Applications of existing biodiversity information: capacity to support decision-making

Fabio Corsi

Applications of existing biodiversity information: capacity to support decision-making

Fabio Corsi

4 October 2004

Promoters:

Prof. Dr. A.K. Skidmore
Professor of Vegetation and Agricultural Land Use Survey
International Institute for Geo-information Science and Earth Observation (ITC), Enschede
and
Wageningen University
The Netherlands

Prof. Dr. H.H.T. Prins
Professor of Tropical Nature Conservation and Vertebrate Ecology
Wageningen University
The Netherlands

Co-promoter:

Dr. J. De Leeuw
Associate Professor, Department of Natural Resources
International Institute for Geo-information Science and Earth Observation (ITC), Enschede
The Netherlands

Examination committee:

Dr. J.R.M. Alkemade
Netherlands Environmental Assessment Agency (RIVM/MNP),
The Netherlands

Prof.Dr.Ir. A.K. Bregt
Wageningen University, The Netherlands

Dr. H.H. de Iongh
Centrum voor Landbouw en Milieu, The Netherlands

Prof. G. Tosi
Università degli Studi dell'Insubria, Italy

Applications of existing biodiversity information: capacity to support decision-making

Fabio Corsi

THESIS

To fulfil the requirements for the degree of doctor on the authority of the Rector Magnificus of Wageningen University, Prof. Dr. Ir. L. Speelman, to be publicly defended on Monday 4th of October 2004 at 15:00 hrs in the auditorium of ITC, Enschede.

ISBN: 90-8504-090-6

ITC Dissertation number: 114

Susan, Barty and Cloclo

Table of Contents

Samenvatting

De crisis veroorzaakt door het versneld uitsterven van soorten heeft geleid tot velerlei wetenschappelijke en politieke reacties. Belangrijk in dit verband zijn de Multilaterale Milieuovereenkomsten (MEAs) die op samenwerking gerichte regionale en globale initiatieven aanmoedigen. Het afsluiten van deze verdragen creëerde een noodzaak om voortgang en resultaten van implementatie te bewaken. Daarmee ontstond een vraag naar gezaghebbende en recente informatie over de status en verspreiding van biologische diversiteit. Eerdere pogingen om aan deze vraag te beantwoorden waren gebaseerd op netwerken tussen overheids en niet gouvernmentele organisaties (NGO's) met als doel een uitwisseling mogelijk te maken tussen informatiehulpbronnen alsmede de status van biodiversiteit te beoordelen en te monitoren.

Deze inspanningen hebben tot nog toe niet aan de verwachtingen voldaan. Dit wordt toegeschreven aan een gebrek aan data en kennis over biodiversiteit. Dit proefschrift stelt dat, hoewel onze kennis en informatie incompleet is, de bestaande informatie in combinate met de juiste analyse technieken het mogelijk maakt om besluitvorming en beheer van natuurlijke hulpbronnen op adequate wijze te ondersteunen:

- informatie is verspreid, maar kan toegankelijk worden gemaakt met een goed gestructureerd en beheerd informatienetwerk ondersteund door elektronische media;
- de waarde van de informatie kan beter begrijpelijk worden gemaakt met toepassing van analytische methodologieën, in het bijzonder ecologische modellering, die gebruik maken van Geografische Informatie Systemen (GIS);
- het potentieel voor informatiegebruik, met inbegrip van het volgen van temporele veranderingen in de status van biodiversiteit, kan worden gemaximaliseerd met een gegevensbeheer infrastructuur die rekening houdt met taxonomische inconsistenties en veranderende gegevensbehoeften en eisen.

Deze thesis stelt dat de bestaande informatie, mits wordt voldaan aan bovengenoemde factoren, de behoefte aan biodiversiteitsinformatie nu en in de directe toekomst op adequate wijze kan ondersteunen. De nu verkrijgbare informatie en analytische hulpmiddelen staan ons toe om schaarse gegevens in besluitvormingshulpmiddelen om te zetten. Specifieke voorbeelden in de thesis illustreren het resultaat van het van het toepassen van data en kennis verkregen van deskundigen in een GIS omgeving.

In het hoofdstuk over de African Mammals Databank worden informatie netwerk technieken en GIS toegepast om distributie van soorten op continentale schaal te voorspellen. Dit levert een context op waarbinnen gedegen onderbouwde besluitvorming plaats kan vinden. Het hoofdstuk Large Scale Model of Wolf Distribution in Italy for Conservation Planning presenteert hoge resolutie milieu-geschiktheidsmodellen voor de wolf voor heel Italië. Het illustreert de toepassing van inductieve modelleringstechnieken aan de hand van een beperkte hoeveelheid veldgegevens, en de potentiële toegevoegde waarde van anderzins verkregen verspreid opgeslagen gegevens. Het hoofdstuk over Expert-based species distribution maps for the assessment of biodiversity conservation status: the Italian Ecological Network toont aan hoe op expert kennis gebaseerde GIS soortsdistributie modellen toegepast kunnen worden voor prioriteitsstelling voor biodiversiteitsbehoud, alsmede hoe voorspellingen van soortenrijkdom kunnen worden gemaakt door meerdere soortendistributie modellen te combineren. Het laat ook het potentieel zien om deze distributiemodellen verder te integreren met bijvoorbeeld patronen van bedreiging en ontwikkelingsplannen.

Gegevensbeheer-infrastructuren die inconsistente gegevens en evoluerende informatiebehoeften toestaan maken het mogelijk om informatie met consistentie over ruimte en tijd te beheren, synthetiseren en te analyseren. De toepassing van dergelijke infrastructuren zal ons toestaan om van kleinere projectanalyses die intensieve gegevensverzameling vereist, te komen naar een informatiehulpbron die veranderingen in biodiversiteitsstatus in de tijd volgt. Met deze stap zal een echte biodiversiteits-monitoring capaciteit van de grond komen.

Summary

The extinction crisis has led to a multitude of scientific and political responses. These include Multilateral Environmental Agreements (MEAs) that encourage consistent and cooperative regional and global initiatives. The need to monitor the progress and results of treaty implementation has increased demand for authoritative and current information about biological diversity, its status and distribution. A number of efforts to build knowledge about biodiversity respond to these demands. Government and NGO networks have been formed to link information resources, and to assess and monitor the status of biodiversity.

The efforts have not yet met expectations. This has frequently been attributed to lack of data and knowledge. However, while the gaps in data and understanding about biodiversity are significant, information and analytical methodologies available today are adequate to support decision-making and natural resource management:

- The information is dispersed but can be made available with a well structured and managed information network supported through the electronic media;
- The value of the information can be better understood with application of analytical methodologies, in particular ecological modelling, which exploit Geographic Information Systems (GIS);
- Potential for information uses, including tracking changes in biodiversity status over time, can be maximised with a data management infrastructure that accommodates taxonomic and data inconsistencies, and evolving data needs and demands.

This thesis argues that with the above factors in place, existing resources can support biodiversity information needs now and in the immediate future. Information and analytical tools currently available allow us to transform scarce data into decision-making tools. Specific examples presented illustrate the result of mobilising experts' data and applying GIS technologies.

The African Mammals Databank applies information networking techniques and GIS to produce species distribution models at the continental scale, a broad-scale result that sets the context within which informed decisions can be made. The Large Scale Model of Wolf Distribution in Italy for Conservation Planning presents high-resolution environmental suitability models for the whole of Italy, illustrating application of inductive modelling techniques to limited locational data, and the potential analytical value of otherwise dispersed data. Expert-

based species distribution maps for the assessment of biodiversity conservation status: the Italian Ecological Network demonstrates how GIS species distribution models derived from expert knowledge can be applied to priority-setting for biodiversity conservation and how predictions of species richness can be made by overlaying multiple species distribution models. It also illustrates the potential to further integrate these models with, for example, patterns of threat and development plans.

Data management infrastructures that accommodate inconsistent data and evolving information needs provide us with the means to manage, synthesise and analyse information with consistency across space and time. Application of such infrastructures would allow us to advance beyond single project analyses that require intensive data-gathering components, towards an information resource that tracks changes in biodiversity status over time. It is this step that will set into motion a true biodiversity monitoring capacity.

Acknowledgments

This thesis summarises the work carried out over many years, including research projects funded from a variety of sources. Many people and organisations have helped to make this work possible. I am particularly pleased that many of the people I have worked with have also become great friends.

I will not provide the detail of a long list of names, because an entire volume would not be enough. However I want each one of you to know that you have my most sincere thanks and gratitude.

Chapter 1:
General Introduction

Documentation of the extinction crisis comes in a number of forms. An example is the IUCN Red List of Threatened Species, a source frequently cited to justify both the extent of the crisis and actions to address it. The 2003 edition reported that of all known species of birds, 99% of the known mammal species and 93% of the known gymnosperm species, almost 17% are threatened with extinction.[1] This figure implies that current extinction rates are between 1,000 and 10,000 times higher than those inferred for most of geological history (May *et al.*, 1995; Pimm *et al.*, 1995; McKinney, 1998).

The response of the political community has been both significant and varied. Most visible are the Multilateral Environmental Agreements (MEAs) initiated over the past decade to encourage consistent and cooperative regional and global responses. The Rio Treaties (Convention on Biological Diversity, Framework Convention on Climate Change and Convention to Combat Desertification) in particular have created a framework within which countries take responsibility for the environmental impacts of human actions while retaining sovereignty over their natural resources.

The need to monitor the progress and results of treaty implementation has been a major factor in the growing demand for authoritative and current information about biological diversity, its status and distribution. A number of efforts to build knowledge about biodiversity respond to these demands. They recognise the need to manage data at the global level, while responding to the fact that policies are set and implementation occurs within the context of national sovereignty. These include data-gathering and analysis, and creation of networks to support movement of information from the gathering point to the points at which decisions about biodiversity conservation are made. Examples of government initiatives intended to ease movement of information include the Clearing-house Mechanism (CHM) of the Convention on Biological Diversity (CBD, 2003) and the Global Biodiversity Information Facility (GBIF, 2003). In the NGO community an example of a cooperative effort to gather and analyse data and information is the IUCN Red List Programme (IUCN-SSC, 2003a), which is supported by a consortium of five information networks and

[1] The 2003 edition reported that over 40% of the 30,000 species assessed are threatened with extinction. The sample is not random, however, as all known species have not been assessed and the IUCN Red List Programme effort to assess full groups of species is not yet complete.

organisations that collectively hold a vast amount of information about species and their conservation.

These initiatives have not yet met expectations, and there is a widely-held belief that existing information about biodiversity is inadequate to support information needs. The gap between biodiversity information needs and existing information is real, as evidenced by estimates of discovered and described species vs. those known to exist. An example that illustrates this gap is the Global Environmental Outlook 2000 (UNEP, 1999), which cites 1.75 million species recognised by the scientific community while an estimated 12.5 million species exist.

However, a wealth of information to support biodiversity assessment and monitoring does exist. This includes a number of lists derived from scientifically rigorous criteria, such as the IUCN Red List of Threatened Species (IUCN-SSC, 2003b) and the CITES Appendices (CITES, 2003). In addition, extensive knowledge about population biology, ecological requirements of species and other pertinent information exists in a variety of sources, including (but not limited to) small databases, grey literature, field notes and the (undocumented) knowledge of research scientists (Cotter and Bauldock, 2000).

Much of the existing information focuses on one unit of biodiversity: species. While many definitions of biodiversity exist, all focus on the three major components of the biosphere: genes, species and ecosystems. Species are the most common units used to measure biodiversity. They are more easily defined than ecosystems and more easily measured then genetic diversity (Gaston, 2000).

Lack of access to and understanding about that information leads to lack of sound, scientific information in decision-making related to biodiversity conservation planning and management. And while the gaps in data and understanding of biodiversity are significant, information and analytical methodologies available today *are* adequate to support decision-making and natural resource management:

- The information is dispersed but can be accessed with a well structured and managed information network supported through the electronic media;
- The true value of the information can be realised with application of analytical methodologies, in particular ecological modelling, which exploit existing technologies (e.g. GIS);
- Potential for information uses, including tracking changes in biodiversity status over time, can be maximised with a data

management infrastructure that accommodates taxonomic and data inconsistencies, and evolving data needs and demands.

This is, admittedly, a broad approach to a circumstance for which resolution requires specificity. In other words, while we must prepare for the uncertainties of the future, *can we support biodiversity information needs of the scientific and political communities right now and in the immediate future?*

This thesis argues that the answer to that question is yes. Existing information and analytical tools currently allow us to transform scarce data into decision-making tools. Specific examples presented illustrate the result of mobilising experts' data and applying Geographic Information Systems (GIS) technologies. Species distribution models and habitat suitability indices are created to support natural resource management and planning. Further, by systematising and standardising existing data in a data management model described here, we can advance a step beyond single project analyses that require intensive data-gathering components and create an information resource that is available for multiple uses and analyses. It is this step that will allow us to track changes in biodiversity status and distribution over time, setting into motion an effective biodiversity monitoring tool.

Chapter 2, *Mobilising Dispersed Information for Biodiversity Conservation,* describes the challenges and solutions to accessing existing but dispersed biodiversity information and ensuring its quality. Mobilising this dispersed information requires management of multiple sources (the experts) and formats (data bases, grey literature, and undocumented knowledge). The experts must be provided with an environment and incentives to motivate their participation, including assurance that the information and products they contribute to are quality assured.

Chapter 3, *Species Distribution Modelling with GIS*, describes use of Geographic Information Systems (GIS) in species distribution modelling. A comprehensive source of biodiversity and its status will remain elusive for many years to come. Emerging methods of species distribution modelling that rely on the increasing availability of geo-spatial data and GIS allow us to apply expert knowledge about species and their ecological requirements to produce predictive models of their distributions. Simply put, GIS supports transformation of limited data into relevant and timely information pertinent to biodiversity decision-making. Chapter 3 reviews the uses, misuses and limitations of GIS applications and proposes approaches to ensure that it is used to best potential.

The chapters that follow demonstrate the full potential of coupling existing information with GIS and other technologies to produce decision-making tools.

In Chapter 4 (*Mobilising biodiversity data in practice: the example of the African Mammals Databank,* AMD), the use of GIS to produce species distribution models at a continental scale is shown. The AMD provides a broad-scale result that sets an important context within which informed decisions can be made. Chapter 5 (*A Large Scale Model of Wolf Distribution in Italy for Conservation Planning*) illustrates application of inductive modelling techniques to limited locational data for the wolf in Italy, and the potential analytical value of otherwise dispersed data. The approach presented produces high-resolution environmental suitability models for the whole of Italy. Management alternatives and implications are assessed. Chapter 6 (*Expert-based species distribution maps for the assessment of biodiversity conservation status: the Italian National Ecological Network*) demonstrates one way that the GIS species distribution models derived from expert knowledge can be applied to support priority-setting for biodiversity conservation. By overlaying multiple species distribution models, prediction of species richness can be made. Potential to further integrate these multi-species models with, for example, patterns of threat and development plans, suggests that the limited data entered into the front-end of the system can produce valuable and relevant analyses for scientifically-based decision-making.

Chapter 7, *Data Management for Biodiversity Conservation*, proposes a data structure and management approach that provides the flexibility and efficiency needed to respond to data inconsistencies and evolving needs. The spectrum of needs and expectations when dealing with biodiversity information are as varied as the number of people interested in or needing to use the information. Further, the basic identifier of species, the taxonomic unit, presents its own set of controversies and inconsistencies. These create a significant barrier to identifying a clear reference point for each piece of information as it enters the system, and for ensuring consistency across time and space.

The data management system proposed here accommodates data inconsistencies, evolving data needs and consistency over time and space. It provides the experts with data management tools that maximise the efficiency of their participation and support their own work. Implementation of this data management system would transform our capacity to assess biodiversity status and distribution into a capacity to monitor changes in that status and distribution over time.

