as well as their tributaries.

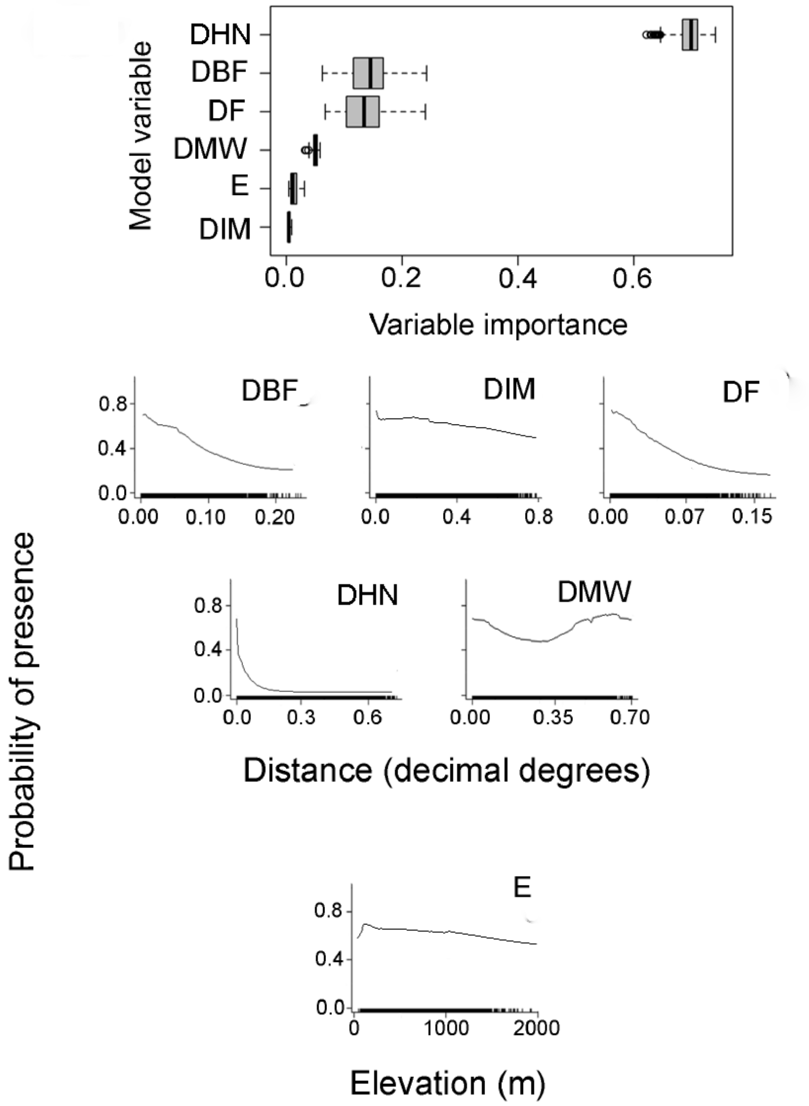


Figure 3. Box plot and response curves for the six variables used to model *Castor fiber* potential distribution in Serbia and Bosnia and Herzegovina with the "biomod2" computer platform. The Box plot (top) expresses variable importance (variables are listed in decreasing order of importance from top to bottom). Response curves (bottom) correlate probability of occurrence (y axis) with values of the explanatory variables (x axis). Each curve represents one variable (DBF = distance from broadleaved forest; DIM = distance from inland marshes; DF = distance from farmlands; DHN, = distance from hydrographic network; DMW = distance from main watercourses; E = elevation). Distances are expressed in decimal degrees, elevation in m a.s.l.

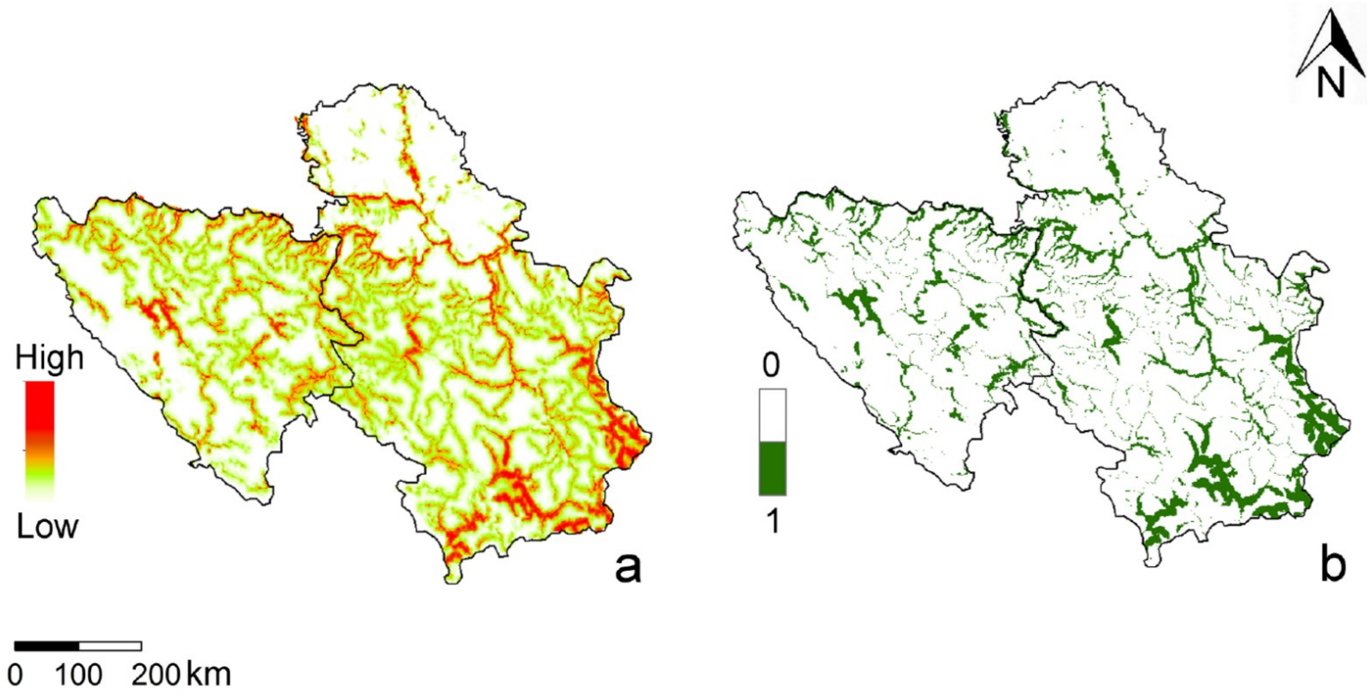


Figure 4. SDM output for *Castor fiber* in Serbia and Bosnia and Herzegovina (**a**) and presence/absence binary map (**b**) obtained using the TSS metrics as a threshold.

3.4.2. Assessment of potentially colonized suitable habitat

From 2004 to 2013 the dispersal distances travelled by the species ranged between 56.61-88.28 km/year with a mean colonization speed of 70.9 ± 12.8 km/year (mean ± SD). According to minimum, mean and maximum dispersal distances potentially covered by beavers since their release up to 2014 (ca. 35% of the all suitable area; Figure 5) individuals from Serbia might have reached the suitable areas of Central Bosnia and Herzegovina since the last occurrence of 2013 had they dispersed at the maximum estimated speed. Within this portion of territory, they would have been more likely to colonize the northern part of Serbia and Bosnia and Herzegovina where more suitable areas occur, especially along the Danube, Tisa and Sava Rivers, exploiting the structural connectivity offered by the hydrographic network. According to the maximum distance buffer, beaver’s populations could have reached large portions of Central-Eastern Serbia, especially along the Great Morava river and its tributaries. By 2014, therefore, the species might have reached the borders of the country in the north-western part of the study area, joining individuals from Croatia. Large parts of South-Eastern Serbia beyond the maximum dispersal distance estimated up to 2014 are also suitable for beavers, so they might be colonized in the future.

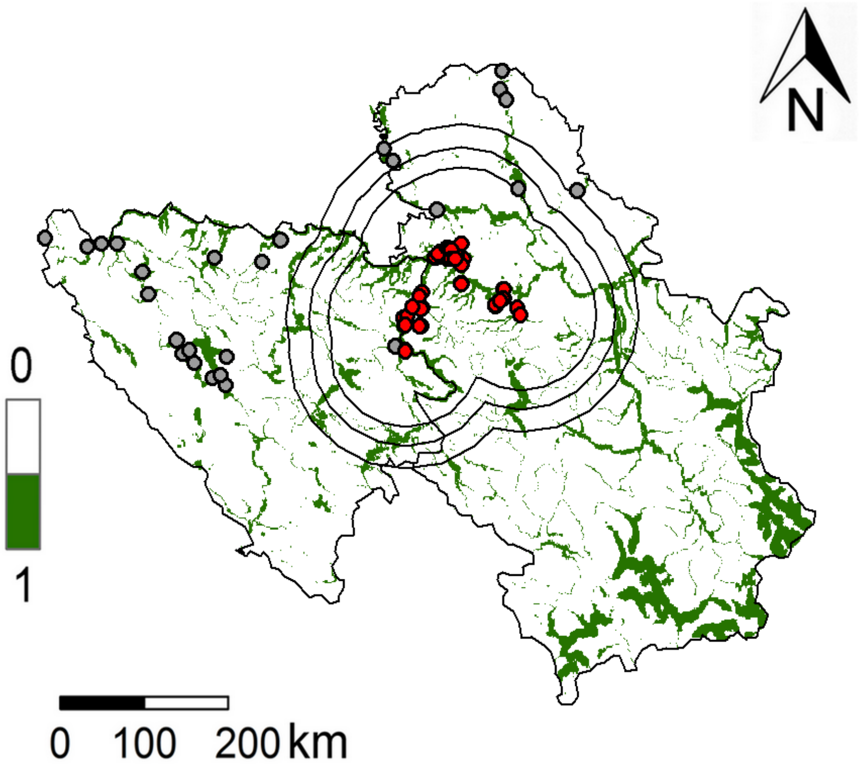


Figure 5. Potential dispersal distances of beavers projected to year 2014 representing the minimum, mean and maximum colonization speeds. The buffers are computed considering only Serbian occurrences (red filled circles) and excluding those derived from Bosnia, Hungary and Croatia (grey filled circles).

3.4.3. Conservation gap analyses

As hypothesised, the country's reserve network offers little protection to the species (Table 2). Specifically, no presence records for Bosnia and 39.06% of those for Serbia lie within the boundaries of the national network of protected areas, while only 2.89% of Bosnian and 9.72% of Serbian suitable habitat fall within this network. In particular, beavers occur in the Serbian Special Natural Reserves of Zasavica and Obedska bara, where they were reintroduced, and in the Fruška Gora National Park in the northern part of the country near Croatia (Figs 2 and 6a). In Serbia, small portions of suitable habitat lie in the Natural Reserves and National Park previously mentioned in the north as well as along river valleys in the Šar planina National Park and Stara Planina Nature Park, respectively in south and eastern Serbia (Figs. 2 and 6b). A very small portion of suitable habitat in Bosnia and Herzegovina falls within three Ramsar sites (a list of wetlands of international importance) and in the Kozara National Park in the country's northern part between Sava and Vrbas Rivers.

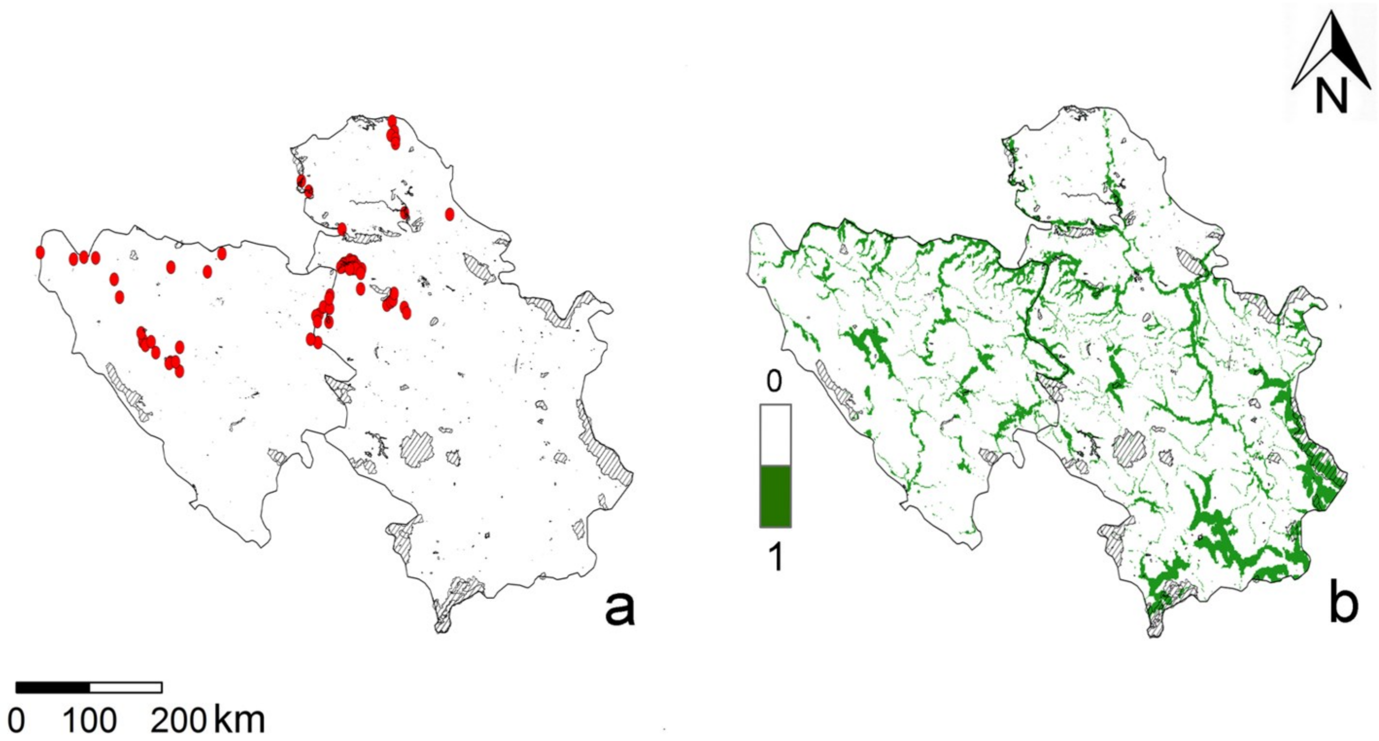


Figure 6. Percentage of *Castor fiber* presence records (red filled circles) (**a**) and suitable habitat (green areas) (**b**) and their overlay with networks of protected areas (simple hatch) for Serbia and Bosnia and Herzegovina.

