BIO_SOS

Project Title:	**BIO_SOS Biodiversity Multisource Monitoring System: from Space TO Species**
Contract No:	FP7-SPA-2010-1-263435
Instrument:	
Thematic Priority:	
Start of project:	1 December 2010
Duration:	36 months

Deliverable No: D6.7

Report on Ecological Niche Modeling (ENM)

Due date of deliverable:	30th September 2012
Actual submission date:	4th October 2012
Version:	v1
Main Authors:	Emilio Padoa-Schioppa, Francesco Ficetola, Anna Bonardi (UNIMIB, P10), Sander (C.A.) Mücher. (ALTERRA, P4); Madhura Nipadkar, Harini Nagendra (ATREE, P5); Vincenzo Leronni, Paola Mairota (UNIBA, P8); Cristina Tarantino e Maria Adamo (CNR, P1)

Project ref. number	**263435**
Project title	**BIO_SOS: Biodiversity Multisource Monitoring System: from Space to Species**

Deliverable title	< Report on Ecological Niche Modeling (ENM)>
Deliverable number	<D 6.7>
Deliverable version	v1
Previous version(s)	v1
Contractual date of delivery	30 September 2012
Actual date of delivery	5 October 2012
Deliverable filename	BIO_SOS_D6.7_v1.pdf
Nature of deliverable	< Choose one of the following: R = Report, P = Prototype, D = Demonstrator, O = Other >
Dissemination level	< Choose one of the following: PU = Public, PP = Restricted to other programme participants (including the Commission Services), RE = Restricted to a group specified by the consortium (including the Commission Services), CO = Confidential, only for members of the consortium (including the Commission Services) >
Number of pages	54
Workpackage	WP6
Partner responsible	UNIMIB
Author(s)	Emilio Padoa-Schioppa, Francesco Ficetola, Anna Bonardi (UNIMIB, P10), Sander (C.A.) Mücher. (ALTERRA, P4); Madhura Nipadkar, Harini Nagendra (ATREE, P5); Vincenzo Leronni, Paola Mairota (UNIBA, P8); Cristina Tarantino e Maria Adamo (CNR, P1)
Editor	Emilio Padoa-Schioppa

COOPERATION

EC Project Officer	Florence Beroud

Abstract	The Deliverable analyses four case studies and builds different approaches of Ecological Niche Models. In all the cases habitat typology is the main variable that emerges as most significant in explaining the model of ecological niche. The results have both theoretical importance (for instance, the use of GHCs improves the prediction of the model for the distribution of some species in Alta Murgia Parco Nazionale) and practical relevance for stakeholders (for instance, the models enable generation of a risk map predicting areas of potential vulnerability to damage by wild boars in Alta Murgia). Presence-absence species distribution data on Alta Murgia were provided by the management Authority of the National Park (i.e., Ente Parco). BIO_SOS project would like to thank Ente Parco Nazionale dell' Alta Murgia.
Keywords	Ecological niche modelling, invasive species, conservation of biodiversity, General Habitat Category, birds, amphibians, wild boar, *Sus scrofa* , *Lantana camara*.

Signatures

Written by	Responsibility- Company	Date	Signature
Emilio Padoa Schioppa	Responsible of D6.7 (P10)	28th Sept.	
Verified by			
Harini Nagendra	WP6 Leader (P5)	1st October	
Approved by			
Palma Blonda	Coordinator (P1)	5th October	
Fifamè N. Koudogbo	QAP (P7)	4th October	

Table of Contents

1. Executive summary

This deliverable forms a component of Task 6.4, wherein the relationship between the GHCs and abundance and composition of some species, both animals and plants, are being investigated by ecological niche modelling. General Habitat Categories (GHCs) were selected in D2.1 as an appropriate surrogate biodiversity measure for habitat description and mapping. Building on this, D6.6 identified a set of *taxa* or functional groups of species that are most suitable to investigate the relationships with the GHCs, in order to relate this with anthropic pressure and landscape features, as well as (possibly) with other indicators detected from EO data and selected in D2.1.

Emerging from the analysis presented in D6.6, the main criteria for the selection of bioindicator taxa were (*i*) the possibility to compare niche models with and without GHCs; (*ii*) the usefulness for stakeholders; (*iii*) the possibility to carry out comparison among different sites; and (*iv*) the actual availability of the distributional data.

Deliverable 6.7 has been divided into seven sections:

- section 1 provides a short literature review of the concept of ecological niche and of ecological niche modelling;
- section 2 links D 6.7 to other BIO_SOS deliverables, and describes the selection of case studies;
- sections 3, 4, 5, and 6present results from case studies analyzed in spring and summer 2012; each case study description has been organized like a scientific paper (introduction + materials and data + results + discussion and conclusions)
- section 7 summarizes the general conclusions that can be deduced from the case studies,and indicates the options to implement and continue the analysis in future

We selected 4 study cases, each of them will allow us to match with one or more criteria selected in previous deliverable.

Case study n° 1 analyses the relationship between an invasive plant species and their habitat and spatial requirements in India (usefulness for stakeholders – invasive species).

Case study n° 2 investigates the habitat requirements for different species of conservation interest and will start a comparison between habitat maps without and with General Habitat Categories (usefulness for stakeholders - conservation, mapping with or without GHC).

Case study n° 3 analyses the possibility to predict wild boar damages in a mosaic of agricultural crops (usefulness for stakeholders – management).

Case study n° 4 compares different habitat maps and bird distribution in Netherlands (usefulness for stakeholders – conservation).

The results of our analysis demonstrate the usefulness and importance of Ecological Niche Models based on habitat and landscape features for further linking habitat maps derived from Earth Observation (EO) data with biodiversity distributions in the field. The habitat maps derived by EO systems play a crucial role in impacting the quality and explanatory power of ENMs. Our results are of special importance for several stakeholders (managers, NGO etc.) of the study areas, because from the results it is clear that ENMs can improve the possibility to find better way of management of species and habitat.

2. Ecological Niche Modeling (ENM)

2.1 Ecological niche: a short history

In ecology **niche** is a term that describes the way of life of a species. Following Krebs (2001) the niche "is the role or the "profession" of an organism in the environment; its activities in the community". Each species has a unique niche, and the ecological niche can describe how an organism (or a population or a species) can react to the distribution of resources and competitors.

The concept has been developed in ecology since 1917, and it was proposed by Joseph Grinnell (1917). Grinnell studied a bird species in California and showed quite clearly that changing the habitat condition will modify the fitness of this species. Grinnell was convinced that an ecological niche was a personal fingerprint of each species, and believed that it was not possible for two species with the same niche to exist in nature.

Grinnel's conclusions were confirmed by models of Lotka, Volterra and Gauss.

Ten years later Elton (in the book Animal Ecology, 1927) used the word "niche" as the role of a species within food chains.

Hutchinson (1957) proposed a model of ecological niche as an n-dimensional hyper-volume, where the dimensions are environmental conditions and the resources that define the requirements of an individual or a species to practice "its" way of life (figure 2.1, as an example). Hutchinson's framework provided an approach to build mathematical models of ecological niche. As explained in Woster (1994), the work of Hutchinson is one of the milestones of ecology between 40 and 60 years in the past.

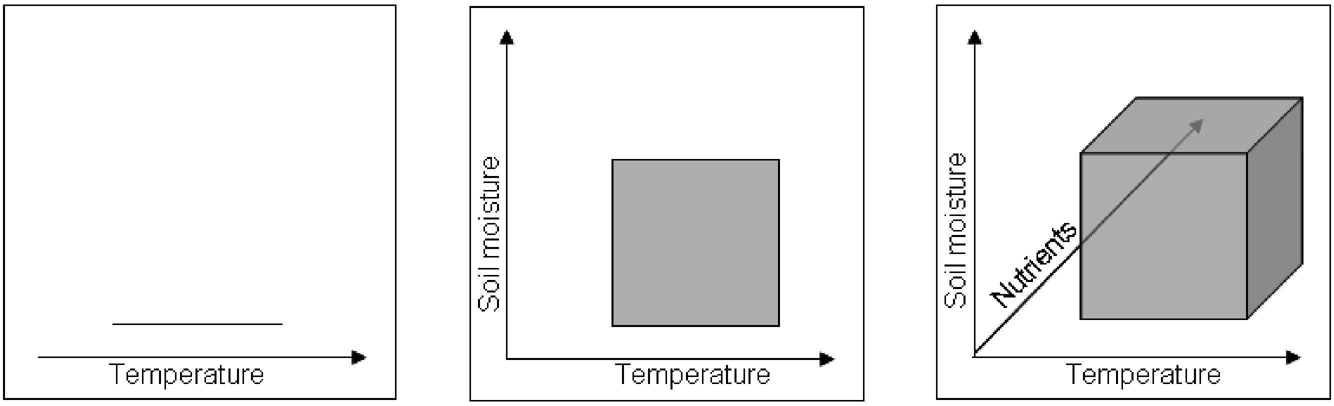


Figure 2.1: Multidimensional concept of ecological niche as proposed by Hutchinson. The example refers to a plant that has a defined requirement for temperature (x axis), degree of soil moisture (y axis) and nutrients (z axis). Adding all other factors (*e.g.* competitors, light, soil chemistry …) will realize the hyper-volume suggested by Hutchinson

Following the concept of Hutchinson, for each environmental factor it is possible to individuate a tolerance range and an optimal range. The development of this concept leads to a distinction between potential niche (or vital or physiological) and realized niche (or ecological): the potential niche is the space that a species would occupy if there were no other limiting factors and/or competitors, while the realized niche is one that is actually occupied, in the presence of limiting factors and competitors. The principle is exposed theoretically in figure 2.2. For example in nature several tropical orchids grow in the

shade, while in a greenhouse can also grow in the light. The critical limiting factor is the temperature: above a certain temperature - these orchids do not grow in tropical areas, and low temperatures occur only in shadow, while in the greenhouses you can keep the temperature in the light and grow orchids. Grace & Wetzel (1981) show the effects of competition between two species: *Typha latifolia* and *Typha angustifolia*. When each species grows in the absence of other, this shows that both species may grow in shallow water, and that only the second species (*T.a*.) can expand into deeper water by 80 cm. When the two species grow together their distribution changes: *T.l.* predominates and can be found where the water is shallower. Because *T.l*. subtracts light, space and nutrients to *T.a.* the result is that *T.l.* limits distribution of *T.a.* only in those spaces where it *T.i* cannot live: i.e. waters deeper than 80 cm.

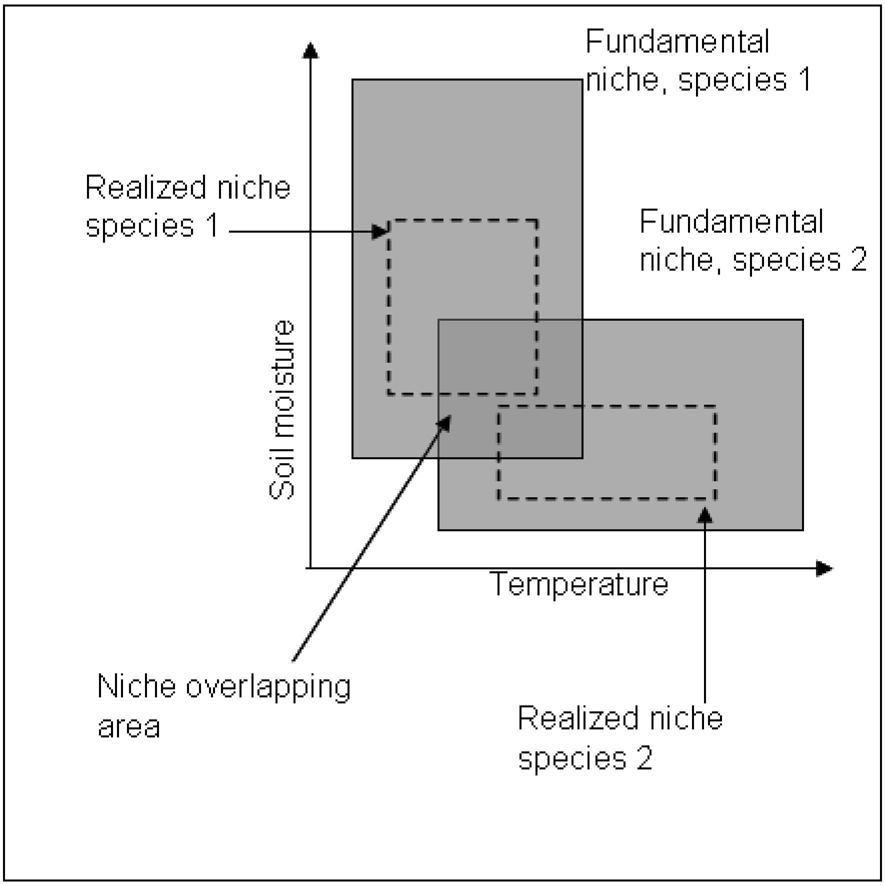


Figure 2.2: Fundamental and realized niches for 2 species

2.2 Ecological Niche Modelling

Ecological Niche Models (ENMs) are methods aiming to identify areas suitable for species survival. ENMs are correlative models based on the environmental features of the sites of presence - and, more rarely, absence of the species (Soberon and Nakamura, 2009). These methods are becoming the tool of choice to identify species' ecological requirements, to assess the impact of environmental changes (habitat, climate, landscape etc.) and to predict how species will respond to changes (Guisan and Thuiller, 2005; Elith and Leathwick, 2009). In practice, ENMs assess the probability that areas with given environmental features are within the species' niche (Godsoe, 2010), and therefore define the environmental requirements of the species.

Hence these models are based on the "ecological niche concept", either realized or potential niche, depending on the statistical methods (Jiménez-Valverde et al., 2008; Soberon and Nakamura, 2009; Godsoe, 2010). Fundamental niche represents all the combinations of environmental variables (including combinations that do not exist in nature) allowing intrinsic positive fitness of the species. Potential niche is constituted by all the parts of the fundamental niche that are available in the real world, in a given region and time. Finally, realized niche represents the part of the potential niche actually used by a species, after taking into account the effect of biotic interactions (e.g. competition and predation) as described in Table 2.1.

Soberon and Nakamura (2009) used a heuristic scheme, the BAM diagram (Figure 2.3), to show how three major causal factors can interact among them in determining the distribution of a species. Their work provides a simple framework to describe the relationship between different niche concepts and different kinds of species distribution, which is useful to better interpret the results of different types of ENMs. The major factors determining species distribution are: environmental conditions, biotic interactions, and movements / dispersal limitation.

In particular, making a simplification, the diagram highlights the fundamental ecological factors allowing us to distinguish between Eltonian and Grinnellian niches. The Eltonian niche represents the requirement for scenopoetic variables, i.e., those on which the species do not have any impact. Scenopoetic variables are mostly abiotic environmental features, such as climatic and topographic parameters. Scenopoetic variables are particularly important at large geographic scales (e.g., country, continental or global scale). The Eltonian niche is represented by the circle A of the BAM diagram. The Grinnelian niche describes the requirement for variables that can be consumed or modified by the species (circle B, biotic interactions). The Grinnelian niche is important to identify areas suitable for species distribution at fine spatial scales (e.g., regional or local). Circle A can represent either the fundamental or the potential niche. Intersection between circle A and B represents the portion of the geographic space suitable for the survival and growth of the species. Circle M represents the area that the species has been able to reach, in a given period of time (dispersal limitation). Areas outside Circle M refer to the places that the species cannot reach, on the basis of their biogeographic history or individual dispersal abilities. The intersection of the three circles (A, B and M) provides G_0, the actual distribution area of the species, i.e. the realized niche. In this framework, G_1 represents the area having suitable biotic and abiotic conditions that may be potentially invaded if the M condition-complex changes (Figure 2.3).

ENMs cannot provide exhaustive estimates of fundamental niche, nevertheless some measures of the fundamental niche may be obtained by physiological data of a species (Kearney and Porter, 2009). On the other hand, ENMs can provide useful estimates of realized and potential niche. ENMs can be built on the basis of both presence and absence data, or using presence-only data. The interpretation of models based on presence-absence data, or based on presence-only data, can be different (Soberon and Nakamura, 2009). This issue can be well illustrated by the help of BAM diagram. In particular, availability of unbiased absence data allows correlative models to directly estimate the area where the species actually occurs, i.e. the realized niche (Jiménez-Valverde et al., 2008), the area G_0 of the BAM

diagram. These model are considered the Species Distribution Model *sensu stricto* (Soberon and Nakamura, 2009; Soberon, 2010). Presence-only ENMs, on the contrary, provide results to be interpreted, in practice, as areas potentially suitable for the species, encompassed between G_0 and A of BAM diagram. In addition, a more accurate estimate of G_0 can be obtained including into the model ancillary information about biotic interactions (circle B) and dispersal constraints (M) of the specie in the region considered. Finally, explicit reference to major ecological or biogeographic factors affecting the distribution of the target species becomes essential to interpret at best ENM outputs (Soberon, 2010).