Chapter 8 (*Applications of existing biodiversity information: capacity to support decision-making now and in the future*), the synthesis, analyses the relevance of biodiversity analyses described in this thesis to current biodiversity management and planning processes. It reviews the conclusions reached regarding our capacity to transform existing information into relevant decision-making tools now and in the immediate future. Looking to the future, it argues

that with the information management infrastructure proposed in Chapter 7, we have the capacity to monitor the status and distributions of species over time. This is the key factor for scientifically sound management and conservation of biodiversity, a widely recognised need by both the scientific and political communities.

Chapter 2: Mobilising Dispersed Information for Biodiversity Conservation[2]

INTRODUCTION

The extinction crisis is well recognised in the scientific and political communities. In 1998, members of the American Biological Society were asked to rank their priority concern for the conservation of the environment: 7 out of 10 biologists agreed that we are facing an extinction crisis. The results of the survey are backed with hard figures. The 2003 IUCN Red List of Threatened Species shows that of the 1.5 million species so far described, more than 12,000 (0.79% of the total) are threatened with extinction. However, if we consider only those groups that have been almost fully assessed (birds, mammals and gymnosperms), 16.8% of the species are threatened with extinction. This figure supports the suggestion that current extinction rates are between 1,000 and 10,000 times higher than those inferred for most of geological history (May *et al.*, 1995; Pimm *et al.*, 1995; McKinney, 1998: Boggs and Roughgarden, 2000).

Such authoritative documentation has lent weight to the growing number of governmental and non-governmental responses at various levels. In response to the political demand, as well as the broader need to understand the dynamics of biodiversity, scientific and government circles are building knowledge about biodiversity, its status, and methods to monitor its dynamic nature.

A number of initiatives have been developed in recent years to address the challenge at the global level, as natural resources do not respect political borders and building knowledge requires consideration of regional (e.g., migrations and water resources) and global (e.g., climate change) issues. At the same time, however, natural resource management, and the policies that support it, are developed at the national level. International bodies must therefore consider national interests and constraints in setting standards and obligations for their member countries, and provide approaches that encourage cooperation.

---

[2] Much of the information and conclusions drawn here come from the author's direct experience: 1) in drawing expert information for analysing species distributions patterns in Africa (Boitani *et al.*, 1999); 2) working on the design of the Italian National Ecological Network (Boitani *et al.* 2002) and 3) playing a central role in design and development of the Species Information Service, the IUCN Species Survival Commission data management infrastructure.

Information systems must draw on diverse expertise and data sources; sharing across organisations and borders is critical. This suggests that a networked system of agencies and organisations is needed to capture and harmonise existing data and information on biodiversity

Conservation organisations and advocacy groups have invested heavily in a variety of initiatives aimed at documenting the status of biodiversity, conserving it and minimising human impacts on the environment. A number of global Conventions, for example, the Convention on Biological Diversity (CBD) and the Ramsar Convention on Wetlands of International Importance, bind their parties to integrate biodiversity considerations into decision making processes inherent to the use of biological resources. Decisions concerning the use of biological resources within a country must take into consideration the effects for sustainability in the other countries.

The global Conventions that address or are related to biodiversity conservation rely on qualitative information to implement and monitor stated objectives. To assess the level of implementation and the effectiveness of the measures taken, the Conventions obligate member States to comply with specific reporting requirements. Reporting requirements include, but are not limited to, trade information (Convention on International Trade in Endangered Species of Wild Fauna and Flora, CITES), *in situ* conservation actions (CBD), status of migratory species (Convention on Migratory Species, CMS), characterisation of sites (Ramsar Convention on Wetlands), and key indicators for monitoring the state of conservation (World Heritage Convention, WHC) (Crain and Collins, 1999). While reporting obligations are precise, the specific types of biodiversity information needed to satisfy them are less clear (WCMC, 1999).

The CBD Clearing-house Mechanism (CHM) has as a primary objective to "develop a global mechanism for exchanging and integrating information on biodiversity". The Global Biodiversity Information Facility (GBIF), aims to make "…the world's biodiversity data freely and universally available" (GBIF, 2003). More localised networks include, for example, the National Biological Information Infrastructure (NBII), the European Network for Biodiversity Information (ENBI), ASEAN Regional Centre for Biodiversity Conservation (ARCBC), and the Inter-American Biodiversity Information Network (IABIN). The recognition by governments that this approach to information exchange and enhancement has value is evident in the number of initiatives and the rising level of cooperation between them.

Concurrently, conservation organisations are increasingly recognising the enhanced value of shared information and are forming cooperative networks to share and exchange biodiversity information. These range from large global generalist initiatives such as Eco-Portal (2003), to consortia of organisations

joining together to enhance the value of their individual information systems. An example of the latter approach is the IUCN Red List of Threatened Species, which has been derived historically from the IUCN Species Survival Commission (a global network of species conservation experts) and BirdLife International (a network of state chapters and networks that, among other objectives, compile status information on bird species), and in recent years has expanded its reach by forming a consortium of organisations that now includes Conservation International-Center for Applied Biodiversity Science (CI-CABS), NatureServe and The Ocean Conservancy. These new partner organisations all collect and manage biodiversity information; their formal participation enhances significantly the biodiversity information available for the IUCN Red List (IUCN-SSC, 2003a).

Nevertheless, information management in the field of biodiversity conservation has not yet provided the expected results (Wilson, 2000). Although several initiatives show great promise, none is yet positioned to provide the support required to respond to Convention obligations or, more generally, to monitor the status of biodiversity. As a result, decision-making in the field of biodiversity either does not occur (the problem is ignored) or political opportunities and/or emotional forces drive the process. This is in contrast to the scenario where decision-making is based on authoritative scientific information.

While there is a widely-held belief that existing information about biodiversity is inadequate to support efforts to respond to biodiversity loss (Stork, 1988; Barnes, 1989; Bennett, 1998; Wilson, 1992), existing information, when coupled with technologies and analytical methodologies that maximise value of limited data, is adequate to support scientifically sound biodiversity conservation planning and management. While waiting for field science to provide more comprehensive hard data, there are significant opportunities to formalise existing "dispersed information" and apply it to biodiversity conservation planning and management.

Examples of successful mobilisation of dispersed biodiversity information range from individual projects, which aim at creating the snap-shot of what we know on specific topics (e.g. the African Mammals Databank, described in *Chapter 4* of this thesis); to the creation of national monitoring tools supported by network of experts (e.g. the Italian National Ecological Network described in *Chapter 6*); to global initiatives that mobilise the information of large networks of experts from around the world such as the IUCN Red List of Threatened Species, which is drawn from the IUCN's Species Survival commission (SSC), a volunteer network of 7000 species conservation experts (IUCN-SSC, 2003b).

This chapter explores these opportunities, in particular the means to capture the dispersed information that rests within the core knowledge of experts. Although

the information is widely dispersed, proven approaches to information networking suggest that mobilising it is not only possible but can be done efficiently and effectively.

BIODIVERSITY INFORMATION: MAXIMISING UTILITY OF WHAT WE KNOW

Many definitions of biodiversity exist (U.S. Congress, 1987; Noss, 1990; WRI, IUCN, UNEP, FAO, UNESCO 1992; Grumbine & Soule, 1993; Wilson, 1997), but all include the three major components of the biosphere: genes, species and ecosystems. Much of the existing information about biodiversity focuses on species, due primarily to the fact that species are more easily defined than ecosystems and more easily measured than genetic diversity (Gaston, 2000). Information about species is therefore the unit of biodiversity upon which the argument about the extent and utility of existing biodiversity information is directed.

The gap between biodiversity information needs and existing information is real. It is well illustrated by the many estimates of discovered and described species vs. existing ones. For example, according to the Global Environmental Outlook 2000 (UNEP, 1999) 1.75 million species are recognised by the scientific community, while an estimated 12.5 million species exist. Alternative numbers are 1.4 million described and 10 million estimated (Daugherty and Allendorf 2002).

Although known vs. unknown species numbers make the knowledge gap unquestionable, much useful information about biodiversity is dispersed amongst the community of thousands of scientists, organisations and government agencies around the world, working in a number of fields related to biodiversity science.

Several authoritative sources list species according to their threat status, the most widely used in political circles being the IUCN Red List of Threatened Species. Other lists complement the IUCN Red List, for example, the CITES appendices include 30,000 species, including 900 species on Appendix I for which international trade is prohibited, but addresses only those species subject to trade (CITES, 2003). Given its specific focus, the CITES list uses different criteria and taxonomic system than those used for the IUCN Red List.

In addition, extensive knowledge about population biology, ecological requirements of species and other pertinent information exists in a variety of sources (Cotter & Bauldock, 2000). While it is held in many forms – including data bases, field notes and the (undocumented) knowledge of research scientists – its most common form is grey literature, which is widely accepted as a robust

source of information. Unfortunately, the source has been thus far neglected by mainstream policy making (IUCN, 2001; CBD, 2001), a fact that is now becoming more widely recognised. For example, in a document on Scientific Assessments circulated in 2001 by the Executive Secretary of the "Subsidiary Body on Scientific, Technical and Technological Advice", the CBD outlines the guidelines for the inclusion of grey literature into CBD scientific assessments and reports. Similarly, during the World Conservation Congress held in Jordan in the year 2000, IUCN prioritised the distribution of grey literature within its ten-year intercessional programme as one of the important aspects of knowledge access.

Information about biodiversity and its status is largely incomplete, but what does exist is sufficient to provide the scientific basis and monitoring capacity to support sound scientific decision-making for biodiversity conservation while we wait for hard data to complete the picture. The unifying aspect of "dispersed information" is in the great opportunity that it offers to build extensive knowledge, in a cost-effective way, on aspects of biodiversity otherwise scarcely known. Techniques to channel this "dispersed information" back into decision making processes are needed. Examples of potential uses of dispersed information to produce biodiversity analyses are found in Chapter 4, *Mobilising biodiversity data in practice: the example of the African Mammals Databank*; Chapter 5, *Modelling species distribution using an inductive approach: the example of the wolf in Italy*; and Chapter 6, *Expert-based species distribution maps for the assessment of the biodiversity conservation status*: *the Italian National Ecological Network.*

The challenge to do this can be met by:

1. Mobilising the wealth of existing but dispersed information in a systematic and quality-assured way;
2. Creating a data management infrastructure that provides the flexibility needed to accommodate the dynamic nature of biodiversity information, diverse and changing information requirements and consistency over time and space;
3. Applying proven and emerging analytical methodologies to produce relevant and timely biodiversity information products.

The three are inter-related; one cannot function effectively without the other two. Accessing dispersed information requires commitment and time from a wide variety of experts holding information in almost as many forms. The experts must be provided with an environment and incentives to motivate their participation, including assurance that the information and products they contribute to are quality assured.

As in all scientific disciplines, advances in both technology and thinking lead inevitably to new approaches for managing biodiversity information (Bowker, 2000). Technological advances, for example, the almost universal access to the Internet, now allow real-time exchange of data and ideas and shifts of data amongst systems. However, data inconsistencies, evolving data needs and assurance of consistency over time and space all present challenges in designing a data management infrastructure and require a high degree of flexibility to be one of its primary characteristics. Data management approaches and structures are explored in detail in *Data Management for Biodiversity Conservation*, Chapter 7 of this thesis, and applications of analytical methodologies are explored in *Species Distribution Modelling with GIS* (Chapter 3).

BOX 2.1. …FOLLOWING AFRICA'S LEAD IN SETTING PRIORITIES

Published as Fonseca G.A.B., A. Balmford, C. Bibby, L. Boitani, T. Brooks, N. Burgess, F. Corsi, E. Dinerstein, C. Gascon, L. Hannah, J. Lovett, R. A. Mittermeier, D. Moyer, S. Olivieri, D. Olson, C. Rahbek, S. Stuart, P. Williams. 2000. …Following Africa's lead in setting priorities. Nature. Vol. 405, 393-394.

It is vital to try to conserve biological diversity, but to do this we must know where it is. New databases mapping African biodiversity now offer this information[1–3]. Quantitative analyses[4] of these data reveal geographical priorities for conservation broadly matching those generated at a global level by international conservation organizations[5–7] (Fig. 2.1, overleaf). Although this approximate overlap with coarse-scale priorities is encouraging, we now need to move tropical priority-setting to finer scales that will enable conservation action. We suggest here three strategies for meeting this challenge: improving data; enhancing collaborations; and working closely with local decision-makers. The main constraint on quantitative conservation priority-setting[8] is that sufficient data are not available at the fine scale required for conservation implementation. The only long-term solution is to collect new biological data, compile and disseminate point locality information, and conduct basic taxonomy. Such work is chronically under-funded, yet is essential for conservation planning[9]. In the short term, the problem can be reduced by deductive modelling.

Environmental data such as elevation, vegetation, rainfall and temperature have been modelled to fine resolutions — for example, 1-km grids — for the whole continent (see http://edcftp. cr.usgs.gov). If we compile information on habitat preferences, we can predict fine-scale species distributions by overlaying environmental data onto species range maps, to identify areas where all of a species' habitat requirements are fulfilled[3].

Another short cut is to base conservation priorities on well-known taxonomic groups[10]. The problem with this is lack of knowledge of cross-taxon congruence[11]— for example, conserving birds may not be enough to protect biodiversity as a whole.

Attempts to address the lack of data on African biodiversity must go hand-in-hand with improved collaboration at all levels — see the letter from Mace *et al.* [15]. Effective collaboration is urgently needed between the biological and social sciences, to incorporate human geography into the quantitative priority-setting process in the

tropics[12]. The development of parallel priority-setting initiatives is another symptom of the lack of effective coordination to date.

Furthermore, difficulties have arisen over data dissemination and public access to information. Such tensions between data providers (for example, museums) and users (such as non-governmental organizations) can be eased by considering mutually beneficial collaborations. For instance, conservation groups could increase their funding for the publication of biological data, and groups with mutual interests could collaborate on fund-raising to pay for data collection.

Finally, much greater use should be made of existing collaborative networks[13]. Effective translation of continental priorities into action depends fundamentally on consensus from local decision-makers. One way of forging this is through expert based priority-setting workshops[14] to assess key regional areas for conservation values in different taxonomic groups and to prioritize these areas across groups. From this synthesis, an integrated set of local priorities can be developed incorporating information on ecology, current and future threats, and landscape-level linkages. Essential components are a commitment to training, empowerment of local specialists, and repatriation of biodiversity information.

Such conservation-prioritization workshops have been held recently for the Upper Guinea region (December 1999) and for the Congo Basin (March 2000). The consensus forged by governmental, non-governmental and academic representatives from the countries in these regions has provided a solid base to translate priorities into action.

We believe that our suggestions will considerably increase the chances of further progress while opportunities for effective conservation in Africa remain.

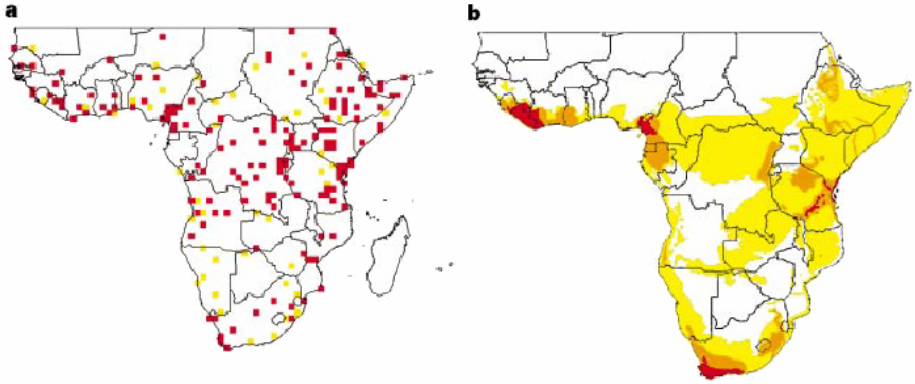


Figure 2.1 Conservation priorities for Sub-Saharan Africa.

a, Quantitatively derived conservation priorities4 for ~4,000 species of bird, mammal, snake and amphibian, mapped on a 1° grid. Coloured cells depict the top 200 areas from which 97.5% of species mapped have been recorded. Red squares are irreplaceable because they contain the entire known distribution of one or more species; orange cells are flexible areas for which alternatives (not mapped) are available.

b, Conservation International's Hotspots5; the World Wildlife Fund US's Global 200 most biologically important ecoregions6; and BirdLife International's Endemic Bird Areas7. Red, orange and yellow show areas of intersection between three, two and one system(s), respectively.