3.5. Discussion

We found that a wide area of Serbia and Bosnia and Herzegovina predicted as suitable for beavers might be colonized in the near future, especially in Central and Southern Serbia. Our analysis highlighted that beavers have remarkable colonization skills, demonstrated by a very high dispersal distance travelled soon after reintroduction, which allows them to reach quickly optimal habitat found far from the release sites. From a conservation point of view, expanding beavers could potentially be at risk in both Serbia and Bosnia and Herzegovina since the national network of protected areas does not grant sufficient protection to both currently occupied areas and those of potential future colonization.

We confirmed the hypothesis that beaver habitat suitability is mostly influenced by the presence of rivers, broadleaved forests and farmlands, as shown by previous studies (Fustec et al., 2001; Maringer & Slotta-Bachmayr, 2006; Vorel et al., 2008; John et al., 2010). For instance, South et al., (2000) developed a spatially explicit model to explore the possibility of reintroducing beavers to Scotland using an individual-based population dynamics module integrated with GIS data on the spatial distribution of habitat. As in our case too, sites with sufficient deciduous woodlands adjacent to rivers were classified as suitable for the beaver. Besides, field observations carried out in the Netherlands showed that territories occupied by a single beaver family include at least ca. 2 km of wooded banks (Nolet & Rosell, 1994). Although some studies concluded that vegetation type contributes little to beaver habitat models and highlighted the importance of geomorphologic river variables like water depth, slope of river banks and interbank distance (Beier & Barrett, 1987; Hartman, 1996; Suzuki & McComb, 1998), John et al. (2010) found that beavers actively sought out areas dominated by willows (*Salix* spp.) during their expansion phase along the Morava River basin (Czech Republic). This result is further confirmed by previous observations (Zurowski & Kasperczyk, 1990) showing that a recently reintroduced beaver population may tolerate environments with extreme water fluctuations to settle close to willow patches. We therefore remark that broadleaved forest habitat should be regarded as an important predictive variable for beaver settlement in both optimal and suboptimal/marginal habitat during expansion phases before a population reaches its carrying capacity.

The importance of farmland is associated with the species' ecological flexibility (Nolet & Rosell, 1998): crops constitute important food sources when close (within c. 20m) to river banks where beavers have settled (Schwab & Schmidbauer, 2003), especially where natural food is scarce (Dewas et al., 2012).

The mean estimated dispersal distance of ca. 70 km travelled by beavers in a year is in agreement with the results of previous studies. For instance, Fustec et al. (2001) estimated a maximum distance travelled by colonizing beavers for the Loire Valley between 0-80 km/year. On the other hand, the annual colonization speed estimated by John et al. (2010) in the Morava River basin from 1995 to 2007 was lower (0-33 km). Our greater values refer to the first ten years following reintroduction, when range expansion typically proceeds at a faster pace (Halley & Rosell, 2002; Hartman, 1995). The long dispersal distances we obtained may also be due to the fact that shortly after reintroduction beavers follow a spatially discontinuous dispersal pattern often colonising optimal sites farther away from closer, less suitable sites (Nolet & Rosell, 1994; John et al., 2009, 2010). This kind of expansion pattern appears particularly relevant for management, as beavers released in suboptimal habitats are likely to cover long distances to settle farther away where better environmental conditions occur (Halley & Rosell, 2002).

The need for appropriate management to favour beavers' expansion in Serbia and Bosnia and Herzegovina is also remarked by further consideration. Beavers are ecosystem engineers in aquatic habitats because they modify riverine and wetland habitats favouring many other species by felling trees (Fustec et al., 2001; John et al., 2010) and building dams that create still water conditions (Rosell et al., 2005; Stringer & Gaywood, 2016; Law et al., 2016). Their impact on forests and cultivations, however, as well as their tendency to settle near human-dominated areas may also generate conflicts with humans (Schwab & Schmidbauer, 2003; Dewas et al., 2012) and increase beaver mortality. Both in Serbia and Croatia beavers are often killed by collision with motor vehicles and entanglement in fishing nets. Most kills occur in spring, when sub-adults are more active in food search and explore new sites, and autumn, when increased agricultural practices expose beavers to be killed by vehicles or farmers (Grubešić et al., 2015). Conflicts might be mitigated by improving the current network of protected areas, which unfortunately is largely insufficient in agreement with our second hypothesis. Appropriate management, such as restoring riparian habitats where needed to reduce crop damage and erecting physical barriers to protect crops would also help reduce conflicts and human-induced mortality (Dewas et al., 2012).

Using occurrence records of a beaver population during colonization (not in equilibrium with the environment), might violate one of the main assumptions of SDMs (Guisan & Zimmermann, 2000), but we overcame this shortcoming by using a hierarchical approach in which weights of pseudo-absences are adjusted according to a large-scale climatic model based also on records of populations that have reached the equilibrium. Moreover, although ca. 30-50 years are needed for complete colonization to occur (Hartman, 1995; Halley & Rosell, 2002), the first areas to be colonized are the most suitable ones, regardless of their distance from the release site (Nolet & Rosell, 1994; John et al., 2009; John et al., 2010). Overall, we are confident that our SDMs were reliable.

We inevitably considered the environment as isotropic with respect to species expansion (i.e. the species has the same probability to expand in all the directions). Recent observations suggest that the area predicted to have been occupied by 2014 has in fact only partly been colonized (D. Ćirović, *pers. obs.*), most likely due to the obviously anisotropic nature of the landscape: for example, gaps between rivers may have slowed down colonization. More accurate forecasts would require the application of a specific connectivity model. Nevertheless, quantifying dispersal distances increased the management value of the SDM, helping to identify which areas might have been already occupied and thus target them with urgent monitoring and protection. This might be especially important to generate an action plan setting priorities for the establishment of conservation actions.

Based on our findings, we urge that the current reserve network is expanded further to assist the colonisation process, reduce mortality and mitigate potential conflicts with people. We therefore remark that accession of the two states to the EU would represent a unique chance to assist the further establishment of beavers through the designation of Natura 2000 sites.

Table 1. List of ecogeographical variables used for the Regional SDMs, their type, index (name used to indicate the variable in the analysis), spatial resolution, and CLC (Corine Land Cover) code. All variables are expressed in m.

Type	Ecogeographical variable	Index	Spatial resolution (km)	CLC raster code
Topographical	Altitude	Elevation	1	-
Habitat	Distance from hydrographic network	dist_Rivers	1	-
Habitat	Distance from main water courses	dist_Water_courses	1	40
Habitat	Distance from farmlands	dist_Land_principally_agriculture	1	21
Habitat	Distance from Broad-leaved forest	dist_Broad.leaved_forests	1	23
Habitat	Distance from inland marshes	dist_Inland_marshes	1	35

Table 2. Degree of protection granted to *Castor fiber* by the network of protected areas of Serbia and Bosnia-Herzegovina based on the current occurrence records and on the binarized map obtained from SDMs. Reserve = Natural, Managed Nature, Strict Nature and Special reserves.

Occurrence records	% records within site			
Geographic area	National parks	Ramsar Networks	Reserves	All
Serbia	6.68	0.00	32.38	39.06
Bosnia and Herzegovina	0.00	0.00	0.00	0.00
Species Distribution Models	% suitable surface area within site			
Serbia	5.88	0.00	3.84	9.72
Bosnia and Herzegovina	0.91	1.89	0.09	2.89

3.6. References

Adhikari, D., Barik, S.K., & Upadhaya K. (2012). Habitat distribution modelling for reintroduction of *Ilex khasiana Purk*, a critically endangered tree species of northeastern India. *Ecological Engineering*, 40, 37-43.

Algar, A.C., Kharouba, H.M., Young, E.R., & Kerr, J.T. (2009). Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography*, 32, 22-33.

Allen, A.W. (1983). Habitat suitability index models: beaver. Western Energy and Land Use Team, Division of Biological Service, Research and Development, Fish and Wildlife Service, US Department of the Interior.

Allouche, O., Tsoar, A., & Kadmon R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, 43, 1223-1232.

Araújo, M.B., Pearson, R.G., Thuiller, W., & Erhard M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, 11, 1504-1513.

Ardestani, E.G., Tarkesh, M., Bassiri, M., & Vahabi, M.R. (2015). Potential habitat modeling for reintroduction of three native plant species in central Iran. *Journal of Arid Land*, 7, 381-390.

Armstrong, D.P., & Seddon, P.J. (2008). Directions in reintroduction biology. *Trends in Ecology and Evolution*, 23, 20-25.

Armstrong, D.P., & Reynolds, M.H. (2012). Modelling reintroduced populations: the state of the art and future directions. *Reintroduction Biology: integrating science and management*, 12, 165.

Barbet-Massin, M., Thuiller, W. & Jiguet, F. (2010). How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography*, **33**, 878–886.

Barbosa, A. M., Real, R., & VARGAS, J. (2010). Use of Coarse-Resolution Models of Species' Distributions to Guide Local Conservation Inferences. *Conservation Biology*, *24*(5), 1378-1387.

Barve N., Barve V., Jiménez-Valverde A., Lira-Noriega A., Maher S.P., Peterson a. T., Soberón J., & Villalobos F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222, 1810–1819.

Beier, P., & Barrett, R.H. (1987). Beaver habitat use and impact in Truckee River basin, California. *The Journal of wildlife management*, 794-799.

Bivand, R., & Rundel, C. (2013). rgeos: interface to geometry engine: open source (GEOS). R package ver 0.3-2.

Bombi, P., & D'Amen, M. (2012). Scaling down distribution maps from atlas data: a test of different approaches with virtual species. *Journal of biogeography*, *39*(4), 640-651.

Bosso, L., Rebelo, H., Garonna, A.P., & Russo, D. (2013). Modelling geographic distribution and detecting conservation gaps in Italy for the threatened beetle *Rosalia alpina*. *Journal for Nature Conservation*, 21, 72-80.

Bosso, L., Di Febbraro, M., Cristinzio, G., Zoina, A., & Russo, D. (2016)a. Shedding light on the effects of climate change on the potential distribution of *Xylella fastidiosa* in the Mediterranean basin. *Biological Invasions*, 18, 1759-1768.

Bosso, L., Russo, D., Di Febbraro, M., Cristinzio, G., & Zoina A. (2016)b. Potential distribution of *Xylella fastidiosa* in Italy: a maximum entropy model. *Phytopathologia Mediterranea*, 55, 62−72.

Bosso, L., Mucedda, M., Fichera, G., Kiefer, A., & Russo, D. (2016)c. A gap analysis for threatened bat populations on Sardinia. Hystrix, the Italian Journal of Mammalogy, doi: 10.4404/hystrix-27.2-117880.

Breitenmoser, U., Breitenmoser-Wursten, C., Carbyn, L.N., & Funk S.M. (2001). Assessment of carnivore reintroductions. *Conservation Biology Series-Cambridge*, 241-281.

Campbell-Palmer R., Gow D., Campbell R., Dickinson H., Girling S., Gurnell J., Halley D., Jones S., Lisle S. (2016). The Eurasian Beaver Handbook: Ecology and Management of Castor fiber. Pp. 358, Pelagic Publishing, UK.

Ćirović, D., Savić, I., & Bjedov, V. (2001). First finding of the beaver (*Castor fiber* L., 1758) on the territory of FR Yugoslavia. The European beaver in a New Millennium, In: Proceedings of the Second European Beaver Symposium, 27-30 September 2000, Białowieza, Poland, 58.

Ćirović, D., Stamenković, S., & Bjedov, V. (2003). Reintroduction of the European beaver (*Castor fiber* L., 1758) into Serbia – Feasibility Study (In Serbian). 48 pp. Faculty of Biology, University of Belgrade.

Ćirović, D., Bjedov, V., & Stamenković S. (2007). Reintroduction of the European beaver (*Castor fiber* L., 1758) into Zasavica – return of an extinct species (In Serbian). In: Zasavica 2007, 107-114. Sremska Mitrovica.

Dewas, M., Herr, J., Schley, L., Angst, C., Manet, B., Landry, P., & Catusse, M. (2012). Recovery and status of native and introduced beavers *Castor fiber* and *Castor canadensis* in France and neighbouring countries. *Mammal Review*, 42, 144-165.

Di Febbraro, M., Roscioni, F., Frate, L., Carranza, M.L., De Lisio, L., De Rosa, D., & Loy, A., (2015). Long-term effects of traditional and conservation-oriented forest management on the distribution of vertebrates in Mediterranean forests: a hierarchical hybrid modelling approach. *Diversity and Distributions*, 21, 1141-1154.

Djoshkin, W.W., & Safonow W.G. (1972). Die Biber der alten und neuen Welt. A. Ziemsen Verlag.

Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.P., & Guisan A. (2011). Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, 17, 1122-1131.

Ducci, L., Agnelli, P., Di Febbraro, M., Frate, L., Russo, D., & Loy, A., Roscioni F., (2015). Different bat guilds perceive their habitat in different ways: a multiscale landscape approach for variable selection in species distribution modelling. *Landscape Ecology*, 30, 2147-2159.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Jin L., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Mcc Overton, J., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., & Zimmermann, N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129 –151.

Elith, J., & Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677.

Fernandes, R. F., Vicente, J. R., Georges, D., Alves, P., Thuiller, W., & Honrado, J. P. (2014). A novel downscaling approach to predict plant invasions and improve local conservation actions. *Biological Invasions*, *16*(12), 2577-2590.

Feuda, R., Bannikova, A. A., Zemlemerova, E. D., Di Febbraro, M., Loy, A., Hutterer, R., ... & Colangelo, P. (2015). Tracing the evolutionary history of the mole, Talpa europaea, through mitochondrial DNA phylogeography and species distribution modelling. *Biological Journal of the Linnean Society*, *114*(3), 495-512.

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

Fustec, J., Lodé, T., Le Jacques, D., Cormier, J.P. (2001). Colonization, riparian habitat selection and home range size in a reintroduced population of European beavers in the Loire. *Freshwater Biology*, 46, 1361-1371.

Gallien L., Douzet R., Pratte S., Zimmermann N.E., & Thuiller W. (2012). Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, **21**, 1126–1136.

Gibson, L.A., Wilson B.A., Cahill D.M., Hill J., 2004. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. Journal of applied ecology 41:213-223.

Grubešić, M., Margaletić, J., Ćirović, D., Vucelja, M., Bjedov, L., Burazerović, J., & Tomljanović, K. (2015). Analysis of beaver (*Castor fiber* L.) mortalityin Croatia and Serbia (In Croatian with English summary). *Šumarski list*, 139, 137-143.

Guisan, A., & Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, 135, 147-186.

Guisan, A., & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, 8, 993-1009.

Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. (2014). Unifying niche shift studies: insights from biological invasions. *Trends in Ecology & Evolution*, 29, 1–10.

Halley, D.J., & Rosell, F. (2002). The beaver's reconquest of Eurasia: status, population development and management of a conservation success. *Mammal review*, 32, 153-178.

