2. THE SPATIAL DIMENSION OF WILDLIFE CONSERVATION

2.1 Introduction

The previous chapter highlighted the importance of conserving biodiversity and the potential offered by the conservation of carnivores. From the review of related variables and threats to biodiversity, the problem emerges of setting targets and selecting specific times and places for conservation action. One of the most basic requisites for applying any conservation practice is the knowledge of the spatial attributes of the target. The simple information of *where* a species lives is vital for any conservation programme to be put in place. Therefore, the spatial distribution of the target is a fundamental characteristic that needs to be known when considering management of natural resources for conservation. This is particularly relevant in programmes that aim to accommodate interests coming from different parties to make the co-existence of wildlife and humans possible, which is the case with carnivore conservation. Generally, when the focus of conservation is animal species, an important factor may be introduced: their ability to move in space. This is relevant in cases of habitat fragmentation, as small fragments may not be enough for conserving species that demand large areas, and the problem may be overcome if connectivity between fragments is present. The analyses of spatial aspects of ecosystems provide the opportunity to consider processes, understand variables that regulate them and develop predictive models.

In this chapter a general introduction of the relevance that the geographical space has in wildlife conservation will be given. This will be followed by the description and discussion of attributes to be considered when dealing with spatial analyses of wildlife. An overview of the mathematical and statistical models used for representing the relationships between wildlife and habitat will be the subject of section 2.4, leading to the integration of GIS and its application in wildlife-habitat models (sections 2.5 – 2.6). The importance of validating models will be discussed in section 2.7. All these aspects are highly relevant to the present study, as each of them was considered during the process of mapping areas suitable for conservation of large carnivores in the Carpathian Mountains.

2.2 The geography of wildlife

Simply stating that the three large carnivores are present in the Carpathian Mountains can have a biodiversity value when considering species distribution at continental scale. However, at regional scale we must be able to say *where* within the Carpathian they actually are and/or may be in order to direct conservation efforts towards the most effective targets. Each species has a characteristic life history, reproduction rate, behaviour and means of dispersion. These determine the way the species interact with the environment and, as a consequence, their spatial distribution.

Stating that a species is distributed over a given area is a limited representation of reality. In fact, our experience of nature represents just one point in a multidimensional and constantly-changing mosaic of animals and plants that are responding to an endless course of social, environmental and climatic changes. There are a series of factors that make the distribution of species highly dynamic. These are of a biological or physical nature, in both the spatial and temporal dimensions (Cox and Cox 1994). Although the *current* distribution of a given species may only be a snapshot of the actual situation, the information is of great importance when conservation plans and monitoring programmes are to be put in place. When the information is further manipulated for predicting distributions and gaps in management practices, the geographical analysis has an increased value. The present work represents a significant contribution to the identification of areas that may need be protected from human activity during the future process of access to the EEC by the Carpathian countries. In order to understand the differences between the depiction of a species' distribution at continental scale and the identification of areas actually and/or potentially occupied by any target species, some concepts related to the species' range and its internal structure will be introduced.

2.2.1 The species' geographic range

The spatial distribution of a species is considered to be its geographic range. A species' range can be defined as the area occupied by its breeding, reproductive populations (Watts 1984). It is characterised by a series of properties of biological, physical and chemical nature and its shape and extent are the result of the interactions of such factors. The influence that external social pressures posed by human populations may have on the specie's range shall not be underestimated, as this can sometimes be fatal. The *actual* range is the region where a species presence has been recorded. This may be different in shape and extension to the *potential* range, which comprises all the areas that have environmental characteristics that make them technically suitable for being occupied by the species (Watts 1984). The actual range may suffer even more severely from the pressures coming from human presence. Gaston (1991) uses the term *extent of occurrence* to represent the areas often considered generally as the species range, as opposed to the *area of occupancy*, which falls within the extent of occurrence but excludes the places not actually occupied by the species (i.e., the actual range *sensu* Watts). The main difference between these two kinds of range, apart from the straightforward one of recorded presence in the area of occupancy, is one of spatial scale. It is very likely that at coarse scale the extent of occurrence is easier to consider than the area of occupancy, which may consist of many isolated patches within it, many of them being undetectable at small geographic scale (Gaston 1991). But more importantly, the area of occupancy is represented by those areas selected by the species among all available ones. It must be noted that the selection can be driven by active preference for some environments over others or it may be induced indirectly by physical or biological factors such as obstacles to free movements (geographical barriers) or competition with other species, including humans (Cox and Cox 1994). In terms of wildlife management the area of occupancy is more relevant than the extent of occurrence, as it brings information about the relationships between the target species and the environment.

Most representations of geographic ranges are maps that only draw the boundaries of areas actually or potentially occupied by a species. Such outputs are very frequently based on data of unknown accuracy and precision, that may be out of date or of a very diverse nature (Karl *et al.* 1999). This technique is a simplification of the actual situation, as it often fails to consider holes within the range boundaries where a species does not occur or islands beyond the perimeter where the species does occur (Brown *et al.* 1996).

2.2.2 The internal structure of the geographic range

The internal structure of a species' geographic range is highly dynamic, being dependent on the relationships between its biological and behavioural characteristics and the biophysical environment both at spatial and temporal levels (Brown *et al*. 1996). A species' range can be continuous or interrupted by so-called *biogeographic barriers* that limit the movement of individuals between places separated by some physical features of the environment. The barrier can be represented by a newly-developed motorway or a wide river, as well as a mountain chain or a cultivated field. Depending on the target species, a spatial feature can represent a limiting barrier that shapes its geographic range or simply a feature that causes the latter to be discontinuous.