Table 2.1: Niche concepts and application. Based on Soberon & Nakamura (2009).

Fundamental niche	All combinations of environmental variables (including combinations that do not exist in nature) allowing intrinsic positive fitness. In most of cases, it is almost impossible to assess the fundamental niche using correlative ecological models. However, physiological data may be used to obtain some measures of the fundamental niche (Kearney and Porter, 2009).
Potential niche	All parts of the fundamental niche available in the real world, in a given region and time. May be approximated by some presence-only correlative models (Soberon and Nakamura, 2009).
Realized niche	The part of potential niche that a species actually uses. Can be described by complex presence-absence correlative models (Jiménez-Valverde et al., 2008).
Grinnellian niche	Species presence is based on abiotic, scenopoetic variables (temperature, precipitation…), more important at large scale (e.g., species range).
Eltonian niche	Species presence is based on biotic interactions, more important at small scale (e.g., landscape).

Multivariate regression methods, such as logistic regression or generalized additive models, can be employed to estimate realized niche when unbiased absence data are available. Recently, more complex methods are emerging for their ability in the analysis of presence-only data, such as MaxEnt, Boosted Regression Trees (BRT) and BIOCLIM (Busby, 1991; Elith et al., 2008; Elith et al., 2011). These approaches are among the best performing tools, and are increasingly becoming the standard of ecological niche modeling e.g., (Elith et al., 2006).

ENMs can be built at multiple spatial scales ranging from local to global, and can describe multiple biological parameters such as the activity of individuals, the response to environmental modification and the species' distribution. Continental and landscape are among the most frequently used scales at which ENMs are built. Large scale models refer to the distributional range of a species and can be also named "biogeographic models". They usually relate species presence to climatic variables that are expected to define the physiological tolerance of organisms to abiotic conditions (e.g., temperature, energy and water availability…, above mentioned scenopoetic variables, circle A in figure 2.3). These models are for instance used to assess the response of a species to climate change (Guisan and Thuiller, 2005; Elith and Leathwick, 2009). Models at the landscape scale usually relate the presence of the species to biotic variables (habitat quality, habitat amount, fragmentation, alien species, competitors, circle B in figure 2.3) to identify suitable areas. At such a finer scale habitat plays a more important role in explaining the species presence. These models can use less coarse descriptors of habitat features than the former, so they can be used to assess the species’ response to threats like habitat loss / degradation (Rushton et

al., 2004; Guisan and Thuiller, 2005). Models at these two scales are complementary and provide essential information. They are based on the two different niche concepts, Grinnellian and Eltonian, above defined (Table 2.3), suggesting that factors A and B of BAM diagram probably have contrasting spatial structures (Soberon and Nakamura, 2009). Therefore, models at the two reported scales allow us to assess the impact of different stressors (e.g., climate change vs. habitat alteration) on organisms. In particular, habitat suitability is pivotal for species conservation. Knowing relationships between species and habitat is a major task to plan suitable management policies, that often are carried out at small scales (i.e., regional or finer). There is therefore a direct link between landscape models for the species and the management for their conservation.

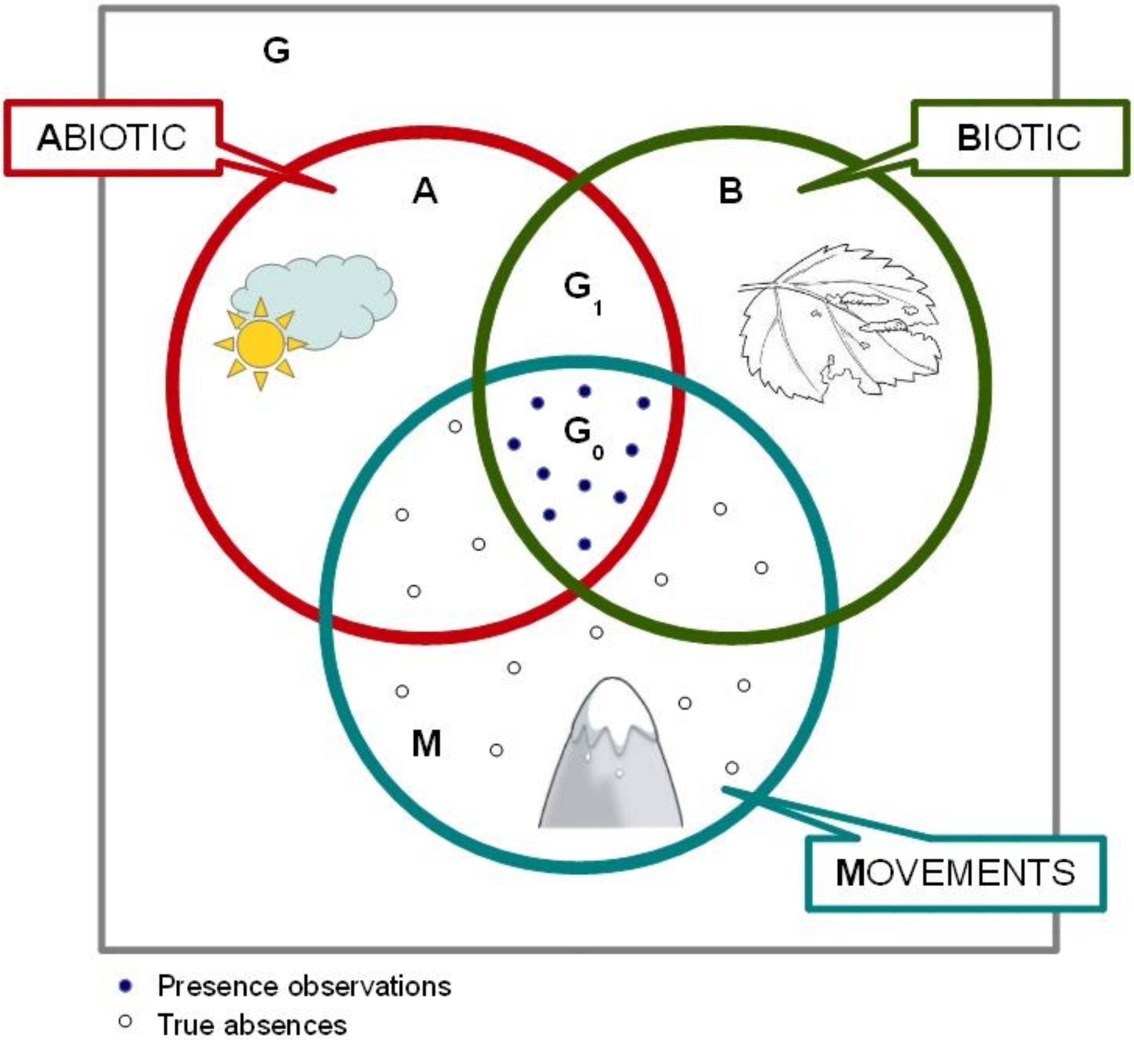


Figure 2.3: BAM diagram, based on Soberon & Nakamura {, 2009 #1031}. G: geographical region under study; A and B: areas where variable scenopoetic and bionomic, respectively, are suitable for the species; M: area that the species has been able to reach in a given time period; G_0: actual area of distribution of the species; G_1: region, having suitable biotic and abiotic conditions, that may be potentially invaded if M conditions change.

3. Ecological Niche Modelling in BIO_SOS project

3.1 ENM and habitat maps in BIO_SOS

The main aim of the BIO_SOS project is the development of an operational ecological modelling system suitable for effective and timely multi-annual monitoring of Natura 2000 sites and their surroundings in areas exposed to different and combined type of pressures. The project will:

1. adopt and develop novel pre-operational automatic High spatial Resolution (HR), Very High spatial Resolution (VHR) and hyper-spectral resolution Earth Observation (EO) data pre-processing and understanding techniques for Land Cover (LC) map and LC Change (LCC) map generation eligible for use in biodiversity monitoring. This is tantamount to saying that BIO_SOS is expected to provide improved pre-operational core service products with respect to state-of-the-art satellite-based LC and LCC mapping systems.
2. develop a modelling framework (scenario analysis) to combine EO and in-situ data in support to the automatic provision of biodiversity indicators and provide a deeper understanding, assessment and prediction of the impacts that human induced pressures may have on biodiversity. This means BIO_SOS aims at developing and integrating new and existing models able to evaluate and predict trends in biodiversity issues. This will lead to the development of new downstream services production.

Maps of habitat are essential to develop new ecological niche models (see paragraph 2.2). For this reason deliverable 6.7 is focused on ecological niche modelling based on map of habitats.

General Habitat Categories (GHCs; Bunce et al. 2008 and 2011) were selected as a surrogate measure of biodiversity (D2.1). GHCs are gathered from Remote Sensing (RS) EO data, therefore they represent the link between detailed site-based level measures and habitat assessments from remote sensing.

Nevertheless, there is a growing need to assess the relationship between the indicators and the entities they are assumed to indicate, e.g. biodiversity (Wiens et al., 2008; Cushman et al., 2010; Lindenmayer and Likens, 2011).

In task D6.7 we will investigate the relationship between the GHCs and the abundance and composition of indicator taxa, both animals and plants.

3.2 Conclusions of D 6.6

The main result of deliverable 6.6 was the selection of criteria to evaluate bioindicators to build ecological niche models for deliverable 6.7..

In D 6.6 we highlighted several criteria to select species that can be used as bioindicators for ENM. In particular 4 criteria are important.

- possibility to compare niche models with and without GHCs
- usefulness for stakeholders
- possibility to carry out comparison among different sites
- availability of data

Comparison of Niche Models with / without GHCs. The possibility to compare niche models with and without GHCs will depend mostly on the availability of different maps with and without GHCs.

Usefulness for stakeholders. One of the aims of this project is to have a link with stakeholders and both in the previous pages and in D2.1 usefulness for stakeholders was underlined as a key criterion for the selection of bioindicators.

According to SEBI report (EEA, 2007), in D2.1 the key criteria to select an indicator were:

- Policy relevant and meaningful
- Biodiversity relevant
- Progress towards 2010 target
- Well founded methodology
- Acceptance and intelligibility
- Routinely collected data
- Cause effect relationship
- Spatial coverage
- Temporal trend
- Country comparison
- Sensitivity towards changes

Invasive alien species are one other group of species that can be useful for stakeholders. In this case we get an indicator of pressure over the environment.

Comparison among different sites. If the list of species (not yet available, see later) shows some overlap on species distribution among different sites, we will choose some of those species (assuming they have the same quality of data).

Availability of data. The real availability of data is an essential condition. We operate our selection choosing among the existing data set described in D4.1 and reported in Appendix 2. In particular we will mostly focus on use data that does not require any further collection.

3.3 Selection of bioindicators and case study

In table 3.1 we summarize the data on biodiversity that were available from pre-existing data sets (criteria n° 4 of D 6.6) and the other criteria listed in deliverable 6.6.

At September 2012 we are able to present 4 different case studies, 2 of them in Italian site "Alta Murgia", one in Indian site and the last one in Netherlands site.

The case studies apply to different species, different processes (i.e. analysis on invasive species in India, study on a game species in Italy, species of conservation concern in Netherlands and in Italy).

There are also some data that we were not able to analyse before the conclusion of deliverable; because of recent collection/provision of data or gaps in the data set (in particular GHC maps still need further implementation). Those data will be added later as supplementary or integrative materials to the deliverable, although in the table 3.1 we consider those case studies ("Integration 1" and "Integration 2").

Table 3.1: Test sites and case study (Cr1 = presence of GHC maps, Cr2 = usefulness for stakeholders, Cr3 = comparison among different test sites, the 4th criterion –availability of biodiversity data is described in column biodiversity data)

BIO_SOS CODE	Biodiversity data	Cr1	Cr2	Cr 3	Comment	Case
IN 1	Plant: *Lantana camara*	No	Yes. Invasive species			Case study 1
IT3	Yes	Yes preliminary version				
	Amphibians	Not suitable for the whole area	Yes: species of conservation values			Case study 2
	Wild boar	Not suitable for this analysis	Yes: game species that require management 8and has an economic impact)		Suitable in a partial area	Case study 3
IT4		Yes				
	Plants	Yes	Yes: species of conservation values			Integration 1 (November 2012)
	Amphibians	Yes	Yes: species of conservation values		Not enough data to build a strong model	
NL	Birds	Yes, preliminary version	Yes: species of conservation values		Preliminay version will be implemented in the next months	Case study 4 (and Integration 2, December 2012)

4. Case study 1 – Logistic Regression modeling for invasive species Lantana camara in Biligiri Rangaswamy Temple Wildlife Sanctuary

4.1 Introduction

Invasive species are those species that are not native to the region in which they establish themselves and grow profusely to cover large tracts in foreign lands. These species are proving to be a problem worldwide as they cause harm to ecosystems as a whole as well as to species in particular (Simberloff, 2006). For invasive species to establish themselves important drivers for adaptation and establishment usually are favourable climatic factors, availability of resources such as light, water, and soil nutrients as well as ecological factors such as propagule pressure (Lockwood et al, 2005) and surrounding habitat types that are conducive to their growth. For wildland managers, it is important that some management actions are put into place for controlling the spread of the invasive species. Amongst many management activities, monitoring and mapping of the occurrence of invasive species is important for control action (Lindenmayer, 2010). Similarly, for deployment of resources and planning ahead, predictive modelling of possible areas of invasion would greatly benefit management.

In India, invasive species have been known to occur in several areas over the whole country, and, being a tropical country with warm climate, the subcontinent plays host to hundreds of different species, especially from the Central American region (Khurroo et al, 2011). Recent inventories of Indian alien flora have identified more than one-third of these to have their origin in South America (Khurroo et al, 2011). Among those particularly dominant are species of the families Verbenaceae (*Lantana camara*), Convolvulaceae (*Ipomoea* spp.) and Asteraceae (*Mikania micrantha, Chromolaena odorata*). Invasive alien plants have established themselves in several types of habitats and in different climatic regimes across India (Khuroo et al, 2011). Their aggressive nature and adaptability has made it easy for them to grow and spread rapidly across regions. The Western Ghats biodiversity hotspot region is particularly vulnerable to these kinds of species due to its similarity with the Central American landscape.

The Biligiri Rangaswamy Temple Wildlife Sanctuary is an area of high biodiversity lying in the Western Ghat hill ranges of the southern Indian peninsula. This 540 km^2 sanctuary is unique in several ways - its heterogeneity of physiographic forms, the climatic regime it experiences, as well as the flora and fauna it supports. The vegetation of the sanctuary has been classified into ten different types from dry scrub thickets to dense wet evergreen forests at high elevations (Ramesh, 1989). In the past decade, studies have reported the occurrence (Murali and Setty, 2001), and subsequently a ten-fold increase (Sundaram, 2011), of invasive alien plant species, with special emphasis on *L. camara*. These studies provide important baseline information about invasion of *L camara* in this landscape indicating increase not only in numbers of stems but also density (Sundaram, 2011). While these studies have identified the potential threat to the local ecosystems, no detailed study has undertaken work on understanding the underlying mechanisms which may be encouraging the growth and spread of *L. camara*. For the management of the park and its inhabitants, it is imperative that the managers develop a fair understanding of the vulnerable areas where this species is more likely to spread. Using the existing knowledge regarding the species occurrence, it is possible to model the possible future occurrences or areas where factors are more conducive to its spread. Thus, this study was undertaken to understand the relationship between the presence of *L. camara* with its environmental and ecological parameters.

4.2 Material and methods

4.2.1 Study area

The Biligiri Rangaswamy Temple Wildlife Sanctuary is an area of high biodiversity nestled between the Western and Eastern Ghat hill ranges of the southern Indian peninsula. A part of the Western Ghats biodiversity hotspot in the state of Karnataka, India, this 540 km^2 sanctuary is unique in several ways - its heterogeneity of physiographic forms, the climatic regime it experiences because of its location and physiography, as well as the flora and fauna it supports. The sanctuary has a hilly terrain varying between 600m to 1800 m above sea level and receives rainfall from both the Southwest and Northeast monsoon winds, with an annual average ranging between 898 – 1750 mm depending on location within the sanctuary. The complexity and diversity in the vegetation of BRT is a function of the spatial variability in topography and climate, along with human activities including fire. The vegetation of the sanctuary has been classified into ten different types (see Table 1) from dry scrub thickets to dense wet evergreen forests and shola-grassland mosaics at high elevations. The area is rich in biodiversity, with at least 1400 species of higher plants, and 254 species of birds. The BRT landscape has 27 species of mammals, including large mammalian herbivores such as the Asian elephant (*Elephas maximus*), gaur (*Bos gaurus*), and carnivores such as the tiger (*Panthera tigris*), leopard (*Panthera pardus*), and dhole (*Cuon alpinus*). The sanctuary has recently been declared a Tiger Reserve, in light of the higher protection it needs for preserving tiger habitat. Though the problem of invasive alien shrub Lantana camara was reported in the past, some recent studies have highlighted an almost ten-fold increase in its density and abundance, extensively affecting the local ecosystem dynamics for the flora as well as fauna.