Halley, D., Rosell, F., & Saveljev, A. (2012). Population and distribution of Eurasian beaver (*Castor fiber*). *Baltic Forestry*, 18, 168-175.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29-36.

Harrell, F.E., Lee, K.L. & Mark, D.B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15, 361–387.

Hartman, G. (1995). Patterns of spread of a reintroduced beaver *Castor fiber* population in Sweden. *Wildlife Biology*, 1, 97-103.

Hartman, G. (1996). Habitat selection by European beaver (*Castor fiber*) colonizing a boreal landscape. *Journal of Zoology*, 240, 317-325.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., & Jarvis A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.

International Union for Conservation of Nature/Species Survival Commission [IUCN/SSC], 2012. Guidelines for reintroductions and other conservation translocations. Version 1. IUCN Species Survival Commission, Gland, Switzerland.

Liu, C., Berry, P.M., Dawson, T.P., & Pearson, R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385-393.

Jarvis, A., Reuter, H.I., Nelson, A., & Guevara E. (2008). Hole-filled SRTM for the globe Version 4. available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org).

Jiguet, F., Barbet-Massin, M., & Henry, P.Y. (2010). Predicting potential distributions of two rare allopatric sister species, the globally threatened *Doliornis cotingas* in the Andes. *Journal of Field Ornithology*, 81, 325-339.

John, F., & Kostkan V. (2009). Compositional analysis and GPS/GIS for study of habitat selection by the European beaver, *Castor fiber* in the middle reaches of the Morava River. *Folia Zoologica*, 58:76.

John, F., Baker, S., & Kostkan, V. (2010). Habitat selection of an expanding beaver (*Castor fiber*) population in central and upper Morava River basin. *European journal of wildlife research*, 56, 663-671.

Kollar, H.P., & Seiter, M. (1990). Biber in den Donau-Auen östlich von Wien: eine erfolgreiche Wiederansiedlung. Verein für Ökologie u. Umweltforschung.

Law, A., McLean, F., & Willby, N.J. (2016). Habitat engineering by beaver benefits aquatic biodiversity and ecosystem processes in agricultural streams. *Freshwater Biology*, 61, 486-499.

Lomba A., Pellissier L., Randin C., Vicente J., Moreira F., Honrado J., & Guisan A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, 143, 2647–2657.

Macdonald, D.W., & Barrett, P. (1993). Mammals of Europe. Princeton University Press.

Macdonald, D.W., Tattersall, F.H., Brown, E.D., & Balharry, D. (1995). Reintroducing the European beaver to Britain: nostalgic meddling or restoring biodiversity?. *Mammal Review*, 25, 161-200.

Maringer, A., & Slotta-Bachmayr, L. (2006). A GIS-based habitat-suitability model as a tool for the management of beavers *Castor fiber*. *Acta theriologica*, 51, 373-382.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and distributions*, 15, 59-69.

McCarthy, M.A., Armstrong, D.P., & Runge, M.C. (2012). Adaptive management of reintroduction. *Reintroduction biology: integrating science and management*, 12, 256.

McPhersonJM, Jetz W, Rogers DJ (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. *Ecological Modelling,* 192, 499–522

Nolet, B.A., & Rosell F. (1994). Territoriality and time budgets in beavers during sequential settlement. *Canadian Journal of Zoology*, 72, 1227-1237.

Nolet, B.A., & Rosell, F. (1998). Comeback of the beaver *Castor fiber*: an overview of old and new conservation problems. *Biological Conservation*, 83, 165-173.

Olson D.M., Dinerstein E., Wikramanayake E.D., Burgess N.D., Powell G.V.N., Underwood E.C., D'amico J.A., Itoua I., Strand H.E., Morrison J.C., Loucks C.J., Allnutt T.F., Ricketts T.H., Kura Y., Lamoreux J.F., Wettengel W.W., Hedao P., & Kassem K.R. (2001) Terrestrial Ecoregions of the World: a new map of life on Earth. *BioScience*, 51, 933.

Olsson, O., & Rogers, D.J. (2009). Predicting the distribution of a suitable habitat for the white stork in Southern Sweden: identifying priority areas for reintroduction and habitat restoration. *Animal Conservation*, 12, 62-70.

Parker, H., Haugen, A., Kristensen, Ø., Myrum, E., Kolsing, R., & Rosell, F. (2001). Landscape use and economic value of beaver Castor fiber on a large forest in southern Norway. In Proceedings of the First European-American Beaver Congress, Kazan, Russia, 24–28 August 1999. Edited by P. Busher and Y. Gorshkov. Volga-Kama National Nature Preserve, Kazan, Russia. pp. 77–95.

Pearson R.G., Dawson T., & Liu C. (2004). Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, **3**, 285–298.

Pebesma E. J., & Bivand R.S. (2005). Classes and methods for spatial data in R. R News 5 (2).

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

Pio, D.V., Engler, R., Linder, H.P., Monadjem, A., Cotterill F.P., Taylor, P.J., & Salamin, N. (2014). Climate change effects on animal and plant phylogenetic diversity in southern Africa. *Global Change Biology*, 20, 1538-1549.

R Development Core Team (2015) *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Radosavljevic, A., & Anderson, R.P. (2014). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of biogeography*, 41, 629-643.

Raes, N. (2012). Partial versus full species distribution models. *Natureza & Conservaçao*, 10, 127–138.

Rosell, F., Bozser, O., Collen P., & Parker H. (2005). Ecological impact of beavers *Castor fiber* and *Castor canadensis* and their ability to modify ecosystems. *Mammal review*, 35, 248-276.

Russo, D., Di Febbraro, M., Rebelo, H., Mucedda, M., Cistrone, L., Agnelli, P., De Pasquale, P.P., Martinoli, A., Scaravelli, D., Spilinga, C., & Bosso, L. (2014). What story does geographic separation of insular bats tell? A case study on Sardinian Rhinolophids. PLoS ONE 9:e110894. doi: 10.1371/journal.pone.0110894. pmid:25340737.

Russo, D., Di Febbraro M., Cistrone L., Jones G., Smeraldo S., Garonna A.P., & Bosso L. (2015). Protecting one, protecting both? Scale-dependent ecological differences in two species using dead trees, the rosalia longicorn beetle and the barbastelle bat. *Journal of Zoology*, 297, 165-175.

Schwab, V.G., & Schmidbauer, M. (2003). Beaver (*Castor fiber* L., Castoridae) management in Bavaria. *Denisia*, 9, 99-106.

Seddon, P.J., Armstrong, D.P., & Maloney, R.F. (2007). Developing the science of reintroduction biology. *Conservation biology*, 21, 303-312.

South, A., Rushton, S., & Macdonald, D. (2000). Simulating the proposed reintroduction of the European beaver (*Castor fiber*) to Scotland. *Biological Conservation*, 93, 103-116.

Stringer, A.P., & Gaywood, M.J. (2016). The impacts of beavers Castor spp. on biodiversity and the ecological basis for their reintroduction to Scotland, UK. *Mammal Review*, 46, 270-283.

Strubbe D., Beauchard O., & Matthysen E. (2015) Niche conservatism among non-native vertebrates in Europe and North America. *Ecography*, 38, 321–329.

Suzuki, N., & McComb, W.C. (1998). Habitat classification models for beaver (*Castor canadensis*) in the streams of the central Oregon Coast Range. *Nofihwest Science*, 72, 102 -110.

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

Vorel, A., Válková, L., Hamšíková, L., Malon, J., & Korbelová, J. (2008). The Eurasian beaver population monitoring status in the Czech Republic. *Natura Croatica*, 17, 217-232.

Warren, D. L., & Seifert, S.N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335-342.

Wilson, M.F., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30, 3-35.

Wilson, C. D., Roberts, D., & Reid, N. (2011). Applying species distribution modelling to identify areas of high conservation value for endangered species: A case study using *Margaritifera margaritifera* (L.). *Biological Conservation*, 144, 821-829.

Zurowski, W., & Kasperczyk, B. (1990). Results of beaver reintroduction in some Carpathian mountain streams. *OchrPrzyr*, 47, 201-213.

Zuur, A.F., Ieno, E.N., & Elphick, C.S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1, 3.

Myotis emarginatus colony (Luca Cistrone)

Hibernating bat (Slaveya Stoycheva)

Chapter IV

4. Species' potential distribution may be affected by phenology: bats as a case study

4.1. Abstract

Phenology is a key feature to consider in the description of species niches to fully capture seasonality in resource use and climate requirements. This is all the more true for temperate bats, whose ecological needs differ markedly across seasons, especially between hibernation and reproduction. Species distribution models (SDMs) are widespread tools to evaluate a species' potential distribution and identify its large-scale habitat preferences. Modelling studies have experienced a recent boost thanks to the increasing amount of data available to implement SDMs: however, despite the chief importance of data phenology to describe a species' niche, the time of year data were collected is often neglected or not controlled for in the process. In this study we tested the hypothesis that the output of SDMs developed for six European bat species will differ according to whether hibernation or reproductive occurrence data are used. We employed a dataset made of 470 independent occurrence records of bat hibernacula and 400 independent records of nursery roosts of selected species and for each species we developed separate winter, summer and mixed (i.e. generated from both winter and summer occurrences) models. Due to the large geographic scale we used (the entire European continent), to generate models we only considered climatic variables. Seasonal and mixed potential ranges differed from each other and the direction of this difference was species specific. In general, we found that mixed model should be preferred over seasonal models for the description of potential ranges. Seasonally-biased models might in fact underrepresent important areas and this might impair their application to management or conservation planning. Overall, our work highlights the importance of considering data seasonality in the development of SDMs in order to encompass comprehensively the different ecological requirements arising from species-specific phenology.

4.2. Introduction

In the last decades, species distribution models (SDMs) have attracted increasing attention (Guisan et al., 2013) as a main tool to make explicit predictions on the potential distribution of organisms (Franklin, 2010; Peterson, 2011). This approach has been successfully applied to support many strategies for species conservation. For instance, SDMs may detect regions of high habitat suitability for species or communities, both in the present and in the future (Beaumont et al., 2007; Rebelo et al., 2010; Rubio-Salcedo et al., 2017; Zhang et al., 2017), estimate the conservation performance of reserve networks (Carvalho et al., 2010; Doko et al., 2011; Tulloch et al., 2015; Bosso et al., 2017c), forecast spatial patterns of biological invasions (Thuiller et al., 2005; Di Febbraro et al., 2016; Bosso et al., 2017a), or inform species reintroduction (Olsson & Rogers, 2009; Ardestani et al., 2015; Smeraldo et al., 2017).

SDMs are calibrated on the observed species distribution (Barbet-Massin et al., 2010) and may rest upon limited datasets (Hernandez et al., 2006). SDMs may, in general, be generated from small datasets, also when these are made of records collected opportunistically, so they perform well with species that are rare, endangered or elusive. SDMs basically correlate species occurrence records with a set of environmental predictors ecologically relevant for the species (Thuiller, 2003; Thuiller et al., 2005; Thuiller et al., 2009). It is therefore evident that an accurate ecological knowledge of target species must be integrated into the modelling process to obtain more realistic outputs (Austin, 2002; Huston, 2002; Wiens, 2002; Guisan & Thuiller, 2005). Many studies have remarked the importance of collecting data, selecting predictors or study areas in a way that best highlights the relationship between a species' occurrence and the environmental variables needed to represent its ecological niche (Guisan & Thuiller, 2005; Van Horn, 2002; Thuiller et al., 2004).

Within this framework, the type and quality of occurrences play an important role (Guisan et al., 2013). Although the distribution of locations in the geographical and/or ecological space may affect model reliability (Reddy & Davalos, 2003; Syfert et al.,2013), the potential effects of sampling biases in the dataset are rarely taken into account (Fourcade et al., 2014). This is a potentially important aspect, since models built using presence records sampled during a short period of time or/and over a limited space may underrepresent a species' potential niche (Guisan & Thuiller, 2005). A common source of variation in occurrence datasets and their ecological correlates is given by phenology, often neglected in ecological niche modelling.

Phenology integrates genetic and physiologic traits of a certain species with climate and environmental variables (Miller-Rushing & Weltzin, 2009), so it logically represents one of the most important traits, if not the most important trait, shaping species distribution (Chuine, 2010). A species' phenological stage is typically associated with seasonally variable ecological requirements whose fulfilment may require selecting different habitats or areas, or even performing long-range movements (migration) to move to more favourable regions (Newton, 2008; Holte et al., 2017). Even many sedentary species may still move periodically on a small-scale to sites matching seasonally variable ecological requirements. Hibernating species often move between summer and winter habitats, often differing considerably in climate or spatial resources (Hansson & Åkesson, 2014). Hibernating mammals cope with food scarcity in winter by decreasing body temperature and metabolic rate to save energy (Lyman, 2013), and to do so they often seek sites that are well protected from predators and provide a stable, cool microclimate. This may imply moving seasonally between summer and winter quarters, sometimes between altitudes or latitudes (Nagel & Nagel, 1991; Webb et al., 1996; Rivers et al., 2006; Papadatou et al., 2008). SDMs built on datasets that do not encompass such seasonal environmental variation may offer an incomplete or biased niche representation. Following this rationale, we set out to test the effects of data phenology on the calibration of SDMs.

Temperate insectivorous bats represent ideal model organisms to analyse the influence of phenology on species distribution because they show markedly different ecological requirements between summer, when they reproduce, and winter, when they hibernate to overcome paucity of prey arthropods caused by low ambient temperatures (Grindal et al., 1992; Lewis, 1993; Hoying & Kunz, 1998; Frick et al., 2010; Bellamy & Altringham, 2015). Summer and winter roosts differ from one another in terms of microclimate, so their often uneven spatial distribution may affect seasonal bat presence. Roost suitability responds to various spatial scales because it is typically influenced by e the characteristics of the roost itself as well as by the bioclimatic conditions of the surrounding environment (Bellamy & Altringham, 2015).

Understanding how climate influences bats in different phenological phases may therefore provide insights into the limitations it poses to bat distributions (Erickson & West, 2002). We developed separate SDMs for six European bat species employing hibernation (winter model), reproduction (summer model) and total occurrences (summer + winter occurrences, hereafter termed “mixed models”). We assumed datasets featuring both phenological categories of presence records to provide a more comprehensive species’ niche representation, and tested two hypotheses: 1) SDMs built using respectively hibernation and reproductive occurrences will differ between each other, and 2) mixed models providing equal weight to the two phenological statuses will offer a more realistic representation of a species’ geographical range.

We tested these hypotheses through an analytical framework that investigated divergences in geographical distribution and extent between seasonal and mixed SDM outputs. We also analysed potential phenological differences in the environmental space through a niche overlap analyses.