References and Notes

1. Burgess, N. *et al. Species* **30,** 16–17 (1998).
2. Lovett, J. C. *et al. Biodiv. Conserv*. **9,** 33–42 (2000).
3. Boitani, L. *et al. AMD African Mammals Databank: A Databank for the Conservation and Management of the African Mammals* (Istituto di Ecologia Applicata, Rome, 1999).
4. Williams, P. H. *WORLDMAP 4.1. Priority Areas for Biodiversity* (Natural History Museum, London, 1996).
5. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403,** 853–858 (2000).
6. Olson, D. M. & Dinerstein, E. *Conserv. Biol.* **12,** 502–515 (1998).
7. Stattersfield, A. J. *et al. Endemic Bird Areas of the World: Priorities for Biodiversity Conservation* (BirdLife International, Cambridge, 1998).
8. Williams, P. H. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 211–249 (Cambridge Univ. Press, Cambridge, 1998).
9. Balmford, A. & Gaston, K. J. *Nature* **398,** 204–205 (1998).
10. Bennun, L. A. & Fishpool, L. D. C. *Ostrich* (in the press).
11. Howard, P. C. *et al. Nature* **394,** 472–475 (1998).
12. Fjeldså, J. & Rahbek, C. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 139–160 (Cambridge Univ. Press, Cambridge, 1998).
13. BCIS *Framework for Information Sharing* (Biodiversity Conservation Information System, Cambridge, 2000).
14. Olivieri, S. *et al. A Participatory Approach to Biodiversity Conservation: the Regional Priority Setting Workshop* (Conservation International, Washington DC, 1995).
15. Mace G. M. et al Nature 405, 393 (2000).

Finally, analytical methodologies for assessing and monitoring biodiversity are needed to ensure that limited data and information is put to the best possible use. For example, the IUCN Red List of Threatened Species is derived from the IUCN Categories and Criteria (IUCN, 2000), a system built around population dynamic theories and designed to accommodate data limitations. Another example is predictive modelling of species distribution, which is now being done at fine scales relevant to decision-making. Box 1 (Fonseca et al. 2000) makes the argument for the importance of scale in biodiversity information. Corsi *et al.* 2000, Boitani *et al.*, 1999 and Corsi *et al.*, 1999 illustrate how applications of the data management approach described here can be applied to analyses at scales relevant to decision-making.

DISPERSED INFORMATION: THE CHALLENGE OF ACCESS

"Dispersed information" lacks formalisation. Here formalisation is defined as any process that standardises the information, tags it with a quality control label and allows it to flow amongst the community of people interested in accessing the information. For example, the scientific publication process, which begins with data collection and leads to the publication of the research results in a scientific journal, is a recognised process for data formalisation that provides a widely accepted "scientifically sound" label.

Lack of formalisation is the major impediment to exploiting existing biodiversity information: lack of access, fragmentation and deterioration.

Lack of access

Accessing relevant information requires an effective connection between the information source and the channel for integrating and distributing it. There is a common two-way barrier: channels for integrating and distributing do not have adequate access to the many sources of information while many of the sources of information are unaware of the channels. Both sources of information and the information formats are widely diverse, a major reason why access is so difficult.

Sources of information are primarily individual experts, scientists for the most part but in some cases amateurs that hold both valid and valuable information (amateur birdwatchers are an example). In many cases the information is held, and to some extent integrated and analysed, by conservation organisations and government agencies (through advisory panels of experts). These sources may be derived from networks of experts, or from other sources such as small data bases or networks of organisations. In addition, what is held comes in multiple formats. Where it has been structured into data bases, few of the data bases are compatible and support information sharing and exchange.

Further, where data sources and channels are known to each other, there is a second barrier to access: proprietary issues (BCIS, 2000c). Data have value and accessing it is seldom as simple as freely sharing it with whoever might ask. The value is measured in a number of ways; the most pertinent to this topic is its value to the scientists who acquire it through their life's work.

Potential economic value can be significant as well and is often a driving force behind difficult negotiations about data movement. An example is the growing concern about bio prospecting, and the use of indigenous knowledge in the discovery of new molecules for the pharmaceutical industry. Indigenous populations contribute their "dispersed information" on curative effects of certain organisms to bio prospectors, who in turn use the expertise of these local groups to identify organisms that produce curative molecules. Bio prospectors channel the information through the accepted formalisation processes. The pharmaceutical industry then uses the result to patent the molecule, synthesise it and produce large revenues, which are seldom (if ever) passed back to the local populations that provided the original information. The controversial article 27.3b of the WTO Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) "de facto" grants the right to patent plant molecules, and strenuously defends the role of developed countries in this process. The economic value of dispersed knowledge is useful to make the "cost-effective"

case in the management of information. This circumstance exemplifies the huge potential of this kind of dispersed knowledge (Moran, 2000).[3]

Fragmentation

Scientific formalisation processes are built to accommodate information structured according to the scientific method: identification of a problem, formulation of a hypothesis, testing of the hypothesis, use of the validated hypothesis to support knowledge. But in many cases individuals hold only portions of the information required to support the entire process.

These pieces of information, although scientifically sound, are omitted from the scientific publication formalisation processes and become "dispersed information". This includes the evidence and the ideas that researchers collect in an informal way during their normal research activity and never formalise into published material (Cotter & Bauldock, 2000). It is precisely this fragmented information that provided the backbone for the studies described in Chapter 4, *Mobilising biodiversity data in practice: the example of the African Mammals Databank*; Chapter 5, *Modelling species distribution using an inductive approach: the example of the wolf in Italy*; and Chapter 6, *Expert-based species distribution maps for the assessment of the biodiversity conservation status*: *the Italian National Ecological Network.*

In the process of collecting data to support their research, scientists acquire a significant amount of information that is somewhat related, although not strictly part of, the primary focus of their research. This information builds a core of expertise that scientists use to support their research hypotheses, to formulate new ideas and to advance their own basic knowledge on the matter under investigation. Although frequently the final research results are shared through one or more scientific publications, the additional, or peripheral, information is seldom shared. On the assumption that sharing (mixing) information is what stimulates new advances in knowledge (QED, 2002), this background information remains in a state of "dispersed information".

In some cases, this information is stored only in the heads of the scientists. Frequently, however, it can be found in scientists' field notes, and sometimes is transcribed into data bases developed by individuals or small groups of cooperating scientists to target specific aspects of science and knowledge. Synthesis of such information faces a number of constraints that prevent ease of data movement: money; efficient and effective connections to information channels; proprietary issues; quality assurance; and competing and conflicting priorities.

[3] Although not discussed in this chpter, the author recognises the significant ethical issues of this topic.

Deterioration

Information that has been partially or fully formalised is frequently lost. Michener *et al.* (1997) describe an empirical data degradation model, which shows the decrease of information content in data, starting from the moment they are collected until they revert back to the status of unknown. The same concept could be applied to monitor the degradation of information from the moment it begins to be formalised to the moment in which it is again in the state of unknown. The outreach of a scientific paper diminishes rapidly a few months after publication (Crane, 1972), when the issue of the journal in which the paper is published is removed from the "latest number" section of libraries. Once the issue is moved to storage only a few papers survive. In most cases the information content reverts back to "dispersed status". While this process may be interpreted as part of the "natural selection" which allows only the "most fit" theories and research results to persevere as common knowledge, it may also result in loss of information and waste of resources.

Information collected by scholars can altogether "deteriorate" back into the realm of "dispersed information". In some cases precious information remains forgotten at the stage of raw data. An historical example of forgotten information dates back to 1898, when Thomas Meehan and Elliott Coues published the book "The 1898 surveys of the Lewis & Clark Herbarium". The book reports the authors' rediscovery and analysis of the plants collected during the expedition in the Western part of North America conducted by Lewis and Clark in the early 1800s. The plants had been neglected for almost 80 years in a storeroom of the American Philosophical Society. Examples of this sort exist in all disciplines of science. Entire periods of human history have been marked by the rediscovery of previously forgotten knowledge. The waste is significant, whether it results from inaccessibility of the source, failure of scholars to adequately investigate the information they have collected, or failure to fully investigate what others have done before them.

One of the driving forces behind the rising interest in electronic publishing, in particular electronic archiving, is the potential to minimise this secondary un-structuring of information by relying on the data mining capabilities of such archives (Ginsparg, 2003).

MOBILISING DISPERSED INFORMATION: ENGAGING THE EXPERTS

The challenge in mobilising dispersed information is to initiate a comprehensive multi-directional information flow in the form of an information network. The information sources, which are primarily individual experts, but also include organisations and government agencies, must have access to (and be willing to

participate in) the networks. The network must be able to accommodate multiple sources of information. Results of integrated information must be shared with contributing experts and others to support a system of quality control. Decision-makers must have access to and confidence in the networks, and the information provided to them must be relevant and timely.

BOX 2.2. DATABASES TAILORED FOR BIODIVERSITY CONSERVATION

Published as Smith A.T., L. Boitani, C. Bibby, D. Brackett, F. Corsi, G.A.B. da Fonseca, C. Gascon, M. Gimenez Dixon, C. Hilton-Taylor, G. Mace, R.A. Mittermeier, J. Rabinovich, B.J. Richardson, A. Rylands, B. Stein, S. Stuart, J. Thomsen, C. Wilson. 2000. Databases Tailored for Biodiversity Conservation. Science. 290(5499):2073-2074.

We were pleased to see the attention afforded bioinformatics for biodiversity in *Science*'s special issue on the topic[1]. It is essential that informatics technology be devoted to increasing our understanding of Earth's taxonomic diversity and to developing means to access the vast resources available in the world's scientific collections, and these are indeed the stated goals of the Global Biodiversity Information Facility (GBIF). However, the articles in the special section contain the potentially perilous implication that historical biodiversity data are sufficient to address contemporary issues in conservation. The section's content emphasizes data from natural history collections, as exemplified by a statement in the section's introduction (p. 2305): "These Web resources will be of greatest use if they can be put into a historical context." We see a clear distinction between a biodiversity database (based on historical data from collections) and a biodiversity conservation database (regularly updated with current information and used to support conservation decisions).
Biodiversity is being lost at an alarming rate. The *2000 IUCN Red List of Threatened Species*[2] and BirdLife International's *Threatened Birds of the World*[3] highlight the recent increase in numbers and severity of change in status of threatened species. Historical data all too frequently reflect long-lost landscapes that are now parking lots or oil palm plantations, or have been otherwise fundamentally altered and are therefore of no use for contemporary analysis or planning. A biodiversity conservation database, based on field studies and scientists' notebooks in addition to museum data, is urgently needed to address the pressing needs for management of what we have left.

In response to this need, the Species Survival Commission (SSC) of the IUCN-World Conservation Union recently announced the development of its Species Information Service (SIS)[4], which will link the SSC's network of more than 7000 species specialists in a distributed data management system, integrating modern principles of data custodianship[5]. The most powerful feature of SIS is that its data will be continuously updated and managed by species experts. This mechanism will link policy-makers with scientists who have first-hand knowledge of the current status of biodiversity. In addition, both scientists and conservationists will have improved access to the data and information they need to assist them in achieving their goals.

References and Notes

1. Bioinformatics for biodiversity. Science **289**: 2305-2314
2. C. Hilton-Taylor (Compiler), *2000 IUCN Red List of Threatened Species* (IUCN, Gland, Switzerland, and Cambridge, UK (www.iucn.org/redlist/2000/index.html, 2000).
3. BirdLife International, *Threatened Birds of the World* (Lynx Edicions, Barcelona, Spain, 2000).
4. IUCN Species Survival Commission Species Information Service (SIS) (www.iucn.org/themes/ssc/programs/sisbrochure.htm). Six years in development, SIS uses an architecture of complementary modules to ensure maximum flexibility [modules on taxonomic information, population status data, geographic information (at variable scales), IUCN Red List classification (the data source for future IUCN Red Lists of Threatened Species), conservation actions, and resources and documentation]. The completed stand-alone SIS software is scheduled for release in early 2001; Web access is under development.
5. J. Busby, Ed., BCIS--Biodiversity Conservation Information System: Framework for Information Sharing (www.biodiversity.org/Handbooks_eng.htm). BCIS is a consortium of international organizations and networks (Conservation International, Association for Biodiversity Information, Wetlands International, Botanical Gardens Conservation International, TRAFFIC International, BirdLife International, World Conservation Monitoring Centre, International Species Information System, IUCN Environmental Law Programme, IUCN World Commission on Protected Areas, IUCN Species Survival Commission, IUCN Commission on Ecosystem Management) dedicated to the cooperative management and provision of biodiversity information.

Mobilisation of the wealth of dispersed information requires an efficient and effective system that minimises the time required of participants (the experts), usually in a volunteer capacity, and maximises the value of information emerging from the process. The system must lower the barriers to information access and ensure current and quality information. This suggests that, along with the traditional means of communication, new ones must be available to the experts *and* to the networks as a whole for collecting, integrating and distributing information.

Examples of effective information networks are abundant; indeed the history of the IUCN Red List of Threatened Species and its success in attracting comprehensive information from a widely dispersed network illustrates the potential of this approach. However, it is not until electronic technologies are introduced to support network functioning that it is possible to achieve the levels of data currency and quality needed to meet current biodiversity information needs (Bisby, 2000).

However, the status of biodiversity can change rapidly, as does the information held by experts. Development projects, natural disasters, field work, recruitment of additional expertise are all circumstances that could push information quickly into the "historic" category. An electronic biodiversity information network must accommodate these changes (Bowker, 2000), both in terms of providing experts with a simple and efficient way to provide new information in a timely manner and in terms of integrating that new information into the larger system. A fully engaged network membership is a key element, as the experts must have sufficient motivation to contribute their information as it evolves.

The importance of currency, and how it distinguishes a biodiversity database (for example, databases on natural history museum collections) from a biodiversity conservation database that accommodates new and changing information, is articulated in the *Science* article "Databases Tailored for Biodiversity Conservation" (Smith *et al.*, 2000) (see Box 2).

The IUCN SSC recently completed the study "Voluntarism in the Species Survival Commission" (IUCN-SSC, 2001). The study found that important motivators for participation include commitment to "the cause", esteem and recognition from peers and personal contact amongst network members. De-motivators include a weak group identity, lack of clarity regarding expectations and failure of the network to make use of the information provided. The de-motivators in particular point to the importance of providing adequate support to the network, especially in terms of personal communication and direction, and results that are shared both within and outside of the network.

Comprehensive information mobilisation therefore requires active (ongoing) participation of a large group of experts and partner organisations. Establishing partnerships amongst organisations requires collaboration agreements that make clear the exchange protocols, proprietary rights and uses (including their limitations). To engage the individual experts to the extent required is the most challenging aspect of the process. While technology offers tools to enhance communication flows, these tools are ineffective if the experts are unwilling to use them. To achieve maximum participation of experts, two elements are needed: 1) a *supporting environment* that ensures efficiency and effectiveness; and 2) *incentives* that motivate a wide variety of experts to participate.