4.2.2 Data

Field data collected for 124 plots distributed uniformly in a 2 km by 2 km grid was used for this exercise (Figure 4.1).

Plots laid were 5 m by 80 m and all woody stems > 1 cm diameter were identified, counted and measured. Thus, presence or absence of *L. camara* was obtained for use as the dependent variable for the modelling exercise. Independent climatic variables used were obtained from the global BIOCLIM dataset version 1.4 (http://www.worldclim.org/bioclim). Of the several climatic variables that are available in this dataset, only two relevant variables - the mean diurnal range (bio2) and annual precipitation (bio12) - were used. A digital elevation model at 30 m resolution was used for obtaining elevation data at each plot location. A vegetation type map developed earlier (1997), with 9 landcover types was used as the base landcover type map. This cover type map has 14 types, but was reclassified into 9 classes for ease of computation and reduction in redundancy. For each plot, polygons of the size of the sample plot were demarcated on the landcover type map, and percentage of area covered by the cover type was noted. To account for land cover type correlations and to test if cover types were producing a spatial influence, we generated three types of buffers around these plot polygons at 25 m, 50 m and 100 m from the plot locations. These buffer polygons were then clipped to the vegetation type map and percentages were recorded for each of the cover types that were included within each of the buffer polygons (Figure 4.2). Thus, for each plot location, we had data on cover type percentages within the plot as well as three buffer polygons indicating progressive distances away from the plot location. A few cover type variables were found to have zero values for all plots, which were subsequently removed. To improve the fit of the models, we then calculated the Arcsin transformation of each of these percentage covertype values. At the end of this data preparation exercise, we had a total of 25 independent variables to be incorporated in the models.

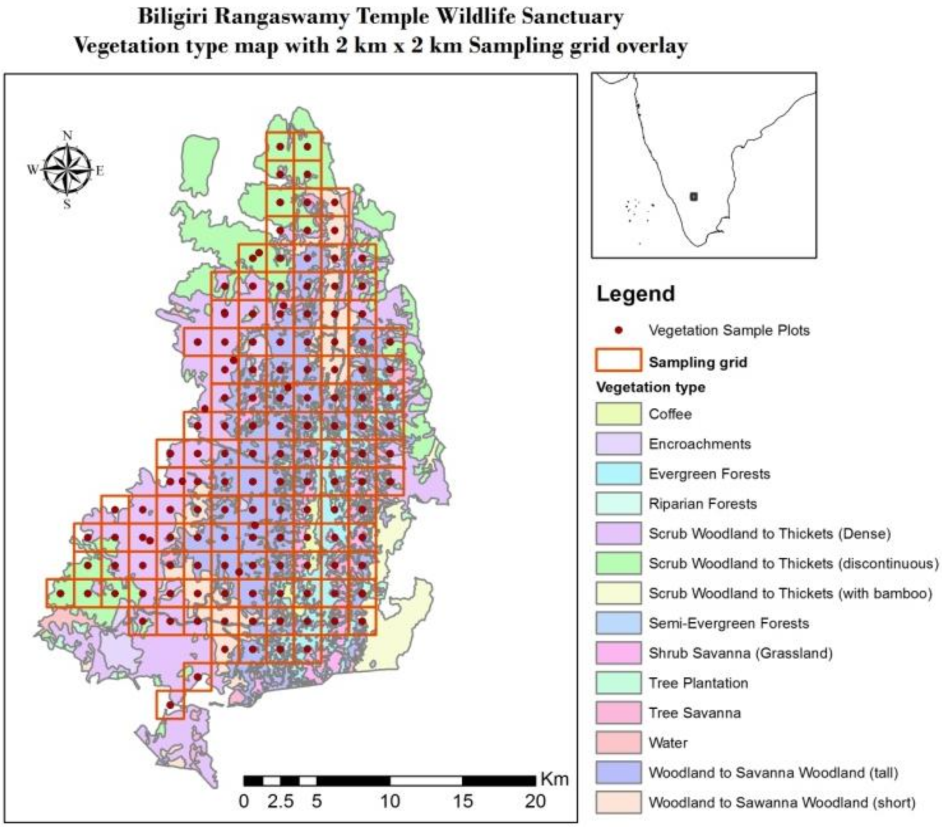


Figure 4.1: Study site IN1 with the sampling scheme overlaid on vegetation type map.

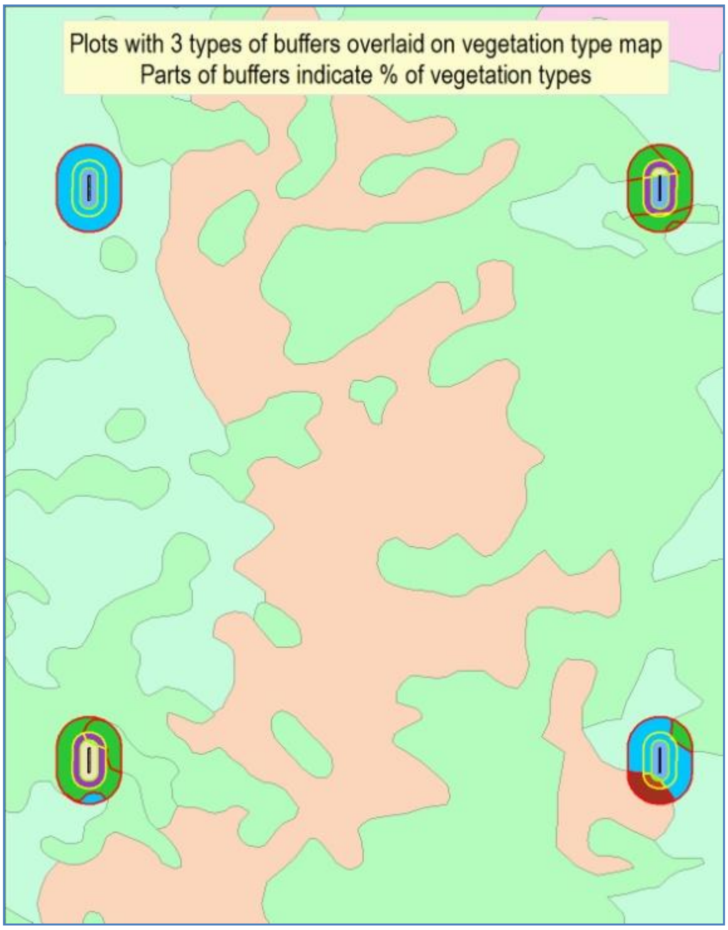


Figure 4.2: Buffer polygons of 3 different distances around plots overlaid onto the vegetation type map.

4.2.3 Methods

The statistical software package R was used for the modelling exercise. All variables described above were combined and generalized linear models were developed for the distribution of *L. camara* using Presence/Absence data as the dependent variable. Since the Arcsin transformation variables should give a better fit of the model, the Arcsin transformation of the data values was used. Individual models were then developed for each distance measure from the actual plot polygon location.

Next, we calculated the AICc values, where AICc is calculated as:

$$AICc = AIC + \frac{2k(k+1)}{(n-k-1)}$$

where n = sample size, and k = n^o of parameters in the model

AICc gives a correction on the AIC value for finite sample sizes; thus AICc is AIC with an additional penalty if the sample size is large.

An ANOVA was then calculated for the 50 m buffer distance model, which would indicate the independent variables most important for *L. camara* occurrence.

4.3 Results

Having tested the use of the independent variables and then using the Arcsin transformation, it was observed that the Arcsin transformation variables gave a better fit of the model. Hence, only models with

Arcsin transformation were considered for further comparison and analysis. A comparison of the AIC values for the models at different distances indicated that the model at 50 m buffer distance had the lowest value, and hence provided the best explanation of the *L. camara* distribution. Table 4.1 is a comparative table of the AIC values for each of these models.

Table 4.1. Comparative table of the AIC values for each of these models

Model	AIC value
Global (all variables)	131.3
Dist0	119.3
Dist25	114.6
Dist50	110.9
Dist100	118.8

A comparison of AICc values again indicated that the model at 50 m buffer distance was the best explanatory model for the *L. camara* distribution data. Table 4.2 indicates the calculated AICc values for all models.

Table 4.2. Calculated AICc values for all models

Model	AICc value
Global (all variables)	14.56531
Dist0	0.55217
Dist25	115.85217
Dist50	112.15217
Dist100	120.7469

Results of the ANOVA are indicated in Table 4.3. The Anova indicates that of the 8 variables included in the 50 buffer distance model, only 3 variables were significant. The climatic variables Bio2 (mean diurnal range of temperature) and Bio12 (annual precipitation) were significant at p=0.05, and the Asinsqrt1_50 (% of moist deciduous forest within 50 m buffer) was highly significant at p=0.001.

Table 4.3. Results of the ANOVA

Variables	Deviance	Pr(>Chi)
Altitude	0.8292	0.36250
Bio2 (Mean diurnal range of temperature)	3.9576	0.04666 *
Bio12 (Annual precipitation)	4.3387	0.03726 *
AsinSqrt1_50 (% Moist Deciduous Forest within 50 m of plot)	18.2216	1.966e-05 ***
AsinSqrt2_50 (%Dry Deciduous Forest within 50 m of plot)	0.2819	0.59543
AsinSqrt3_50 (%Scrub forest within 50 m of plot)	0.4625	0.49646
AsinSqrt4_50 (%Evergreen forest within 50 m of plot)	0.1184	0.73080
AsinSqrt5_50 (%Grassland/Shola forest within 50 m of plot)	0.7888	0.37446

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

4.4 Discussion & conclusion

Invasion by *L. camara* is a long-established (Bhagwat, et al. 2012) and well-recognized problem in the Indian subcontinent (Prasad, 2009; Ramaswami and Sukumar, 2011; Sundaram, 2011).

The above results indicate the importance of the Moist Deciduous forest type in the occurrence of *L. camara* along with the two climatic variables of temperature and rainfall. Field observations have indicated that the presence of *L. camara* is high in the Moist Deciduous type of forest (Figure 4.3), where the habitat is moister than the Dry deciduous forests and more open than the Semi-evergreen or Evergreen forest type.

Figure 4.3: *L.camara* in a patch of Moist deciduous forest

Although it is also fairly common in the Dry Deciduous forest, as well as the Scrub types of forest, the dominance and vigour is visibly higher in the Moist Deciduous type of forest than the shorter and thinner form in the Dry Deciduous forest type (Figure 4.4).

Figure 4.4: Lantana shrubs in the understorey of Dry Deciduous forest

The shrub is also seen growing with much vitality along riparian tracts in dry forest patches, where moisture plays a major role. This study has helped to highlight the importance of these parameters in the spread and possible areas of occurrence of this species. The use of buffer zones of different

distances around the plot areas helped to factor in effects of forest types much more than would have been possible with using only the forest types encountered at the sample plot locations.

These results also point to the fact that deeper studies need to be undertaken in different forest types to understand the role that they play in encouraging the establishment and occurrence of invasive species which occur in understorey habitats. It may also prove worthwhile to explore other climatic variables from the repertoire of the BIOCLIM dataset to test whether they have an impact on promoting the spread of invasive species.

5. Case study 2: Land-use and distribution of amphibians in the Alta Murgia national park

5.1 Introduction

In the last two decades, many studies pointed out a widespread decline of amphibians at global scale (Blaustein & Wake 1995, Stuart et al. 2004), with hundreds of amphibian species now facing concrete extinction risk (Stuart et al. 2004, Beebee & Griffiths 2005). Diseases, global warming, alien species and complex interactions among factors are often key causes of the decline (Blaustein & Wake 1995, Stuart et al. 2004, Beebee & Griffiths 2005). However, habitat changes caused by human activities probably are the most important driver of population decline worldwide (Gardner et al. 2007). An increasing number of studies evaluated the relationship between amphibian distribution and habitat across all the continents, to provide scientific information helping to set up conservation plans. These studies improved our knowledge of mechanisms influencing the patterns of species abundance and occurrence (Van Buskirk 2005), increased our understanding of the ecological processes that underlie the effects of habitat change (Gardner et al. 2007), and helped to individuate the factors determining variation in species distribution (Skelly et al. 1999, Dalbeck et al. 2007, Werner et al. 2007).

Traditionally, analyses of relationships between amphibians and habitat focused on breeding environments (i.e., pond features) more than on the features of the surrounding environment (Beebee 1981, 1985). Nevertheless, for most amphibians the terrestrial phase is longer than the aquatic phase. In the last decade studies are increasingly stressing importance of analyses performed at the landscape scale: the terrestrial habitats are necessary for the post-reproductive life of amphibians, and can also influence the features of aquatic breeding wetlands (Pope et al. 2000, Porej et al. 2004, Van Buskirk 2005, Denoël & Lehmann 2006, Piha et al. 2007, Zanini et al. 2008, Ficetola et al. 2009, Ficetola et al. 2011, Plaiasu et al. 2012). Nevertheless, our knowledge of the effect of environmental modification on amphibians is far to be complete, and more studies are required on the relationships between amphibian distribution and habitat, before generalizations can be performed (Gardner et al. 2007).

The aim of this analysis was evaluating the importance of the surrounding landscape for amphibians within the Alta Murgia National Park. We assessed whether the features of surrounding landscape are important for the presence / absence of amphibians in this area, and identified which land-use variables are more important for amphibians. These results can provide useful indications to management strategies aiming at the conservation of amphibian communities within the study area.

5.2 Methods

5.2.1 Species and landscape data

Amphibian distribution in the Alta Murgia National Park was assessed by repeated surveys in 2010-2011 over 63 wetlands comprised within the park or in the surrounding buffer zone. Potentially suitable wetlands were identified on the basis of bibliographic information and direct surveys. Each wetland was monitored repeatedly during the amphibian breeding season, and multiple techniques (visual encounter surveys, dip netting, audio point counts) were performed to ascertain species presence absence. Additional details on the study area and sampling methods are reported by Marcone et al. (2012)

Land use variables were extracted from the 1:5000 vector UDS land use map of the Puglia Region. In the original UDS, land use typologies are represented by 68 distinct categories. To avoid over-parametrization of models, the original categories were aggregated; for analyses, we considered eight

land use classes: forest, shrubland, cropland, vineyard, olive groves, woody cultures, herbaceous cultures and grassland. We used the ArcView GIS to measure the land cover in the landscape surrounding each wetland, on the basis of the UDS land-use map. Land cover was measured in radius of 500 m from each wetland (Gustafson et al. 2011).

5.2.2 Statistical analyses

We used an information theoretic approach, based on Akaike's Information Criterion corrected for small sample size (AICc) (Burnham & Anderson 2002), to identify the combination of variables best describing the distribution of amphibians detected in > 10 sites. We built generalized linear models (GLM) assuming binomial error, including all possible combinations of variables. For each model, we calculated AICc and Δ-AICc, which is the difference in AICc between a candidate and the model with lowest AICc (i.e., the best model). The use of AICc as sole selection criterion may select overly complex models therefore, as an additional criterion, we considered a complex model only if it had a Δ-AIC lower than the Δ-AIC of all its simpler nested models (Richards et al. 2011). Per each model *i*, we calculated the AIC weight w_i, which is the probability for a model to be the best one among the candidates (Richards et al. 2011). Models were compared among them and with the null model (i.e., the model non including any of the land-use variables). AICc values lower than the AICc of the null model indicate that landscape variables are likely to be important predictor of amphibian distribution. We also used a lilelihood ratio test to test the significance of variables. We performed analyses using the software R (R Development Core Team 2012); we extracted and compared the AICc values using the package MuMIn (Barton 2011).

5.3 Results and discussion

Five amphibians were detected in the 63 ponds: the Italian newt *Lissotriton italicus* (occurrence = 13 sites); the Italian crested newt *Triturus carnifex* (occurrence = 2 sites); the green toad *Bufo viridis* (occurrence = 14 sites); the common toad *Bufo bufo* (occurrence = 9 sites); the pool frogs *Phelophylax* synklepton *hispanicus* (occurrence = 32 sites). Therefore, analyses were performed on three species: Italian newt, green toad and pool frogs.