Geographic ranges represented by *disjunct* areas may result from long-term climate changes or continental drifts as well as environmental catastrophes (e.g., volcanic eruptions), and human disturbances (e.g., urban developments, motorways, dams).

The risk of isolating populations by habitat fragmentation is directly linked to the risk of extinction, as the size of the patches and the connectivity among patches are the primary variables upon which the viability of a population depends (Gilpin and Soulé 1986).

2.2.3 Corridors

Patches of suitable habitat can be connected by *biocorridors* (hereafter called corridors). A corridor is ‘a linear two-dimensional landscape element that connects two or more patches of wildlife habitat that have been connected in historical time...’ (Soulé and Gilpin 1991).

Corridors give habitat patches the property of connectivity, which can affect both demographic and genetic processes. There are different reasons why connectivity between patches can be vital for wildlife populations, and these are a function of the species characteristics. For example, the size of a single patch can be too small for the population to be viable (Gilpin and Soulé 1986). Examples of corridors in the landscape include hedges, shelterbelts, roads and powerlines (Lavers and Haines-Young 1993).

2.3 Issues of scales in ecosystem management

The degree of habitat fragmentation has reached alarming levels in some areas, sometimes threatening the viability of original wildlife populations (Noss 1987). As the process of fragmentation increases, so too does the conflict between human activities and the preservation of natural habitats. This has driven forward efforts to study the spatial distribution of species and habitats. Particular emphasis has been placed on modelling the species-habitat relationships and the spatial representation of degree of habitat fragmentation. For the management of natural resources to be effective in the long-term, the understanding of species-habitat relationships should be at least at the landscape level (Petch and Kolejka 1993: 42).

The description, study and management of landscapes become difficult when considering the multitude of components that contribute to their functioning (Perez-Trejo 1993). Within the extent of landscapes, the identification of *priority* areas (e.g., those areas with high degrees of biodiversity, see Myers *et al.* 2000), or high development potentials, or where rare species have been recorded, or where economic and political interests are determinant), together with their size and location, has become of paramount importance in any management plan (Stoms and Estes 1993), so as to focus on specific sites that are representative of the landscape's biodiversity. Nevertheless, the focus of a large number of studies during the last few decades has been the understanding of species-habitat relationships at local level. Their contribution to wildlife conservation is vital for the design of any management measure, but the information they provide are seldom applicable to areas other than the one they were described for (Scott *et al*. 2002).

Considering a species across its whole distribution range as the target of conservation practices can be very difficult if the range is a broad one and extends over diverse landscapes, because individual populations may adapt to locally-prevailing environmental conditions. This makes generalisations across the entire species' range difficult (Primack 1998). For this reason, the operational unit for wildlife conservation is often an individual population. Complete wildlife studies should include the processes that take place within the populations' boundary. Although this may seem to be the best possible spatial resolution for conservation activities, in the real world it is very rare to deal with geographically-isolated populations that show clear spatial boundaries.

Therefore, a population may be considered to be either a functional group of individuals (a herd of deer of a pack of wolves), or a completely arbitrary designation dictated by administrative boundaries or landscape units (Soulé 1987). The application is further complicated by the set of biological interactions that produce non-homogeneous spatial distributions even against apparently homogenous landscape units (Gilpin 1987). The selection of spatial scales that are needed to describe a particular aspect of species' range is vital to achieve the objectives of a study. It is very important to note that a *right* scale does not exist, as different scales are suited for different purposes (Trani 2002). The fundamental concept to recognise is that processes occur at all scales at the same time and that phenomena taking place at different scales may interact to produce the final picture that we see at the selected resolution (Levin 1992). Morrison and Hall (2002) define scale as ‘the resolution at which patterns are measured, perceived, or represented’. This implies the fact that scales exist because there is an observer who sets them (Maurer 2002). It must not be forgotten that different factors may drive the processes that occur at different scales (Wiens 1989), so that the variables identified for a particular process may be dependent on the spatial resolution selected.

There are two attributes of scale: the spatial resolution and the extent (Morrison and Hall 2002). The spatial resolution refers to the smallest identifiable unit on the ground, while the extent is the size of the study area. In the process of selecting an appropriate scale for studying a given species, attention should be paid to the species’ biological characteristics (Trani 2002) as different wildlife species perceive the environment at diverse spatial scales.

The trophic level that a species occupies within a community is usually associated with its range size such that species at higher trophic levels have larger ranges. Within the mammals, carnivores’ ranges are the largest (Watts 1984), although this varies according to social organisation and population structures. Thus the same area may be occupied by species with very different requirements and patterns of habitat use, living at different spatial scales, and having very different perceptions of the very same geographical space. Some of them inhabit areas that may include the ranges of many other species, as in the case of large carnivores. Defining the right scale for describing patterns and predicting species occurrence should consider the target species’ home range, seasonal area use and landscape patterns influencing their ecology. Not least, the appropriate scale depends on the study’s objectives (Trani 2002), and applying the outputs of a model developed at one scale at another scale may lead to misleading results (Heglund 2002).