Supporting environment

A supporting environmental for participants includes: 1) minimising the workload; 2) complementing competing priorities; 3) efficient means of information transfer; and 4) protection of intellectual property. It is assumed that any such supporting environment must include an adequate support structure in the form of basic management, and those human supports that ensure participants assistance with questions or problems within the system.

Minimising the additional burden on the expert's workload is a critical aspect of the supporting environment in which the expert can carry through with network activities and expectations. From the perspective of scientists, contributing information from grey literature, files, small data bases, or even information from their heads, requires time that they do not have to spare. In a 2001 Conservation Biology editorial, G.K. Meffe, editor of the journal, analyses the increasing peer-review turnover time. He identifies the problem in the over commitment of the experts involved in conservation biology research: "…they

are swamped with teaching, research, grant writing, student support, book and paper writing, administration, and service activities." The same analysis holds true if we look at experts' participation in networks that aim to share data, information and ideas.

Competing priorities will always exist, therefore activities expected of participating experts should, to the extent possible, complement and support his or her every-day (paid) work. Participation in information-sharing networks is generally done on a voluntary basis, therefore this latter point is especially important. The fact that competing demands from within the participants' organisations are a major factor must be recognised. In the biodiversity conservation arena, where a significant percentage of information network participants are academics, the "publish or perish" environment is a particularly fitting example. Therefore, if sharing information requires experts to organise data for their research activities, and the return of value-added information further supports that research and corollary publications, then the burden on the experts' workload is diminished or even eliminated.

A number of initiatives develop around this approach by building on top of existing activities to reduce the overload. Information networks can build small databases and networks into larger networks with a broader scope, providing participating experts with value added services like integrated datasets, specialised analyses, and the use of software that would be otherwise too expensive for the individual expert to acquire.

An efficient means of information transfer is achieved, in part, through the use of tools that standardise common language, taxonomy, definition of threats and other elements that ensure consistency and clarity for those contributing information and for those collecting and synthesising it. Supporting tools also include methods of analysis, and the capacity to carry out analyses that produce results of a quality and scale that are adequate to support decision-making processes. For example, ecological modelling tools are increasingly being used to assess and predict species distributions at a variety of scales. Ecological modelling relies on the capacity to integrate comprehensive information from experts with other data sets such as land use and vegetation maps. Geographic Information Systems (GIS) and remote sensing technologies are at the heart of this integration. Examples of the utility of this approach, and the potential return for network participants, can be found in Corsi *et al.* (1999; 2002) and in Boitani *et al.* (1999).

Tools must also support movement of information, including information gathering from the experts, allowing experts to share information with each other, and returning value-added information to those who have contributed information. Within the network, this suggests the need for simple, easy-to-use

software available at appropriate information-gathering points within the network and, at a minimum a standardised form (hard copy or electronic) for expert inputs. The key is to ensure easy data transfer that minimises both the work of the expert and the work required at the first point of integration (at a central or sub-network point of entry into the system).

Links between networks add a further layer of integration and potential value-added for contributed information. Internet based biodiversity portals, or online communities, have been developing over the past few years. A search on Google (May 2003) for "biodiversity portal" returned 31,200 hits, 17% of them distinct sources[4]. As mentioned above, these can be large global generalist initiatives such as Eco-Portal (2003), to more specialised ones like the Global Biodiversity Information Facility (GBIF), and can be initiated by governments or conservation organisations. By linking existing web sites and Internet databases, these initiatives generate value added information. Lesser known information is aggregated into more accessible and topic driven frameworks. No specific study has yet been carried out on the viability of these biodiversity portals. However, the fact that the impact of these initiatives on experts' workloads is generally low, and that they build on top of existing initiatives, suggests that they may be viable over the long term.

A second and more complex service approach, which relies in Internet technologies, is one adopted by many organised networks of experts. Those with large membership bases, such as The Nature Conservancy, BirdLife International, NatureServe, and the IUCN SSC, whose main assets lay in the information provided by the networks of experts, have been developing specialised data management systems to support their members. The systems that are currently being used or developed are evolutions of the original databases that many of the above organisations established in the early days of mass diffusion of IT. While conserving their original purpose of collecting data from the members of the network, these systems have evolved into multidirectional information flows which not only return processed information back to the experts, but also support direct links between experts. All are evolving; therefore measures of their success are not yet apparent.

Protection of proprietary rights is a fundamental element of an information network. Formalised agreements about use and exchange of data help to establish the trust needed to encourage their free and flexible movement through a large system while providing appropriate checks to ensure their quality. The

[4] To assess the number of distinct hits (e.g., excluding pages that refer to the same initiative) a random sample of ten pages, each with ten links, was selected from the pages returned by Google. For each one of these ten pages, the number of individual (distinct) hits was counted, and the number extrapolated to the total number of hits.

tension between absolute data "ownership" and the argument that data about the environment is a part of the "global commons" is one that should be acknowledged and recognised within any data use policy (BCIS, 2000a). The challenge of constructing appropriate data use and exchange policies can build on different levels of access, rights and responsibilities, as described in the BCIS "Framework for Information Sharing: Data Custodianship" (2000b). The concept of data custodianship within a dispersed network of data and information providers can be the key to building the trust needed for efficient data flow: the custodian has responsibility for building and managing the data set, while concurrently holding responsibility for regulating access and assuring that intellectual property rights are respected.

Regardless of the chosen formula of rights and access, an individual contributing his or her data needs some assurance that they will be used only for the purpose agreed and will not be sold without fair compensation. Such agreements are also fundamental when the networking includes collaboration amongst organisations, and links information networks within them.

Incentives

The true motivations for participating in information networks lie in the larger issues. These include: 1) commitment to the cause; 2) ownership in the process; 3) a high-profile result; 4) assurance that the best information influences decision-making.

Participants must be *committed to "the cause",* that is they must clearly recognise the importance of the network's purpose and believe that it is a priority.

Experts are more willing to invest time in initiatives in which they feel some *ownership*. The large number of web sites developed on specific topics that reflect the work of one single expert or a small group of experts (e.g. one single project, one small group of species, a specific conservation topic) testify to the time devoted to these types of information formalisations and exchanges. There is an intrinsic need for scientists to show the results of their work and outreach to their scholar community and other fellow scientists (Harnad, 1996). These types of initiatives are not perceived as impacting on the workload of the experts; the maintenance of the web sites is either included in the expert's institutional work, or is done within the time that is normally dedicated to that outreach activity.

The "challenge approach" exploits the ownership concept. It influences experts otherwise unwilling to participate, or who, for any number of reasons, are not meeting information deadlines, to respond. Experts are provided with

alternative information (gathered or inferred) and are requested to correct or approve the findings for publication. The expert is thus "challenged" to make sure that the information circulated about the topic of his/her expertise reflects the most current knowledge. An advantage for the expert is that this approach may require only validation. The time burden for the expert is significantly reduced, however, the approach comes with the cost of a support environment with sufficient resources to gather and synthesise information.

A corollary benefit to the challenge approach is the "forgotten information" that may be uncovered in the information-gathering process. The basic bibliographic search required to assess information and pose the challenge to the expert often leads the researcher to valuable sources of information and data that have been sitting unused.

A high-profile initiative motivates experts to participate. It gives the expert exposure and visibility while contributing to something that is recognised to have a direct and generally immediate impact on biodiversity conservation. The IUCN Red List of Threatened Species is one example. It is promoted by a well-established organisation, has global reach and is capable of influencing policymaking. Conservation decision-making frequently acknowledges the IUCN Red List categorisations as criteria for determining the importance of proposed projects or policies. The members of the IUCN Species Survival Commission (SSC), who provide the backbone of information for the IUCN Red List, have been freely providing their expertise for several decades.

Priority setting workshops organised by conservation organisations such as Conservation International and the World Wide Fund for Nature (WWF) have a similar appeal to experts. Brainstorming sessions, which bring together experts from different fields of knowledge, build information sources about specific regions. The demand on the expert is limited and concentrated in time, while the results can have a significant impact on decision-making. For instance, the government of Brazil has adopted the Priority Setting workshop approach for its conservation policy and is sponsoring its own workshops (CABS, 2003). Similarly, the Critical Ecosystems Partnership Fund (CEPF) bases its funding strategy on ecosystem profiling which is largely the result of priority setting workshops for the various regions of the world (CEPF, 2003). CEPF is a funding mechanism jointly created by Conservation International, the Global Environment Facility, the Government of Japan, the John D. and Catherine T. MacArthur Foundation and the World Bank, to enhance conservation of critically endangered ecosystems throughout the world.

Assurance that the best available information is used for decision-making is a motivation for experts to participate. An information network draws on experts for their most current information; providing them with a channel to apply that

information, in an ongoing fashion, assures that the best information available is influencing decision-making. This not only provides a strong incentive to provide information, but also counters the risk of de-motivation that results when an expert's information is *not* put to good use (wasting the expert's time).

QUALITY ASSURANCE

A validation system to assure quality of biodiversity information is essential to creating an authoritative information source. Integrating dispersed information suggests, however, the need to quality-assure pieces of information, rather than the traditional approach which focuses at the publication (finished product) level.

Peer review is the traditional approach used to validate scientific papers and filter out those without sufficient quality or originality to merit publication. The process evolved along with the scientific societies of the 18th and 19th centuries, which served to bring scholars together, establish communication, and exchange and discuss ideas. The journals of these societies were the repository of information.

Over time, the efficacy of the societies as general repositories of knowledge and as networking facilities has greatly diminished. Increasing specialisation and diffusion of knowledge away from the traditional cultural centres of the 18th and 19th centuries pushed scholars towards other, less direct, forms of communication (von Foerster, 2001). Instead of meeting and discussing within the facilities of the societies, scholars began to engage in discussions through journals, setting the basis for today's peer-review process. Thus the society journals, and the many others born from their example, remained the primary repository of scientifically sound knowledge.

The soundness of that knowledge was supported by the authoritative approval of the peer-review process as we know it now. Peer-review remains the basic method for guaranteeing data quality and filtering out spurious information. Years of scientific publishing have established peer-review as the standard way by which the soundness of any piece of scientific information can be "certified". This process continues to work for theoretical sciences, where the pace of advances can accommodate lengthy turn-around time inherent to traditional peer review processes.

The same cannot be said for applied sciences, such as conservation biology. Information and ideas needed to support conservation biology must move quickly through the system if it is to be relevant (Meffe, 2001). Further, the most relevant and valuable information does not come in the form of a final

product, but instead is a piece or pieces of information that grow in value as they are integrated with other pieces of information.

Traditional characteristics of peer-review do not respond to the time-sensitive and flexible nature inherent to information networks. While a few suggestions have been made to adapt the traditional peer-review process to electronic publishing, these suggestions still focus on peer-review of the full publication (Harnad, 2000; Ginsparg, 2003).

Quality assurance of dispersed information therefore requires a new approach. Dispersed information is collected piece by piece. For example, an expert might provide information about a single ecological requirement of a species. This does not mean that the ecological requirements of the species are fully understood; only that one contribution has been made toward building that knowledge. To make sense out of this information, it must be integrated with other, related pieces of information.

New concepts like peer commentary (Harnad, 1996; 2000) respond more directly to these core concepts of mobilising dispersed information. Peer commentary is defined as an open process by which a scientific paper (or a scientific idea) is freely circulated among peers for comments and suggestions. The process starts when a scientific paper or an idea (or, in this case, a simple piece of information useful to the understanding of biodiversity) is made public. The entire community of peers contributes to building comments and opinions on the newly published paper, generating a whole, more informative, piece of scientific literature than the original paper.

While identifying and "weeding out" those pieces of information that are inaccurate, peer commentary also stimulates the appearance of additional information on the same topic and the assembly of the individual pieces into the larger picture. In addition, the peer commentary approach supports information validation at various stages of the collection and synthesis process, thus allowing the cycle to repeat itself as new information is introduced.

This informal approach offers the flexibility needed for efficient information networks while concurrently contributing to an interesting evolution of the traditional peer-review process. The approach also emphasises the importance of electronic tools that support rapid and constant information flow, and the human interventions needed to ensure the information is openly available to those experts best positioned to review it.

The IUCN Red List provides an example of the move toward more interactive solutions to quality control and peer review. The IUCN SSC and its IUCN Red List partners are organised into a series of Specialist Groups each of which

focus on a species, group of species or specific conservation issue. The process of assigning a category of threat builds in quality control and a "peer review" element, and allows for external response to accommodate new information or differing views. Within the network individual experts are assigned to assess the status of a species using the IUCN Criteria of Threat (IUCN, 2000). An evaluator then "peer reviews" the assessment. This process integrates into the larger information networking process; the role of providing peer commentary is part of the expectation placed upon the participating experts. Once the IUCN Red List is published, a formal appeals process is available to accommodate challenges on the basis of new or inaccurate information. This enables inputs from *outside* of the network.

The adaptation of peer review to peer commentary, and the IUCN SSC's own adaptation of the concept, provides a first glimpse at methods of validation within the process of integrating pieces of information. Adoption of a peer commentary system is fundamental both to assure the veracity of the biodiversity information analyses promoted to decision-makers, and to assure the participating experts that their information is integrated with other quality-assured information. As in any feedback process, it can also unearth additional knowledge. What must be recognised here is that the validation process is fundamental, but the methods to carry it out have not been fully developed or tested, especially in light of the possibilities offered by Information Technology.

There are two considerations in implementing the process into a widely dispersed information network: 1) the structure of the network must accommodate this process; 2) the points of validation must be clear, and balance the need for data quality while recognising the practical issues related to the massive amount of information that could potentially move through the system. An example is, again, the IUCN Red List, which employs a review process at a point where experts' data has been synthesised into an assessment of species status. Only when the results of that assessment are questioned are the data behind it closely examined; acceptance of the assessment assumes that the data held by the expert are valid.

CONCLUSIONS

The call for biodiversity information to support decision-making, in the form of implementation of the environmental Conventions and a variety of biodiversity conservation initiatives, is a daunting challenge and one that many believe cannot be met with the current paucity of information.

While huge gaps in information are inarguable, sufficient information does exist to support decision-making. This information is spread across multiple sources (thousands of individuals and organisations) and in many different forms. The first step toward ensuring that what information does exist is put to good use is to mobilise it through information networks supported by information technology.

Effective networks must efficiently link the sources of information to the decision-makers. This can be done by ensuring that a supporting environment and adequate incentives are in place to motivate experts and organisations to participate. A supporting environment minimises the participants' workloads, complements competing priorities, provides an efficient means of information transfer and protects intellectual property. Incentives include commitment to the cause, a high-profile result and assurance that the best information influences decision-making.

Finally, the experts must know that their information becomes part of a whole that is quality assured. Quality assurance again relies on the experts, but with a focus on validating pieces of information rather than final products (publications) that are characteristic of the traditional peer-review process. The validation process must therefore be woven within the larger information networking process, and applied with an efficiency that meets the "supporting environment" and "incentives" standards necessary to motivate full and continuing participation of the experts.

A key element to the functioning of information networks is a data management infrastructure that accommodates the dynamic nature of biodiversity information, and maintains data consistency over time and space. See Chapter 7, *Data Management for Biodiversity Conservation,* for a discussion of the challenges and solution to this element.

With these elements in place, existing information can be mobilised and use to create timely and relevant information products to support biodiversity conservation. Comprehensive understanding of the status of a variety of plant and animal species, coupled with a common knowledge of species distributions, can be achieved, and provide decision-makers with information fundamental to their charge.