For all the study species analyzed, land cover variables were significantly related to the presence / absence in breeding wetlands (Table 5.1, Table 5.2). Furthermore, for all study species the models including landscape variables showed AICc values much lower than the null models. Actually, for all species the AICc weight of the null model was 0.01 or lower (Table 5.1). This indicates that the surrounding landscape is extremely important for the distribution of aquatic breeding amphibians: on the basis of AIC weights, the probability that the landscape is not important is 1% or less (Lukacs et al. 2007).

For the Italian newt, three variables were included in the best model: cover of shrublands, vineyards and woody cultures. The newt was associated with wetlands surrounded by shrubby landscapes, and with limited cover by woody cultures (mostly orchards in the study area) or vineyards, although the relationship with vineyards was not significant (Table 5.1, Table 5.2). All the other candidate models included the same variables (Table 5.1), confirming the robustness of the best model. Among the amphibians considered in this analysis, the Italian newt is the species with the highest conservation concern, as it is included in the Appendix IV of the EU habitat directive (directive 43/1992). As other species of small-bodied European newts, it is not associated to forested areas, still it takes advantage of relatively natural landscapes, like shrublands (Denoël & Lehmann 2006, Denoël & Ficetola 2007, 2008). The dependency of this species from shrublands may make it complex the conservation of this species, as shrublands often undergo encroachment, and ecological succession often transforms them into forests. Actually, populations of this newts are declining across the species' range (Arntzen et al. 2009), indicating that actions are needed for the conservation of this species in the long term: the preservation of landscape features needed for its terrestrial activity will certainly be an important step.

Table 5.1: Candidate generalized linear models for amphibians. K: number of parameters included in the model; AICc: Akaike's information Criterion, corrected for small sample size; ΔAICc: difference in AICc values with the best model; *w*: AICc weight of the model.

Species	Model rank	Variables	K	AICc	ΔAICc	*w*
Lissotriton italicus	1	Shrublands, vineyard, woody cultures	4	56.69	0.000	0.453
	2	Shrublands, woody cultures	3	57.11	0.417	0.368
	3	Vineyard, woody cultures	3	60.06	3.374	0.084
	4	Woody cultures	2	60.27	3.583	0.076
	5	Vineyards	2	63.36	6.671	0.016
	6	Null model	1	66.21	9.520	0.004
Bufo viridis	1	Herbaceous cultures, vineyards	3	60.25	0.536	0.838
	2	Vineyards	2	64.36	4.644	0.107
	3	Herbaceous cultures	2	66.18	6.463	0.043
	4	Null model	1	68.81	9.090	0.012
Pheophylax s. *hispanicus*	1	Crops	2	80.61	0	0.906
	2	Grassland, olive groves	3	87.83	7.228	0.024
	3	Grassland, woody cultures	3	88.41	7.804	0.018
	4	Grassland, vineyards	3	88.79	8.186	0.015
	5	Woody cultures	2	89.15	8.546	0.013
	6	Olive groves	2	89.22	8.616	0.012
	7	Null model	1	89.39	8.781	0.011

Conversely, the green toad was positively associated with vineyards and herbaceous cultures (Table 5.2). This toad is an opportunistic amphibian that can take advantage by the small water reservoirs associated to traditional agricultural activities. In this case, species conservation would require very different strategies, compared to the conservation of the Italian newt. For green toad, the preservation in the park would be favoured by the maintenance of traditional agriculture, of the small elements where the toad can find shelter (e.g., stone walls), and ideally by small, temporary wetlands in sunny areas (Bologna & Giacoma 2006). Also for the green toad, the best model was quite robust, as all the other candidate models included a similar set of variables.

Finally, the situation for pool frogs was more complex. Pool frogs are among the more aquatic amphibians. Nevertheless, also for this frog landscape features were extremely important, as the support of the null model, non-including any landscape variable, was very limited (Table 5.1). The best model indicated a strong, positive relationship with crops. Pool frogs are termophilous, and prefer sunny wetlands (Ficetola & De Bernardi 2004). Therefore small waterbodies associated to open agricultural fields are often occupied by this species. Pool frogs are the commonest amphibians within the park, and this can be easily explained by their ecological plasticity, and by the ability to survive in areas exploited by human activities.

Table 5.2: Variables included in the best models for the three species included into analyses.

Species	Estimate	SE	χ^2	d.f.	*P*
Lissotriton italicus					
Vineyards	-20.144	29.908	2.700	1	0.100
Woody cultures	-425.460	33653.273	10.980	1	0.001
Shrubland	22.471	13.688	5.657	1	0.017
Bufo viridis					
Vineyards	4.304	1.609	8.134	1	0.004
Herbaceous cultires	27.183	14.866	6.315	1	0.012
Pelophylax s. *hispanicus*					
Crops	2.050	0.677	10.915	1	0.001

5.4 Conclusion

This analysis suggests that landscape features, obtained by land-use maps, are extremely important to explain the distribution of amphibians. All the study species breed in water, and two of them (Italian newt and pool frog) require water for long periods. The importance of landscape variables for these species is therefore striking, and confirms the necessity to perform management over broad spatial scales. In other words, the conservation of breeding wetlands is not enough for these semiaquatic species, and wide areas with suitable landscape (e.g., shrublands for the newt, traditional agriculture for the green toad) need to be maintained (Gibbons 2003, Ficetola et al. 2009, Ficetola et al. 2011)

The present analyses used land-use maps obtained through traditional technologies. However, new technologies will allow to obtain maps better representing the amphibian terrestrial habitats. This will probably permit to better identify the landscape requirements for amphibians, and therefore obtain management indications that allow to better target the need of these threatened vertebrates.

5.4.1 Do General habitat Categories improve the performance of species distribution models?

For part of the Alta Murgia National Park, General Habitat Category maps have been recently developed (Figure 5.1). It was therefore possible to assess whether the GHC maps would help to improve habitat models. The GHC map currently covers an area of about 33 x 15 km, therefore it is not possible to include all the amphibian wetlands in the analyses (Figure 5.1). Furthermore, the analysis was limited to the pool frogs Pelophylax s. hispanicus, which is the species with more presence data in the area.

We used an information-theoretic approach to assess whether the GHC maps allow to improve the performance of species distribution models (SDMs). First, we used the procedure detailed above (paragraph 5.2.2) to identify the best AICc model relating the presence / absence of pool frogs to:

- GHC categories
- Land use categories, obtained from the UDS map.

The best models identified through the two sets of land use were then compared as explained before, using AIC.

5.4.2 GHC versus traditional maps: results and discussion

21 amphibian breeding wetlands were located within the area covered by GHC. Pool frogs were present in 9 out of these wetlands (43%).

For the GHC maps, the best-AIC model suggested a positive relationship between pool frogs and the GHC category CUL_WOC (i.e., woody cultures). This model was similar to models 2-5 built for pool frogs for the whole park (see Table 5.1, pool frog models). For the UDS map, the best-AIC model for this subset of wetlands confirmed the positive relationship with crops (Table 5.1, Table 5.3).

The best GHC model showed an AICc value lower than the best UDS model (Table 5.3). The difference between the two models was limited (0.7 AICc units). Still, on the basis of AICc weight, the support of the GHC model was 1.4 times higher than the support of the UDS model. The limited differences between the models probably arise because of the limited number of wetlands falling within the area currently covered by GHCs. Actually, the number of occupied wetlands (nine) was below the threshold of 10 occupied sites used in the analysis of the whole park.

Nevertheless, these preliminary analyses suggest that GHC provide a better description of the terrestrial habitat, compared with the traditional land use maps. The expansion of the area covered by GHC will certainly allow to perform more accurate comparisons of the performance of models, and to obtain better models describing species distribution that can be used for the management of the park.

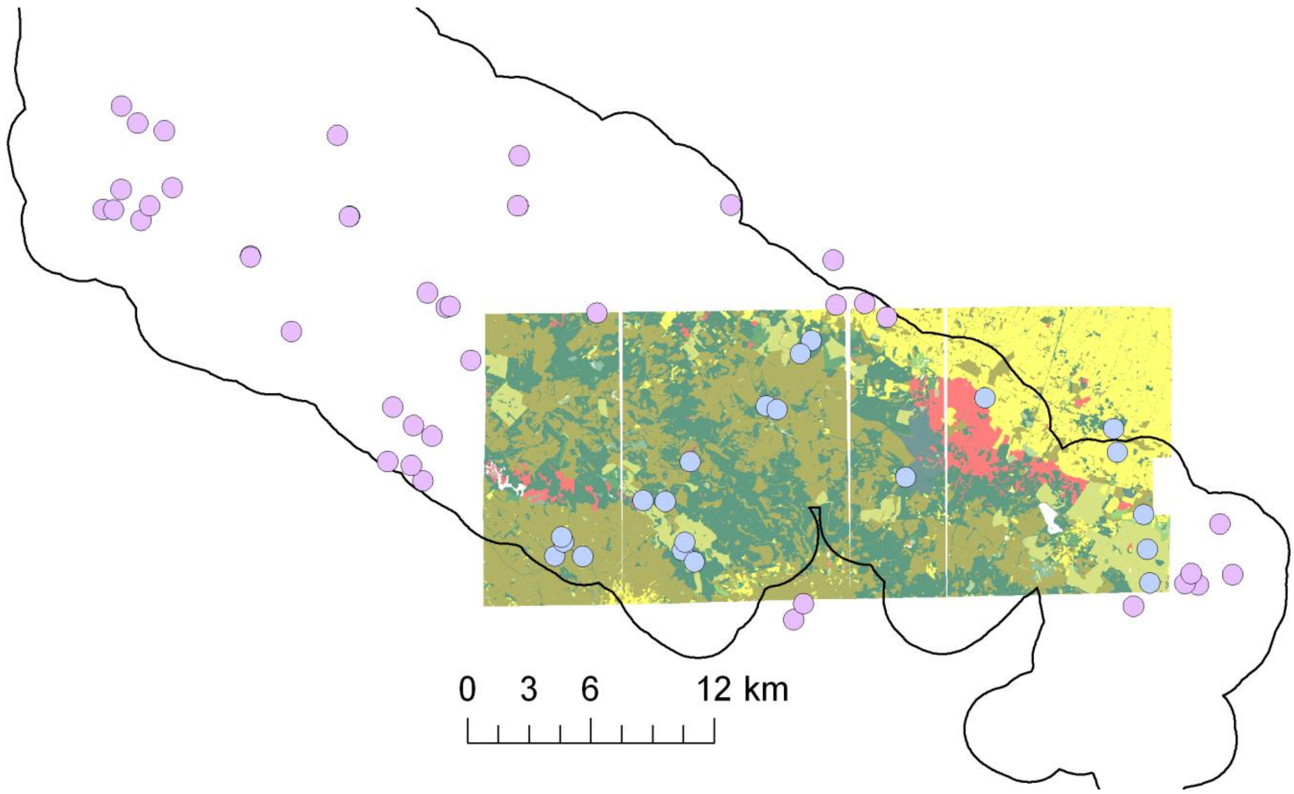


Figure 5.1: Distribution of the amphibian breeding localities within the Alta Murgia National Park. Blue dots are amphibian breeding wetlands for which GHC data are available, pink dots represent wetlands for which GHC data were not available. The black line represents the borders of the park, including a 2-km buffer zone. The map also shows the area for which the GHC map is available.

Table 5.3: Comparison of best models including at least one land use variable, obtained using GHC and UDS maps. The best model is in bold. K: number of parameters included in the model; AICc: Akaike's information Criterion, corrected for small sample size; ΔAICc: difference in AICc values with the best model; *w*: AICc weight of the model.

Source	Variables	K	AICc	ΔAICc	*w*
GHC	**Woody coltures**	**1**	**30.7**	**0.000**	**0.59**
UDS	Crops	1	31.5	0.7	0.41

6. Case study 3: Predictions of wild boar damages to croplands in a mosaic of agricultural and natural areas

6.1 Introduction

Crop damage by wildlife is a frequent form of human-wildlife conflict, particularly within or nearby protected areas, and can have an important impact on the functioning of protected areas. On the one hand, damages can make communities intolerant and antagonistic towards wildlife, and this may limit the effectiveness of conservation strategies (Naughton-Treves 1998, Linkie et al. 2007). On the other hand, to limit human-wildlife conflicts, properties suffering wildlife damages are often financially compensated, particularly within or nearby protected areas (e.g. Cozza et al. 1996, Schley et al. 2008, Hegel et al. 2009). However, compensation may have negative consequences in the long term (Bulte & Rondeau 2005); furthermore, as funding available for protected areas is limited, the additional cost of compensation may reduce resources available for other management activities (e.g., conservation of threatened species or habitats). The mitigation of human-wildlife conflict can take advantage from the identification of factors determining crop damages, and from the development of predictive model that identify the areas that are most at risk of damages. These information can allow to set up strategies to prevent damages (Linkie et al. 2007, Hegel et al. 2009, Lavelle et al. 2011, Schlageter & Haag-Wackernagel 2012), or to target culling activities toward specific areas, improving their efficacy (Honda & Kawauchi 2011).

Wild boar populations have dramatically increased all over Europe in the past decades, because of the increase of forest cover and repeated releases for hunting. The wild boar expansion determines increasing activity into agricultural land, with intensification of damages and conflicts with humans (Toso & Pedrotti 2001, Thurfjell et al. 2009, Schlageter & Haag-Wackernagel 2012). The identification of factors increasing the risk of damages can provide important information for the control of population and the limitation of damages (e.g., fences, odours…) (Schley et al. 2008, Honda & Kawauchi 2011, Schlageter & Haag-Wackernagel 2012). Wild boar damages depend not only from the abundance of wild boars, but also from topographical and landscape features (Cocca et al. 2007, Schley et al. 2008, Thurfjell et al. 2009, Honda & Kawauchi 2011). Spatially explicit species distribution models (SDM) allow to identify relationships between species and environmental features measured at multiple spatial scales, and on the basis of a variety of environmental features (e.g., climate, landscape features…). SDM are usually used to identify the areas with the highest likelihood for the presence of a species (Elith & Leathwick 2009, Jiménez-Valverde et al. 2011). In addition, recent studies showed that SDM can accurately predict other important parameters, such as reproductive success (Ficetola et al. 2009, Brambilla & Ficetola 2012) and risk of damages by target species (Honda & Kawauchi 2011). The validation of SDM predictions can be challenging, and requires the use of independent data collected in different areas or time (Nogués-Bravo 2009, Jiménez-Valverde et al. 2011). However, these data are rarely available and used for an accurate validation of model predictions.

In this study, we analysed wild boar damages in the area of the Alta Murgia National Park in Southern Italy, and related them to landscape features. We used SDM to identify areas where the risk of wild boar damages is highest. Furthermore, we use data on wild boar damages collected over five years to evaluate the capability of our model to accurately predict damages in the future.

6.2 Methods

6.2.1 Data

The study area corresponds to the Alta Murgia National Park (AMNP) (surface; 680 km^2) and a 2 km buffer surrounding the park. The AMNP is located in Southern Italy and is partly coincident with the Natura 2000 “Alta Murgia” IT9120007 SCI/SPA (Figure 6.1). This Natura 2000 site, spanning over ~1,259 km^2 and a 300 m to 679 m a.s.l. altitudinal range is one of the study sites of the BIO_SOS project. It is characterized by a typical dry karst landscape, and by a Mediterranean with a pluviseasonal-oceanic Mediterranean bioclimate; the ombrotypes range from dry to sub humid and thermotype is mesomediterranean (Forte et al. 2005). The sub natural vegetation type mostly represented and of greater conservation concern in this site are Mediterranean steppe grasslands. Other sub- and semi- natural vegetation types include garrigues and shrubs, Macedonian oak (*Quercus trojana*) and Downy Oak (*Q. pubescens*) woodlands, and pine plantations. These vegetation are embedded within an agricultural matrix mostly dominated by cereal crops which have substituted large chunks of the grasslands mainly over the last century, and to lesser extent comprised of olive groves, almond orchards and vineyards (Dimopoulos et al. 2011). The AMNP with the surrounding buffer zone was chosen instead of the whole Natura 2000 site as damage records were only available for the former.