2.4 The species-habitat relationships

The spatial distribution of species depends on their requirement and their response to environmental characteristics (Elton 1927). Traditional wildlife management considers three fundamental physical variables of the environment, which represent basic vital requirements: food, water and shelter (Dasmann 1964). The local conditions and available resources contribute to define the *habitat* of a given species (Morrison *et al*. 1998), and the presence and survival of species are directly dependent on environmental conditions (Anderson and Gutzwiller 1994). The identification of key habitats for wildlife species is essential for development programmes, where drastic land changes could cause the disappearance of some environmental structure (Litvaitis *et al*. 1994)

When sampling wildlife environmental preferences for management purposes, the consideration of all areas occupied by a species can be difficult if basic information about the species' life history are not available. Behavioural studies give a fundamental contribution to conservation and little management action can be successful without knowledge of species' habitat requirements, range size, mating system, and inter-specific relationships (Curio 1996, Sutherland 1998).

Although the description and understanding of all requirements of any species is often impossible because of costs and time involved, some key factors that strongly influence the distribution of many species may be identified and used for modelling the relationships between the species and the environment they live in. The quantification of the relationships between a species and the environment it inhabits represents one of the key aims of environmental management. Species-habitat relationship models give a representation of *goodness* of habitat patches for any target species, and predict the probability of detecting a species, given a set of environmental conditions (Stauffer 2002). They can be developed as binary system (i.e., suitable/unsuitable), ordinal (i.e., high, medium, low) or ratio (i.e., index scores) values (Stoms and Estes 1993). The fundamental assumption underlying these models is that once the key environmental variables have been identified, the distribution of a species can be estimated by knowing the distribution of such variables (Scott *et al.* 1993).

With the relatively recent development of geographic information tools, it has become easier to represent the spatial distribution of environmental variables and produce visual presentation of spatial models as maps of habitat suitability or probability of occurrence of the species. Geographically-explicit models have become extremely powerful tools for representing the species-habitat relationship and they are extensively used in applied contexts for management. A variety of statistical and mathematical models have been developed in the last decades for the representation of species-habitat relationships (Guisan and Zimmerman 2000) and in section 2.6.2 a review of the most commonly used ones will be given.

The present study will make full use of the management and analytical abilities of geographic information systems for modelling the spatial distribution of suitable areas for large carnivores in the Carpathians by adopting a series of techniques for graphically representing wildlife-habitat models.

2.5 Geographic Information Systems and wildlife conservation

Geographically-explicit models are represented graphically as maps showing the distribution of areas associated with different intensities of the relationship being modelled. A map can be considered as an analogue depiction of the Earth's surface that links features in a spatial context. The mapping sciences (i.e., geodesy, cartography and photogrammetry) have developed highly sophisticated tools for accurately recording and representing the location of physical, natural and anthropogenic features and processes, and the rapid development of computing tools has enabled the handling of vast amounts of information coming from remote sensors. One of them is represented by the software that enables the development of Geographical Information Systems.

2.5.1 Geographic Information Systems

Geographical Information Systems (GIS) provide the opportunity for analysing large data sets with quantitative and qualitative approaches. They are nowadays computerised, though the first GIS were nothing other than the superimposition of maps on transparencies for the combination of target variables.

Technically, in a GIS spatial information is stored in a numerical format, allowing a wide range of mathematical and statistical analyses of various degrees of complexity and sophistication (Star and Estes 1990). GIS are tools that can be personalised according to the operator's needs and interests. They can be thought of as static systems if their only purpose is the representation of variables in the spatial dimension, i.e. for producing cartographic maps, or dynamic systems when their analytical potentials are applied to modelling processes in space and time (Stow 1993).

A GIS is typically a system able to store and access data, and can be thought of as a set of working practices, management structures and data organised to use the spatial data-handling functions of a software/hardware system so as to solve a users' problem. Its peculiarity lies in its ability to analyse the set of spatial information so as to obtain output maps that are not simply descriptions of the single elements, but rather are the result of spatial correlation among them (Johnston 1990). This particular attribute of GIS, together with the possibility of combining information coming from diverse sources using the spatial component as relational element, makes it a suitable tool for developing models and simulating processes in diverse contexts.

2.5.2 Relevance of GIS in wildlife conservation

The landscape approach to environmental conservation and management has greatly benefited from the development of efficient GIS, as large areas can now be covered with relative ease offering a synoptic view. The possibility of overlaying maps and developing spatial models that identify environments at various scales has contributed to the progress made in biogeography and environmental management. This led to the development of a trend in predictive geographical modelling in ecology, the core of which is represented by the quantification and geographical representation of species-environment relationships (Guisan and Zimmerman 2000).

GIS has been used extensively in the last two decades for representing the distribution of conservation targets. The identification of biodiversity *hot spots* is one example of a successful application of GIS at global scale (Myers 2000). During the 1980s GIS was used in wildlife conservation mainly for descriptive purposes and production of maps of species geographic range (Johnston 1998), but in the last decade its use has been focused on a more analytical approach, which has then been used, for example, to plan reserve networks (Groves *et al.* 2000) or devise management strategies (Gurnell *et al*. 2002). The analytical system has been improved by the development of software complements that enable the users to address specific issues like habitat fragmentation and wildlife migration (Akcakaya 1995), potential vegetation mapping, and environmental impact assessment (Corsi *et al.* 2000).