Chapter 3:
Species distribution modelling with GIS[5]

INTRODUCTION

The power of a Geographic Information System (GIS) resides in its capability to handle large amounts of spatial data and supporting analysis of existing spatial relationships. It increases the number of variables that can be considered in an analysis and the spatial extent to which the analysis can be carried out (Burrough 1986; Haslett 1990). This capability supports development of species distribution models from limited data on the species, either by analysing the species-environment relationship through expert analysis of ecological requirements or by extrapolating the species-environment relationship from a small number of observed point localities. Examples of these latter two approaches are found in Chapter 4*, Mobilising biodiversity data in practice: the example of the African Mammals Databank* and Chapter 5, *Modelling species distributions using an inductive approach: the example of the wolf in Italy.*

In this chapter the uses, misuses and limitations of GIS applications are explored, as is the potential of using GIS to transform limited data on species into effective decision-making tools.

Distribution ranges are seldom absent in comprehensive descriptions of species, as evidenced in the wide variety of checklists, atlases and field guides available around the world. Distribution ranges are used to better understand species biology, carry out inventory assessment of a geographic region and even define specific management actions. In the latter case, knowledge of the area in which a species occurs is fundamental for the implementation of adequate conservation strategies. Recognising that species' range dimension is correlated to population size (Gaston 1994; Mace 1994), conservation is concerned with fragmentation and/or reduction of the distribution as an indication of population viability (Maurer 1994).

Most animals move and this poses a challenge in mapping their occurrence. Traditional methods used to store information on species distributions are

[5] Based on: Corsi F., I. De Leeuw, A.K. Skidmore. 2000. Modeling species distribution with GIS. In: Research Techniques in Animal Ecology. Boitani L. and T.K. Fuller (eds.) Columbia University Press, New York: 389-434.

generally poor (Stoms and Estes 1993). Distributions have been described by drawing polygons on a map (the "blotch") to represent, with varying approximations, species' ranges (Gaston 1991; Miller 1994). The accuracy of the polygons relies on the empirical knowledge of specialists, and encloses the area in which the species is considered likely to occur. The probability level associated with this "likelihood" concept is seldom specified. A more sophisticated approach divides the study area into sub-units (e.g. administrative units, equal size mesh grid), with each sub-unit then associated with information on the presence or absence of the species. In this case the distribution range of a species is defined by the total of all sub-units in which presence is confirmed. However, blank areas are ambiguous; they may indicate absence of the species or that no records are available (Scott *et al.,* 1993).

New approaches tend to shift from the concept of distribution range towards one of area of occupancy[6]. This concept is particularly useful for conservation action and has, therefore, been included in the IUCN red list criteria (IUCN 2000).

A biologist who needs to locate zebras will serve as an example to describe the basic concepts. Intuitively, the odds of finding zebras in Scandinavia are very low. In Kenya the odds become relatively high. This process relies on basic assumptions, including the fact that zebras live in warm places with, for example, an average annual temperature from 13° to 28°. There are many other ecological requirements and reasons, such as historical constraints (see Morrison *et al.,* 1992 for a review) and species behavioural patterns (Walters 1992), that contribute to defining the distribution of the zebra. Therefore the observer won't expect to find zebras in every place on Earth that has an average annual temperature between 13° and 28°. Nevertheless, if the biologist extends the same process, taking into account the preferred ranges of values of various environmental variables, the probability of finding the species in the areas in which these preferences are simultaneously satisfied increases.

If the aim of the researcher is to map the areas in which the species is most likely to be found, rather than to find an individual, the entire process can be seen as a way of describing the species' presence in terms of correlated environmental variables. If relatively inexpensive and broadly acquired environmental data (e.g. vegetation indices maps derived from satellite data) are

[6] Area of occupancy is defined as the area within the species' extent of occurrence which is occupied by a taxon, excluding cases of vagrancy. The measure reflects the fact that a taxon will not usually occur throughout the area of its extent of occurrence, which may, for example, contain unsuitable habitats. The area of occupancy is the smallest area essential at any stage to the survival of existing populations of a taxon (e.g. colonial nesting sites, feeding sites for migratory taxa) (IUCN, 2000).

utilised to define species probability of presence, then maps of species distribution can be quickly and efficiently produced.

To provide a formal approach to species distribution modelling the process can be divided into two phases. The first phase assesses the species' preferred range of values for the environmental variables taken into account. The second identifies all locations in which these preferred ranges of values are fulfilled. The first phase is referred to as Habitat Suitability Index (HSI) analysis, Habitat Evaluation Procedures (HEP) (Williams 1988; Duncan *et al.,* 1995) or, more generally, species-environment relationship analysis. The potential of the second, which involves the true distribution model, has been greatly enhanced in the past 10 years through use of GIS. GIS can extrapolate the results of the first phase to large portions of territory.

GIS is a convenient tool to address the multidimensional nature of the species-environment relationship (Shaw and Atkinson 1990) and the need to integrate large portions of land (eventually the entire biosphere) into the analysis (Sanders *et al.,* 1979; Klopatek *et al.,* 1983; Flather and King 1992; Maurer 1994), with the aim of producing robust conservation oriented models.

This chapter presents a review of models and methods used in GIS-based species distribution models. It is based on a literature review carried out on GEOBASE[7] with the following keywords: GIS, Remote Sensing (RS), wildlife, habitat and distribution. The 82 papers collected were classified according to the main tool used (GIS or RS), the modelling approach, the analysis technique, the discussion of the assumptions and the presence of a validation section. At the same time information was gathered on the use of the term "habitat", on the number of variables used for modelling and on the type of output produced.

The review is the starting point for a tentative taxonomy of GIS species distribution models. This taxonomy is presented in the following paragraphs. At the same time, the review focused attention on issues that are important for appropriate use of the GIS tool in species distribution modelling.

Though offering powerful tools for spatial analysis, GIS has been largely misused and lacks a clear framework to enable users to fully exploit its potential. These issues range from unspecified objectives in the process of model building to lack of adequate support for the assumptions underlying the models themselves. A large part of the chapter is also devoted to the problem of validation which is crucial throughout the entire process of model building and which is seldom taken into account.

[7] Bibliographic database of literature in earth sciences, ecology and geography, published by Elsevier Ltd.

Terminology inconsistencies are also addressed. This problem extends throughout ecology and well beyond the realm of species distribution modelling. In this context, the belief is that the problem results from misleading utilisation of the same term in the different disciplines which have come to coexist under the wide umbrella of GIS.

TERMINOLOGY

Multi-disciplinary fields of science are appealing as they bring together people with different experience and backgrounds, and whose constructive exchange of ideas may generate new solutions. In fact, many solutions that have been successfully developed and used in one field of science may be adapted for use in other fields. The very nature of GIS makes it essential that specialists in different scientific disciplines contribute to the general effort of developing and maintaining common data sets.

One drawback is that in the early phases of tool development (such as GIS), people who master the new tool tend to become generalists, invading other fields of science without having the necessary specific background. This may cause problems both in the solutions provided, which tend to be too simplistic, and in terminology, as the same term and/or concept can be used with different meanings in different disciplines. This is the case, for example, with use of the concept of "scale" and the term “habitat”.

Scale definitions and use

For the cartographer, large scale pertains to the domain of detailed studies covering small portions of the Earth’s surface (e.g. Butler *et al.,* 1986). For the ecologist large scale means an approach that covers regional or even wider areas (e.g. Edwards *et al.,* 1994). This derives from the fact that cartographers use “scale” to mean the ratio between a unit measure on the map and the corresponding measure on the Earth's surface, while the ecologist uses it in the sense of proportion and/or extent. As an example, the relationship between the geographical scale and the extension of ecological studies supplied by Estes and Mooneyhan (1994) highlights that "large" scale in ecology should be associated with “small” geographical scale, and provide the following table for comparison between ecological scale and geographic scale:

site = 1:10,000 or larger
local = 1:10,000 to 1:50,000
national/regional = 1:50,000 to 1:250,000
continental = 1:250,000 to 1:1,000,000
global = 1:1,000,000 or smaller

In ecology it would be better to use the terms "fine" or "broad" (Levin 1992) which places the term scale more in the context of its second meaning.

Habitat definitions and use

Similar to the term “scale”, the different uses of the word “habitat” give rise to major misunderstandings and thus need to be clarified (Hall *et al.,* 1997).

	BIOTA		**LAND**
	Species	*Species and Communities*	
Cartesian space	Begon *et al.,* 1990 Krebs 1985 Odum 1971 Webster’s 1981	Zonneveld 1995	
Cartesian space and environment	Morrison *et al.,* 1992 Mayhew & Penny, 1992	Encycl. Britannica 1994 Yapp 1922	Stelfox and Ironside, 1982 Kerr 1986 USFWS 1980a, 1980b Herr and Queen 1993
Environment	Collin, 1988 Webster’s 1981 Whittaker *et al.,* 1973 Moore 1967		

Table 3.1. Classification scheme of the term habitat. Overview of the various meanings of the term habitat, grouped according to whether the term relates to biota (species or species and communities) or land, and whether it relates to Cartesian space, environmental space or both.

The term habitat[8] forms a core concept when dealing with wildlife management and the distribution of plant and animal species. The fact that the actual sense in which it is used is rarely ever specified suggests that its meaning is taken for granted. However, Webster’s dictionary (1981) provides two different definitions and Morrison *et al.,* (1992) observed that use of the word habitat remains ambiguous. The latter distinguished two different meanings: one concept which relates to units of land homogeneous with respect to environmental conditions and a second concept for which habitat is a property of species.

Our literature review provided us with a variety of definitions and uses of the term “habitat” that are wider than the dichotomy suggested by Morrison *et al.,*

[8] Habitat is also used in connection to mankind. In this article the term habitat refers to plant and animal species excluding human beings.

(1992). These various meanings were arranged according to two criteria: first, whether the term relates to biota (either species and or communities) or to land; second, whether it relates to Cartesian (e.g. location, such as a position defined by a northing and easting) and/or environmental space (e.g. the environmental envelope defined by factors such as precipitation, temperature, land cover etc).

Although the classification in table 3.1 allows us to partition the different definitions of habitat that were traced, in reality this partition is hazy. Definitions range from the place where a species lives (Begon *et al.,* 1990; Webster's 1981; Odum 1971; Krebs 1985), which is a Cartesian space related concept, to the environment in which it lives (Collin 1988; Moore 1967; Webster's 1981; Wittaker *et al.,* 1973). In this latter case habitat is seen as a portion of the environmental space. At both extremes of the range of definitions, the subtle differences in the terms used allow us to define a continuous trend between the Cartesian and the environmental concept. This is further supported by the definitions which combine both the Cartesian and environmental space (Morrison *et al.,* 1992; Mayhew and Penny 1992). These authors define habitat as the area which has specific environmental conditions required for the survival of a species. Note that all of these definitions relate habitat to a species and some describe it as a property of an organism.

With a similar range of definitions, another group relates habitat to both species and communities. For instance, Zonneveld (1995) in accordance with a Cartesian concept, defines it as "the concrete living place of an organism or community". Others relate it both to Cartesian and environmental space, defining it as the place in which an organism or a community lives including the surrounding environmental conditions (Encyclopaedia Britannica 1994; Yapp 1922).

All of the definitions cited so far define habitat in terms of biota. Zonneveld (1995) remarked that the term habitat may only be used when specifying a species (or community). Alternatively the term habitat has been used as an attribute of land. Riparian habitat, for instance, refers to a specific environment, with no relation to biota. Utilisation of habitat in this sense is widespread throughout the ecological literature; e.g. old-forest habitat (Lehmkuhl and Raphael 1993) or woodland habitat (Begon *et al.,* 1990). The concept predominates in ecology applied to land management such as habitat mapping (Stelfox and Ironside 1982, Kerr 1986), habitat evaluation (USFWS 1980a, 1980b; Herr and Queen 1993) and habitat suitability modelling (USFWS 1981). A similar meaning of habitat is used in a review of habitat-based methods for biological impact assessment (Atkison 1985). Although it has been used very frequently in this sense a single definition was not found. A closely related concept, the habitat type which is used in habitat mapping, has been defined as "an area, delineated by a biologist, that has consistent abiotic and biotic

attributes such as dominant or sub-dominant vegetation" (Jones 1986). Daubenmire (1976) noted that this meaning of habitat type corresponds to the land unit concept (Walker *et al.,* 1986; Zonneveld 1989). In articles dealing with habitat evaluation the term is used in a similar sense.

Ambiguous terms cause confusion between scientists. The ambiguity of habitat, however, is also observed within the same publication; for instance in Lehmkuhl and Raphael (1993) where "old-forest habitat" and "owl habitat" are used simultaneously. Even ecological textbooks are not free from ambiguity. Begon *et al.,* (1990) defined habitat as "the place where a micro-organism, plant or animal species lives" suggesting that they consider habitat as a property of a species. However, when outlining the difference between niche and habitat, they later describe habitat in terms of a land unit (p. 78) viz.: "a woodland habitat for example may provide niches for warblers, oak trees, spiders and myriad other species". Confusion arises with respect to habitat evaluation as well. When defined as a property of a species, unsuitable habitat does not exist since habitat is habitable by definition. In this case some land may be classified as habitat and all of this will be suitable. When defined as a land property all land will be habitat, be it suitable or unsuitable for a specific species.

Why is the term habitat used in these various senses? The word originates from habitare, to inhabit. According to Webster's (1981) the term was originally used in old natural histories as the initial word in the Latin descriptions of species of fauna and flora. The description generally included the environment in which the species lives. This leads to the conclusion that habitat was originally used as a species-specific property. It is interesting to note that both geography and ecology have diverted from the original definition of habitat, which suggests that the confusion was not the result of separate developments in two fields of science.

At some time habitat started to be used as a land-related concept, most likely in conjunction with habitat mapping. A possible explanation for the change is given by Kerr (1986), who remarks that mapping habitat[9] individually for each species would be an impossible job. He argues that a map displaying habitat types and describing the occurrence of species in each type would be more useful to the land manager. This suggests that the land-related habitat concept arose because it was considered more convenient to map habitat types rather than the habitat of individual species.

We suggest that there was a second reason for the popularity of habitat type maps. In general the distribution of species will be affected by more than one environmental factor. Until a decade ago it was virtually impossible to display

[9] Here habitat relates to species, while it refers to land in the next sentence.

more than one environmental factor on a single map. The habitat type, defined as a mappable unit of land "homogeneous" with respect to vegetation and environmental factors, circumvented this problem, and was the basis of land systems (land concept) maps developed in the 1980s (Walker *et al.,* 1986; Zonneveld 1989). Land systems mapping is based on the assumption that environmental factors show an interdependent change throughout the landscape, and that the environmental factors are constant within the "homogeneous" area. Thus, to a certain extent the land unit meaning of the term habitat arose as a way to overcome operational difficulties in species distribution mapping. Nevertheless, given that the variation of one environmental factor affecting the distribution of a species tends to be independent of the other environmental factors, "homogeneity" is seldom the case. If "homogeneity" cannot be assumed, then also the simple relationship between species and habitat types is very difficult to accept. It is difficult to defend the assumption that certain habitat types host certain species when it cannot be assumed that they are homogeneous as regards the environmental variables that effect the species distribution.

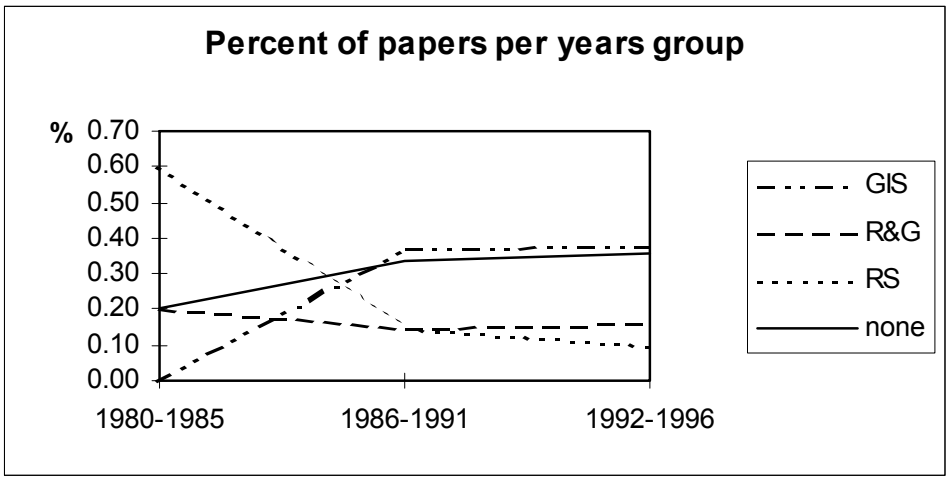


Figure 3.1 Percentage of the papers reviewed which used Remote Sensing (RS), Geographical Information Systems (GIS), both (R&G) or none as their investigation tools.