Records of wild boar damages to agriculture were provided by the AMNP Authority for 2007 (20 records), 2008 (23), 2009 (38), 2010 (29) and 2011 (47). These records were referred to cadastral parcels. Within the study area, the average surface of parcels is 5.1 ha, so we used a grid with 220 x 220 m cells (i.e., cell size roughly corresponding to the average surface of parcels)

Land use variables were extracted from the thematic land cover / land use (LC/LU) map of the Puglia region (2006, 1:5000 nominal scale), classified in the Corine Land Cover taxonomy (III and IV level) of the Puglia Region. In the original LC/LU map, land use typologies are represented by 68 distinct categories. To avoid overparametrization of models, the original categories were aggregated in 8 land use classes: forest, shrubland, croplands, vineyard, olive groves, permanent crops, grassland and urban. As a measure of the landscape composition in each cell, we calculated the cell % covered by each of the eight land use classes. Furthermore, we calculated the Euclidean distance between the edge if each cell and the edge the nearest patch of forest/shrubland, as wild boards use forest and shrubland as shelter (hereafter: distance to shelter) (e.g., Meriggi & Sacchi 2001, Cocca et al. 2007, Thurfjell et al. 2009, Honda & Kawauchi 2011).

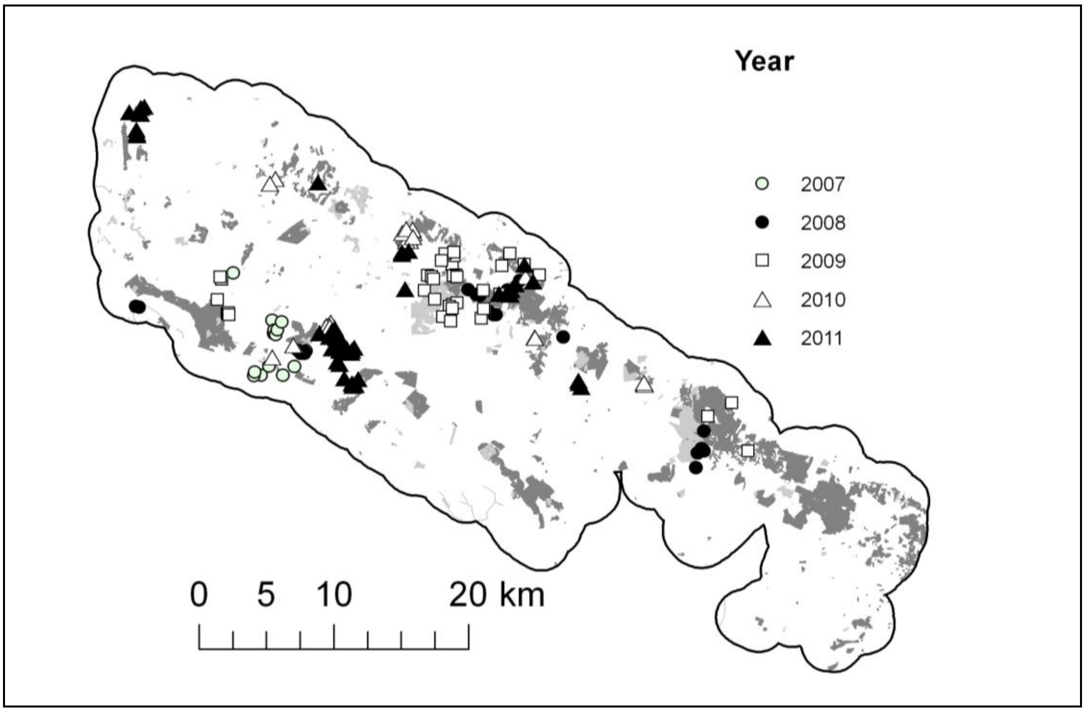


Figure 6.1: Wild boar damages reported in the Parco Nazionale dell'Alta Murgia in 2007-2011. Open circles: 2007; filled circles: 2008; open squares: 2009; open triangles: 2010; filled triangles: 2011. Pale gray: shrublands; dark gray: forests.

6.2.2 Analyses

We used maximum entropy modelling (MaxEnt) to identify the areas with highest risk of damages (Phillips et al. 2006, Elith et al. 2011). MaxEnt is a machine-learning approach that assesses the probability of presence in a given cell on the basis of environmental features in that cell; it is considered one of the most efficient approaches to SDM using presence-only data (Elith et al. 2006, Elith et al. 2011). The program establishes flexible relationships between the dependent and independent variables, and is therefore well suited to evaluate complex or non-linear relationships. Several studies demonstrated the ability of MaxEnt to build models that can be extrapolated into new contexts (e.g., in other regions or in other temporal steps) with high predictive performance (e.g., Pearson et al. 2007, Rödder et al. 2009, Ficetola et al. 2010, Fouquet et al. 2010, Reshetnikov & Ficetola 2011). MaxEnt provides a relative index of likelihood of presence instead than the actual probability of presence. Nevertheless, MaxEnt outcome is very strongly correlated with the output of models producing the probability of presence (Li et al. 2011), while making a lower number of assumptions (Phillips 2012).

We built a series of MaxEnt models by using the location of wild boar damages data. We performed analyses by measuring landscape features at different radii, and with different values of the regularization multiplier β.

[*landscape radius*] Species distribution can be affected by landscape features at different spatial scales. We built models at four scales, by using as independent variables (*i*) the landscape cover within the cell (i.e., within a radius of 0 km from the focal cell), or the average cover of each landscape category within a radius of (*ii*) 2, (*iii*) 4 and (*iiii*) 6 km from the focal cell. These values represent distances that can be travelled by wild boards, particularly during foraging activities s the (e.g., Janeau et al. 1995, Cocca et al. 2007, Honda & Kawauchi 2011). Distance from shelter was included into models built at all the spatial scales.

[*regularization multiplier*] MaxEnt models can have varying degrees of complexity, expressed by the regularization multiplier β (Phillips et al. 2006, Warren & Seifert 2011). Models with increasing values of β have higher smoothing and therefore less complexity (Elith et al. 2010, Warren & Seifert 2011). For

each landscape radius, we build seven models with increasing values of β: 1, 2, 3, 5, 7, 10, 15 (Warren & Seifert 2011).

Overall, we built 28 MaxEnt models (4 scales X 7 values of β). We then used Akaike's Information Criterion, corrected for small sample size (AICc) (Burnham & Anderson 2001, 2002) to identify the best combination of β and landscape radius. AICc trades off explanatory power versus model complexity. Parsimonious models explaining more variation have the lowest AICc value and are considered to be the "best-AICc model". Simulations suggested that this approach allows to identify the models with highest generality and transferability to other time periods (Warren & Seifert 2011). The best-AICc model also tends to estimate well the relative importance of environmental variables. We used bootstraps (100 replicates) to build response curves for the best model, as well the associated standard deviation.

After the identification of the best-fitting values of β and landscape radius, we built a series of MaxEnt models using only a subset of wild board data, to assess the actual capability of our models to predict damages in the future. First, we trained a model using the 2007-08 data, and tested its ability to correctly predict the location of the 2009 damages (test data). Similarly, we trained a model using the 2007-09 data and used the 2010 data as test, and we trained a model using the 2007-2010 data and used the 2011 data as test. In each model, we assumed that a cell has high risk of damages if its suitability score was greater than the 10th percentile of training presence points (Pearson et al. 2007). We then examined the omission error of models to evaluate their predictive performance (Jiménez-Valverde et al. 2008); we used a binomial test to assess the significance of predictions. In these three tests, the test data are temporally distinct from the training data, therefore they constitute an independent validation of the model predictive performance, and permit to assess whether our model allows to predict the location of damages in the future (Nogués-Bravo 2009, Ficetola et al. 2010).

Models were built using MaxEnt 3.3.3 (Phillips et al. 2006, Elith et al. 2011); AICc of models was calculated with ENMTools 1.3 (Warren et al. 2010, Warren & Seifert 2011).

6.3 Results

We used 20 records of wild boar damages in 2007, 23 in 2008, 38 in 2009, 29 in 2010 and 47 in 2011 (Figure 6.1). The number of records of damages significantly increased over the study period (Spearman's correlation, r_S = 0.90, P = 0.037). Damages were often clustered in specific areas of the park, and in certain sectors damages were reported only in one of the study years (Figure 6.1). Cereals, legumes and vineyards were the crops of cadastral parcels more frequently damaged by wild boards (Figure 6.2).

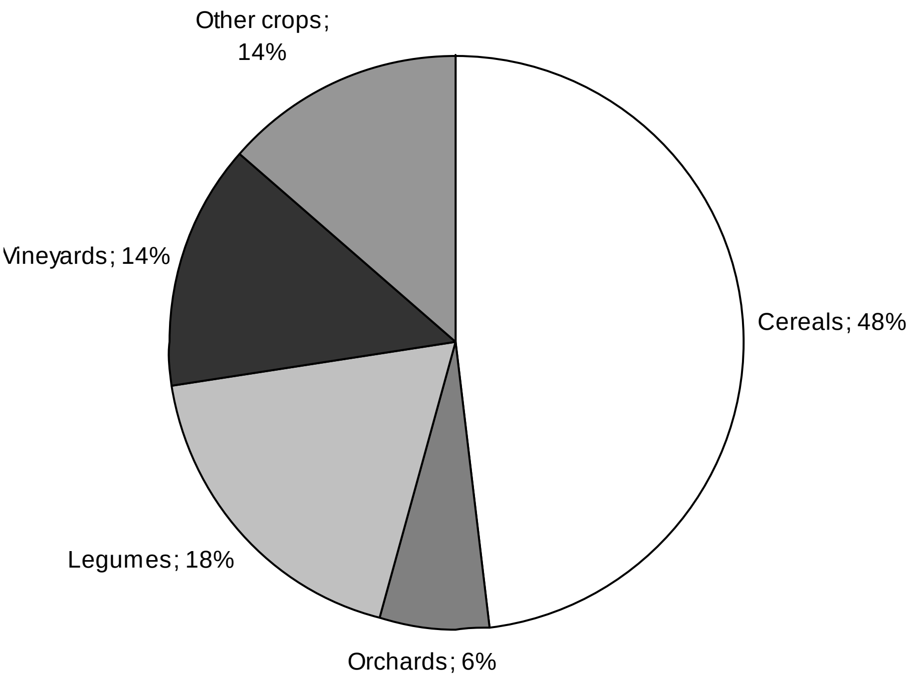


Figure 6.2: Types of crops damaged by wild boars in 2007-2011. It should be remarked that, within cereals, 71% of damages only were attributed to a specific culture (wheat, barley or oats); for the remaining cereals the damaged was attributed to undefined cereals.

The model with regularization multiplier β = 3 and landscape measured in a radius of 2 km showed the lowest AICc (AICc = 2887), and was thus considered to be the best-AICc model (Table 6.1). No model had AICc values very close to the best model. The model with the second lowest value of AICc (β = 5, radius = 2 km) had AICc = 2901 (Table 6.2). The difference in AICc values between the two models was 14, thus indicating a very strong support in favour of the best model (see e.g., Burnham & Anderson 2001).

The cover of forest, urban, olive groves and shrublands within 2 km were the variables most important to explain the location of damages (Table 6.2). The risk of damages was highest in cells with low cover of urban areas or olive grows, intermediate values of forest cover and relatively high values of shrubland cover (Figure 6.3).

As reported in the study from Corriero et al. 2010, based on a discriminant function analysis, wide areas of the park have high – to very-high risk of wild boar damages This is confirmed by our results (Figure 6.4), yet the spatial distribution of the risk of damage is different.

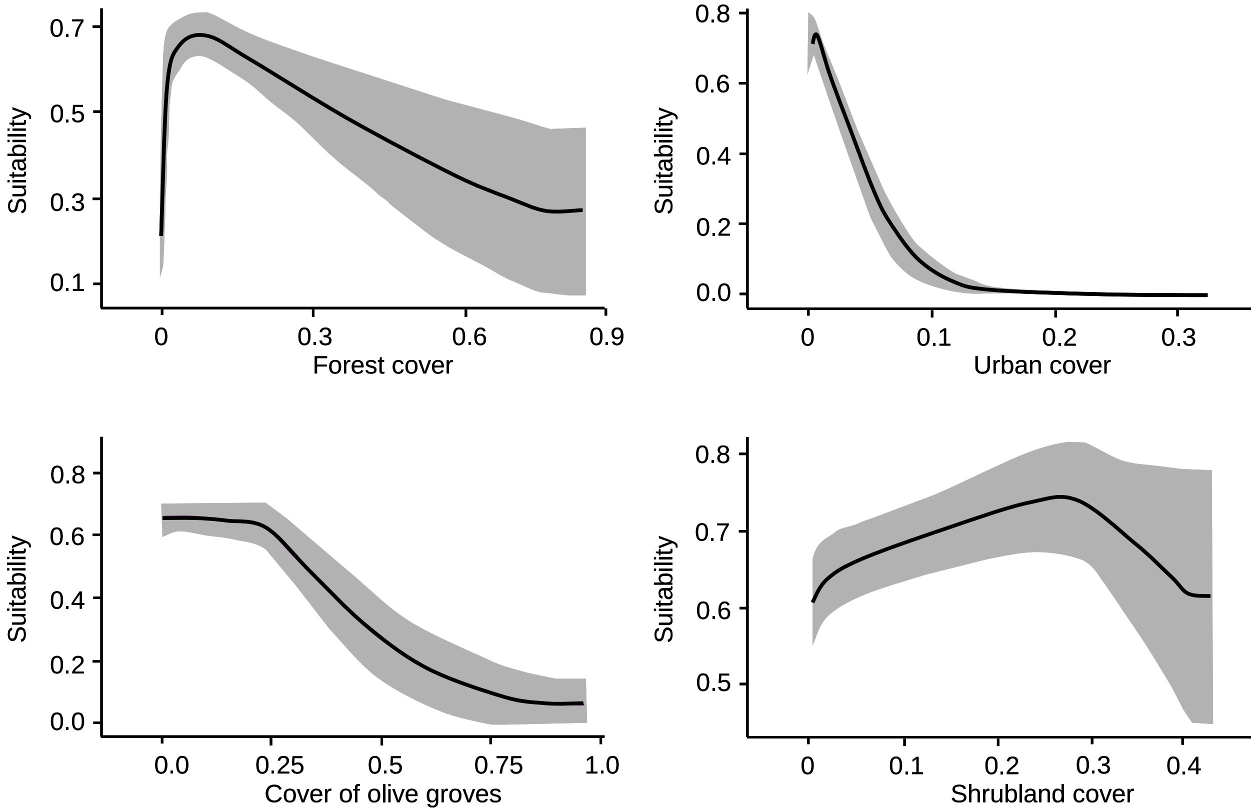


Figure 6.3: Response curves showing the relationship between landscape features and the risk of wild boar damages. Plots represent the four landscape features most important to explain the distribution of damages: a) forest cover; b) urban cover; c) cover of olive groves; d) shrubland cover. Shaded bands represent one standard deviation estimated on the basis of 100 bootstraps.

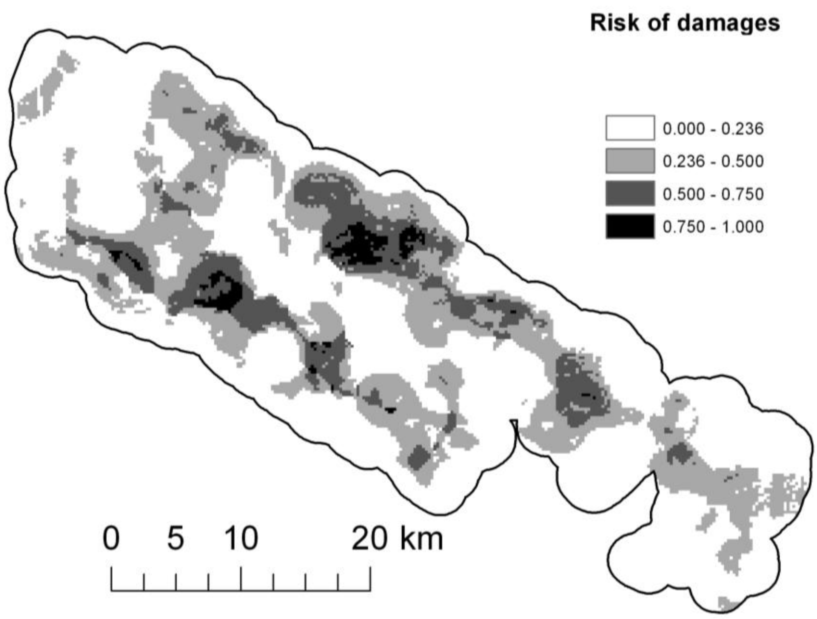


Figure 6.4: Risk of wild boar damages in the Parco Nazionale dell'Alta Murgia, on the basis of data collected from 2007 to 2011.

Table 6.1: Comparison of models built using different values of the regularization multiplier, β, and measuring landscape cover at different spatial scales. Models are compared using the Akaike Information Criterion, corrected for small sample size (AICc). The lowest-AICc model is shown in bold.