2.6 Wildlife-habitat models

The relevance of mapping environmental suitability for some key species for the conservation of biodiversity appears evident and the practice of developing wildlife-habitat models has given a significant contribution to such practice. Wildlife-habitat models can be of different nature and be produced adopting different approaches, depending on the objectives of the study and the available data. The present study uses GIS-based modelling for mapping conservation areas for carnivores in the Carpathians and the technique adopted has been selected among a variety of available ones. In the following sections I will review some of the most common types of models.

2.6.1 The modelling approaches

A wide variety of statistical models is currently in use to represent and simulate the spatial distribution of terrestrial animal and plant species as well as biomes and processes, and GIS are being increasingly used to model wildlife-habitat relationships (Davis *et al.* 1990, Scott *et al.* 1993). The selection of significant parameters is essential for successfully model any process, thus the identification of causal variables is the most critical step in model development (Guisan and Zimmerman 2000).

Models used in ecology have been classified repeatedly using different approaches, but the main differences are consistent throughout different classifications. Stoms *et al.* (1992) define both deductive and inductive approaches for modelling wildlife habitat with GIS. Their classification is comparable to the one offered by Loehle (1983), which recognises two types of models: calculating tools (inductive *sensu* Stoms *et al.*), empirical models intended to describe the configuration of the real world (Guisan and Zimmerman 2000), and theoretical models (deductive *sensu* Stoms *et al.*), synonymous with the mechanistic ones in Levins' (1969) classification, that are capable of predicting scenarios from knowledge of causal variables and relations (Guisan and Zimmerman 2000).

In the deductive approach, the environmental requirements for the target species are known *a priori* and suitable habitat is identified through mapping the distribution of environmental characteristics. The output is a map of the distribution of *potential* habitat for the species, as the actual presence of the species at each location needs to be verified after modelling (Stoms *et al.* 1992). The inductive approach describes the relationship between the target species and the environment by direct observation of environmental characteristics of locations where the species' presence has been recorded. This approach is frequently used in empirical models *sensu* Levins (1969), the major aim of which is to understand the actual geographic distribution of a species rather than project its potential one according to established parameters. In this case the importance of having a measure of accuracy of the species' locations used in the model-building process appears evident (Stoms *et al.* 1992). This is not a trivial issue and will be discussed in section 2.7.

2.6.2 Spatial models for wildlife conservation

Wildlife-habitat models are developed upon the assumption that predictable relations exist between the occurrence of a species and some selected variables of the environment (Heglund 2002). In addition to being inductive or deductive, models can be descriptive or predictive (O'Connor 2002). The former aim at estimating a qualitative/quantitative relationship between the species' presence and some environmental variables in known areas and use such relationship in order to predict the species presence in unsurveyed areas. The predictive models are expressed in probabilistic terms and aim at predicting changes in specie's distribution under variable environmental conditions (O'Connor 2002). The majority of wildlife-habitat models fall into the first group. Such models can be used for habitat evaluation in areas subject to development, with the aim of minimising the impact on wildlife and selecting areas essential for wildlife survival (Scott *et al.* 1993).

2.6.2.1 Habitat Suitability Indices

In the last decades there has been a considerable amount of work done towards the identification of relationships that would express how species relate to the environment, thus guiding towards the conservation of particularly valuable areas. One approach used extensively for this purpose is the development of *Habitat Suitability Indices* (HSI). They were developed by the U.S. Fish and Wildlife Service (USFWS) in the attempt to establish linear relationships between species and environmental variables (Conway and Martin 1993, Donovan *et al.* 1987, Duncan *et al.* 1995, Thomasma *et al.* 1991) in a standardised way across all the United States (U.S. Fish and Wildl. Serv. 1981).

HSI are models that incorporate a number of environmental variables considered to be important for the presence of a given species. They are related to the species presence in a quantitative way using data from field studies and combined spatially in a GIS (Donovan *et al.* 1987). Despite representing a step towards standardisation and objectivity, the HSI contain a great deal of subjectivity in various steps of their development. The selection of significant variables is left to subjective decision by scientists and experts who have conducted field studies (Scott *et al.* 1993). More importantly, the weight that each variable is assigned can be highly subjective and strongly location-dependent (Heinen and Lyon 1989). The applicability of the models is therefore restricted to specific areas and generalisation can hardly be done unless the model is developed upon information coming from a large number of sites throughout the species' range (O'Neil *et al.* 1988).

Habitat Suitability Indices are combinations of regressions between an environmental variable thought to be a limiting factor for the species' survival and the species presence according to scores associated with each variable by experts studying the target species. Although HSI are only simple models that tackle the multivariate nature of wildlife-habitat relationships as a series of univariate components, their development has represented a starting point for a whole set of models recently developed for mapping potential distribution of species. The major limitation of conventional HIS is that they assume linearity between the species presence and each environmental variables, condition rarely met in nature (Heglund 2002). Their use has strongly been criticised because of the lack of a validation process that could assess their reliability. They are not usually represented as maps, but rather as regression lines and/or equations.

Coupling GIS and HSI enables the extrapolation of environmental characteristics over inaccessible areas and the interpolation and representation of the models across extensive regions, assuming the model is suitable at such a scale. Furthermore, the integration of GIS in the development of HSI has enabled the consideration of environmental heterogeneity over large areas and the incorporation of biophysical variables that are otherwise difficult to measure (i.e., topography, distance from landscape features etc.). This has represented an advance in the consideration of the whole species' range and its internal structure, in an exercise called *Habitat Evaluation Procedure* (HEP) that includes basic information on a number of species in a generalised way so as to model environmental changes following the alteration of few variables (Williams 1988).