4.3. Methods

4.3.1. Study area and presence records

The study area covered mainland Europe (west of Caucasus), United Kingdom and Ireland and all major Mediterranean islands spanning from latitudes 77° N - 34° N and longitudes 24° 31'W - 69° 5'E (corresponding to an overall surface area of ca. 4,030,720 km^2 and elevation range between 0 - 4954 m a.s.l.; Fig. 1). Therefore, the study area included the European geographical range of all bat species (Mitchell-Jones et al., 1999) we considered for our study. Overall, the climate of Europe can be characterized as temperate although considerable variations exist, with a Mediterranean climate dominating in the south (with hot summers and mild winters) and a considerably colder and more humid climate in the north and in mountainous regions (Huntley et al., 2007).

Presence records for European bat species were provided by several countries within the Eurobats agreement (http://www.eurobats.org/). Additional data were retrieved for Spain, Germany and Italy. Each record included coordinates of winter or summer roosts. As the study aimed to develop seasonal SDMs, for each species we split the datasets in two groups according to records' phenological statuses. Subsequently, only bat species for which ≥ 20 presence records for each phenological stage were used for analysis (Wisz et al., 2008).

Based on such criteria we selected six bat species, i.e. *Myotis emarginatus*, *Myotis myotis*, *Miniopterus schreibersii*, *Rhinolophus euryale*, *Rhinolophus ferrumequinum* and *Rhinolophus hipposideros* (Table 1).

Records were screened in ArcGIS (version 10.2.2) for spatial autocorrelation using average nearest neighbour analyses to remove spatially correlated data points and guarantee independence (Dormann et al., 2007; Russo et al., 2014; Kwon et al., 2016; Smeraldo et al., 2017; Bosso et al., 2017b). The final presence dataset varied between a minimum of 132 locations for *R. euryale* and a maximum of 157 for *M. myotis* including a total of 870 presence records covering most of the European distribution known for our study species. Using seasonal datasets of differing size might have potentially biased the resulting models. Therefore, we obtained equally large sample size of winter and summer locations as follows. For each species, the larger phenological dataset (whether the winter or the summer one) was sampled randomly to extract a number of occurrences equaling that of the less numerous dataset, and the sampling procedure was repeated 10 times. Mixed models were also developed using the same number of presence records used for the corresponding seasonal models and considering equal numbers of winter and summer data to avoid all risks of bias. Each of the 10 independent sets of occurrences was used to train a separate SDM.

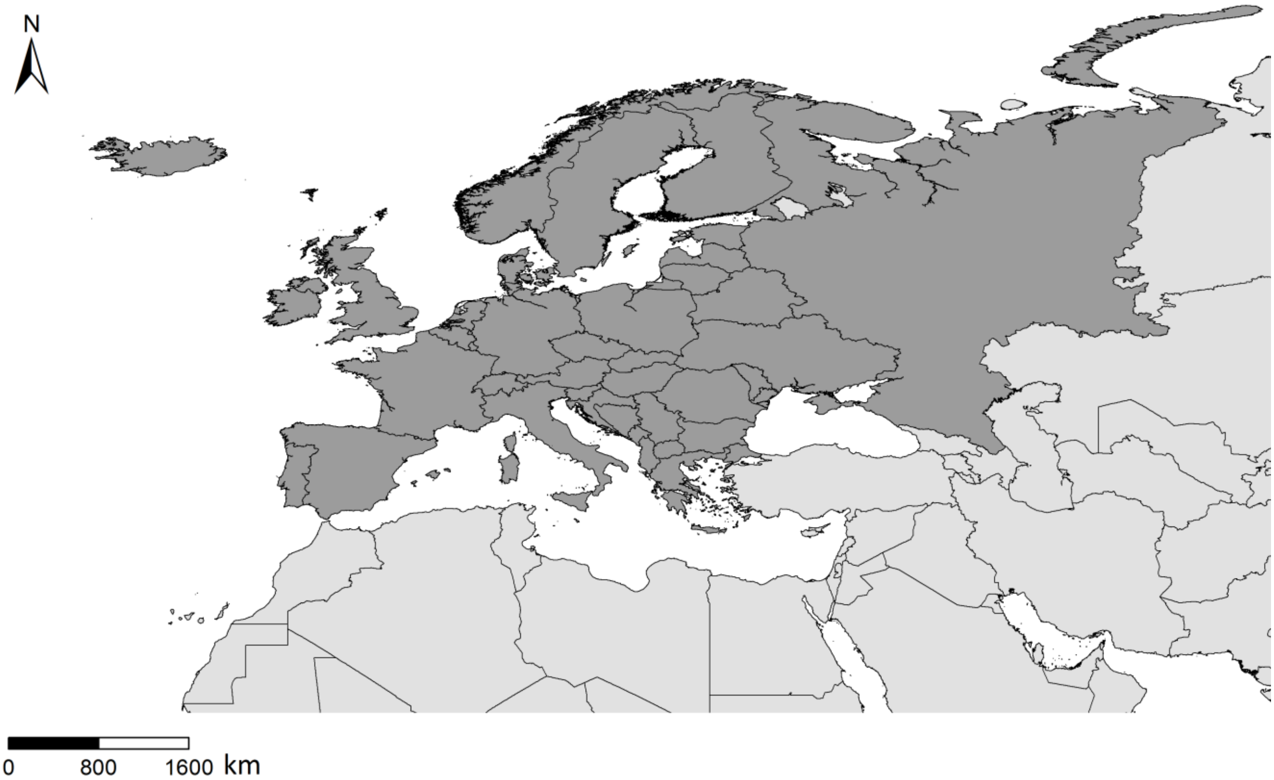


Figure 1. Representation of the study area (dark grey) corresponding to the European territories on which the potential distribution of six bat species (*Myotis emarginatus, M. myotis, Miniopterus schreibersii*, *Rhinolophus euryale*, *R. ferrumequinum*, *R. hipposideros*) was projected for the aims of the present study.

Table 1. Initial dataset of species' presence records collected for both phenological stages used to implement seasonal and mixed SDM models.

Species	Summer records	Winter records
Myotis emarginatus	128	140
Myotis myotis	480	813
Miniopterus schreibersii	229	99
Rhinolophus euryale	152	70
Rhinolophus ferrumequinum	169	343
Rhinolophus hipposideros	134	389

4.3.2. Environmental variables

As initial set of environmental predictors for SDM training, we considered elevation and 19 bioclimatic variables derived from the WORLDCLIM dataset (www.worldclim.org/current). Because our models were developed for the whole Europe, given such a large scale, we refrained from using land cover variables (Huntley et al., 1995; Pearson & Dawson, 2003; Thuiller et al., 2004; Araujo & Guisan, 2006). All variables were provided as rasters at a spatial resolution of 2.5 arc–minutes (≈5 km; Hijmans et al., 2005; Di Febbraro et al., 2016). To take into account the pairwise correlation between the predictors, the final set of variables was subselected considering a Pearson's correlation coefficient $|r|<0.75$ (Hernandez et al., 2006; Lobo et al., 2010; Russo et al., 2015; Bosso et al., 2016). This procedure was carried out in the R environment (R Development Core Team, 2012). From this set of predictors, we further subselected only the most relevant variables according to the species' ecological requirements during the two phenological stages (hibernation and maternity) based on prior knowledge of the ecology of temperate bats (Erickson & West, 2002; Speakman & Thomas, 2003; Austin & Van Niel, 2011), i.e. : annual mean diurnal range (BIO2), isothermality (BIO3), minimum temperature of coldest quarter (BIO6), mean temperature of warmest quarter (BIO10), precipitation seasonality (BIO15), precipitation of warmest quarter (BIO18) and elevation (Table 2).

Table 2. List of bioclimatic variables used to develop seasonal and mixed SDMs for six bat species in Europe (*Myotis emarginatus, M. myotis, Miniopterus schreibersii, Rhinolophus euryale, R. ferrumequinum, R. hipposideros*)

Type	Bioclimatic variable	Unit	Source
Topographical	Altitude	m	WorldClim database
Climatic	Annual mean diurnal range	°C	WorldClim database
	Isothermality	%	WorldClim database
	Minimum temperature of the coldest month	°C	WorldClim database
	Mean temperature of the warmest quarter	°C	WorldClim database
	Precipitation seasonality	%	WorldClim database
	Precipitation of the warmest quarter	mm	WorldClim database

4.3.3. Species distribution modelling

For each bat species we developed three SDMs: two seasonal models from separate winter and summer datasets respectively, and a mixed model generated as described above. SDMs were built through an ensemble forecasting approach, as implemented in the R package "*biomod2*" (https://cran.r-project.org/bin/windows/base/; Thuiller et al., 2009). We considered the following eight modelling techniques (Thuiller et al., 2009; Jiguet et al., 2010; Ducci et al., 2015): 1) generalized linear models (GLM); 2) generalized additive models (GAM); 3) generalized boosted models (GBM); 4) random forests (RF); 5) multivariate adaptive regression spline (MARS), 6) maximum entropy models (MAXENT), 7) Surface range envelope (SRE) and 8) artificial neural network (ANN; for further details, see Thuiller et al., 2009). Following Pio et al. (2014), the modelling settings were tuned as follows. GLMs and GAMs were calibrated using a binomial distribution and a logistic link function. GBMs were calibrated with a maximum number of trees set to 5000, threefold cross-validation procedures to select the optimal numbers of trees to be kept and a value of seven as maximum depth of variable interactions. RF models were fitted by growing 750 trees with half the numbers of available predictors sampled for splitting at each node. MARS models were fitted with a maximum interaction degree equal to 2, while MAXENT models were fitted with the default settings apart with a maximum value of 1000 iterations. Default parameters were used to fit ANN and SRE models. Each occurrence dataset was randomly split into a 70% sample, used for the calibration of the model, and the remaining 30%, used to evaluate model performance. Because our datasets contain only presence data, a set of 10,000 background points were randomly placed over a region identified by all the WWF terrestrial ecoregions (Olson et al., 2001) were species records occurred (Hirzel et al., 2002; Barve et al., 2011; Barbet-Massin et al., 2012; Smeraldo et al., 2017). Predictive performances of SDMs were assessed by measuring the area under the receiver operating characteristic curve (AUC; Hanley & McNeil, 1982) and the true skill statistic (TSS; Allouche, Tsoar, & Kadmon, 2006). These validation methods have been widely used (e.g. Bosso et al., 2016b; Feuda et al., 2015; Breiner et al., 2015; Balestrieri et al., 2016; Smeraldo et al., 2017) and offer excellent performances. The data splitting procedure was repeated 10 times and the evaluation values averaged. For each species' dataset, we ran a total of 80 SDMs (eight algorithms × 10 splitting replicates for model evaluation). After excluding models with AUC<0.8, model averaging was performed by weighting the individual model projections by their AUC scores, a method shown to be particularly robust (Marmion et al., 2009). The relative importance of variables was also calculated from the ensemble model using the specifically devoted functionality available in the biomod2 package (Jiguet et al., 2010). For winter and summer models, the final potential distributions were obtained by averaging the projections from the10 replicated ensemble models coming from the subsampling procedure (see above). Final maps were also transformed into presence–absence values using a threshold maximizing sensitivity (the percentage of correctly predicted presence) and specificity (the percentage of correctly predicted absence; Fielding & Bell, 1997). Such threshold has been widely used (e.g. Algar et al., 2009; Dubuis et al., 2011; Di Febbraro et al., 2015; Smeraldo et al., 2017) and constitutes one of the most accurate (Liu et al., 2005).

4.3.4. Models comparison

We assessed the differences between seasonal and mixed SDMs by comparing the resulting distributions both visually and quantitatively. In the former approach, we compared visually the degree of matching between potential distributions obtained using seasonal vs. mixed datasets. We also measured the spatial agreement in location as the Euclidean distance between geographical centroids of each ranges (Herkt et al., 2017). For each species, we then measured the absolute (in km^2) and percent surface of both the entire study area and the corresponding IUCN species' range covered by seasonal and mixed models. We assumed the IUCN species range (obtained from www.iucnredlist.org) to express the observed species' geographic range for our study species.

For each species we also overlaid the binary maps from the two seasonal models and reclassified the output in two main categories: "areas of agreement", where suitable habitat predictions generated by both models matched, and "areas of disagreement" where the models disagreed with each other. All procedures were carried out with ArcGIS (version 10.2.2).Finally, a generalized linear mixed model (GLMM) was employed to test the differences in surface areas of predicted ranges generated by seasonal vs. mixed models. We used the percent surface area of predicted range as the response variable, the model type (summer, winter and mixed) as fixed term and the species as random term, allowing the model to vary its intercept according to these factors. GLMMs were run with R using the "*lme4"* package (Bates et al., 2008).

4.3.5. Niche overlap between seasonal models

We performed niche overlap analyses using the analytical framework proposed by Broennimann et al. (2012) and applied to several recent mammal studies (e.g. Russo et al., 2014; 2015; Loy et al., 2017). Within this framework, the available environmental space of the six species (i.e. all of the environmental conditions occurring in the study area) was used to train a principal components analysis (PCA). Species records associated with the two phenologies and environmental conditions were projected into this PCA space. Subsequently, the density of the background environments and species records across the first two PCs were calculated by a kernel density smoother (Broennimann et al., 2012). The resulting density grids of r × r cells in the environmental space were used to compute niche overlaps between the two phenologies in terms of Schoener's D (Schoener, 1970), a metric that ranges from 0 (no overlap) to 1 (complete overlap). We performed the niche similarity test *sensu* Warren et al. (2008) to evaluate if the two phenologies are more similar/different in their niche space than expected by chance. The test proceeds by comparing the niche overlap values (Schoener's D) to a null distribution of 100 overlap values, yielding a significant outcome if the observed Schoener's D value between the two phenologies is higher (significant similarity) or lower (significant dissimilarity) than the 95% of the null distribution ($P < 0.05$). In addition, the similarity between niches was tested in both directions, i.e. the amount of phenology 1 niche included in phenology 2 niche, and vice versa, following Broennimann et al. (2012). Schoener's D values for the ten replicated occurrence datasets were then averaged, and the percentage of tests yielding significant outcomes was considered (i.e. 100% of significant similarity tests indicates that a significant outcome arises from each of the ten replicated occurrence datasets). The procedure was carried out in R using the "*ecospat*" package (Di Cola et al., 2017).