		Radius		
Regularization multiplier	0 km AICc	2 km AICc	4 km AICc	6 km AICc
1	3088.9	2991.3	3483.4	5649.8
2	3077.5	2946.6	2988.7	3079.9
3	3068.8	**2887.0**	2943.4	2994.3
5	3071.0	2901.0	2935.7	3010.3
7	3071.8	2927.7	2949.3	2987.4
10	3079.3	2977.2	2973.6	2997.1
15	3093.1	2996.6	2987.5	3006.0

Table 6.2: Relative contribution of environmental variables to the best AICc model

Variable	Contribution (%)
Forest cover	31.6
Urban cover	26.7
Olive groves	16.0
Shrrubland	11.5
Wineyard	4.5
Distance from shelter	3.4
Woody cultures	3.2
Cropland	2.3
Grassland	0.8

Table 6.3: Predictive performance of models, assessed by testing the model outcome with the data of the year subsequent to the training period.

Training years	Test year	omission rate	*P*
2007-2008	2009	0.237	<0.0001
2007-2009	2010	0.310	<0.0001
2007-2010	2011	0.511	0.005

The predictive performance of model was always good, when we tested them against the data collected during the year subsequent to the training period. In all cases, models predicted damages in the subsequent year significantly better than expected by chance (Table 6.3).

6.4 Discussion

As in other areas of Europe, wild boar damages are quickly augmenting in the Alta Murgia National Park, with a two-fold increase in just five years. Despite the number of damaged parcels at present may

seem limited, if compared with other regions (e.g., Schley et al. 2008), the quickly growing trend raises concern for the situation in the next future. It should also be remarked that the damage records available were limited to the reports performed by farmers, and a number of damages are probably non reported, therefore the available figures certainly represent an underestimation of the current damages. In our study area, it seems that wild boars thrive on natural trophic resources and switch on cultivated ones during the summer, when they mostly damage vineyards and almond groves (Corriero et al. 2010). Cereals are the most frequent crop within the study area, and probably represent the landscape matrix which the animals most frequently have to cross to reach both shelters and other crops. This probably explains the very high number of damages due to trampling on cereals.

Our model identifies landscape features that increase the risk of crop damages: the risk was highest in areas with very low urban cover, intermediate values of forest and shrub cover, and intermediate to low cover of olive groves (Fig. 3). These relationships allow to map the areas of the AMNP where damage risk is highest (Fig. 4). The negative relationship with human presence can be explained by the avoidance of these areas, and has been observed in other regions of Europe (Schley et al. 2008). Similarly, the partially positive relationship with forests and shrublands likely occurs because of the high suitability of these vegetation typologies for wild boars (Meriggi & Sacchi 2001, Schley et al. 2008, Honda & Kawauchi 2011). Shrubland cover was the variable most positively related to damages (Fig. 3d), still in the whole study area no cells had more than 40% of shrub cover within 2km, and the highest suitability was associated with shrubland cover around 30%. Distance from potential shelter areas (i.e., shrublands or forests) was not among the most important predictors in our models. However, our model included the cover of forests and shrublands, and these variables are partially correlated to the distance from shelters (correlation between shrub cover and distance from refugia: $r = 0.28$, $P < 0.001$; correlation between forest cover and distance from refugia: $r = 0.56$, $P < 0.001$).

It should be remarked that damage risk was not associated with very high values of a particular landscape features. Actually, damage risk was high in areas having low values of a combination of several landscape features (e.g., forest, olive groves, shrubland…; Fig. 3). This probably occurs because we did not model the suitability of cells for wild boars, which is expected to be highest in areas with the most natural vegetation (Meriggi & Sacchi 2001). Instead, we tried to identify the areas with highest risk of damages. The risk of damages can be challenging to model, particularly for species whose habitat requirements are manifold due to both phenological traits and climatic constraints, and even more so when the species is has great dispersal and adaptive abilities.. Overall, our analysis suggests that the risk of damages is highest in areas with very heterogeneous landscape far from human settlements, comprising some patches of natural or semi-natural vegetation (i.e., forests or shrubland), and agricultural areas the wild boars can exploit for foraging. This makes it complex the management of wild boar damages, because very large areas of the AMNP share these features (Fig. 4).

One basic assumption of SDMs is that species are at equilibrium with the available environmental features. In our study area, wild boars are expanding their distribution: in 2007 damages were limited to specific, small areas; despite they now occupy a much larger range (Fig. 1), the ongoing positive trend suggest that they are probably still absent from potentially suitable localities, particularly in the southern sectors (Fig. 4). Nevertheless, SDM were able to yield useful predictions that identified well the areas most at risk of damages in the subsequent years (Table 3). Other studies have demonstrated the ability of SDMs to predict distribution changes of species not at equilibrium. The fact that species have not reached the equilibrium may limit the predictive ability of SDMs – yet, I models may be able to identify relationships with habitats with sufficient accuracy to perform useful predictions. Equilibrium is a key assumption of SDMs, yet it is very difficult to test. The possibility to obtain useful SDMs even when this assumption is not fully met is good news for models aiming at predicting the potential distribution of species in quick expansion, such as many species causing human-wildlife conflicts, invasive species, or species undergoing the effects of climate change. At present, these are some of the situations for which the predictive ability of SDMs is most frequently called as an important step toward management (Jeschke & Strayer 2008).

Such potentials of SDMs could be further exploited within the framework of regular monitoring programmes based on both damage recording and landscape features change assessment. More realistic indications could be obtained if frequently updated and much detailed LC/LU maps were fed in the model, such as those produced by means of Earth observation techniques which provide accurate detection and quantification of landscape features dynamics over broad geographic scales (Nagendra et al. 2013).

7. Case study 4: Land-use variables, remote sensing data, and the performance of species distribution models for birds

7.1 Introduction

Correlative species distribution models (SDMs) analyse relationships between species distribution data and environmental features. SDMs allow to evaluate the suitability of a given area for one or multiple species, and provide important information on ecological factors determining species distribution. The output of SDMs is increasingly used for multiple purposes, including the identification of conservation priorities, prediction of species invasions and analyses of the impact of environmental changes on biodiversity (Elith and Leathwick 2009; Jiménez-Valverde et al. 2011).

SDMs can be performed at many spatial scales, including local (e.g., one single reserve) regional, continental and global. SDMs require both species distribution data (e.g., presence/absence, abundance or presence-only data, depending on the technique used) and relevant environmental data. Among environmental data, abiotic variables (e.g, climate) tend to be more important in models analysing distribution at the biogeographical scale (particularly continental or global), while variables representing landscape, vegetation or biotic interactions can be more important at finer spatial scales (Soberon and Nakamura 2009; Jiménez-Valverde et al. 2011; Boulangeat et al. 2012). For instance, a wide region can have a suitable climate for a given species but, within this region, the target species can attain positive fitness only in the areas with certain landscape features, or with appropriate resources (Soberon and Nakamura 2009; Ficetola et al. 2010; Jiménez-Valverde et al. 2011; Boulangeat et al. 2012; Brambilla and Ficetola 2012). Models performed at the landscape scale are particularly relevant to guide management planning in protected areas. For instance, the land use features positively associated to species of conservation concern can be favoured, while those increasing the risk of invasion by alien species can be limited (Brambilla et al. 2010; Ficetola et al. 2010).

Nevertheless, the identification of relevant environmental variables for modelling species distribution is a complex task that is frequently overlooked. Both at broad and at landscape scale, many analyses include a very large number of predictors. For example, a large number of studies analysing climatic suitability used the 19 "bioclimate" variables of the WorldClim dataset as predictors (Hijmans et al. 2005). However, including too many predictors into models determines the risk of models over-fitting the data, and may limit the ability of models to perform predictions under different conditions (i.e., model transferability) (Peterson and Nakazawa 2008; Rödder et al. 2009). The situation may be even more complex for models considering landscape variables. Frequently, landscape variables are derived from land-use / land-cover digital maps: Researchers identify the land-use typologies present within the study area; the percentage cover of a given land-use variable is then used as environmental predictor in the models. However, maps can have a very large number of land-use categories, and it is not easy to identify *a priori* how many (and which) land-use categories are relevant and should be included into analyses. Furthermore, recent remote sensing techniques provide improve additional information on ecological structure, that can be can be added to SDMs in addition to standard land-use categories, or perhaps might even substitute the traditional data (Tattoni et al. 2012).

In this study, we used an information-theoretic approach to identify the landscape composition variables that are more appropriate to build SDMs for birds in a protected area of The Netherlands. The study area is relatively small and without major reliefs (Figure 7.1), therefore it is possible to rule out the possible impact of climatic or topographic heterogeneity on the study species. Specifically, we compared five approaches to the inclusion of environmental variables as predictors into SDMs at landscape scale. 1): Models built using a relative large number of traditional land-use variables; 2) models built using a small number of land-use variables; 3) models excluding land-use variables, and considering only vegetation height data, obtained through remote sensing (LIDAR: light detection and ranging; Lefsky et al. 2002; Vierling et al. 2008). Furthermore, we built models combining traditional land-use variables and

LIDAR data. 4): Models using both LIDAR and a large number of land-use variables; 5) Models using both LIDAR and a small number of land-use variables. The use of an information-theoretic approach allows to explicitly assess the relative likelihood of multiple models, and to identify the models with the highest exportability (Symonds and Moussalli 2011; Warren and Seifert 2011).

7.2 Material and methods

7.2.1 Study area, environmental and species distribution data

The Dutch study area for Bio-SOS is located within the Natura 2000 site the Veluwe (site codes: NL9801023+ NL3009017) in the Province of Gelderland, and falls under the Habitat Directive as well as Bird Directive. The Veluwe is the largest end moraine in the Netherlands, an undulating sandy landscape that was created during penultimate glacial period, about 150,000 years ago. The final landscape of alternating sand dune areas, heathlands and dry forests were created by a long history of intensive land use (starting from Celtic fields in Iron Age). Drifting inland sand dunes caused by anthropogenic overexploitation were a serious threat and were battled around 1900 by massive forest plantations. Since then the sand dune area has severely diminished. However, the inland sand dunes of the Veluwe still belong to the largest of Europe.

Although the entire Veluwe has a total surface of 91.200 ha, the selected sites are much smaller. The heathland area Ginkelse and Ederheide covers an area of approximately 1000 ha in size and is known for its large area covered by Calluna heath vegetation. The main Annex Habitat types for this study area are: i) 4030 European dry heaths; ii) 2310 Dry sand heaths with Calluna and Genista; iii) 2330 Inland dunes with open Corynephorus and Agrostis grasslands. This terrain Ginkelse and Ederheide is managed by Ministry of Defence. The Wekeromse Zand, an active inland sand dune area, 3 km North of the Ginkelse & Ederheide, has a total area of approximately 500 ha and is managed by Geldersch Landschap. About 100 ha of this area is covered by open space with active inland sand dunes. Less digital data are available for this site, but the active inland sand dunes are of European importance. The main Annex Habitat types for of the Wekormse Zand are: i) 2310 Dry sand heaths with Calluna and Genista; ii) 2330 Inland dunes with open Corynephorus and Agrostis grasslands; iii) 6230 Species-rich Nardus grasslands; iv) 9120 Atlantic acidophilous beech forests with Ilex and sometimes also Taxus in the shrublayer; v) 9190 Old acidophilous oak woods with Quercus robur on sandy plains. Most important pressure to the biodiversity of the above mentioned habitats are the nitrogen deposition caused by intensive agriculture in the region, which is causing moss (*Campylopus introflexus*), grass (*Molinia caerulea*), shrub (*Rubus fruticoses spp.)* and tree encroachment (*Pinus sylvestris* and *Betula pendula)*. Specific pressures for Ginkelse & Ederheide are also caused by recreation. The conservation status is of all Annex I habitat types (17 in total) of the Veluwe is bad (-) or really bad (--). Main management objectives are: to obtain a good heathland structure; to prevent encroachment; to prevent loss of sand dunes; to avoid fragmentation; and to provide optimal environmental conditions for heath and sand dune fauna. Annex II habitat species are amongst others: Nightjar (*Caprimulgus europaeus*), Eurasian Wryneck (*Jynx torquilla*), Tawny Pipit (*Anthus campestris*), Stonechat (*Saxicola torquata*), Wheatear (*Oenathe oenathe*) and Red-backed shrike (*Lanius collurio*).

Figure 7.1: Location of the Dutch study area Ginkelse - Ederheide & Wekeromse Zand within Natura 2000 Veluwe in the centre of the Netherlands.

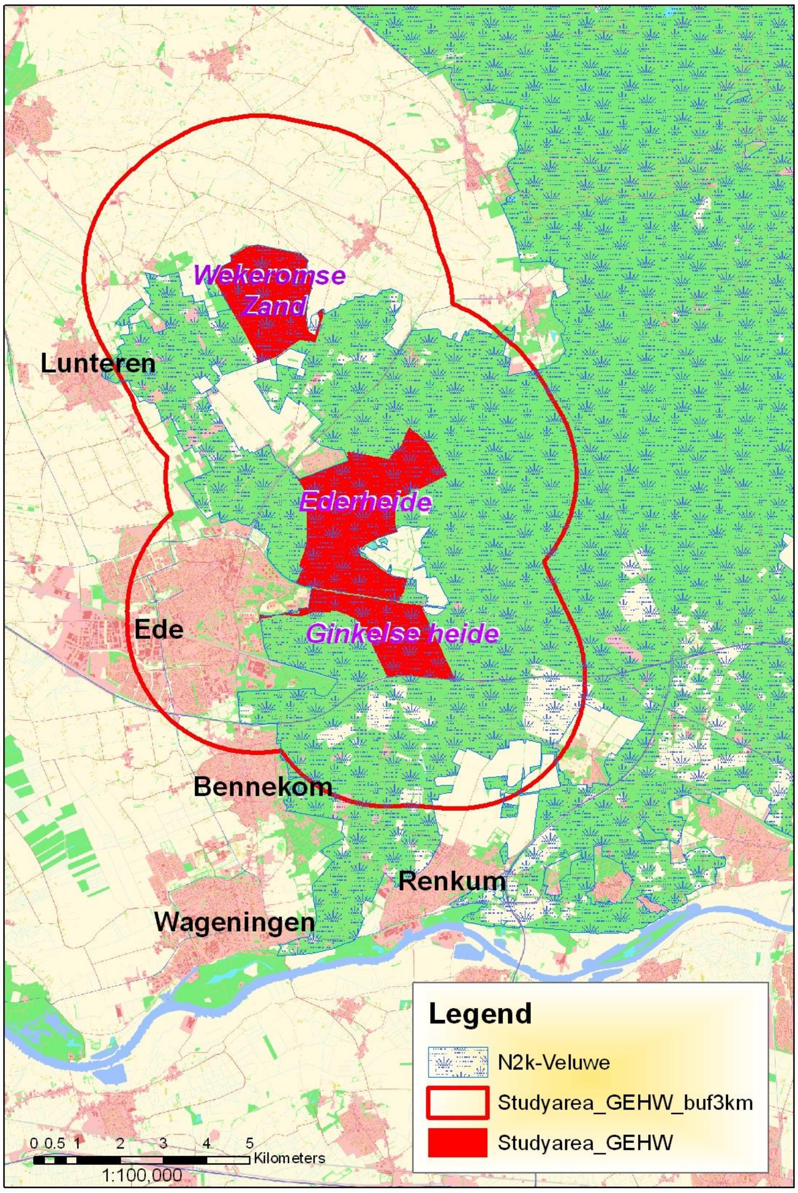


Figure 7.2: Exact location of the Dutch study area Ginkelse - Ederheide & Wekeromse Zand with a buffer zone (redline) of 3 kilometres.

7.2.2 LIDAR Data

LIDAR is a remote sensing technology that measures the tree dimensional distribution of plant canopies, and can estimate the structural features of vegetation.

7.2.3 Bird point data

Avifauna data were collected by bird point counts at unlimited distance (Bibby et al. 2001).

7.2.4 Statistical analyses

We used maximum entropy modelling (MaxEnt) (Phillips et al. 2006; Elith et al. 2011) to build models relating bird occurrence data to the land-use and LIDAR data. MaxEnt is a machine-learning approach that assesses the probability of presence in a given cell on the basis of environmental features in that cell; comparative analyses showed that MaxEnt is among most efficient approaches to SDM using presence-only data (Elith et al. 2006; Elith et al. 2011). The program establishes relationships between the dependent (species presence) and independent variables, and is well suited to evaluate complex or non-linear relationships. MaxEnt provides a relative index of likelihood of presence instead than the actual probability of presence. Nevertheless, MaxEnt outcome is strongly correlated with the results of models producing the probability of presence (Li et al. 2011), while making a lower number of assumptions (Phillips 2012).