Including HSI in management plans at regional level is very important, because they can guide the identification of critical areas for the survival of endangered species. For example, Conway and Martin (1993) suggested that the best management practice of woodlands in Arizona should consider the selective conservation of large dead tree trunks, required for nesting by Williamson's sapsucker (*Sphyrapicus thyroideus*), as revealed by development of the HSI.

The spatial scale at which the HSI are applied within a GIS is important because inappropriate scales would fail in the identification of small patches of highly-suitable habitat for species that use small areas (Duncan *et al.* 1995). Baker *et al.* (1995) noted that the investigation of habitat suitability should be made at different spatial scales. Their study on sandhill crane (*Grus canadensis tabida*) habitat benefited from the use of GIS tools at all stages, from selecting sampling points to spatially interpolating habitat characteristics and modelling the HSI at five resolutions. The authors pointed out that restricting the analysis to the areas known *a priori* as being potentially suitable (i.e., wetlands in this case) would save time and contribute to a better understanding of the species' ecology.

Habitat Suitability Indices can be used for modelling the impact that environmental changes may have on the target species. Pereira and Itami (1991) used HSI for red squirrel in Mt. Graham, Arizona, to assess the potential impact a project for constructing an observatory in the area would have had on the squirrel' habitat. The study was also an opportunity to test the HSI with data collected in the field rather than through expert knowledge. The authors initially used univariate analysis for selecting those variables to be included in the model. Logistic multiple regression was then used for developing an environmental model that was subsequently integrated with a trend surface one in a Bayesian approach, adding a probabilistic component to the modelling procedure, and tackling the issue through non-linearity approach. Kliskey *et al.* (1999) simulated the impact of different forest management scenarios on the landscape using HSI for marten (*Martes americana*) and caribou (*Rangifer tarandus*) over a period of 120 years. The process included an HSI testing phase that suggested some additional variables may be included in the model, although the quality of data used for testing was not optimal.

Another significant limitation of HSIs lies in their inappropriateness for generalisation. They are usually built on information collected over restricted areas and the resulting wildlife-habitat relationship may well be highly site-specific. The application of HSI on large areas, at the regional or state level, is often difficult and leads to large commission errors (areas where wildlife presence is predicted but not detected; Block *et al.* 1994). Evaluation of HSIs is essential before they can be used by wildlife managers, and testing should be done using as many study sites as possible, in order to account for most of the species' ecological valence (O'Neil *et al.* 1988).

2.6.2.2 Gap Analysis

Notwithstanding the limitations of the HSI, they were extensively used by the USGS Biological Resources Division when the innovative Gap Analysis Programme (GAP) was launched in the early 1990s as a mean for identifying areas with high biodiversity lacking appropriate protection (Scott *et al.* 1993). GAP uses GIS mapping procedures for giving a quick overview of the biodiversity conservation status of large areas by overlaying vegetation maps and distributions of some biodiversity indicator species (i.e., butterflies and some vertebrates). The objective of the programme was to identify gaps of protected land within areas of high biodiversity at a regional scale (Scott *et al.* 1993). Known wildlife distributions were interpolated in areas not surveyed relying on the assumption that vegetation composition and structure can be used as proxies for wildlife presence under the relationships established by the HSI (Edwards *et al.* 1995).

The approach was new in its consideration of landscape-sized samples, as opposed to the more common local views where single species were considered over small areas. Although appropriate for the magnitude of the project (state-wide across the U.S.A.), it must be noted that, at a local level, the analysis may fail in considering the internal structure of ranges that fall entirely within the Minimum Mappable Unit (Tamis and Zelfde 1998), which was set accordingly to the objective of continental mapping at 100 ha. Nevertheless, the advantage of the whole exercise is to provide a general baseline map that will serve as starting point for detailed studies at local scales and the spatial resolution and the extent adopted satisfied the objectives set (Scott *et al.* 1993). In this context, the present study represents a valuable contribution in this direction, through the production of baseline maps that will represent the spatial distribution and character of areas associated with different degrees of suitability of the presence of large carnivores, used as proxies for biodiversity.

The product of GAP is a map of areas with high biodiversity that are lacking legal protection. The visual representation is extremely powerful and valuable for management purposes. Notwithstanding the many limitations it has, GAP still represents one of the greatest attempts in the assessment of biodiversity conservation status across broad areas, because it integrates GIS with habitat models and remotely-sensed data to provide the basis for setting up conservation projects at local and regional scales (Jennings *et al.* 1997). The benefits of such an approach were pointed out by Davis *et al.* (1990), who underlined the urgent need for a biodiversity conservation information system. The integration of information from different sources and in different formats can sometimes prove difficult, but it represents the only way that such sparse information can be possibly used (Davis *et al.* 1990). This is also addressed in the present study.

2.6.2.3 Continuous models

The representation of habitat suitability has often been modelled with a deductive approach producing deterministic discrete models (e.g., HSIs). The real world is very rarely structured in a discrete manner and, particularly when considering environmental characteristics, the continuous distribution and regionalised behaviour of variables should be taken into account. For this reason, continuous models based on probability theory and fuzzy logic tend to be preferred by ecologists (Hill and Binford 2002).

In the process of developing a model to assess the distribution of wildlife habitat, the initial steps that contribute to the establishment of rules to follow in the decision-making phase are particularly critical because they drive the interpretation process once the output is produced.