The advent of GIS has made it possible to store the variation of environmental factors independently and subsequently integrate these independent environmental surfaces into a map displaying the suitability of land as a habitat for a specific species.

The first examples of such GIS-based habitat mapping were published in the second half of the 1980s (e.g. Hodgson *et al.,* 1988). Since then there has been a steady increase of the number of GIS-based habitat models (figure 3.1). The increase illustrates a move away from the general "habitat type" mapping applicable for multiple species towards more realistic species-specific habitat maps. At the same time the habitat type loses its usefulness because of the disappearing necessity to classify land in homogeneous categories. In other words, species-specific habitat mapping is increasingly incorporating independent environmental databases processed using information on the preferences of the species concerned. In view of the anticipated development towards species-specific habitat models, the original species-related concept of habitat instead of a land-related concept is preferred, and to avoid confusion, in this chapter the terms species-environment relationships and ecological requirements are used instead of the terms species habitat and habitat requirements.

GENERAL STRUCTURE OF GIS-BASED MODELS

The rationale behind the GIS approach to species' distribution modelling is straightforward: the database contains a large number of data sets (layers), each one of which describes the distribution of a given measurable and mappable environmental variable. The ecological requirements of the species are defined according to the available layers. The combination of these layers and the subsequent identification of the areas that meet the species' requirements identify the species' distribution range, either actual, if there is evidence of presence, or potential, if the species has never been observed in that area.

This basic scheme can be implemented using different approaches. A few classifications based on different criteria have been attempted. For example, Stoms *et al.,* (1992) classify models based on the conceptual method used to define the species-environment relationship, while Norton and Possingham (1993) base their classification on the result of the model and its applicability for conservation. Accordingly, Stoms *et al.,* (1992) classify GIS species distribution models into two main groups: deductive and inductive, while Norton and Possingham (1993) give a more extensive categorisation of modelling approaches.

We have tried to define logical frameworks that can be used to classify species distribution models based on the major steps needed to build them. To this end, the deductive-inductive categorisation is the most suitable starting point as it focuses attention on the definition of the species-environment relationship, which is the key point for the implementation of distribution models.

The deductive approach uses known species' ecological requirements to extrapolate suitable areas from the environmental variable layers available in the GIS database. Analysis of the species-environment relationship is delegated to the synthesising capabilities and wide experience of one or more specialists who decide, to the best of their knowledge, which environmental conditions are the most favourable for the existence of the species. Once the "preferences" are identified, a logical (e.g. Breininger *et al.,* 1991; Jensen *et al.,* 1992) or arithmetic map overlay operation (e.g. Donovan *et al.,* 1987; Congalton *et al.,* 1993) is used to merge the different GIS environmental layers to yield the combined effect of all environmental variables.

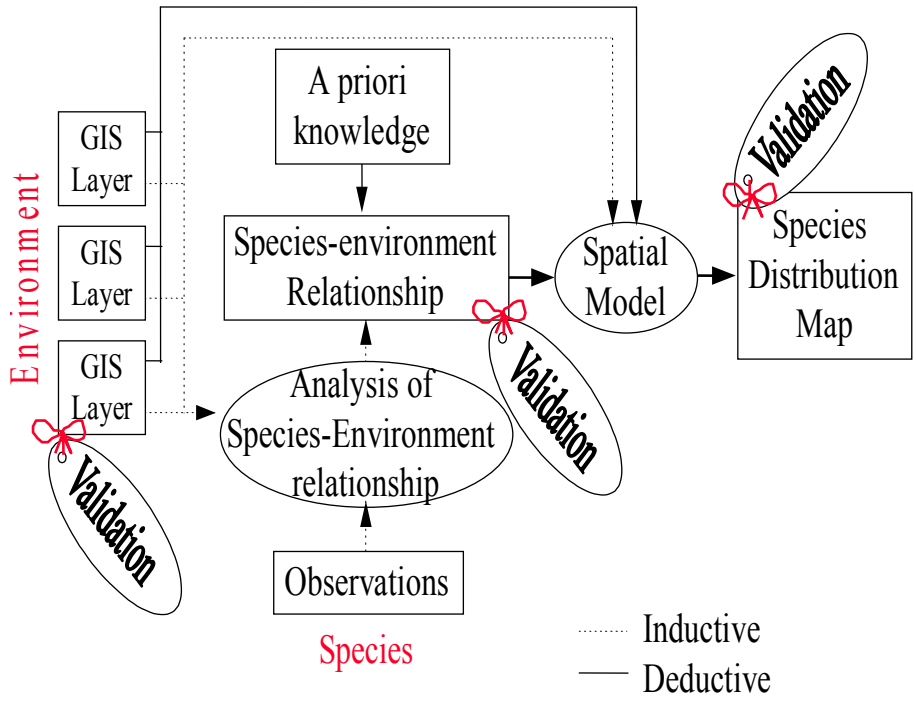


Figure 3.2 Flow diagram of the inductive and deductive approach for species distribution modeling, evidencing the different stages at which GIS layers are involved and the places in which validation can, or should be performed.

When the species-environment relationships are not known *a priori*, the inductive approach is used to derive the ecological requirements of the species from locations in which the species occurs. A species' ecological "signature" can be derived from the characterisation of these locations. Then, with a process similar to the one used in deductive modelling, but which is more objectively

driven by the type of analysis used to derive the "signature", the distribution model is extrapolated (e.g. Pereira and Itami 1991; Aspinall and Matthews 1994).

In Figure 3.2 the data flow of GIS-based species distribution models both for the deductive and the inductive approach are summarised. While in the deductive approach GIS data layers enter the analysis only to create the distribution model, in the inductive approach they are used both to extrapolate the species-environment relationship and the distribution model. Along with the data flow, the steps that need validation are also evidenced in the figure. Validation will be addressed later in this chapter. Nevertheless it is interesting to note here that validation procedures are needed at many different stages in the flow diagram.

Both inductive and deductive models can be further classified according to the kind of analysis performed to derive the species-environment relationship. Essentially these can be subdivided into two categories: the descriptive and the analytical. Models pertaining to the first category use either the specialists' *a priori* knowledge (deductive-descriptive) or the simple overlay of known location of the species with the associated environmental variable layers (inductive-descriptive) to define the species-environment relationship. Descriptive models are generally based on very few environmental variable layers, most often one single layer. They tend to describe presence and absence in a deterministic way. Each value or class of the environmental variable is either associated with presence or with absence (e.g. the species is known to live in savannah with an annual mean temperature of 15° to 20° C, thus savannah polygons falling within the adequate temperature range are to be included as suitable environments). No attempt is made to define confidence intervals to the individual estimate nor is any information provided on the relative importance of one variable (e.g. to follow the previous example, vegetation types vs. temperature) compared with another. Moreover no estimate of the degree of association or of its variability is provided with the relationship.

On the other hand, models that fall into the analytical group introduce variability. For example, advice from different specialists is combined to define species-environment relationships, thus introducing variability in terms of different opinions of the experts (deductive-analytical). Species observation data are analysed in a way that takes into account the range of acceptability of all environmental variables measured, their confidence limits and their correlation. Both the deductive-analytical and the inductive-analytical approach tend to estimate the relative importance of the different environmental layers considered in the analysis, thus moving towards an objective combination of environmental variable layers.

Examples of deductive-analytical models are based on techniques such as Multi Criteria Decision-Making (MCDM) (Pereira and Duckstein 1993), Delphi (Crance 1987), and Nominal Group Technique (NGT) (Allen *et al.,* 1987). Generally speaking these techniques use the advice of more than one specialist as independent estimates of the "true" species-environment relationship and evaluate its variability based on these estimates.

Inductive-analytical techniques rely on samples of locations which are analysed with some sort of statistical procedure. Different techniques have been used ranging from generalised linear models (GLMs) (McCullagh and Nelder 1988; for applications see Akçakaya *et al.,* 1995; Bozek and Rahel 1992; Pausas *et al.,* 1995; Pearce *et al.,* 1994; Pereira and Itami 1991; Thomasma *et al.,* 1991; Van Apeldoorn *et al.,* 1994), Bayes Theorem approach (e.g. Aspinall 1992; Aspinall and Matthews 1994; Pereira and Itami 1991; Skidmore 1989), classification trees (e.g. Walker 1990; Walker and Moore 1988; Skidmore *et al.,* 1996), and multivariate statistical methods such as discriminant analysis (e.g. Dubuc *et al.,* 1990; Flather and King 1992; Haworth and Thompson 1990; Livingston *et al.,* 1990; Verbyla and Litvaitis 1989), discriminant barycentric analysis (e.g. Genard and Lescourret, 1992;), PCA (e.g. Lehmkuhl and Raphael 1993; Picozzi *et al.,* 1992; Ross *et al.,* 1993), cluster analysis (e.g. Hodgson *et al.,* 1987), and Mahalanobis distance (Clark *et al.,* 1993; Knick and Dyer 1997, Corsi *et al.,* 1999).

Further differences should be outlined for models that rely on the interpolation of density or census estimates to extrapolate distribution patterns. Although these models have been included in the inductive-analytical group, the geo-statistical approach (e.g. Steffens 1992) on which they are generally based suggests they might be categorised in a different sub-group.

Another method of classifying GIS distribution models is based on their outputs. Essentially, these can be distinguished into categorical-discrete models and probabilistic-continuous models. Most often the products of the first type of models are polygon maps in which each polygon is classified according to a presence/absence criterion or a nominal category (e.g. frequent, scarce, absent). The products of the second type of model are continuous surfaces of an index that describes species presence in terms of the relative importance of any given location with respect to all the others. Indices that have been used are: suitability index (Akçakaya *et al.,* 1995; Pereira and Itami 1991), probability of presence (Agee *et al.,* 1989; Skidmore 1989; Aspinall 1992; Clark *et al.,* 1993, Walker 1990), ecological distances from “optimum” conditions (Corsi *et al.,* 1999), and species densities (Palmeirin 1988; Steffens 1992). Each of these indices can be mapped as a continuous surface throughout the species’ range.

Generally, discrete models are built associating the presence of a species to polygons of land unit types (e.g. vegetation categories) most often with a deductive approach. This method transfers into the realm of GIS the traditional way of producing distribution maps (*see also* Habitat definitions and use). There are also some examples of binary classifications of continuous environmental variables (e.g. slope, aspect, elevation) using statistical techniques such as logistic regression (Pereira and Itami 1991) or discriminant analysis (Corsi *et al.,* 1999). Categorical-discrete models do not account for species mobility and tend to give a static description of species distribution. Nevertheless this approach can be used to address the problem of defining areas of occupancy (Gaston 1991), and thus can be successfully used for problems of land management and administration. Probabilistic models can describe part of the stochasticity typical of locating an individual of a species and can be used to address problems of corridor design and meta-population modelling (Akçakaya 1993), thus introducing the geographical dimension into analysis of species viability.

	Deductive					**Inductive**					
Descriptive	GIS	GIS & RS	RS	non spatial		GIS	GIS & RS	RS	non spatial		**36**
	9	8	7	8	**32**	3	1	0	0	**4**	
Analytical	GIS	GIS & RS	RS	non spatial		GIS	GIS & RS	RS	non spatial		**46**
	3	0	1	4	**8**	14	4	4	16	**38**	
					40					**42**	

Table 3.2 Classification of reviewed paper. Papers are classified according to the approach used to define the species-environment relationship and whether their approach was simply descriptive or analytical. Further subtopics indicate whether the author considers its research pertaining to the domain of remote sensing, GIS or both. Non-spatial is used for those papers that do not contain an explicit distribution model but define species-environment relationship in terms of variables that can be mapped.

LITERATURE REVIEW

Table 3.2 indicates the results of the bibliographic review. Papers are classified according to the categories described in the previous paragraph. We have considered both GIS and RS as two different views of the same tool, the former being devoted more to spatial-correlation analysis and the latter more concerned with basic data production. In fact, the two families of software tools share a great number of basic functions and are evolving towards integration into a

single system. It should be noted that included in the review are not only papers that use GIS and/or RS but also some that deal with HSI, HEP and more general assessment of species' ecological requirements. The papers in this last group do not generally represent examples of spatial models (Scott *et al.,* 1993) in the sense that their products are not distribution maps, but they have been included because they are considered to be just a few steps away from a real distribution model. In fact they describe the ecological requirements of the species in terms of mappable environmental conditions.

Most of the papers that use the deductive approach consider the *a priori* knowledge sufficient to define the ecological requirements of the species under investigation. This is especially true of papers that model distribution on the basis of interpretation of remotely sensed data. In fact 15 out of 16 papers pertaining to the deductive group, and which used remotely sensed data to model species distributions, fall within the descriptive group. In these papers, image classification techniques tend to receive more emphasis, while the ecological application is most often seen as an excuse to apply a specific classification algorithm.

The time trend of the papers published shows stable use of RS technology and increasing use of GIS. Before 1986, no paper makes explicit reference to the term GIS, even though some of the papers dealing with the utilisation of RS use raster GIS-style overlay procedures to define their distribution models (e.g. Lyon 1983) and others use a spatial approach but do not mention the term "GIS" (e.g. Mend *et al* 1981).

Little is generally said about model assumptions. Of the 82 papers reviewed, only 21 discuss their assumptions to some extent. The ones that do generally limit their discussion to the statistical assumptions of the technique used to perform the analysis. Very few deal with the biological and ecological assumptions and tend to take them for granted. When dealing with ecological modelling both biological and methodological assumptions must be taken into account, along with some general assumptions which may limit the applicability of the results produced (Starfield 1997).

When performed, validation, a step which is evidenced at different levels in the data flow diagram (figure 3.2), is generally limited only to the accuracy of the result of the analysis (e.g. distribution map). Nothing is said about the accuracy of the original data sets (e.g. GIS data layers, observation locations etc.), and no consideration is given to issues such as error propagation in GIS overlay (e.g. Burrough 1986).

Only 15 papers validate of the accuracy of their results based on an independent estimate of the distribution (either through comparison with an independent set of observations or through comparison with the known distribution of the species). Interestingly, 50% of these papers are based on the deductive approach. In fact it should be noted that since observation data sets are the most expensive data to be collected within the general framework of setting up a GIS species distribution model, the deductive approach is the most cost effective if seen from the validation point of view. In fact, to avoid bias, a model developed with an inductive approach cannot be validated using the same data set used to derive the species-environment relationship. Thus validation can be performed either with a second, independent data set or by dividing the original data set into two subsets, one of which is used to derive species-environment relationships and the other to validate the resulting model.

It is interesting to note that the multidimensional power of GIS is yet to be supported by adequate quantity and quality of geographical data sets (Stoms *et al.,* 1992). This is reflected in the number of environmental variables used in analysis. In the papers reviewed, the average is just below 4.8, and only 9 out of 82 analyse more than 9 environmental variables, whereas 23 papers base their distribution models upon only one environmental variable, generally vegetation.