We built models for the nine bird species for which we obtained >25 species occurrence data (Table 7.1). For each species, we built five MaxEnt models, using different sets of independent variables: 1): using all the nine land-use variables; 2) using only the three land use variables more represented in the study area (i.e., heathland, coniferous and broadleaved forest cover); 3) using LIDAR vegetation height data only; 4): using the nine land-use variables and LIDAR; 5) using the three more represented land-use variables and LIDAR. We then used an information-theoretic approach, based on Akaike's Information Criterion corrected for small sample size (AICc) to identify the best model for each species. AICc trades off explanatory power versus model complexity; parsimonious models explaining more variation have the lowest AICc value and are considered to be the "best-AICc model". Simulations suggested that AICc allows to identify the models with highest generality and transferability better than using other approaches such as cross-validation (Warren and Seifert 2011). Models were then ranked on the basis of their Δ-AICc, which represents the difference in AICc units between the best model and the model of interest. For each model, we also calculated the AICc weight *w*, which represents the average support of the model, given the set of candidate models and the data.

We also used cross-validation to test the predictive performance of models. For each model we split the data in 5; the model was built with 80% of data and the predictive performance was estimated for the remaining 20% of data. We calculated the area under the curve of the receiver operator plot (AUC) for the test data, averaged over the five replicated runs, as a measure of predictive performance. Models with AUC = 0.5 do not perform better than random; AUC > 0.7 indicate useful performance, AUC > 0.8 indicate good performance and AUC > 0.9 indicate excellent performance.

7.3 Results

The species with at least 25 presence points were: sky lark *Alauda arvensis*; eadow pipit *Anthus pratensis*; tree pipit *Anthus trivialis*; stonechat *Saxicola rubicola*; song thrush *Turdus philomelos*; crested tit *Lophophanes cristatus*; short-toed treecreeper *Certhia brachydactyla*; linnet *Carduelis cannabina* and yellowhammer *Emberiza citronella*. For each species, we obtained 26-214 presence points (Table 7.1). Out of these species, the sky lark is listed in the annex II of the EU Bird Directive and is a SPEC3

species (not concentrated in Europe but with an unfavourable conservation status); the crested tit and the lined are SPEC2 species (concentrated in Europe and with an unfavourable conservation status, see Bird Life International, 2004).

For seven out of the nine species (i.e., 78% if species), LIDAR data were included in the best AICc model. For 56% of species the best AICc model included LIDAR data only, and did not include any land use variable. The reduced set of land-use variables was included in the best model for 22% of species, while for 22% of species the best model included both the full set of land-use variables, and LIDAR data.

Overall, the model including LIDAR only tended to be those with lower rank across species, and had higher average weight. Conversely, the model including all land use variables but excluding LIDAR was consistently the model with poorest performance (Figure 7.3).

Cross validation indicated that the models for all species showed performance ranging from useful to excellent (Table 7.2). The suitability maps for all species are reported in Figure 7.4.

Table 7.1: Species with more than 25 presence points

Species	*N*	Variables	AICc	ΔAICc	*w*
Alauda arvensis	214	All variables + LIDAR	3941.00	0.00	0.999
		All landuse variables	3954.47	13.47	0.001
		Reduced landuse + LIDAR	3957.78	16.78	0.000
		Reduced landuse	3964.42	23.42	0.000
		LIDAR only	4019.19	78.19	0.000
Anthus pratensis	154	All variables + LIDAR	2866.74	0.00	0.643
		Reduced landuse	2867.94	1.20	0.353
		All landuse variables	2877.18	10.43	0.003
		Reduced landuse + LIDAR	2879.84	13.09	0.001
		LIDAR only	2934.75	68.00	0.000
Anthus trivialis	156	LIDAR only	2927.81	0.00	0.784
		Reduced landuse + LIDAR	2930.38	2.58	0.216
		All variables + LIDAR	2950.00	22.19	0.000
		Reduced landuse	2956.09	28.28	0.000
		All landuse variables	2963.01	35.20	0.000
Saxicola rubicola	75	LIDAR only	1393.56	0.00	0.997
		Reduced landuse + LIDAR	1405.01	11.45	0.003
		Reduced landuse	1420.11	26.56	0.000
		All variables + LIDAR	1420.57	27.01	0.000
		All landuse variables	1437.10	43.55	0.000
Turdus philomelos	26	LIDAR only	1393.56	0.00	0.997
		Reduced landuse	1405.01	11.45	0.003
		All landuse variables	1420.11	26.56	0.000
		Reduced landuse + LIDAR	1420.57	27.01	0.000
		All variables + LIDAR	1437.10	43.55	0.000
Lophophanes cristatus	39	Reduced landuse	717.32	0.00	0.789
		LIDAR only	719.96	2.64	0.211
		Reduced landuse + LIDAR	731.92	14.59	0.001
		All landuse variables	742.77	25.45	0.000

		All variables + LIDAR	761.92	44.60	0.000
Certhia brachydactyla	33	LIDAR only	556.73	0.00	0.999
		Reduced landuse	571.24	14.52	0.001
		Reduced landuse + LIDAR	583.87	27.14	0.000
		All landuse variables	602.37	45.65	0.000
		All variables + LIDAR	619.07	62.35	0.000
Carduelis cannabina	33	LIDAR only	641.61	0.00	0.906
		Reduced landuse + LIDAR	646.64	5.03	0.073
		Reduced landuse	649.21	7.60	0.020
		All variables + LIDAR	657.26	15.65	0.000
		All landuse variables	686.98	45.37	0.000
Emberiza citrinella	58	Reduced landuse	1107.41	0.00	0.997
		All landuse variables	1119.26	11.85	0.003
		LIDAR only	1129.01	21.60	0.000
		Reduced landuse + LIDAR	1132.15	24.74	0.000
		All variables + LIDAR	1140.66	33.26	0.000

Table 7.2: Predictive performance of best AICc models, obtained using cross-validation

Species	Test AUC	SD
Alauda arvensis	0.800	0.025
Anthus pratensis	0.777	0.029
Anthus trivialis	0.752	0.012
Saxicola rubicola	0.777	0.037
Turdus philomelos	0.906	0.019
Lophophanes cristatus	0.785	0.027
Certhia brachydactyla	0.920	0.027
Carduelis cannabina	0.746	0.088
Emberiza citrinella	0.716	0.085

7.4 Conclusion

This case study showed the usefulness of remote-sensing data to build ecological models. Traditional land use maps can certainly provide important information. However, for the majority of species the LIDAR images greatly improve the performance of SDMs. Actually, for more than 50% of species the LIDAR vegetation height image alone performed better than any of the traditional land-cover maps. The performance of models may be further improved by the availability of very high quality habitat maps.

A further step that can be carried out in the next months will be a comparison with GCH maps. A preliminary version of GHC maps has been provided in September 2012. A preliminary data inspection, show that some on-going problem in labelling GHC (as an example some patches labelled as water or aquatic vegetation are herbaceous vegetation) will bias the general results. Instead to provide now a result that is unrealistic we will perform additional analysis when a definitive map of GHC can be available.

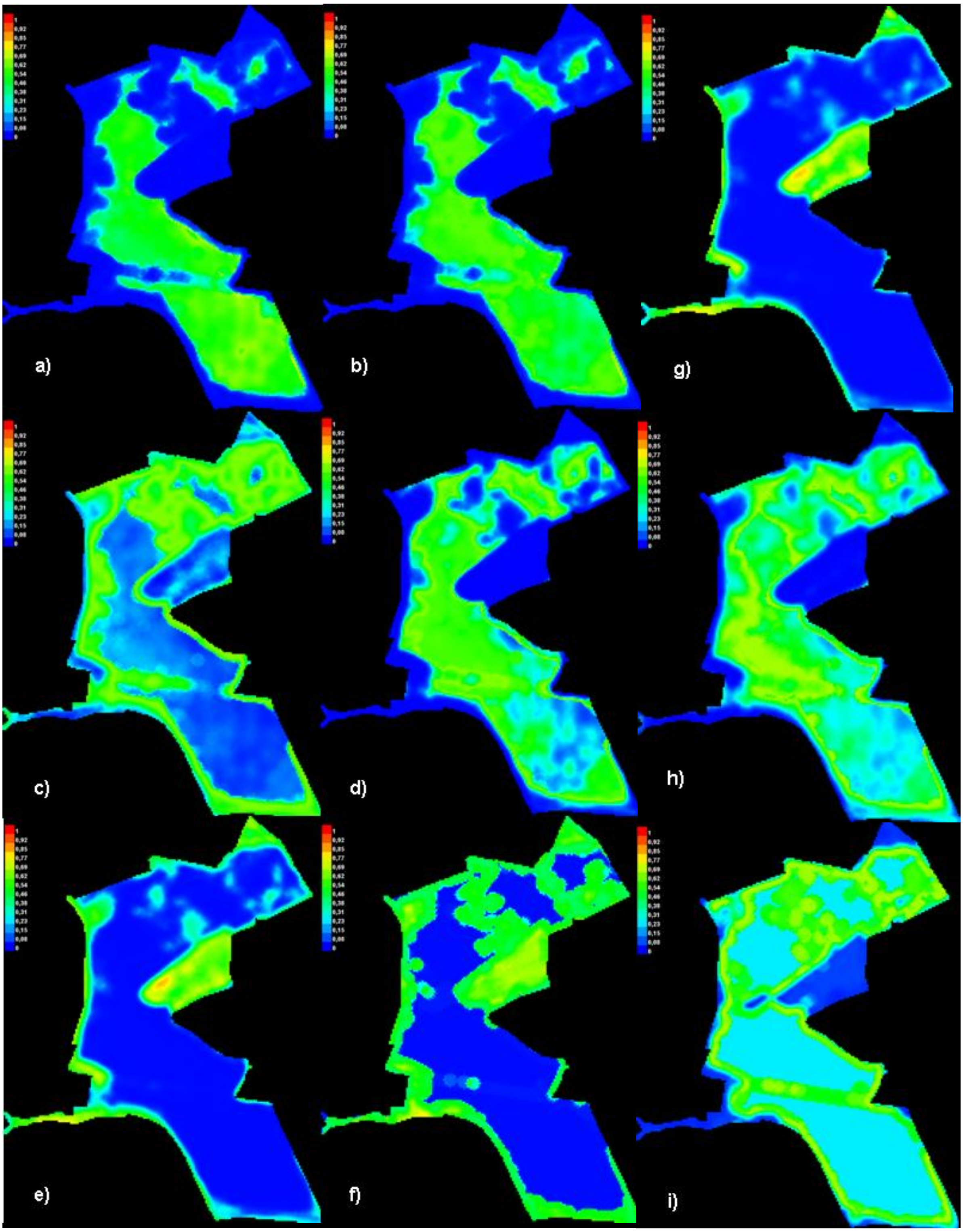
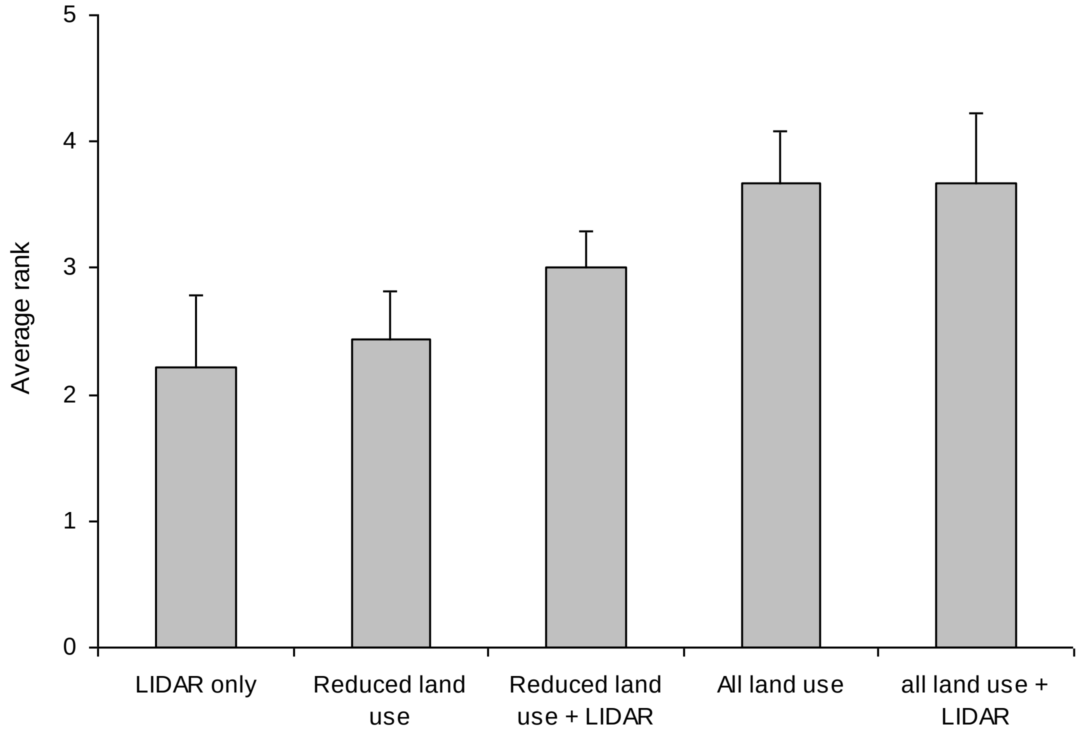


Figure 7.3: Suitability maps for the nine bird species. Warmer colours indicate higher suitability; suitability ranges from zero (no suitability) to 1 (very high suitability). Species are: *a) Alauda arvensis*; b) *Anthus pratensis*; c) *Anthus trivialis*; d) *Saxicola rubicola*; e) *Turdus philomelos*; f) *Lophophanes cristatus, g) Certhia brachydactyla; h) Carduelis cannabina; i) Emberiza citronella*.



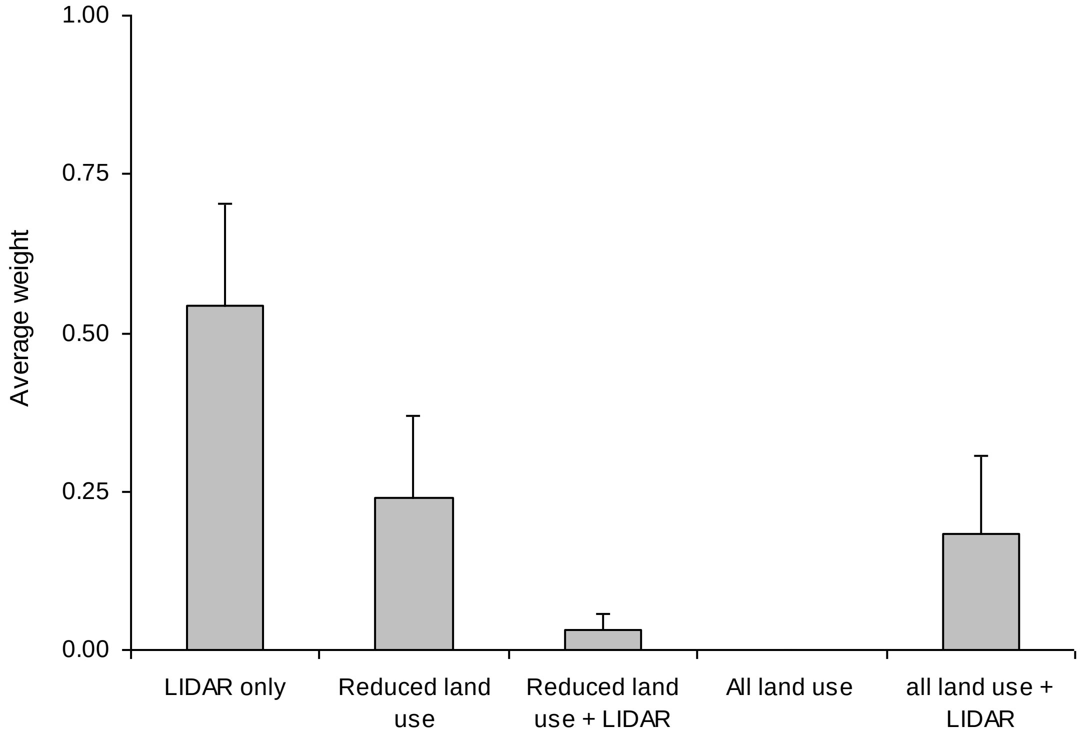


Figure 7.4: Relative importance of the five candidate models, averaged across species. In a) the importance is assessed as the rank of the model (lower is better); in b) the importance is assessed as the average AICc weight of the model (higher is better). Error bars are standard errors of the mean.