As already mentioned for the development of HSI, the first problem to tackle in any modelling process is the selection of environmental variables to include in the model. These are often chosen based on the knowledge of field experts and sometimes a set of univariate regressions are used for eliminating redundant variables or to establish the relationships between each variable and the species' presence or abundance (Schamberger and O'Neil 1986). The implication with this procedure is that correlated variables may yield significant regression results and are incorporated in the model as independent variables, while the contribution of one of them is redundant because the variates are actually co-variates (i.e., they are correlated). The use of multivariate statistics is more appropriate for tackling problems that obviously are of multivariate nature (Manly 1994, Sokal and Rohlf 1995). Nevertheless, it must be noted how multivariate statistical inference can sometimes give misleading results. Rexstad *et al.* (1988) indicated that the interpretation of coefficients from multivariate procedures is often arbitrary and meaningful *a priori* ecological thinking can very rarely be replaced by sophisticated multivariate statistics techniques. Multivariate statistical methods often require that data meet rigid assumptions, but in some cases the biological justifications of some broken assumptions may be used (O'Connor 2002).

Methods for selecting variables used in multivariate modelling include both inductive and deductive approaches, as well as mixed approaches. Herr and Queen (1993) considered expert knowledge-based information and calculated some variables in an inductive manner from plotted nest locations of sandhill crane in a GIS. A set of chi-square tests were performed for checking significant differences in the distributions of expected and observed sandhill crane nest locations with respect to each variable considered at a time. The authors recognised that this step-wise process would not be appropriate as some of the variables that were considered were correlated, thus they performed a general chi-square using only a combination of few selected variables.

Once the variables are selected, the modelling approach can be of probabilistic nature or fuzzy logic, depending on the approach adopted, and more often, on the data available for building the model. Augustin *et al.* (1996) developed an autologistic model for predicting the presence of red deer in Scotland using data scattered across the area and calculating the probability of presence in each cell of the digital data set by interpolating from data in neighbouring cells in a deductive manner. The authors present two supplementary methods for estimating probability of occurrence of deer in each cell, including the Gibbs sampler - which gives estimates in a reiterative way for each cell - and a combination of the autologistic and Gibbs methods. The main advantage of such an approach is the consideration of autocorrelation between locations where the presence of wildlife is recorded. This is particularly important for species that live in big groups, such as deer herds. The advantage of the Gibbs sampler (as well as any other re-iterative sampling method) is the ability to amplify the data set, thus offering the opportunity to estimate the accuracy of the model.

An inductive approach assumes that the environmental variables are not known *a priori*. A GIS can be used for extracting such variables from known locations of the target species. A principal component analysis could successively guide the reduction of dimensionality of variables to be modelled (Buckland and Elston 1993).

The availability of wildlife data in a presence/absence format sometimes makes their use for statistical analyses difficult. Methods often used, such as that generalised linear models (GLM) are not usually able to estimate satisfactorily the distribution of wildlife populations as successfully as a simple count of individuals would (Guisan and Zimmerman 2000). Nevertheless, Buckland and Elston (1993) successfully modelled red deer (*Cervus elaphus*) census data in Scotland using a GLM with a logistic link function, that could be built with data of presence/absence.

Walker (1990) used logistic regression to map the distribution of three kangaroo species in Australia against climate parameters. The model was built with an inductive approach using a Classification and Regression Tree (CART) function within a GIS that established decision-rules, as well as a probability function (logistic regression model) for mapping the probability distribution of the species. A comparison between the two methods revealed that commission error (areas where presence of kangaroos was predicted but not recorded) was higher with logistic regression than with CART.

Multivariate techniques are more frequently used in inductive approaches rather than deductive, as the extraction of basic information from wildlife location is relatively easy. Aspinall (1992) used deductive spatial modelling based on Bayes' theorem for describing the distribution of red deer in Scotland. The method assumed *a priori* probability of the presence of particular environmental characteristics by estimating the probability of their occurrence in observed red deer presence/absence locations. The variables to be included in the model were selected using chi-square tests for presence against random locations. The model produced a probabilistic habitat suitability map with an overall accuracy of 70%, when compared with census data for red deer. The main limitation of this approach is a methodological one in that Bayes' theorem assumes independence between covariates, a condition rarely met in ecology.

A number of wildlife studies that have successfully applied a multivariate method based on the concept of similarity to some optimal conditions, classifying regions according to the distance to a given set of environmental conditions. One such method is the Mahalanobis distance, which will be used in the present study. The method is commonly used to assess environmental suitability as indicated by the distance of each point in the area considered from a reference point that represents the *optimum*. It also can be used for guiding the selection of predictor variables by the maximisation of the Mahalanobis distance between known *good* and *bad* areas (Johnson *et al.* 1998). Hill and Binford (2002) define the models based on distance classifiers as belonging to the fuzzy logic family, as they do not use the probability theory, but rather the decision rule is based on the concept of similarity.

The Mahalanobis distance was successfully used for mapping potential suitable habitat for black bear in Arkansas (Clark *et al.* 1993), grey wolf in Italy (Corsi *et al.* 1999) and large carnivores in the Alps (IEA 1998) and in the Scandinavian Peninsula (Støbet-Lande *et al.* 2003). The method builds an inductive model from wildlife observations plotted on maps of environmental variables within a GIS. The areas actually occupied by the species are assumed to represent the optimum combination of environmental parameters for that species. This is a strong assumption that will be discussed in chapter 5, when evaluating the strengths and weaknesses of the method adopted in the present study. The Mahalanobis distance then uses this *optimal* combination to predict species distribution. The training data required to build the model do not have to be in the format of presence/absence. This is an advantage, as absence data are difficult to record. The values obtained are dimensionless as they are a function of standardised variables. The technical details of this method will be given in chapter 5.