MODELLING ISSUES

Based on the results of the literature review, five major issues have been identified that need to be addressed to allow a sound GIS modelling of species distributions: 1) uncertainty in the objectives of the research; 2) lack of adequate support for assumptions underlying the implementation of GIS models; 3) scale, in both time and space dimensions; 4) lack of data availability that limits the types of models that can be developed; 5) inadequate validation and accuracy of GIS models.

Clear Objectives

When setting up an ecological model, the first step is to make a clear statement of the model's objective (Starfield 1997). There is great confusion about the objectives of many published papers. This may be due either to over-qualification of the tool, in the sense that use of the tool becomes the objective of the paper, or to uncertainty in defining the model's goals, along with coexisting purposes of predicting and/or understanding (Bunnell 1989). For example, most of the papers based on the inductive approach deal with the definition of a species-environment relationship without specifying whether they intend to analyse the relation of cause and effect or simply use the relation

as a functional description of the effect. In the first case, the goal would be to evidence the limiting factors that are related to the species' biological needs and that drive the distribution process; in the second, it would be the simple use of correlated variables whose distribution is functional to the description of the species' distribution.

Basically species needs in a) food, b) shelter and c) adequate reproduction sites can be summarised (e.g. Flather *et al.,* 1992; Pausas *et al.,* 1995). When using the distribution of an environmental variable to describe the species' distribution it is implicitly assumed that there is a correlation between these basic needs and the environmental variables used. This correlation can be causal, that is, it describes the species' basic needs. In such cases a function is identified that, within a reasonable range of values, associates each value of the environmental variable to a measure of the fulfilment of the species' basic needs (e.g. reproductive success). But it can also be a functional description; it is unknown why some ranges of values of the environmental variable are preferred by the species but observation shows that the species tends to occur more frequently within those ranges. The variable might influence all the species' basic needs simultaneously, or be correlated to another variable that describes one or another of the species' needs.

Generally speaking the quantity and quality both of the locational data and of the GIS layers used in analyses are not sufficient to assess cause-effect relationships which determine the species' distribution. Furthermore, cause-effect relationships spring from the interactions of biophysical factors that range through different time and space scales (Walters 1992). Few papers take scale dependency into account in their analysis. Moreover, in this kind of analysis causal effects can be hidden by independent interfering variables (Piersma *et al.,* 1993) or by the unaccounted stochasticity of natural events such as weather fluctuations, disturbance and population dynamics (Stoms *et al.,* 1992), and should be assessed in controlled environments.

Such uncertainties can be addressed by defining the overall goal as the assessment of the relationship which best describes the species distribution. In other words, even if the causal understanding of a relationship is not clear, whenever the species-environment relationship is able to describe the distribution of a species satisfactorily, the overall goal is achieved (Twery *et al.,* 1991).

Obviously this approach has some drawbacks. Without an adequate description of the cause-effect relationship between the species and environmental variables, models lose in transferability - both in space and time - and this limits their predictive capabilities (Levin 1992).

Assumptions

All models analysed extrapolate their results to an entire study area on the assumption of space independence of the phenomenon observed at a given place. That is, in the case of both a deductive and an inductive approach, the species-environment relationship is built on evidence that a certain species occurs somewhere and that the values of the environmental variables at those locations are known. Obviously it is known that a species occurs at locations where it has been observed. Only part of these locations have measurements of the environmental variables and usually these measurements are collected only for the limited time range during which the investigation was carried out. Thus, when building distribution models, evidence collected within a portion of the range is extrapolated to the entire range of occurrence of a species. In order to do so, it is assumed that the species-environment relationship used to build the model is invariant in space and time. Usually this is not the case, especially for species with a wide range and/or for generalist species. In fact, the higher the variance of the species-environment relationship, the higher the number of locations required to provide an adequate ecological profile for the species.

Secondly, it is generally implicitly assumed that variables which are not included in the analysis have a neutral effect on the results of the model. Either it is assumed that the species' ecological response to these environmental variables is constant or that the response is highly correlated with the other variables included.

Even though both of these general assumptions are difficult to test, they should be discussed on a case-by-case basis since the result of the violation of these assumptions is species-specific, and cannot be generalised. Errors may be negligible in certain cases but can introduce major interpretation problems in other cases.

Biological assumptions

Biological assumptions are direct consequences of the general assumptions discussed in the previous paragraph. While they have received minimal attention in the literature, they are probably the most critical.

The first assumption, which follows from the general assumption of space and time independence, states that observations reflect distribution. This implies that observation data also indicate absence (Rexstad *et al.,* 1988, Clark *et al.,* 1993), which is not the case. In fact any time there is a record for a species it is certain that the species occurs, at least occasionally, at that location. In contrast, if there is no observation for a species, it can only be assumed that there is a record of absence if there is no bias in the sampling scheme, and that observations have

been conducted over a sufficiently long period. Even then there is no way to evaluate the random effects that are intrinsic in observing animals.

These assumptions can have statistical relevance when dealing with inductive-analytical approaches, but must hold true also for the deductive models. If there is a constant bias in the visibility of a species' individuals, for instance because part of their range is less accessible than others to researchers and thus cannot be as carefully investigated, the species-environment relationship will reflect this bias. For instance observation data are often gathered through sightings carried out by volunteers (e.g. Stoms *el at.* 1992; Hausser 1995) which do not follow a pre-defined (e.g. random) sampling scheme. Or habitat cover may limit observations to where the species is visible (Agee *et al.,* 1989). This may create an artificial response curve that associates a positive relation to the values of the environmental variables measured in the locations were the species is more visible and a negative one in those measured in areas were the species has been less investigated. In such cases, the areas where the species and the observers are most likely to meet would be mapped, and not the true distribution of the species.

The example is tailored to inductive-analytical models but can easily be extended to deductive ones, both descriptive and analytical. The deductive approach is based on the *a priori* knowledge of specialists who rely on series of observations to gain experience and define the species-environment relationship. Again, these observations can suffer from accessibility or visibility biases.

A further assumption is that observations reflect the environmental selection of the species. Obviously this is not always true as, for example, when observations include occurrences of migrant and/or vagrant individuals whose presence in a given location is occasional. The bible [Exodus 10:19] gives us an extreme example of locusts swarms blown by strong winds into the middle of the desert. Clearly, their presence does not reflect an ecological preference. Nevertheless if only the observation *per se* is considered, it would be concluded that high densities of locusts are found in the desert, thus that locusts do prefer (with all the limitations that this term carries in such an analysis) desert environments. Obviously the strong wind should be regarded as a stochastic event and thus be treated as an outlier in the definition of a possible GIS distribution model. In other words, observations should be analysed for their content of unconstrained selection by the species.

When dealing with the issues of scale, GIS distribution models tend to describe only the deterministic components that drive a species' distribution pattern. Thus stochastic events need to be either averaged over the long term and/or eliminated as outliers. When observations are carried out for a limited time and

the biology of the species under investigation is scarcely known, this problem can become increasingly important as the identification of outliers will become virtually impossible.

Statistical assumptions

Most of the statistical techniques used to define species-environment relationships rely on the identification of two observation sets: one that identifies locations in which the species is present and the second in which it is absent. Even though this cannot be properly identified as a statistical assumption, it is probably the most important factor limiting the applicability of the statistical techniques that rely on the two groups of observations.

The most common procedure to define the two subsets is to compare locations of known presence with a random sample of locations not pertaining to the previous set. Obviously some of the random locations can represent a suitable environment for the species, thus introducing, for that particular environment, a bias which underestimates the species-environment association.

To overcome this problem, data sets can be screened for outliers (Jongman *et al.,* 1995), using for instance a scatter plot of the variables taken two by two. Once an outlier is identified it can be checked to identify possible reasons for the absence of the species and, if necessary, removed from the analysis. Similar results can be achieved through analyses such as decision trees, where additional rules can be introduced to predict outliers (Walker 1990; Skidmore *et al.,* 1996).

Another way of getting around the problem is by eliminating the absence subgroup. Skidmore *et al.,* (1996), for example, have used both the BIOCLIM approach and the supervised non-parametric classifier, which use only observation sites to derive distribution patterns. The same result can also be achieved by using distance (or similarity) measures from the environmental characteristics of locations in which the species has been observed. A measure of distance which seems particularly promising for this application is the Mahalanobis distance (Clark *et al.,* 1993; Knick and Dyer 1997). It has many interesting properties as compared to other measures of similarity/dissimilarity, the most appealing of which is that it takes into account not only the mean values of the environmental variables measured at observation sites, but their variance and covariance as well. Thus the Mahalanobis distance reflects the fact that variables with identical means may have a different range of acceptability, and eliminates the problem that the use of correlated variables can have in the analysis.

Along with the identification of presence/absence data sets, it is important to recall that each statistical method has some specific assumption which must be

satisfied for correct application of the technique. For example, nonparametric statistical tests may assume that a distribution is symmetrical, while a parametric test may assume that the test data are normally distributed. The assumptions of the different statistical methods will not be further discussed as it is beyond the scope of this chapter. The reader is referred to more specific books and journal papers on statistical methods.

Spatial and temporal scale

Scale is a central concept in developing species distribution models with GIS. As explained in the paragraph devoted to terminology, this concept is common to both geography and ecology, the two main disciplines involved in the development of GIS species distribution models. The concept of scale evolves from the representation of earth surface on maps, and is the ratio of map distance to ground distance. Scale determines the following characteristics of a map (Butler *et al.,* 1986): 1) the amount of data or detail that can be shown; 2) the extent of the information shown; and 3) the degree and nature of the generalisation carried out. This group of characteristics determines the quality of the layers derived, that is, the quality of the environmental variables stored in the GIS database and the type of species-environment relationship that can be investigated (Bailey 1988; Levin 1992; Gaston 1994) using the capabilities of the GIS.

The scale of the analysis influences the type of assumptions that need to hold true for sound modelling. To clarify this concept, the fact that species distribution is the result of both deterministic and stochastic events needs to be considered. The former tend to be described in terms of the co-existence of a series of environmental factors related to the biological requirements of the species, while stochastic processes are regarded as disturbances caused by unpredictable or unaccountable events (Stoms *et al.,* 1992). Generally distribution models are built upon deterministic events and are averaged over wide spatial and temporal ranges to minimise the error related to the unaccounted stochasticity.

As has been seen, GIS distribution models rely on species-environment relationships to extrapolate distribution patterns based on the known distribution of the environmental variables. It has also been seen that the relationships reflect the biological needs of the species. The extent to which temporal and spatial scales need to be coarsened depends on the stochastic events that need to be minimised, which in turn depend essentially on the dynamics of the species under investigation. To this extent, it is important to note that major population dynamics events happen at different scales in both time and space. In figure 3.3 (modified from Wallin *et al.*, 1992) the two axes indicate the increasing

temporal and spatial scale at which population dynamics events happen. In accordance with the hypothesis formulated by other authors (O'Neill *et al.,* 1986; Noss 1992), the figure shows a positive correlation between space and time scales; that is, events that happen on a broader spatial scale are slower and thus take more time.

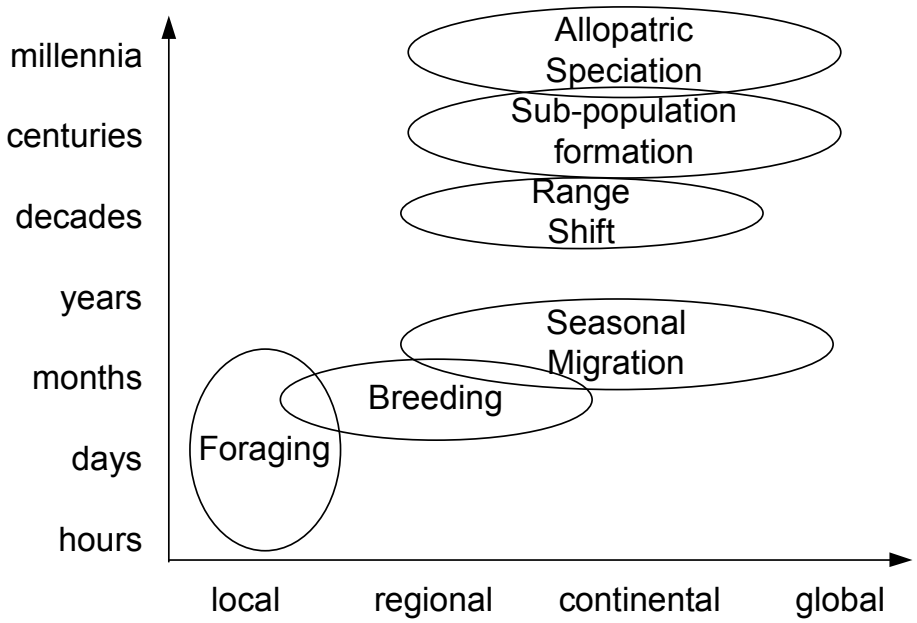


Figure 3.3. Population dynamics event in relation to time and space scales (modified from Wallin *et al.,* 1992).

As a tool for distribution modelling this graph can be of great help in defining scale thresholds both towards a minimum and a maximum scale for an analysis. For instance, when considering cause-effect species-environment relationships, the processes involved (e.g. feeding behaviour) must be analysed at an adequate scale (e.g. in the example, very detailed scale both in time and space). On the other hand, if the stochasticity introduced in the observation scheme needs to be overcome by, for instance, individual foraging behaviour, the results must be averaged on a coarser scale both in time and space.

Thus, in GIS distribution models, both temporal and spatial scales are generally broadened so that stochastic events can average to a null component and thus be ignored. For example, the stochasticity associated with the individual selection of a particular site, which greatly influences the distribution at a local scale, is

overcome when dealing with distributions at regional scale averaging the selection of different individuals. Stochastic events such as local fires, which influence regional distributions when measured over a short time interval (e.g. 5 to 10 years), are considered outliers in an analysis that takes into account the average vegetation cover over a longer time or a wider spatial span. Similarly, it is known that in short time intervals the population dynamics status of a population is highly unpredictable, whereas it may be more easily averaged on longer time scales (Levin 1992), to become scarcely predictable again at even longer intervals.

A similar consideration is intrinsic in the minimum mapping unit (MMU), a concept largely used to address spatial scale issues in GIS species distribution models (e.g. Stoms 1992; Scott *et al.,* 1993) and which can be readily extended to the time scale. MMU can be seen from two points of view. On one hand, it is a property of the data set that is being analysed, that is, the minimum dimension of an element (e.g. a polygon representing vegetation types of a given category, the time span between successive manifestations of a given ecological event) that can be displayed and analysed. On the other hand, it indicates the kind of averaging which needs to be carried out to smooth noise introduced by stochasticity. In fact, in the case of local fires, if the MMU is defined as larger than the extent of the fires both in time and in space, the fire events are excluded from the analysis.

When dealing with scales on a practical basis, it should be noted that the structural complexity of distribution modelling can be simplified according to the hierarchical hypothesis (O'Neill *et al.,* 1986) that states that at any given scale there are particular environmental variables that drive the ecological processes. Thus weather becomes important at very broad spatial scales (e.g. continental scale). This is the basis of approaches behind models such as BIOCLIM (Busby 1991), that of Walker (1990) and that of Skidmore *et al.,* (1996); all of them describe species distribution at a continental scale in terms of their direct relation to climatic data. At successively finer scales such as regional landscapes, land form and topography play an important part (e.g. Haworth and Thompson 1990; Aspinall 1992; Flather *et al.,* 1992; Aspinall and Veitch 1993), while at the most local scales, indigenous land-use structures become increasingly significant (e.g. Thomasma *et al.,* 1991; Picozzi *et al.,* 1992; Herr and Queen 1993) to the extent that even an individual stand of timber (Pausas *et al.,* 1995) or a single pond (Genard and Lescourret 1992) can play a role. The factors that are important vary according to scale, meaning that factors which are important at one scale level can greatly diminish their importance at others (Noss 1992).