4.4. Results

4.4.1. SDMs predictive performance and comparison

Both seasonal and mixed models showed high levels of predictive performance. The greatest seasonal model performances were achieved for *R. euryale* (summer, AUC = 0.982 ± 0.004, TSS = 0.891 ± 0.024; winter, (AUC = 0.978 ± 0.003, TSS = 0.901 ± 0.011). Mixed models for both *R. hipposideros* and *R. euryale* showed the highest AUC values (respectively 0.984 ± 0.007 and 0.984 ± 0.005) and TSS (0.917 ± 0.027 and 0.905 ± 0.023) (Table 3). For all species, the same bioclimatic variables drove both seasonal and mixed models, i.e. minimum temperature of the coldest quarter, isothermality, elevation, mean temperature of the warmest quarter and annual mean diurnal range (Table 4). However, the contribution of a certain variable differed among seasonal models: for instance, the minimum temperature of the coldest month (importance = 0.487±0.122) and the mean temperature of the warmest quarter (0.11±0.073) influenced especially summer distribution, whereas isothermality (0.258±0.236) and annual mean diurnal range (0.074±0.052) contributed more to winter than to summer models (Table 4 and Fig. 2). For all species, occurrence in both seasons was more likely for minimum temperatures of the coldest month > -10 °C. However, in winter bats were more likely to occur in areas where this variable was comprised between -10°C – 0°C, whereas in summer occurrence was more likely at higher values, as well as where mean temperatures of the warmest quarter were >15°C (Fig. 3).

Table 3. Mean AUC and TSS scores of the SDMs for the six analyzed bat species. Standard deviations are reported in parentheses.

Species	Mixed model		Summer model		Winter model	
	AUC	TSS	AUC	TSS	AUC	TSS
Myotis emarginatus	0.976 (±0.005)	0.889 (±0.02)	0.975 (±0.058)	0.880 (±0.068)	0.973 (±0.002)	0.871 (±0.013)
Myotis myotis	0.983 (±0.007)	0.893 (±0.026)	0.970 (±0.023)	0.862 (±0.012)	0.976 (±0.005)	0.888 (±0.023)
Miniopterus schreibersii	0.976 (±0.003)	0.886 (±0.015)	0.975 (±0.003)	0.876 (±0.013)	0.975 (±0.004)	0.895 (±0.012)
Rhinolophus euryale	0.984 (±0.005)	0.905 (±0.023)	0.982 (±0.004)	0.891 (±0.024)	0.978 (±0.003)	0.901 (±0.011)
Rhinolophus ferrumequinum	0.971 (±0.008)	0.890 (±0.023)	0.961 (±0.015)	0.860 (±0.022)	0.957 (±0.007)	0.836 (±0.033)
Rhinolophus hipposideros	0.984 (±0.007)	0.917 (±0.027)	0.974 (±0.013)	0.890 (±0.040)	0.969 (±0.011)	0.859 (±0.038)

Table 4. Average variable importance for seasonal, mixed and total SDMs developed for six bat species in Europe (*Myotis emarginatus, M. myotis, Miniopterus schreibersii*, *Rhinolophus euryale*, *R. ferrumequinum*, *R. hipposideros*)

Bioclimatic Variable	Mixed SDMs	Summer SDMs	Winter SDMs
Elevation	0.085 (±0.055)	0.117 (±0.059)	0.111 (±0.069)
Annual mean diural range	0.047 (±0.043)	0.029 (±0.016)	0.074 (±0.052)
Isothermality	0.170 (±0.088)	0.156 (±0.082)	0.258 (±0.236)
Minimum temperature of the coldest month	0.470 (±0.068)	0.487 (±0.122)	0.405 (±0.286)
Mean temperature of the warmest quarter	0.053 (±0.044)	0.110 (±0.073)	0.079 (±0.113)
Precipitation seasonality	0.009 (±0.007)	0.027 (±0.011)	0.016 (±0.010)
Precipitation of the warmest quarter	0.029 (±0.020)	0.025 (±0.008)	0.041 (±0.027)

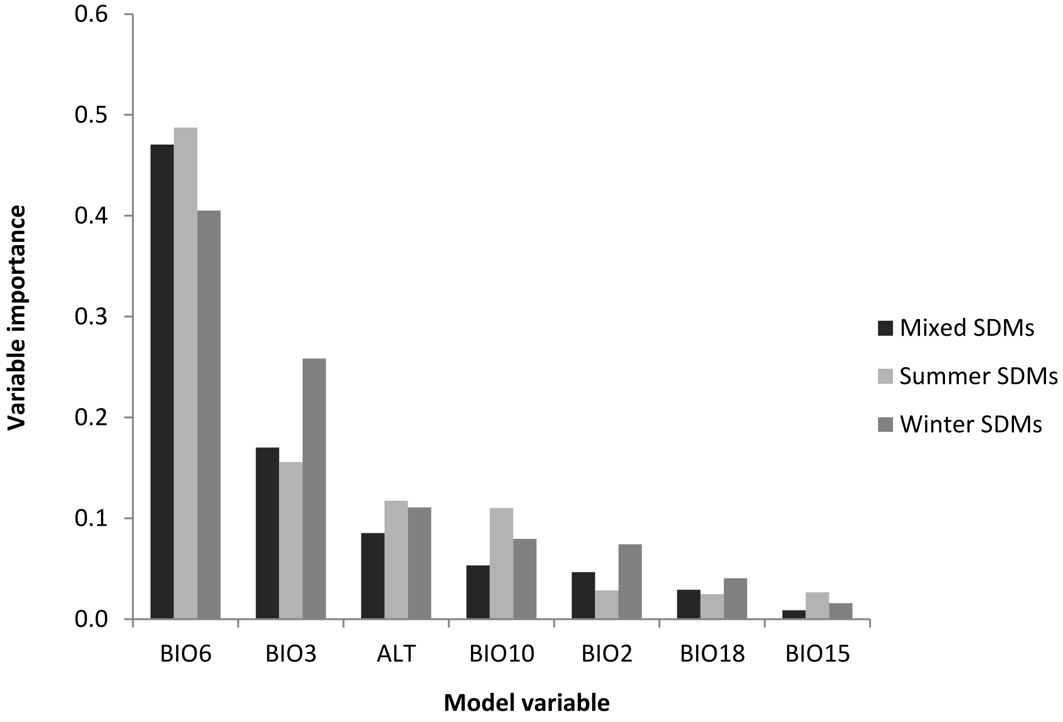


Figure 2. The histogram expresses the contribution of each variable to summer, winter and mixed models' predictive performance (variables are listed in increasing order of importance from top to bottom). Each bar represents one variable (BIO15 = precipitation seasonality; BIO18 = precipitation of the warmest quarter; BIO2 = annual mean diurnal range; BIO10 = mean temperature of the warmest quarter; ALT = elevation; BIO3 = isothermality; BIO6 = minimum temperature of the coldest month).

Summer and winter potential distribution differed considerably from each other, the greatest overlap between seasonal models being shown by *M. schreibersii* and *R. euryale* (61.30 % and 54.11 % respectively; Table 5). Summer suitable areas were localized mainly in south-western Europe while winter potential ranges extended towards higher latitudes and eastern European regions (Figs 4-5). Seasonal models also predicted smaller portions of species ranges than did mixed models (Table 6). For all species, potential distribution as predicted by mixed SDMs offered a more comprehensive representation of the range as they correctly predicted ca. 90% of both summer and winter suitable habitats (Table 5). Moreover, for all species mixed SDMs predicted a larger suitable surface of the study area than did seasonal models (Figs 4-5; Table 6) and ca. 57 % of the IUCN range for four out of six bat species, whereas seasonal SDMs in most cases underrepresented it (Table 7). The extent of predicted species' range differed significantly among the three (winter, summer and mixed) SDMs: both seasonal SDMs predicted smaller ranges than did mixed SDMs (Table 8 and Table 9). Finally, for almost all the species (except *M. schreibersii* and *R. euryale*), the range centroids of mixed models were closest to those of IUCN ranges centroids, the smallest distance being 80.9 km for *M. emarginatus* (Table 10).

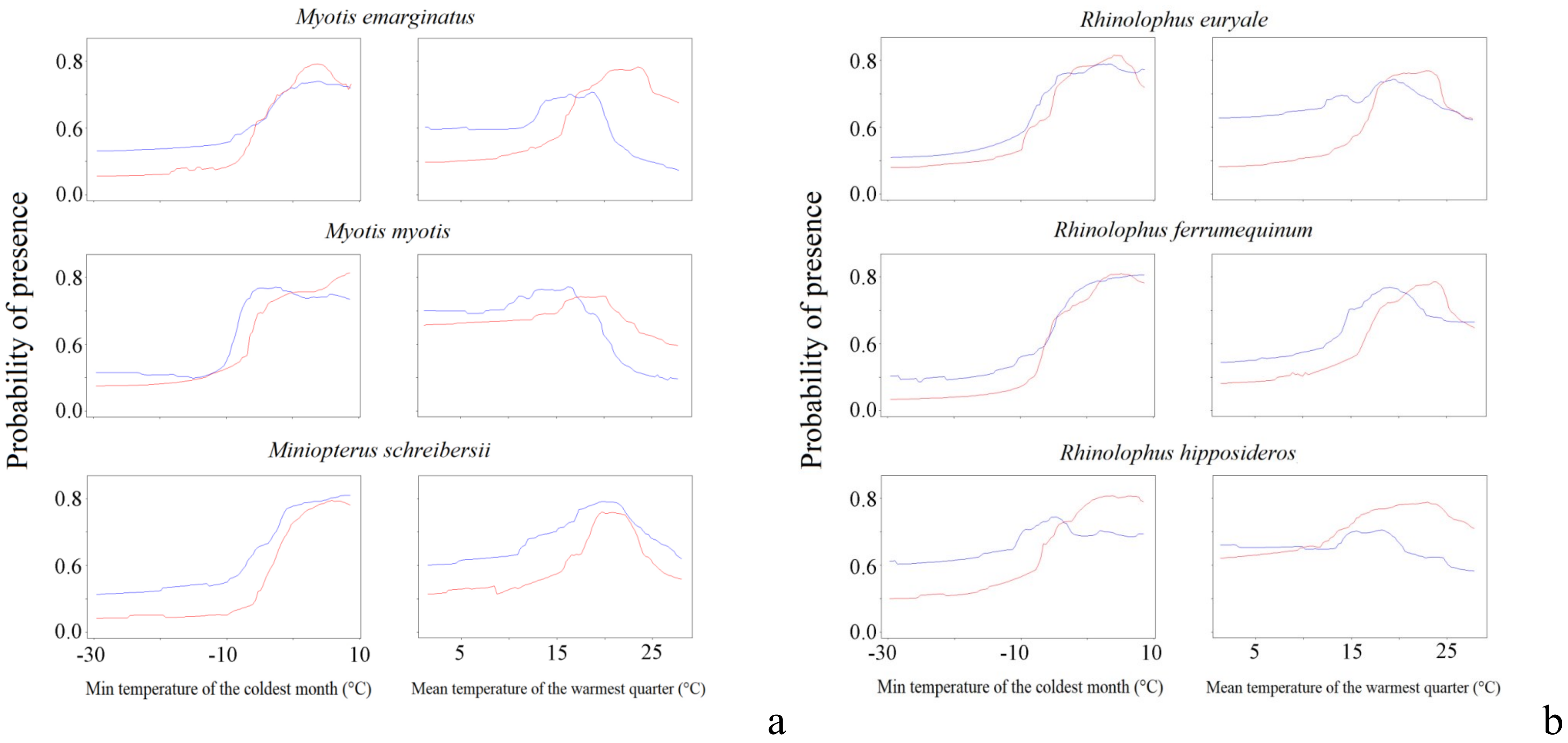


Figure 3. a-b. Response curves for two out of the ten variables used to model the potential distribution of six European bat species. The minimum temperature of the coldest month (left column) and the mean temperature of the warmest quarter (right column) determine a different probability of presence for each species depending on the seasonal model, i.e. the summer (red curve) or the winter (blue curve) model.

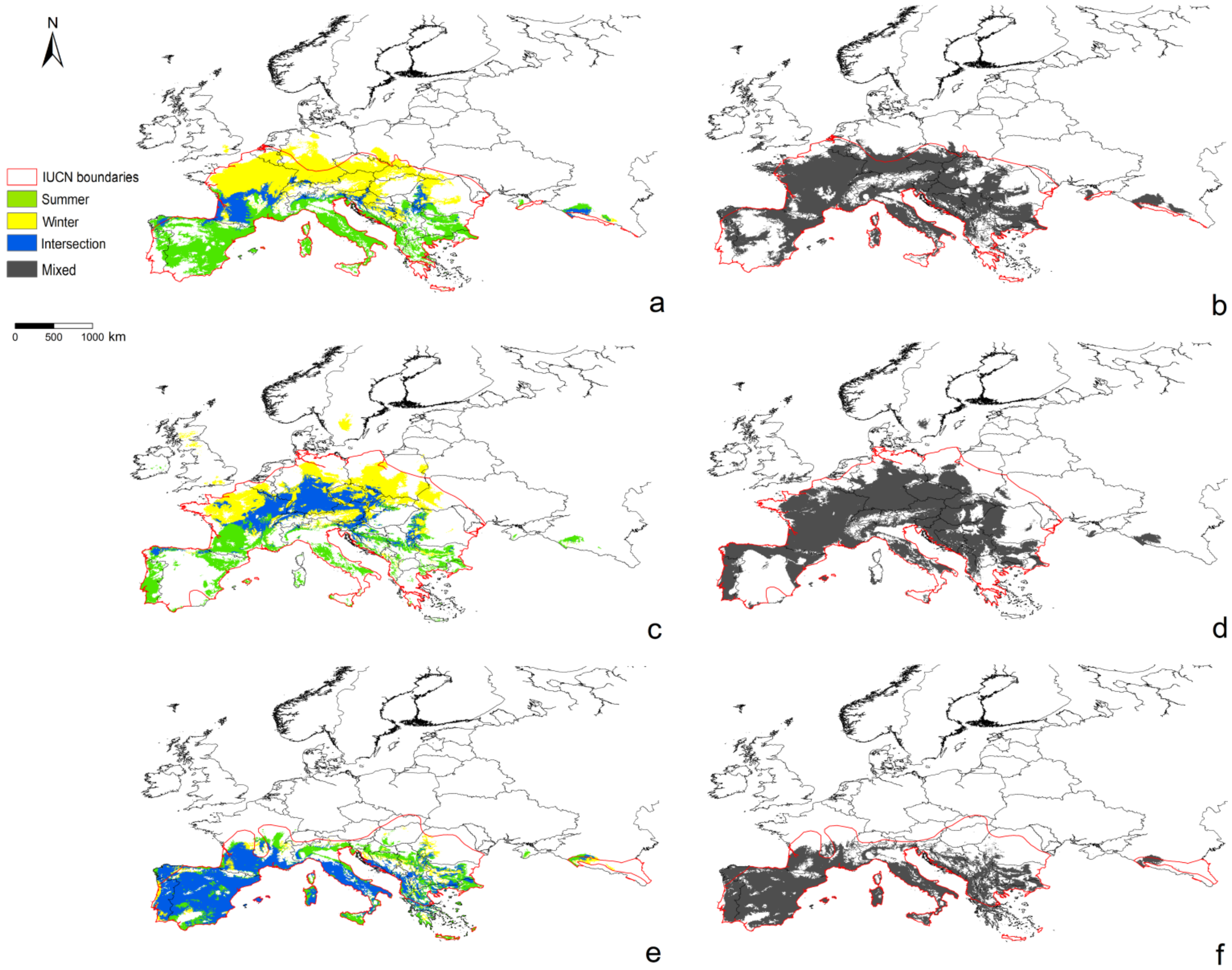


Figure 4. Presence/absence maps generated with the *biomod2* package for: a-b) *Myotis emarginatus*; c-d) *M. myotis*; e-f) *Miniopterus schreibersii.* The left column shows the overlap between summer and winter predicted distributions. The right column represents the habitat suitability map obtained with the mixed model approach.