A deductive-analytical model was developed by Pereira and Duckstein (1993) using multiple-criteria decision-making (MCDM) techniques for assessing land suitability. The authors found the methodology to be suitable for overcoming the limitations often present in land suitability approaches: inappropriate data scaling and lack of independence among factors. Compromise programming was used to identify the set of optimal conditions. The Mahalanobis distance was then used for estimating degrees of suitability as represented by distance values from the optimum point. The method was used to assess habitat quality for the endangered Mt. Graham red squirrel in Arizona.

Knick and Dyer (1999) used the Mahalanobis distance for estimating habitat suitability of black-tailed jackrabbit (*Lepus californicus*) in South-western Idaho. The results obtained were validated with an independent data set and found to well represent the distribution of areas where jackrabbits were found. The authors emphasise some of the limitations associated with the method, which were highlighted in a previous work (Knick and Rotenberry 1998). Such limitations are generated by the use of a set of presence data that is assumed to represent the optimum combination of the variables considered. This is a strong assumptions that may produce misleading results when data on species' presence are collected over areas that are not optimal. This will be further discussed in chapter 5. In this case, the outputs of Mahalanobis distance may consider optimal areas as being different from the mean 'optimum' vector, resulting in large omission errors (Knick and Rotenberry 1998, Rotenberry *et al.* 2002).

In spite of such limitations, the method has proved suitable for applications with data sets recorded at different scales and it is particularly useful for species with large ranges and generalised habitat requirements such as large carnivores (IEA 1998).

Hirzel *et al.* (2002) recently proposed a multivariate method that takes into consideration the n-dimensional niche and estimates that a habitat suitability probability map is the one based on the comparison between the characteristics of presence areas and the whole study area. In principle, it is very similar to the

Mahalanobis distance method, but it gives a weight to the different environmental factors through their vectorial dimensionality, thus accounting for the ecological characteristics of the species.

New techniques for the estimation of wildlife habitat suitability are continuously being explored, sometimes borrowed from other disciplines. One of them, very little explored in this field, is the Artificial Neural Network (ANN). The main strength of ANN is that they assume neither data normality, nor linearity in the response of wildlife presence to environmental variables (Lek and Guégan 1999). Nevertheless, their ability to produce accurate predictions is very much dependent on the training data, and Lusk *et al.* (2002) found that a high variability in the training data may cause the ANN to perform poorly when compared, for example, with a multiple regression model.

2.7 Model evaluation

Modelling wildlife habitat in a GIS has some intrinsic limitations that are mainly related to data quality. Although a range of models has been developed for a large number of species, their application for management and conservation practices is still restricted by lack of testing and validation (Schroeder and Vangilder 1997).

2.7.1 Sources of error

Input error associated with every information layer is kept in a GIS and it contributes to the model output overall error. Sources of error in GIS modelling are present at different levels and in various forms (Burrough and McDonnell 1998). The accuracy of the base maps is particularly important for highly developed areas, where the landscape changes rapidly as a consequence of human impact. Depending on the variables extracted from the base maps, the dating of maps can be of variable relevance (Guisan and Zimmermann 2000). Maps are frequently used as source data and the process of digitising paper maps brings a considerable error that results from the distortions caused by the physical nature of the paper and the digitising process (Burrough 1986).

Other errors in the data used for building models include the density of observations and their positional accuracy; as well as the variation brought by different people collecting the field data. Data input is frequently associated with errors and cross-checking by different people may be advisable. Map scale and spatial resolution are sources of error that need to be taken into consideration, as any spatial model will provide an output with a spatial accuracy equal to the smallest scaled input data layer (Star and Estes 1990). Geographic projection and coordinate systems are associated with errors, and when data layers originally created in different coordinate systems are overlaid, the coordinate conversion process is made through an interpolating algorithm, bringing in some error. When layers at different spatial scales are used, they need to be interpolated to the same scale and in order to be used for modelling, they frequently need to be in raster format. Re-scaling and rasterising are forms of interpolation that have usually some combined errors (Burrough and McDonnell 1998).

Single operations such as map overlays are associated with errors, and the operation itself assumes that the input maps are perfect representations of the real world, a condition rarely met (Arbia *et al.* 1998). Wildlife models are often built upon expert knowledge, itself a source of subjective interpretation and error. The presence of error is therefore inevitable in environmental modelling with a GIS, but, as long as this error is defined and is acceptably low, it is a minor disadvantage greatly out-weighed by the advantages a model output represents. However, it is important to consider as many sources of error as possible and quantify the input error associated with each one of them, possibly including the potential multiplication and propagation effects that GIS operations are associated with (Heuvelink 1998).

2.7.2 Model Validation

Once the model is applied to an area, and the habitat patches are rated for their suitability, field surveys can be made to check for the species' presence and only at that stage should any management decision be taken (Stoms *et al.* 1992).