8. Discussion of D 6.7: results and future integration

In this deliverable we carried out different examples of ecological niche modelling. Our results show how scientific research can help stakeholders in managing protected areas like Natura 2000 sites.

Invasive species are a well-recognized problem worldwide. The results in case study 1 indicate the importance of a habitat (moist deciduous forests) and an environmental variable (moisture) in influencing the occurrence of *Lantana camara*. The use of buffer zones of different distances around the plot areas helped to factor in effects of forest types much more than would have been possible with using only the forest types encountered at the sample plot locations, helping us to understand that it is not just the local occurrence of a particular habitat type, but the influence of other locally surrounding habitat types that additionally influence the distribution of this important invasive species. These results also point to the fact that deeper studies need to be undertaken in different forest types to understand the role that they play in encouraging the establishment and occurrence of invasive species which occur in understorey habitats.

The analysis on amphibians (case study 2) suggests that landscape features, obtained by land-use maps, are extremely important to explain the distribution of amphibians. All the study species breed in water, and two of them (Italian newt and pool frog) require water for long periods. The importance of landscape variables for these species is therefore striking, and confirms the necessity to perform management over broad spatial scales. In other words, the conservation of breeding wetlands is not enough for these semiaquatic species, and wide areas with suitable landscape need to be maintained.

In Alta Murgia National Park, General Habitat Category maps have been recently developed and it was therefore possible to assess whether the GHC maps would help to improve habitat models. The GHC map currently covers an area of about 33 X 15 km, therefore it is not possible to include all the amphibian wetlands in the analyses. Furthermore, the analysis was limited to the pool frogs *Pelophylax s. hispanicus*, which is the species with more presence data in the area. These preliminary analyses suggest that GHC provide a better description of the terrestrial habitat, compared with the traditional land use maps. The expansion of the area covered by GHC will certainly allow to perform more accurate comparisons of the performance of models, and to obtain better models describing species distribution that can be used for the management of the park.

In case 3 the model identifies landscape features that increase the risk of crop damages from wild boar. These relationships allow to map the areas of the AMNP where damage risk is highest. The results of our model Such potentials of SDMs could be further exploited within the framework of regular monitoring programmes based on both damage recording and landscape features change assessment. More realistic indications could be obtained if frequently updated and much detailed LC/LU maps were fed in the model, such as those produced by means of Earth observation techniques which provide accurate detection and quantification of landscape features dynamics over broad geographic scales.

The last case study showed the usefulness of remote-sensing data to build ecological models. Traditional land use maps can certainly provide important information. However, for the majority of species the LIDAR images greatly improve the performance of SDMs. Actually, for more than 50% of species the LIDAR vegetation height image alone performed better than any of the traditional land-cover maps. The performance of models may be further improved by the availability of very high quality habitat maps.

The Deliverable will be updated in next months. There are at least 2 possibilities to expand the deliverable:

- On the area Le Cesine GHC map and endangered plants distribution became available in the second half of September (too late for a complete analysis before the end of October).

- On Netherland site a further step that can be carried out in the next months will be a comparison with GCH maps. A preliminary version of GHC maps has been provided in September 2012. A preliminary data inspection, showed some on-going problems in labelling GHC (as an example some patches labelled as water or aquatic vegetation are herbaceous vegetation) that can bias the general results. Instead of utilizing this map to provide a result that is unrealistic, we will perform additional analyses when a definitive map of GHC becomes available.

9. Appendix

9.1 Appendix 1: Acronym list

AIC Akaike Information Criteria
AMNP Alta Murgia National Park
ANOVA Analysis of Variance
BAM Biotic, Abiotic and Movements
BRT Boosted Regression Trees
ENM Ecological Niche Modelling
ENMs Ecological Niche Models
EO Earth Observation
GHC General Habitat Category
HR High Resolution
LC Land Cover
LC/LU land cover / land use
LCC Land Cover Change
LIDAR light detection and ranging
NGO non-governmental organization
SDMs Species distribution models
SEBI Streamlining European 2010 Biodiversity Indicators
SPEC Species of European Conservation Concern
UDS Uso del suolo (Land use)
VHR Very High spatial Resolution

10. References

Bhagwat, Shonil A., Elinor Breman, Tarsh Thekaekara, Thomas F. Thornton, and Katherine J. Willis. “A Battle Lost? Report on Two Centuries of Invasion and Management of Lantana Camara L. in Australia, India and South Africa.” Edited by Anna Traveset. *PLoS ONE* 7, no. 3 (March 5, 2012): e32407.

Bibby, CJ; Hill, DA; Burgess, ND; Mustoe, S; 2001 Bird Census Techniques. Academic Press

Bird Life International (2004) Birds in Europe: population estimates, trends and conservation status Cambridge University Press

Boulangeat, I., Gravel, D., Thuiller, W., 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. Ecology Letters 15, 584-593.

Brambilla, M. Ficetola, G.F. (2012): Species distribution models as a tool to estimate reproductive parameters: a case study with a passerine bird species. Journal of Animal Ecology, 81: 781–787.

Brambilla, M., Casale, F., Bergero, V., Bogliani, G., Crovetto, G.M., Falco, R., Roati, M., Negri, I., 2010. Glorious past, uncertain present, bad future? Assessing effects of land-use changes on habitat suitability for a threatened farmland bird species. Biological Conservation 143, 2770-2778.

Bunce, R.G.H.; Bogers, M.M. B. , Roche, P., Walczak, I. Geijzendorffer, R. , Jongman, R.H.G. 2011. Manual for Habitat and Vegetation Surveillance and Monitoring: Temperate, Mediterranean and Desert Biomes. First edition. Wageningen, Alterra report 2154 . 106 pp.; 15 fig.; 14 tab.; 35 ref.

Bunce, R.G.H.; Metzger, M.J.; Jongman, R.H.G.; Brandt, J.; Blust, G. de; Elena-Rossello, R.; Groom, G.B.; Halada, L.; Hofer, G.; Howard, D.C.; Kovár, P.; Mücher, C.A.; Padoa-Schioppa, E.; Paelinx, D.; Palo, A.; Pérez-Soba, M.; Ramos, I.L.; Roche, P.; Skånes, H.; Wrbka, T. 2008. A Standardized Procedure for Surveillance and Monitoring European Habitats and provision of spatial data. Landscape Ecology, 23:11-25

Bulte, E.H. Rondeau, D. (2005): Why compensating wildlife damages may be bad for conservation. Journal of Wildlife Management, 69: 14-19.

Burnham, K.P. Anderson, D.R. (2001): Kullback-Leibler information as a basis for strong inference in ecological studies. Wildlife Research, 28: 111-119.

Burnham, K.P. Anderson, D.R. (2002): Model selection and multimodel inference: a practical information-theoretic approach. Springer Verlag, New York.

Busby, J.R., 1991. BIOCLIM - a bioclimate analysis and prediction system, In: C.R. Margules, M.P. Austin (Eds.), Nature conservation: cost effective biological surveys and data analysis, CSIRO, Melbourne, pp. 64-68.

Cocca, G., Sturaro, E., Dal Compare, L. Ramanzin, M. (2007): Wild boar (*Sus scrofa*) damages to mountain grassland. A case study in the Belluno province, eastern Italian Alps. Italian Journal of Animal Science, 6: 845-847.

Cozza, K., Fico, R., Battistini, M.L. Rogers, E. (1996): The damage-conservation interface illustrated by predation on domestic livestock in central Italy. Biological Conservation, 78: 329-336.

Dimopoulos, P., Kallimanis, A., Honrado, J.P., Mucher, C.A., Lucas, R.M., Mairota, P., Tomaselli, V. Terzi, M. (2011): Sites’ Description. (BIO_SOS) Deliverable D6.10.

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.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77, 802-813.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura,

M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberon, 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., Kearney, M. Phillips, S. (2010): The art of modelling range-shifting species. Methods in Ecology and Evolution, 1: 330-342.

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.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. Yates, C.J. (2011): A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, 17: 43–57.

Ficetola, G.F., Maiorano, L., Falcucci, A., Dendoncker, N., Boitani, L., Padoa-Schioppa, E., Miaud, C. Thuiller, W. (2010): Knowing the past to predict the future: land-use change and the distribution of invasive bullfrogs. Global Change Biology, 16: 528-537.

Ficetola, G.F., Thuiller, W. Padoa-Schioppa, E. (2009): From introduction to the establishment of alien species: bioclimatic differences between presence and reproduction localities in the slider turtle. Diversity and Distributions, 15: 108-116.

Forte, L., Perrino, E.V. Terzi, M. (2005): Le praterie a *Stipa austroitalica* Martinovsky ssp. *austroitalica* dell'Alta Murgia (Puglia) e della Murgia Materana (Basilicata). Fitosociologia, 42: 83-103.

Fouquet, A., Ficetola, G.F., Haigh, A. Gemmell, N. (2010): Using ecological niche modelling to infer past, present and future environmental suitability for *Leiopelma hochstetteri*, an endangered New Zealand native frog. Biological Conservation, 143: 1375-1384.

Godsoe, W., 2010. I can't define the niche but I know it when I see it: a formal link between statistical theory and the ecological niche. Oikos 119, 53-60.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993-1009.Hegel, T.A., Gates, C.C. Eslinger, D. (2009): The geography of conflict between elk and agricultural values in the Cypress Hills, Canada. Journal of Environmental Management, 90: 222-235.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. High resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978.

Honda, T. Kawauchi, N. (2011): Methods for constructing a wild boar relative-density map to resolve human-wild boar conflicts. Mammal Study, 36: 79-85.

Janeau, G., Cargnelutti, B., Cousse, S., Hewison, M. Spitz, F. (1995): Daily movement pattern variations in wild boar (*Sus scrofa* L.). Ibex - Journal of Mountain Ecology, 3: 98-101.

Jeschke, J.M. Strayer, D.L. (2008): Usefulness of bioclimatic models for studying climate change and invasive species. Annals of the New York Academy of Sciences, 1134: 1-24.

Jiménez-Valverde, A., Lobo, J.M. Hortal, J. (2008): Not as good as they seem: the importance of concepts in species distribution modelling. Diversity and Distributions, 14: 885-890.

Jiménez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. Lobo, J.M. (2011): Use of niche models in invasive species risk assessments. Biological Invasions, 13: 2785–2797.

Kearney, M., Porter, W., 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology Letters 12, 334-350.

Khuroo, Anzar A., Zafar A. Reshi, Akhtar H. Malik, Ewald Weber, Irfan Rashid, and G. H. Dar. "Alien Flora of India: Taxonomic Composition, Invasion Status and Biogeographic Affiliations." *Biological Invasions* 14, no. 1 (April 9, 2011): 99–113.

Lavelle, M.J., Vercauteren, K.C., Hefley, T.J., Phillips, G.E., Hygnstrom, S.E., Long, D.B., Fischer, J.W., Swafford, S.R. Campbell, T.A. (2011): Evaluation of fences for containing feral swine under simulated depopulation conditions. Journal of Wildlife Management, 75: 1200-1208.

Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar remote sensing for ecosystem studies. Bioscience 52, 19-30.

Li, W., Guo, Q. Elkan, C. (2011): Can we model the probability of presence of species without absence data? Ecography, 34: 1096-1105.

Lindenmayer, D. and G. Likens (2010). Effective Ecological Monitoring. Collingwood, CSIRO

Linkie, M., Dinata, Y., Nofrianto, A. Leader-Williams, N. (2007): Patterns and perceptions of wildlife crop raiding in and around Kerinci Seblat National Park, Sumatra. Animal Conservation, 10: 127-135.

Lockwood, Julie L., Phillip Cassey, and Tim Blackburn. “The Role of Propagule Pressure in Explaining Species Invasions.” *Trends in Ecology & Evolution* 20, no. 5 (May 2005): 223–228.

Meriggi, A. Sacchi, O. (2001): Habitat requirements of wild boars in the northern Apennines (N Italy): a multi-level approach. Italian Journal of Zoology, 68: 47-55.

Murali, K. S., and S. R. Setty. “Effect of Weeds Lantana Camara and Chromelina Odorata Growth on the Species Diversity, Regeneration and Stern Density of Tree and Shrub Layer in BRT Sanctuary.” *Current Science* 80, no. 5 (2001): 675–678.

Nagendra, H., Lucas, R.M., Honrado, J.P., Jongman, R.H.G., Tarantino, C., Adamo, M. Mairota, P. (2013): Remote sensing for protected area assessment: monitoring habitat area, condition, biodiversity and threats. Ecological Indicators, under review.

Naughton-Treves, L. (1998): Predicting patterns of crop damage by wildlife around Kibale National Park, Uganda. Conservation Biology, 12: 156-168.

Nogués-Bravo, D. (2009): Predicting the past distribution of species climatic niche. Global Ecology and Biogeography, 18: 521-531.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. Peterson, A.T. (2007): Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography, 34: 102-117.

Peterson, A.T., Nakazawa, Y., 2008. Environmental data sets matter in ecological niche modelling: an example with Solenopsis invicta and Solenopsis richteri. Global Ecology and Biogeography 17, 135-144.

Phillips, S. (2012): Inferring prevalence from presence-only data: a response to 'Can we model the probability of presence of species without absence data?'. Ecography, 35: 385-387.

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

Prasad, Ayesha E. “Effects of an Exotic Plant Invasion on Native Understory Plants in a Tropical Dry Forest.” *Conservation Biology* 24, no. 3 (January 11, 2010): 747–757.

Ramaswami, Geetha, and R. Sukumar. “Woody Plant Seedling Distribution Under Invasive Lantana Camara Thickets in a Dry-forest Plot in Mudumalai, Southern India.” *Journal of Tropical Ecology* 27, no. 04 (May 31, 2011): 365–373.

Ramesh, B. R. (1989). Flora of Biligiri Rangan Hills. Chennai, Madras University. PhD thesis.

Reshetnikov, A.N. Ficetola, G.F. (2011): Potential range of the invasive fish rotan (*Perccottus glenii*) in the Holarctic. Biological Invasions, 13: 2967-2980.

Rödder, D., Schmidtlein, S., Veith, M. Lötters, S. (2009): Alien invasive slider turtle in unpredicted habitat: A matter of niche shift or of predictors studied? PLoS ONE, 4: e7843.

Rushton, S.P., Ormerod, S.J., Kerby, G., 2004. New paradigms for modelling species distribution? Journal of Applied Ecology 41, 193-200.

Schlageter, A. Haag-Wackernagel, D. (2012): Evaluation of an odor repellent for protecting crops from wild boar damage. Journal of Pest Science, 85: 209-215.

Schley, L., Dufrene, M., Krier, A. Frantz, A.C. (2008): Patterns of crop damage by wild boar (*Sus scrofa*) in Luxembourg over a 10-year period. European Journal of Wildlife Research, 54: 589-599.

Simberloff, D. (2006). "Invasional meltdown : 6years later: important phenomenon, unfortunate metaphor, or both?" Ecology Letters 9(8): 912-919.

Soberon, J.M., 2010. Niche and area of distribution modeling: a population ecology perspective. Ecography 33, 159-167.

Soberon, J., Nakamura, M., 2009. Niches and distributional areas: Concepts, methods, and assumptions. Proceedings of the National Academy of Sciences of the United States of America 106, 19644-19650.

Sundaram, B. (2011). Patterns and processes of Lantana camara persistence in Southern Tropical Dry Forests. ATREE. Bangalore, Manipal Academy of Higher Education. PhD thesis.

Symonds, M.R.E., Moussalli, A., 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behavioral Ecology and Sociobiology 65, 13-21.

Tattoni, C., Rizzolli, F., Pedrini, P., 2012. Can LiDAR data improve bird habitat suitability models? Ecological Modelling 245, 103-110.

Thurfjell, H., Ball, J.P., Ahlen, P.A., Kornacher, P., Dettki, H. Sjoberg, K. (2009): Habitat use and spatial patterns of wild boar *Sus scrofa* (L.): agricultural fields and edges. European Journal of Wildlife Research, 55: 517-523.

Toso, S. Pedrotti, L. (2001): Linee guida per la gestione del cinghiale (*Sus scrofa*) nelle aree protette. Quaderni di Conservazione della Natura, 2, Ministero dell'Ambiente, Modena.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., Clawges, R.M., 2008. Lidar: shedding new light on habitat characterization and modeling. Frontiers in Ecology and the Environment 6, 90-98.

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.

Warren, D.L., Glor, R.E. Turelli, M. (2010): ENMTools: a toolbox for comparative studies of environmental niche models. Ecography, 33: 607-611.

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.

Worster, D.; 1994. Nature's Economy: A History of Ecological Ideas (Studies in Environment and History) Cambridge Univesity Press