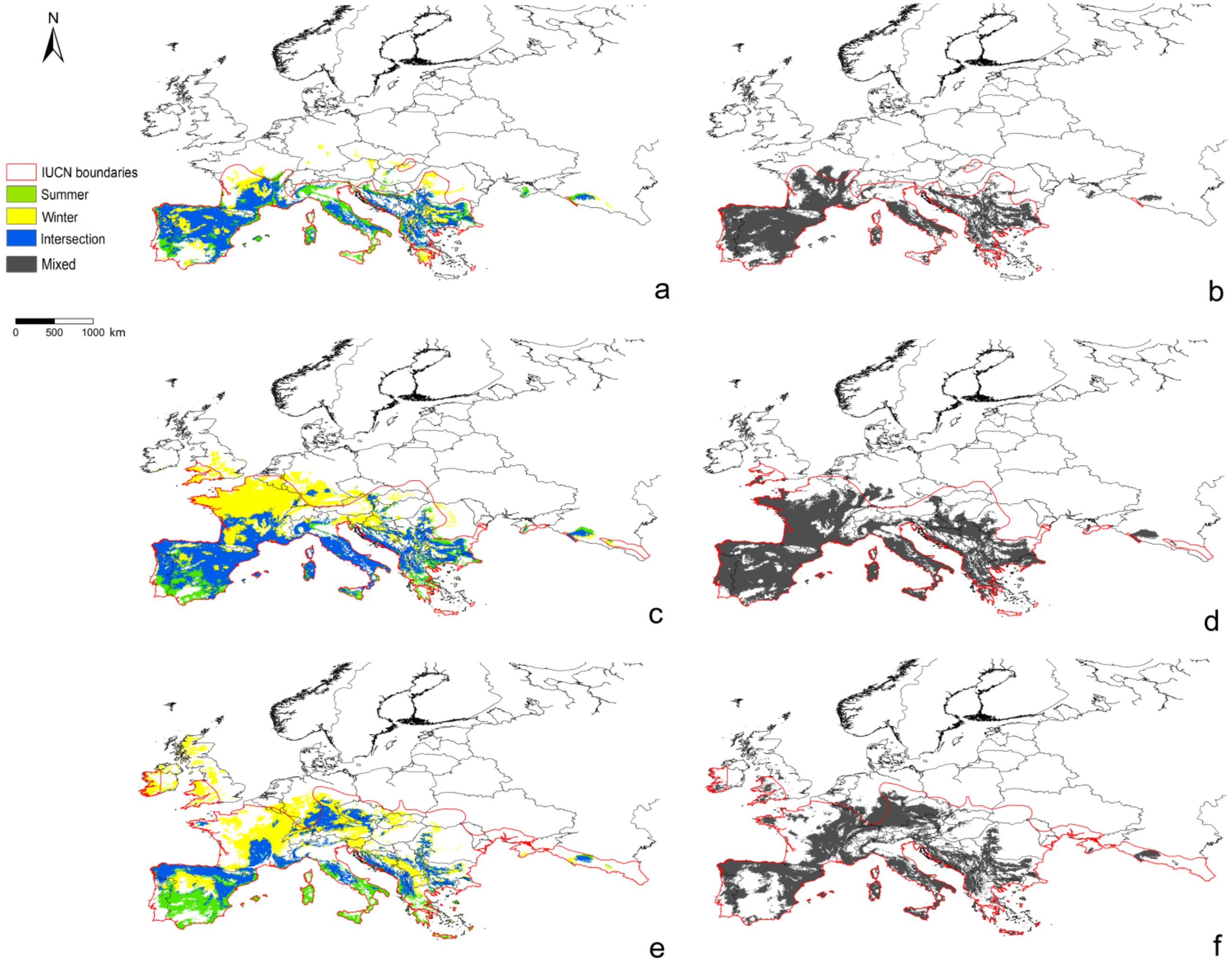


Figure 5. Presence/absence maps generated with the *biomod2* package for: a-b) *Rhinolophus euryale*; c-d) *R. ferrumequinum*; e-f) *R. hipposideros*. The left column shows the overlap between summer and winter predicted distributions. The right column represents the habitat suitability map obtained with the mixed model approach

Table 5. Percentage of agreement between predicted summer and winter ranges (S-W), mixed and summer models (Mix-S) and between mixed and winter models (Mix-W) developed for six bat species in Europe. The values shown in table express the percent overlap between binary maps derived from the species' ensemble models. S = summer model, W = winter model, Mix = mixed model

Species	S-W	Mix- S	Mix-W
Myotis emarginatus	12.82	73.23	84.31
Myotis myotis	28.21	94.71	85.11
Miniopterus schreibersii	61.30	94.27	98.48
Rhinolophus euryale	54.11	94.51	86.68
Rhinolophus ferrumequinum	46.92	97.81	84.39
Rhinolophus hipposideros	33.23	87.35	77.08

Table 6. Percentage and size of the study area predicted as suitable for six European bat species by seasonal and mixed SDMs

Species	Range size					
	Mixed model		Summer model		Winter model	
	%	km²	%	km²	%	km²
Myotis emarginatus	13.97	1,410,317	8.20	827,817	8.38	846,392
Myotis myotis	14.87	1,501,175	7.62	769,264	8.53	860,626
Miniopterus schreibersii	8.90	898,484	8.61	869,208	7.35	742,007
Rhinolophus euryale	9.22	930,789	7.14	720,706	7.09	715,759
Rhinolophus ferrumequinum	13.96	1,409,307	9.11	919,684	12.50	1,261,613
Rhinolophus hipposideros	10.58	1,068,085	7.42	749,073	10.51	1,060,514

Table 7. Percentage and size of the IUCN range predicted as suitable for all bat species by seasonal and mixed SDMs respectively.

Species	Range size						
	IUCN	Mixed model		Summer model		Winter model	
	km²	%	km²	%	km²	%	km²
Myotis emarginatus	2,807,211	52.25	1,466,768	34.94	980,840	33.72	946,592
Myotis myotis	3,309,502	55.13	1,824,528	28.63	947,510	28.42	940,560
Miniopterus schreibersii	3,101,089	59.61	1,848,559	58.30	1,807,935	46.90	1,454,411
Rhinolophus euryale	1,782,141	66.42	1,183,698	56.00	997,999	52.54	936,337
Rhinolophus ferrumequinum	2,574,230	64.91	1,670,933	47.22	1,215,551	66.20	1,704,140
Rhinolophus hipposideros	3,199,009	40.82	1,337,826	30.41	972,819	41.53	1,296,558

Table 8. Summary of the dataset used to develop the GLMM tests. For each species we calculated the percentage of IUCN ranges predicted respectively by mixed, summer and winter SDMs. Because 10 replicates were developed for each SDM, in this table we indicated only mean value and standard deviation of the IUCN ranges size.

Species	Predicted % of IUCN range size		
	Mixed model	Summer model	Winter model
Myotis emarginatus	36.7 (±7.26)	34.94 (±0.01)	33.33 (±7.97)
Myotis myotis	40.87 (±5.06)	40.64 (±0.01)	29.62 (±4.25)
Miniopterus schreibersii	49.61 (±7.79)	47.32 (±6.19)	46.43 (±0.01)
Rhinolophus euryale	54.16 (±11.32)	48.94 (±9.81)	49.19 (±0.02)
Rhinolophus ferrumequinum	56.66 (±5.00)	43.42 (±0.02)	54.43 (±8.06)
Rhinolophus hipposideros	30.61 (±7.12)	28.14 (±0.01)	35.53 (±11.77)

Table 9. Results from generalized linear mixed models testing the effect of summer, winter or mixed SDMs on predicted range sizes of six European bat species (*Myotis emarginatus, M. myotis, Miniopterus schreibersii*, *Rhinolophus euryale*, *R. ferrumequinum*, *R. hipposideros*). Asterisks denote significant results.

GLMM response	Variable	Estimate	Error	Pr(>|z|)
Range size	Summer model	-0.09563	0.03751	0.01079 *
	Winter model	-0.10839	0.03753	0.00387 **

Table 10. Euclidean distance (km) measured between the centroids of our predicted SDMs and IUCN ranges of six European bat species

Species	Distance from IUCN centroids		
	Mixed model	Summer model	Winter model
Myotis emarginatus	80.9	299.4	337.8
Myotis myotis	155.9	202.9	327.5
Miniopterus schreibersii	432.9	422.5	547.3
Rhinolophus euryale	139.2	46.1	80.7
Rhinolophus ferrumequinum	195.8	274.8	219.5
Rhinolophus hipposideros	266.4	451.5	373.6

4.4.2. Phenology-dependent niche differences

The first two PCA axes featured ca. 60 % of bioclimatic variable' contribution. The niche similarity tests highlighted the existence of three main groups of bat species according to the degree of overlap between summer and winter niches, i.e. between reproduction and hibernation ecological requirements (Figure 4). Specifically, *M. schreibersii* and *R. euryale* showed the highest Schoener's D (> 0.7) with a symmetric overlap of the two niches in the environmental space (Figure 4). *R. ferrumequinum* and *R. hipposideros* belonged to a second group characterized by a high niche similarity (Schoener's D index > 0.6) and a moderately symmetric overlap between seasonal niches (Figure 4). Finally, *M. myotis* and *M. emarginatus* showed the lowest niche similarity and an asymmetric overlap of the two niches (Schoener's D index ≤0.4) although for *M. emarginatus* no replicates of the niche similarity test were significant (Fig. 6). Furthermore, in case of difference between the percentages of the two-directions similarity tests yielding significant outcomes (i.e. winter → summer and summer → winter) we may identify which phenological niche includes the other one. In particular, *R. euryale* and *R. hipposideros* summer niches included winter niches, and the opposite was observed for *R. ferrumequinum* and *M. myotis*.

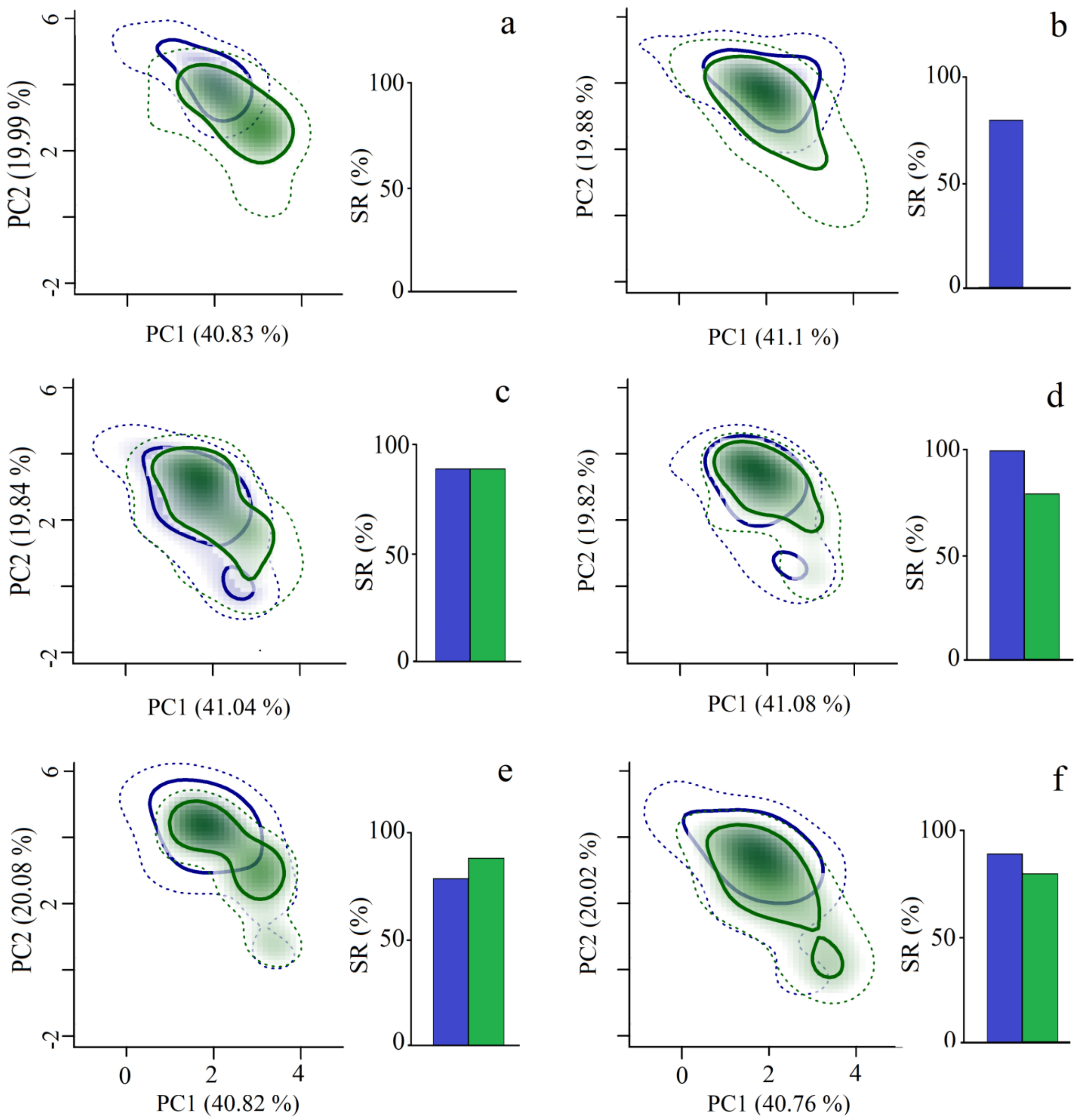


Figure 6. Graphical representation of the environmental niches of: a) *Myotis emarginatus;* b) *M. myotis;* c) *Miniopterus schreibersii*; d) *Rhinolophus euryale*; e) *R. ferrumequinum*; f) *R. hipposideros*. Each panel represents summer (green) and winter (blue) niches of a species along the first two axis (PC1 and PC2) of PCA (left); the percentage of replicates (SR) for which the similarity test of winter to summer niche (blue bars) and of summer to winter niche (green bars) was significant (right).

4.5. Discussion

All models showed an AUC value > 0.9 and a TSS values > 0.8 representing a very high predictive performance (e.g.: Rebelo et al., 2010; Domíguez-Vega et al., 2012; Smeraldo et al., 2017; Bosso et al., 2017b).