HSIs have been developed for over 160 species across the USA, but they have rarely been validated. Donovan *et al.* (1987) proposed a test for validating HSI for the fisher (*Martes pennanti*) in Michigan. The authors made a field survey and calculated a Preference Index (PI) by ratioing the percentages of habitat use and availability within each HSI class. The HSI and PI were overall positively correlated, although some differences emerged according to habitat types (i.e., pine plantations had high HSI but were poorly used by fishers). This suggests how site-specific the HSI are. Because of the nature of HSI, where the variables that determine the outcome need to be known *a priori*, its cartographic modelling should be considered as a hypothesis-formulation process rather than a scientific truth achieved by a hypothesis-testing procedure (Johnston 1998).

Schroeder and Vangilder (1997) tested a set of HSI for five different species, all based on oak-mast production. Field surveys and monitoring of oak-mast production in two areas in Oregon were used as testing data sets. The authors concluded that the 5 HSIs were positively correlated with the oak-mast production model they developed, but suggested the HSIs may be modified with the inclusion of oak canopy closure variable.

Duncan *et al.* (1995) validated the HSI for Florida scrub jay using data on the species' demography from three successive years, acknowledging the potential stochastic variability associated with wildlife demography. The areas associated with high suitability correlated well with the demography of the jay, and the authors suggested that the use of demography data could be useful for identifying potential population sources and sinks[1], thus guiding the wildlife managers towards selective protection of areas that may have different demographic roles.

Detailed data on a species' demography are often not available to model developers and wildlife managers, making this kind of validation process sometimes difficult. Particularly when models are developed with an inductive approach, the presence/absence data are often the only available ones and they are used for extracting the information about wildlife-habitat relationships. The validation of such models can then be performed either using newly collected field data or expert knowledge. The former option is usually very expensive and time consuming, while the latter may be site-specific and very subjective. In the present work it was possible to collect independent data for validating the outputs produced through an intensive field campaign. The methods and results of validation will be explained and presented in chapters 4 and 5.

Model validation should always be performed because the recognition and quantification of errors is a vital requisite for robust model development. The validation phase should be part of the modelling process, ideally at all stages (Morrison *et al.* 1998). When applied for management, models should be used considering their ability to reproduce the real situation and their major limitations

[1] The concept of population sources and sinks refers to areas where the population is healthy and can successfully reproduce (a source). This produces individuals that can locally disperse into areas with lower habitat quality, and potentially undergoing processes of local extinction (sinks), where the dead individuals are replaced by new ones coming from the source instead of being replaced by new locally-born ones (Pulliam 1988).

(Guisan and Zimmermann 2000, Morrison *et al.* 1998). This ensures no misleading interpretations are made (Block *et al.* 1994).

2.8 Conclusions

The consideration of the spatial distribution of priority areas for conserving biodiversity is particularly relevant when efforts are being made for preserving and integrating wilderness needs with human economic and social demands on the land. The EU has highlighted the need for a pan-European approach to Biodiversity conservation that is based on transparent, consistent and accountable approaches, thus encouraging the development of management tools (EU 2001). These are continuously calling for rules that enable managers to make decisions within an ecologically sound and scientifically robust approach. This has stimulated the development of a series of mathematical and statistical models that may contribute to the production of pragmatic management plans.

Models are a simplification of the real world and as such they should be used with caution and with cognition of their limitations and associated errors. Notwithstanding their limitations, they still represent a very effective tool for understanding the processes and generating hypotheses to be tested. A plethora of techniques has been developed and each one of them may present limitations that make it more or less suitable for any application, depending on the data set available and the purpose of the modelling process. Reliability of models depend on many factors, starting from the spatial scale at which they are developed and applied, up to their ability to represent the real world, assessed through a validation phase.

In the light of the issues discussed in the previous sections, the aim of the present study is to contribute to the selection of priority areas for the conservation of Carpathian biodiversity. This will be done through the development of a model for the distribution of conservation areas for large carnivores in the Carpathian Mountains, assessing its robustness at different spatial scales as well as its reliability for predicting the real situation. These will be achieved by validating the outputs against an independent data set. The accomplishment of such aim was reached through the achievement of a number of objectives throughout the study. They are outlined below:

- *Establishing contacts with collaborators and interest groups in the target countries.*
- *Being integral part of the effective network of scientists and experts within the Carpathian Ecoregion.*
- *Populating the data base on environmental variables and carnivore presence*
- *Carry out a field campaign in the area and collaborate with local partners in the production of geographical data.*
- *Standardising the data across the Ecoregion to obtain one ecoregional cover for each variable.*
- *Reclassify and validate Land Cover data.*
- *Create and validate species-specific reception regions using techniques of map algebra.*
- *Define environmental suitability classes for each species using the Mahalanobis Distance classifier.*
- *Acquire and pre-process satellite images at 1km and 30m resolution.*
- *Organise and carry out Validation field campaign (The Carpathian Expedition).*
- *Define environmental suitability across spatial resolution and estimate differences.*
- *Pre-process time series of images for unclassified land cover data.*
- *Map and compare results obtained with classified images and vegetation index.*
- *Analyse distribution of existing protected areas in relation to the estimated suitability areas for the three carnivores.*
- *Estimate potential conflicts with human activities across the whole Ecoregion.*
- *Estimate potential conflicts with shepherds over a subset of the study area at 30m resolution.*
- *Discuss results for management purposes.*

Once the objectives will be achieved, the study will represent a contribution to the understanding of the relationship between large carnivores and the environment in an area poorly studied and where wildlife is conserved in relatively good conditions. The next chapter will describe the study area and the methods used for analyses, and introduce to the ecology of the three target species.