For almost all the species our first hypothesis was confirmed, i.e. model outputs differed greatly between seasons depending on the dataset used. The results also supported our second hypothesis, i.e. SDMs developed using records from winter or summer roosts separately were less accurate than the corresponding mixed models in predicting a species' potential range. Such a serious bias should be taken carefully into account since SDMs relying on incomplete datasets that underrepresent phenological variation may lead to erroneous outputs and potentially affect conservation strategies. Carretero & Sillero (2016) evaluated the robustness of SDMs against the use of spatially incomplete datasets for the threatened Iberian lizard *Podarcis carbonelli* and found that the resulting models were biased because models provided a partial representation of the environmental niche, which may vary across the distribution range, thus violating the assumption of niche equality (e.g. Guisan & Zimmermann, 2000; Wiens et al., 2009; Peterson, 2011). Our study adds a further dimension to this problem since occurrence datasets may also underrepresent temporal variation in the use of space, and this is especially likely when modelling the ecological niche of species showing pronounced phenological patterns. It is also important to remark that for our study we conservatively included species that are sedentary or only carrying out small-scale migration movements such as *M. myotis* (Rodrigues et al., 2003). Using long-distance migrants would have probably led to even larger discrepancies among models. A thorough knowledge of a species' year-round physiological requirements is therefore essential to select suitable occurrence datasets and generate faithful representations of potential distribution (Feng & Papes, 2017).

Many studies confirmed that a deep knowledge of species' biology and ecology represents the initial and most critical phase of SDMs' development before giving it concrete support with the use of modelling techniques (Austin, 2002; Guisan & Thuiller, 2005; Franklin, 2010; Mathewson et al., 2017). Raes (2012) advised that SDMs calibrated across the broadest spatial, environmental and/or temporal extent are biologically and biogeographically justifiable to capture a species' niche in its broadest sense. Neglecting the full environmental variation under which a species may occur, i.e. in our case using phenologically biased collections to fit SDMs may result in under-predicted species' distributions, and essentially similar to modelling a partial niche.

In our study, winter SDMs failed to predict many potentially suitable regions of southern Europe and, on the contrary, summer SDMs excluded most of the north-eastern part of species' ranges. These findings are related to the influence of different climatic predictors over seasonal ranges (Morganti et al., 2017; Ortega & Pérez-Mellado, 2016). Summer potential distribution was mostly limited by the minimum temperature of the coldest month: this is explained by the negative response of bat reproduction and lactation towards cold summer temperature, which decrease food availability, in turn delaying embryo development and reducing milk production (Lewis, 1993; Hoying & Kunz, 1998; Arlettaz et al., 2001). Colder roost temperatures, especially those of poorly insulated roosts that respond promptly to drops in environmental temperatures (Russo et al., 2017), may slow down juvenile growth and ultimately affect their winter survival (Racey & Swift, 1981; Tuttle & Stevenson, 1982; Erickson & West, 2002). Bats may therefore select for warmer areas in summer (Erickson & West 2002; Frick et al., 2010) and this was reflected in our models. On the other hand, in winter hibernating bats need low ambient temperatures and small thermal fluctuations to minimize energy expenditure (Nagel & Nagel, 1991; Speakman & Rowland, 1999; Speakman &Thomas, 2003; Masing & Lutsar, 2007). Since temperatures of hibernacula are generally a consequence of the mean annual surface temperature (Perry, 2013), in our winter models low temperature areas were more important than in summer models.

The geographical divergence existing between seasonal SDMs was confirmed by the niche overlap analyses for three species, i.e. *M. myotis*, *R. ferrumequinum* and *R. hipposideros*, which showed either an asymmetric or a moderate symmetric similarity between winter and summer ranges in the environmental space. Instead, *M. schreibersii* and *R. euryale* showed a higher degree of similarity between seasonal niches which reflected a greater overlap between their habitat suitability maps than in the previously mentioned species group. Both species have a more southerly range than the remaining we analysed and share similar roosting requirements (Russo et al., 2002), are associated with mild Mediterranean winters and frequently use warm underground sites year-round. Based on our results, and given the noticeable differences found between seasonal SDMs, we propose that mixed SDMs developed using the same numbers of winter and summer occurrences will portray a more realistic potential distribution.

To evaluate the spatial prediction performances of our models, we considered the IUCN distribution maps as "observed" species' ranges. However, Herkt et al. (2017) found that IUCN expert maps of African bats differ substantially from carefully crafted SDMs designed to depict species' complete ranges. IUCN boundaries are developed from spatial information on species occurrences and do not include information on species' environmental preferences (de Castro Pena, Kamino, Rodrigues, Mariano-Neto, & de Siqueira, 2014). In severely undersampled regions, all the more in the speciose tropics including those of Africa, such maps will tend to fall short of estimating the complete species' range. Nevertheless, IUCN expert maps may still be appropriate for ecological analyses when they focus on easily recorded taxa and well-sampled regions (Breiner et al., 2017; Herkt et al., 2017) such as in our work.

We found that for all species mixed SDMs were able to predict a larger area of the IUCN species' range than did seasonal models so the former provided a better estimation of the observed species' distribution. Nevertheless, some areas within the IUCN ranges were left out, especially some regions at the edge of the range. These findings may depend on the inability of IUCN maps to exclude unsuitable habitats due to discontinuities in ecological variables necessary to the species, or their sensitivity to extreme outliers (Burgman & Fox, 2003; de Castro Pena et al., 2014). For the aims of our work, i.e. to emphasize the phenological significance of the occurrences we used, we considered only hibernation and maternity roost as species' presence records. This deliberate restriction of the dataset, which excluded records from bat captures or acoustic detection at foraging or drinking sites, might have underestimated species' realized niche. However, this type of dataset has been also used successfully to generate summer roosting habitat maps for bats species in England at different scales (Bellamy & Althringham, 2015). Moreover, given the typical (< 30 km: e.g. Biscardi et al., 2007) scale of foraging trips made by bats from their roosts to the feeding sites, and considering the scale at which our modelling exercise was done, we rule out that omitting foraging records from our dataset may have generated discrepancies between our predicted distributions and the IUCN maps.

To our best knowledge, our work is the first to test the effects of data phenology on SDMs. Our conclusion is likely to apply to a broad range of organisms besides bats and question to validity, or at least the comprehensiveness, of the findings of many previous studies. While we encourage future studies on other species showing a more or less pronounced phenological spatial variation, we strongly recommend that for conservation and management applications, models based on seasonally mixed datasets in which the different phenologies are evenly represented should be preferred over models encompassing only one phenological stage or built on seasonally biased datasets.

4.6. References

Algar, A.C., Kharouba, H.M., Young, E.R., & Kerr, J.T. (2009). Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography*, *32*, 22-33.

Allouche, O., Tsoar, A., & Kadmon R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, *43*, 1223-1232.

Araujo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of biogeography*, *33*(10), 1677-1688.

Ardestani, E. G., Tarkesh, M., Bassiri, M., & Vahabi, M. R. (2015). Potential habitat modeling for reintroduction of three native plant species in central Iran. *Journal of arid land*, *7*(3), 381-390.

Arlettaz, R., Christe, P., Lugon, A., Perrin, N., & Vogel, P. (2001). Food availability dictates the timing of parturition in insectivorous mouse-eared bats. *Oikos*, *95*(1), 105-111.

Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling*, *157*(2), 101-118.

Austin, M. P., & Van Niel, K. P. (2011). Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, *38*(1), 1-8.

Balestrieri, A., Bogliani, G., Boano, G., Ruiz-González, A., Saino, N., Costa, S., & Milanesi, P. (2016). Modelling the distribution of forest-dependent species in human-dominated landscapes: Patterns for the pine marten in intensively cultivated lowlands. *PloS one*, *11*(7), e0158203.

Barbet-Massin, M., Thuiller, W. & Jiguet, F. (2010). How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography*, *33,* 878–886.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in Ecology and Evolution*, *3*(2), 327-338.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*(11), 1810-1819.

Bates, D., Maechler, M. & Dai, B. (2008). Lme4: Linear-Mixed Effects Models Using S4 Classes. R Package Version 0.999375-27. Available at: http://lme4.rforge.r-project.org/.

Beaumont, L. J., Pitman, A. J., Poulsen, M., & Hughes, L. (2007). Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, *13*(7), 1368-1385.

Bellamy, C., & Altringham, J. (2015). Predicting species distributions using record centre data: multi-scale modelling of habitat suitability for bat roosts. *PloS one*, *10*(6), e0128440.

Biscardi, S., Russo, D., Casciani, V., Cesarini, D., Mei, M., Boitani, L. (2007). Foraging requirements of the endangered long-fingered bat: the influence of micro-habitat structure, water quality and prey type. *Journal of Zoology,* 273,372–381.

Bosso, L., Mucedda, M., Fichera, G., Kiefer, A., & Russo, D. (2016)a. A gap analysis for threatened bat populations on Sardinia. *Hystrix, the Italian Journal of Mammalogy*, *27*(2), 212 – 214.

Bosso, L., Di Febbraro, M., Cristinzio, G., Zoina, A., Russo, D. (2016)b. Shedding light on the effects of climate change on the potential distribution of *Xylella fastidiosa* in the Mediterranean basin. *Biological Invasions*, *18*,1759-1768.

Bosso, L., De Conno, C., & Russo, D. (2017)a. Modelling the Risk Posed by the Zebra Mussel Dreissena polymorpha: Italy as a Case Study. *Environmental Management*, *60*,304–313.

Bosso, L., Luchi, N., Maresi, G., Cristinzio, G., Smeraldo, S., & Russo D. (2017)b. Predicting current and future disease outbreaks of *Diplodia sapinea* shoot blight in Italy: species distribution models as a tool for forest management planning. *Forest Ecology and Management, 400*, 655-664.

Bosso, L., Smeraldo, S., Rapuzzi, P., Sama, G., Garonna, A.P., & Russo D. (2017)c. Nature protection areas of Europe are insufficient to preserve the threatened beetle *Rosalia alpina* (Coleoptera: Cerambycidae): evidence from species distribution models and conservation gap analysis. *Ecological Entomology*, DOI: 10.1111/een.12485.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Breiner, F. T., Guisan, A., Nobis, M. P., & Bergamini, A. (2017). Including environmental niche information to improve IUCN Red List assessments. *Diversity and Distributions*, *23*(5), 484-495.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* 21, 481–497.

Burgman, M. A., & Fox, J. C. (2003, February). Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. In *Animal Conservation forum* (Vol. 6, No. 1, pp. 19-28). Cambridge University Press.

Carretero, M. A., & Sillero, N. (2016). Evaluating how species niche modelling is affected by partial distributions with an empirical case. *Acta Oecologica*, *77*, 207-216.

Carvalho, S. B., Brito, J. C., Pressey, R. L., Crespo, E., & Possingham, H. P. (2010). Simulating the effects of using different types of species distribution data in reserve selection. *Biological Conservation*, *143*(2), 426-438.

Chuine, I. (2010). Why does phenology drive species distribution?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*(1555), 3149-3160.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., D'Amen, M., Randin, C., ... & Pellissier, L. (2017). ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography*, *40*(6), 774-787.

Di Febbraro, M., Roscioni, F., Frate, L., Carranza, M.L., De Lisio, L., De Rosa, D., & Loy, A., (2015). Long-term effects of traditional and conservation-oriented forest management on the distribution of vertebrates in Mediterranean forests: a hierarchical hybrid modelling approach. *Diversity and Distributions*, *21*, 1141-1154.

Di Febbraro, M., Martinoli, A., Russo, D., Preatoni, D., & Bertolino, S. (2016). Modelling the effects of climate change on the risk of invasion by alien squirrels. *Hystrix, the Italian Journal of Mammalogy*, *27*(1).

Doko, T., Fukui, H., Kooiman, A., Toxopeus, A. G., Ichinose, T., Chen, W., & Skidmore, A. K. (2011). Identifying habitat patches and potential ecological corridors for remnant Asiatic black bear (*Ursus thibetanus japonicus*) populations in Japan. *Ecological Modelling*, *222*(3), 748-761.

Domíguez-Vega H, Monroy-Vilchis O, Balderas-Valdivia CJ, Gienger CM, & Ariano-Sánchez D (2012) Predicting the potential distribution of the beaded lizard and identification of priority areas for conservation. *Journal for Nature Conservation, 20*,247–253.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M. & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography,* 30, 609–628.

de Castro Pena, J. C., Kamino, L. H. Y., Rodrigues, M., Mariano-Neto, E., & de Siqueira, M. F. (2014). Assessing the conservation status of species with limited available data and disjunct distribution. *Biological Conservation*, *170*, 130-136.

Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.P., & Guisan A. (2011). Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, *17*, 1122-1131.

Ducci, L., Agnelli, P., Di Febbraro, M., Frate, L., Russo, D., & Loy, A., Roscioni F., (2015). Different bat guilds perceive their habitat in different ways: a multiscale landscape approach for variable selection in species distribution modelling. *Landscape Ecology*, *30*, 2147-2159.

Erickson, J. L., & West, S. D. (2002). The influence of regional climate and nightly weather conditions on activity patterns of insectivorous bats. *Acta Chiropterologica*, *4*(1), 17-24.

Feng, X., & Papeş, M. (2017). Can incomplete knowledge of species' physiology facilitate ecological niche modelling? A case study with virtual species. *Diversity and Distributions*, *23*, 1157–1168.

Feuda, R., Bannikova, A. A., Zemlemerova, E. D., Di Febbraro, M., Loy, A., Hutterer, R., ... & Colangelo, P. (2015). Tracing the evolutionary history of the mole, Talpa europaea, through mitochondrial DNA phylogeography and species distribution modelling. *Biological Journal of the Linnean Society*, *114*(3), 495-512.

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

Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PloS one*, *9*(5), e97122.

Frick, W. F., Reynolds, D. S., & Kunz, T. H. (2010). Influence of climate and reproductive timing on demography of little brown myotis *Myotis lucifugus*. *Journal of animal ecology*, *79*(1), 128-136.

Franklin, J. (2010). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Grindal, S. D., Collard, T. S., Brigham, R. M., & Barclay, R. M. (1992). The influence of precipitation on reproduction by Myotis bats in British Columbia. *American Midland Naturalist*, 339-344.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2), 147-186.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, *8*(9), 993-1009.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., ... & Martin, T. G. (2013). Predicting species distributions for conservation decisions. *Ecology letters*, *16*(12), 1424-1435.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29-36.

Hansson, L. A., & Åkesson, S. (Eds.). (2014). *Animal movement across scales*. Oxford University Press.

Herkt, K. M. B., Skidmore, A. K., & Fahr, J. (2017). Macroecological conclusions based on IUCN expert maps: A call for caution. *Global Ecology and Biogeography*, *26*(8), 930-941.

Hernandez, P.A., Graham, C.H., Master, L.L., & Albert, D.L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29, 773-785.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, *25*(15), 1965-1978.

Hirzel, A. H., Hausser, J., Chessel, D., & Perrin, N. (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?. *Ecology*, *83*(7), 2027-2036.

Hoying, K. M., & Kunz, T. H. (1998). Variation in size at birth and post-natal growth in the insectivorous bat *Pipistrellus subflavus* (Chiroptera: Vespertilionidae). *Journal of Zoology*, *245*(1), 15-27.

Holte, D., Köppen, U., & Schmitz-Ornés, A. (2017). A comparison of migratory strategies of partial migratory raptors from Germany. *Journal of Ornithology*, *158*(2), 579-592.

Huntley, B., Berry, P. M., Cramer, W., & McDonald, A. P. (1995). Special paper: modelling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography*, 967-1001.

Huntley B, Green R, Collingham YC et al. (2007) A Climatic Atlas of European Breeding Birds. *Lynx Edicions*, Barcelona.

Huston, M. A. (2002). Introductory essay: critical issues for improving predictions. *Predicting species occurrences: issues of accuracy and scale*, 7-21.

Jiguet, F., Barbet-Massin, M., & Henry, P.Y. (2010). Predicting potential distributions of two rare allopatric sister species, the globally threatened *Doliornis cotingas* in the Andes. *Journal of Field Ornithology*, 81, 325-339.

Kwon, H.S., Kim, B.J., Jang, G.S. (2016). Modelling the spatial distribution of wildlife animals using presence and absence data. *Contemporary Problems of Ecology,* 9, 515-518.

Lewis, S. E. (1993). Effect of climatic variation on reproduction by pallid bats (Antrozous pallidus). *Canadian Journal of Zoology*, *71*(7), 1429-1433.

Liu, C., Berry, P.M., Dawson, T.P., & Pearson, R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*, 385-393.

Lobo, J.M., Jimenez-Valverde A., Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography, 33*, 103-114.

Loy, A., Cassini, M.H., Colangelo, P., & Di Febbraro, M. (2017). Distribution, spatial interaction and niche analysis in three species of European moles (*genus Talpa*, Soricomorpha: Mammalia) in Italy. *Biological Journal of the Linnean Society*, online version.

Lyman, C. P. (2013). Hibernation and torpor in mammals and birds. Elsevier.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and distributions*, *15*, 59-69.

Masing, M., & Lutsar, L. (2007). Hibernation temperatures in seven species of sedentary bats (Chiroptera) in northeastern Europe. *Acta Zoologica Lituanica*, *17*(1), 47-55.

Mathewson, P. D., Moyer-Horner, L., Beever, E. A., Briscoe, N. J., Kearney, M., Yahn, J. M., & Porter, W. P. (2017). Mechanistic variables can enhance predictive models of endotherm distributions: the American pika under current, past, and future climates. *Global change biology*, *23*(3), 1048-1064.

Miller-Rushing, A.J., & Weltzin J. 2009. Phenology as a tool to link ecology and sustainable decision making in a dynamic environment. *New Phytologist, 184*, 743–745.

Mitchell-Jones AJ, Amori G, Bogdanowicz W et al. (1999) The Atlas of European Mammals. *Poyser Natural History*, London.

Morganti, M., Preatoni, D., & Sarà, M. (2017). Climate determinants of breeding and wintering ranges of lesser kestrels in Italy and predicted impacts of climate change. *Journal of Avian Biology*. DOI: 10.1111/jav.01179

Nagel, A., & Nagel, R. (1991). How do bats choose optimal temperatures for hibernation?. *Comparative Biochemistry and Physiology Part A: Physiology*, *99*(3), 323-326.

Newton I (2008) The migration ecology of birds. Elsevier, Amsterdam.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., ... & Loucks, C. J. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), 933-938.

Olsson, O., & Rogers, D. J. (2009). Predicting the distribution of a suitable habitat for the white stork in Southern Sweden: identifying priority areas for reintroduction and habitat restoration. *Animal Conservation*, *12*(1), 62-70.

Ortega, Z., & Pérez-Mellado, V. (2016). Seasonal patterns of body temperature and microhabitat selection in a lacertid lizard. *Acta Oecologica*, *77*, 201-206.

Papadatou, E., Butlin, R. K., & Altringham, J. D. (2008). Seasonal roosting habits and population structure of the long-fingered bat Myotis capaccinii in Greece. *Journal of Mammalogy*, *89*(2), 503-512.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?. *Global ecology and biogeography*, *12*(5), 361-371.

Perry, Roger W. 2013. A review of factors affecting cave climates for hibernating bats in temperate North America. Environmental Reviews 21: 28-39.

Peterson, A. T. (2011). *Ecological niches and geographic distributions (MPB-49)*(No. 49). Princeton University Press.

Pio, D.V., Engler, R., Linder, H.P., Monadjem, A., Cotterill F.P., Taylor, P.J., & Salamin, N. (2014). Climate change effects on animal and plant phylogenetic diversity in southern Africa. *Global Change Biology*, 20, 1538-1549.

Racey, P. A., & Swift, S. M. (1981). Variations in gestation length in a colony of pipistrelle bats (Pipistrellus pipistrellus) from year to year. *Journal of Reproduction and Fertility*, *61*(1), 123-129.

Raes, N. (2012). Partial versus full species distribution models. *Nature Conservation*, *10*(2), 127-138.

Rebelo, H., Tarroso, P., & Jones, G. (2010). Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology*, *16*(2), 561-576.

Reddy S., Davalos L.M. (2003) Geographical sampling bias and its implications for conservation priorities in African Journl of Biogeography, *30,* 1719–1727.

Rivers, N. M., Butlin, R. K., & Altringham, J. D. (2006). Autumn swarming behaviour of Natterer's bats in the UK: population size, catchment area and dispersal. *Biological conservation*, *127*(2), 215-226.

Rodrigues, L., Zahn, A., Rainho, A., & Palmeirim, J. M. (2003). Contrasting the roosting behaviour and phenology of an insectivorous bat (*Myotis myotis*) in its southern and northern distribution ranges. *Mammalia*, *67*(3), 321-336.

Rubio-Salcedo, M., Psomas, A., Prieto, M., Zimmermann, N. E., & Martínez, I. (2017). Case study of the implications of climate change for lichen diversity and distributions. *Biodiversity and Conservation*, *26*(5), 1121-1141.

Russo, D., Jones, G., & Migliozzi, A. (2002). Habitat selection by the Mediterranean horseshoe bat, Rhinolophus euryale (Chiroptera: Rhinolophidae) in a rural area of southern Italy and implications for conservation. *Biological Conservation*, *107*(1), 71-81.

Russo, D., Di Febbraro, M., Rebelo, H., Mucedda, M., Cistrone, L., Agnelli, P., De Pasquale, P.P., Martinolo, A., Scaravelli, D., Spilinga, C. & Bosso, L. (2014). What story does geographic separation of insular bats tell? A case study on Sardinian Rhinolophids. *PLoS One*, *9*(10), e110894.

Russo, D., Di Febbraro, M., Cistrone, L., Jones, G., Smeraldo, S., Garonna, A.P., Bosso, L. (2015). Protecting one, protecting both? Scale-dependent ecological differences in two species using dead trees, the rosalia longicorn beetle and the barbastelle bat. *Journal of Zoology, 297*, 165-175.

Russo, D., Cistrone, L., Budinski, I., Console, G., Della Corte, M., Milighetti, C., ... & Ancillotto, L. (2017). Sociality influences thermoregulation and roost switching in a forest bat using ephemeral roosts. *Ecology and Evolution*, *7*(14), 5310–5321.

Schoener, T.W. 1970. Non-synchronous spatial overlap of lizards in patchy habitats. *Ecology*, *51*,408-418.

Smeraldo, S., Di Febbraro, M., Ćirović, D., Bosso, L., Trbojević, I., & Russo, D. (2017). Species distribution models as a tool to predict range expansion after reintroduction: A case study on Eurasian beavers (*Castor fiber*). *Journal for Nature Conservation*, *37*,12-20.

Speakman, J. R., & Rowland, A. (1999). Preparing for inactivity: how insectivorous bats deposit a fat store for hibernation. *Proceedings of the Nutrition Society*, *58*(1), 123-131.

Speakman, J. R., & Thomas, D. W. (2003). Physiological ecology and energetics of bats. *Bat ecology*, 430-490.

Syfert M.M., Smith M.J., Coomes D.A. (2013) The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS One* 8: e55158.

Thuiller, W. (2003). BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. *Global change biology*, *9*(10), 1353-1362.

Thuiller, W., Araujo, M. B., & Lavorel, S. (2004). Do we need land-cover data to model species distributions in Europe?. *Journal of Biogeography*, *31*(3), 353-361.

Thuiller, W., Richardson, D. M., PYŠEK, P., Midgley, G. F., Hughes, G. O., & Rouget, M. (2005). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, *11*(12), 2234-2250.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M.B., 2009. BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Tulloch, A. I., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P., ... & Rhodes, J. R. (2016). Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation*, *199*, 157-171.

Tuttle, M. D., & Stevenson, D. (1982). Growth and survival of bats. In 'Ecology of Bats'.(Ed. TH Kunz.) pp. 105–150.

Van Horn, B. (2002). Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. Á In: Scott, JM et al.(eds), Predicting species occurrences: issues of accuracy and scale.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, *62*, 2868–2883.

Webb, P. I., Speakman, J. R., & Racey, P. A. (1996). How hot is a hibernaculum? A review of the temperatures at which bats hibernate. *Canadian Journal of Zoology*, *74*(4), 761-765.

Wiens, J. A. (2002). Predicting species occurrences: progress, problems, and prospects. *Predicting species occurrences: issues of accuracy and scale*, 739-749.

Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19729-19736.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763-773.

Zhang, J., Nielsen, S. E., Chen, Y., Georges, D., Qin, Y., Wang, S. S., ... & Thuiller, W. (2017). Extinction risk of North American seed plants elevated by climate and land-use change. *Journal of Applied Ecology*, *54*(1),303-312.

Chapter V

5. Concluding remarks

The aim of this thesis was to evaluate the effectiveness of SDM's technique in supporting species conservation actions. We used SDMs as a methodological approach to test different hypotheses regarding ecological, biological and conservation issues. In particular, within the case studies addressed in this thesis, SDMs allowed us obtain the following results:

1. to detect new potential suitable areas where rare and threatened species were not yet surveyed or where they have not been observed before due to their elusive behaviour. In this perspective, the resulting habitat suitability maps provided a valuable support to researchers who may focus monitoring efforts in precise areas avoiding useless and time consuming actions.

2. to investigate the degree of protection guaranteed by the network of protected areas both to the current presence sites and to the all potential suitable habitat. Such analyses allow to detect potential gaps in species' conservation and may suggest both to extend the surface of protected areas to new suitable habitat and to increase the degree of connectivity between those already existing.

3. to highlight which environmental variables mostly determine species' probability of presence and allow to develop a niche overlap analyses in order to compare ecological niches of two species. However, broad scale SDMs revealed limited differences in preferred environmental predictor variables while failed to detect differences in microhabitats occupied. Only a small scale niche analyses provided us with detailed ecological differences characterizing the two species and gave information on the type of management that the species need. Our findings highlighted that it is always advisable to carry out analyses at multiple scales due to limitations of SDMs related to the choice of the scale of analyses which should be selected according to the aim of the study.

4. to evaluate the effectiveness of SDMs in managing the conservation of reintroduced species during the post-release phase. We were able to predict suitable areas which might be colonized in the near future by the reintroduced species and to evaluate the potential risk posed to the expanding population by the very low degree of protection offered by the national network of protected areas. Developing effective strategies aimed to increase the current network of protected areas with the support of SDMs may lead to reduce conflicts and human-induced mortality of the species. Using occurrence records of a reintroduced population during its colonization phase, might violate one of the main assumptions of SDMs which says that species must be in equilibrium with the environment

in the study area. We suggested to apply a hierarchical approach in which weights of pseudo-absences are adjusted according to a large-scale climatic model based also on records of populations that have reached the equilibrium.

5. to focus on the importance of considering species' phenology as a fundamental trait when collecting presence records to develop SDMs. Until now, modelling exercises mostly used presence data collected in multiple ways, including all occurrence points in a unique dataset. This approach may lead to indirectly neglect the ecological requirements of a species during a certain phenological phase when the latter is underrepresented within the species' presence dataset. We demonstrated that SDMs developed using different seasonal data separately, i.e. records of sites used by species for hibernation or reproduction, may predict only partial species' ecological niches. Then, we suggested a more valuable method for data collecting aimed to obtain a dataset in which seasonal records are equally represented in order to develop SDMs which take into account for all the potential ecological requirements of the species during its complete life cycle and predict a more realistic species' geographical range.

Further research may be build up on our findings in order to predict future distribution patterns under climate or land use change. For instance, we might apply phenology- dependent SDMs to predict the impact of climate change on summer and winter ranges of migratory species and evaluate which seasonal range might be at risk or expand in the future.

This thesis has clear implications for conservation and management. Predicting distribution ranges by modelling may offer important guidance in the development of conservation plans for example by allowing more exhaustive gap analyses or by helping detect corridors, or low-suitability areas in need of restoration to improve connectivity as mentioned above. However, linkage between SDM science and practice is still weak. We strongly encourage to involve SDMs in real decision-making processes that will benefit from their technical input as this modelling approach have the potential to better bridge theory and practice, and to contribute to improve both scientific knowledge and conservation outcomes.

Acknowledgements

I would like to thank all the people who have contributed to achieve my goal. First of all, I thank my tutor Prof. Danilo Russo for believing in me and his support over my academic studies. Since I attended his first lectures I developed a high esteem for him and for the way he conveys to students his love for his research work. Finally I had the extraordinary opportunity to do my PhD under his supervision. Special thanks to Luciano Bosso who has been not only a very important reference person for my project but especially a friend whose support and advice have been essential countless times. I had the luck to be advised by Mirko Di Febbraro and learn from him much of in the modelling procedures I used. I thank him very much for the patience and availability in sharing the "secrets" of modelling. I also would like to thank Dr. Hugo Rebelo from the CIBIO research center in Lisbon who hosted me during my research stay in Portugal with remarkable kindness and availability. He provided a very precious contribution to my training and greatly helped to develop my project. Many other researchers and experts helped by providing data for my project, more than I could mention here – thanks to all of you. I am very grateful to my family for supporting and sharing every goal of my life providing strength and acting as a guiding light. Last but not least, I wish to thank Silvano who coped along with my family - being in fact part of it - with my nerve attacks in some difficult moments. I feel lucky because I also have to thank many friends who supported me and made my life easier – I am not mentioning you here, but you know how important you are!