As with any type of classification, the relationship between scale and environmental variables that drive ecological processes should not be taken too

rigidly. While most authors agree that for broader scales climate is the most important factor, the same cannot be said when trying to identify the driving forces at finer scales. For example, variables considered useful at coarser scales are used in detailed studies, as in the cases of Pereira and Itami, (1991) and Ross *et al.,* (1993), which use topography to explain species distribution at a much finer scale than the regional one. And the same consideration applies to Aspinall and Matthews (1994), which use climatic data on a regional scale. On the other hand, land use/land cover is often used in distribution models developed at regional scale (e.g. Livingston *et al.,* 1990; Flather and King 1992).

Finally, it must be considered that distribution is the result of the interaction of many different biological events. An ecological event cannot be exhaustively described on any single specific scale, but is the result of complex interactions of phenomena happening at different scales (Levin 1992; Noss 1992). The challenge is to integrate different scales in the description of the species' distributions. Buckland and Elston (1993) give an example of the integration of environmental variables stored at different resolutions within the same distribution model.

It is important to note that the concept of scale not only determines the biological extent to which a distribution model can be applied, but that it also affects the use that can be made of such a model for conservation. Conservation actions can be seen as having a hierarchical approach (Kolasa 1989). For example, Scott *et al.,* (1987) identify six different levels of intervention: landscape, ecosystem, community, species, population and individual. Not surprisingly, conservation actions tend to become more effective and less expensive when the assessment moves towards broader scales, that is, when one moves from the individual to the landscape approach (Scott *et al.,* 1987). Obviously this relates only to the extent of the analysis and not to its resolution. Nevertheless it can be stated that, on a cost/benefit basis, it is generally more efficient to address conservation-related issues at a coarser scale that enables a landscape approach, than it is to concentrate on a more detailed scale (e.g. individual, and/or population level) that requires high resolution data to be analysed.

Economics suggest that conservation science take a broader view of phenomena. A broad-scale approach and the possibility of predicting the potential dynamics of spatial patterns are needed to manage fragmentation of suitable environments and the inevitable meta-population structure of the resulting population (Noss 1992). May (1994) indicates that when multiple levels of biological organisation are concerned, as in a typical conservation action, the best management approach can be achieved on the regional landscape scale (10^3 to 10^5 km^2). This scale level has suffered historically from

limitations in the tools available for consistent analysis and is the one that has gained the most from the evolution of GIS; in fact most of the distribution models based on GIS address problems at the regional landscape level.

Availability of environmental data

While use of GIS in species distribution modelling can transform limited data about species into decision-making tools, availability and quality of environmental data are two of the three limiting factors in the development of GIS-based species distribution models (the other being reliability of the models themselves [Stoms *et al.,* 1992] which will be discussed later in this chapter). The problem of developing extensive data sets of environmental variables is currently limited by economic and political rather than technical constraints. Estes and Mooneyhan (1994) list a number of different attitudes of governments throughout the world that limit the availability of high resolution, "science-quality"[10] environmental data sets. These range from military classification of the data, thereby precluding the use of the data to the scientific community, to the low political priority that certain governments give to environmental issues. Moreover even when policy is not an obstacle to the production and availability of data sets, entire nation-wide data sets are sometimes lost during revolutions, wars and/or civil disturbances. To this it should be added that some governments (e.g. the European Union countries) ask high prices for data sets, which are generally acquired with tax money, actually preventing their broad use in any type of activity and more specifically in environmental research.

In many cases, high quality site-specific data sets are generated for a particular research project but are compiled with non-standard techniques, thus rendering them unsuitable for combination and the achievement of more extensive knowledge of an area.

In the past few years there has been an increasing effort to develop meta-databases of available data sets throughout the world. Meta-databases are being developed by national and international organisations (e.g. UNEP, World Bank, USGS, EEA). These initiatives still do not address the problem of producing high quality data sets, but at least it is a start in collating existing information on geo-spatial data. As for data set production, an important example is given by the joint efforts of the US Geological Survey (USGS), the University of Nebraska-Lincoln, and the European Commission's DG Joint Research Centre which are generating a 1-km resolution Global Land Cover Characterisation

[10] '... "science-quality" means that, in so far as both practical and possible, the errors inherent in the overall production of [these] maps have been documented.' (Estes and Mooneyhan, 1994).

(GLCC) database suitable for use in a wide range of environmental research and modelling applications from regional up to continental scale. All data used or generated during the course of the project (source, interpretations, attributes, and derived data), unless protected by copyrights or trade secret agreements, are distributed through the Internet. This effort goes in the direction of producing and distributing homogeneous medium resolution quality data sets, with known standards of accuracy.

Further aspects on raw data sets will be discussed in the validation paragraphs in which issues on the quality of the data used to build models will be analysed. Here this issue will not be further discussed as it is not believed to be a problem that can be addressed directly by conservation biologists or ecologists, even though they can contribute to develop awareness of the need for standardisation of the data sets and for the production and dissemination of those that they require.

Validation and accuracy assessment

The main function of a GIS-based species distribution model is to produce a map or its digital analogue for assessment of management and conservation actions. Possibly the most important question to be asked by a user is 'how accurate is the distribution map that has been produced?'.'

Many articles have been written on the sources of error in the data layers that may be included in a GIS. Nevertheless few authors of papers dealing with animal distribution include an assessment of the accuracy of their model and a validation of the product. As we believe this issue to be central for the entire process of species distribution modelling, the aim of this section is to review sources of error in GIS, to discuss methods of assessing mapping accuracy and to evaluate the accumulation of thematic map errors in GIS, thus providing a framework for assessment of the accuracy of distribution models developed with GIS.

Sources of errors
GIS data layers are traditionally classified according to their data structure, either raster or vector. To a certain extent, both error sources and accuracy evaluation methods have been investigated following this traditional classification.

Raster images may be obtained from remote sensing instruments carried by aircraft or spacecraft platforms, or by converting an existing line map (vector data structure) to a raster data structure. Two types of error are inherent in remotely sensed images, viz. geometric and radiometric. These error sources are

addressed in detail in numerous monographs and papers including Colwell (1983) and Richards (1986).
A raster image is usually made up of a regular grid of adjacent rectangular cells or pixels (i.e. a rectangular tessellation).

Geometric error in a remotely sensed image is caused by: 1) movement in the remote sensing platform; 2) distortion due to the earth's curvature and terrain; 3) different centrifugal forces from earth affecting spacecraft movement; 4) earth rotational skew; 5) distortions introduced by the remote sensing device itself, including systematic distortions caused by sampling sequentially from each detector, and non-linear scanning (Adomeit *et al.,* 1981); and 6) errors introduced by the georeferencing process. Geometric error causes a point on the remotely sensed image to occur in the wrong position relative to other points in the image.

Correction of geometric errors in remotely sensed data is now a routine aspect of their pre-processing. The map or image is usually "rubber sheet stretched" to fit it to an appropriate map projection. Corrected images with geometric errors of less than 0.5 pixel are now obtainable and acceptable[11] (Ford and Zanelli 1985; Ehlers and Welch 1987; Skidmore 1989b). However, the base maps from which control point information is derived may be of poor quality. Bell (1986) reported that maps used to geometrically correct images of the Great Barrier Reef contained errors of up to 1 km. The accurate selection of control points is crucial in obtaining acceptable results.

Points within a rubber sheet stretched image will no longer be on a regular grid as they have been warped to fit into the projection defined by the Ground Control Points (GCP). To obtain a regular grid, an interpolation method is employed to nominate a value for a regular grid point which falls between the points in the rubber sheet stretched image. Lam (1983) provides an excellent review of other interpolation methods including splines, finite difference and kriging.

Radiometric errors occur as a result of differential scattering of electromagnetic radiation at varying wavelengths, sensors that have poorly calibrated multiple detectors within a band, sensor calibration error, signal-digitisation error, and scene-specific error such as off-nadir viewing, irradiance variation and terrain topography (Richards 1986). Correction of band-to-band distortion is performed using image histograms (shifted to the origin to remove atmospheric scattering

[11] Note that the accuracy of the geometric correction is sometimes expressed as 'root-mean-square' (rms) error, which is the standard error (of the difference between the transformed GCPs and the original GCPs) multiplied by the pixel size.

effects), while line striping effects are reduced by calibration of detectors or by matching detector statistics during computer processing (Teillet 1986).

A final type of error may be caused by a time lag between ground truthing and image collection. In this case, pixels may be noted as incorrect in the error matrix (see below for description of an error matrix), when they may be actually correct at the time of image acquisition.

Vector "images" have been traditionally recorded and stored as maps. Maps are subject to many errors. Some errors are introduced during the creation of the map, such as the original line smoothing by draughtsmen who may not follow the true isolines on the ground (Chrisman 1987). Other errors may be associated with the physical medium used to "store" the map (for example paper stretch and distortion).

Maps may be represented in computer GIS by a variation of the vector data structure (Peuquet 1984), or converted to a raster data structure. In its simplest form, the vector data structure has map lines approximated to a set of points (nodes), which are linked by lines (or arcs). Vector data may be obtained by digitisation.

Digitisation introduces a number of errors. Varying line thickness on the original map requires automatically scanned vector lines to be thinned. During manual digitisation the centre of the map line must be carefully followed if the map lines vary in thickness (Peuquet and Boyle 1984). This requires very careful hand digitising or high accuracy automatic scanners. The number of vertices (points) used to approximate a curve is also critical (Aldred 1972). Too few vertices will result in the line appearing stepped, while too many vertices create large data volumes. Thus, even with extreme care, error is introduced during digitisation.

As for raster images, the main method of correcting geometric error in vector images is by using ground control points from a cartographically correct map to transform the vector image to a known projection.

Data layers error quantification

Methods for quantifying error in a raster data layer are based on the error matrix (also called a contingency table or confusion matrix) concept, first expounded for remotely sensed data in the 1970s (e.g. Hoffer 1975).

The aim of the error matrix is to estimate the mapping accuracy (i.e. the number of correctly mapped pixels) within an image. An error matrix is constructed from points sampled from the image. The reference (or verification) data are normally represented along the columns of the matrix, and are compared with

the classified (or image) data represented along the rows. The major diagonal of the matrix represents the agreement between the two data sets (Table 3.3).

To check every cell for correctness would be impossible except in the smallest map area, so various sampling schemes have been proposed to select pixels to test. The design of the sampling strategy, the number of samples required, and the area of the samples have been debated by the remote sensing and GIS community.

	Class	**Classification of pixels (ground thruth)**									
		I	II	III	VI	V	VI	VII	VIII	IX	Total
	I	14	3	7	2		5		7		38
	II		14	3				3	4	1	25
	III		1	16	1		1	1	1		21
Classification	IV	4		8	3		2	1	3		21
of	V	2				16				1	19
Pixels	VI			1		1					2
(model)	VII		1	3				3			7
	VIII								2		2
	IX										
Total no. of pixels		20	19	38	6	17	8	8	17	2	135

Overall classification accuracy* 50.4%

Table Legend: I = Yertchuk; II = Gum/Stringybark; III = Silvertop Ash; IV = Blue-leaved Stringybark; V = Clearcut/road; VI = Tea Tree; VII = Gum/Silvertop Ash; VIII = Black Oak; IX = unclassified

* Ratio of the sum of correctly classified pixels in all classes to the sum of the total number of pixels tested.

Table 3.3: Example of error matrix drawn from the classification of vegetation types from a satellite image

As with any sampling problem, one is obviously trying to select the sampling design which gives the smallest variance and highest precision for a given cost (Cochran 1977). A number of alternative designs have been proposed for sampling the pixels to be used in constructing the error matrix. Berry and Baker (1968) recommended the use of a stratified systematic sample[12]. The advantage of systematic sampling over random sampling is that sample units are distributed equitably over the area. The disadvantage is that the resulting sample is weighted in favour of the class covering the largest area, and those classes with a small area may not be sampled at all.

[12] Each stratum has an unaligned systematic sample.

Simple random sampling in land evaluation surveys emphasises larger areas and undersamples smaller areas (Zonneveld 1974). Zonneveld (1974) suggested that a stratified random sample was preferable, and Van Genderen *et al.,* (1978) agreed that a stratified random sample is "the most appropriate method of sampling in resource studies using remotely sensed data".

Rosenfield *et al.,* (1982) suggested a stratified systematic unaligned sampling[13] procedure (i.e. an area weighted procedure) as a first stage sample to assist in identifying categories occupying a small area, followed by further stratified random sampling for those classes with fewer than the desired minimum number of points. Todd *et al.,* (1980) argued that single stage cluster sampling is the cheapest sampling method, as multiple observations can be checked at each sample unit on the ground.

Congalton (1988) simulated five sampling strategies (viz. simple random sampling, stratified random sampling, cluster sampling, systematic sampling and stratified systematic unaligned sampling) using a different number of samples over remotely sensed images of forest, rangeland and grassland. The aim of the study was to ascertain the effect of different sampling schemes on estimating map accuracies using error matrices. He concluded that great care should be taken in using systematic sampling and stratified systematic unaligned sampling because these methods could overestimate population parameters. Congalton (1988) also stated that cluster sampling may be used, provided a small number of pixels per cluster are selected (he suggested a maximum of 10 sample pixels per cluster). Stratified random sampling worked well and may be used where small but important areas need to be included in the sample. However, simple random sampling may be used in all situations.

The number of samples may be related to two factors in map accuracy assessment: 1) the number of samples that must be taken in order to reject a map as being inaccurate; and 2) the number of samples required to determine the true accuracy within some error bounds for a map.

Van Genderen *et al.,* (1978) pointed out that for a given number of sample pixels, we wish to know the probability of accepting an incorrect map. In other words, when high mapping accuracy is obtained with a small sample (e.g. 10 items), there is a chance that no pixels which are in error may be sampled (i.e., a type II error[14] is committed). The corollary as stated by Ginevan (1979) is also

[13] According to Berry and Baker (1968).

[14] Type I errors have been termed "consumer's risk" and type II errors "producer's risk" by Fung and LeDrew (1988), and others. These terms are taken from a branch of

important, that is, the probability of rejecting a correct map (i.e. committing a type I error) must also be determined.

Van Genderen and Lock (1977) and Van Genderen *et al.,* (1978) argued that only maps with 95 per cent confidence intervals (i.e. b = 0.05) should be accepted and proposed a sample size of 30. Ginevan (1979) pointed out there is no allowance made by Van Genderen *et al.,* (1978) for incorrectly rejecting an accurate map. The trade-off one makes using Ginevan's more conservative approach (1979) is to take a larger sample, but in so doing the chance of rejecting an acceptable map is reduced. Hay (1979)[15] concluded that minimum sample size should be 50, greater than that of Van Genderen and Lock (1977).

The number of samples must be balanced against the area covered by a sample unit, given a certain quantity of money to perform a sampling operation. "Should many small-area samples or a few large-area samples be taken?" was a question posed by Curran and Williamson (1986). The answer is that it depends on the cover type being mapped. A highly variable cover type such as rainforest is better suited to many small-area samples, while for relatively homogeneous cover types it is more efficient to take fewer large-area samples. The reasons for this conclusion are discussed below.

Generally, mapping of heterogeneous classes such as "forest" and "residential" is more accurate at 80 m resolution than at finer resolutions such as 30 m (Toll, 1984). However more homogeneous classes such as agriculture and rangelands are more accurately mapped at 30 m than at 80 m (Toll, 1984). The reason for this is the trade-off between ground element size and image pixel resolution.

Based on the error matrix, different measures of accuracy can be derived. A commonly cited measure of mapping accuracy is the "overall accuracy" which is the number of correctly classified pixels (i.e. the sum of the major diagonal cells in the error matrix) divided by the total number of pixels checked (Table 3.3). Anderson *et al.,* (1976) suggested the minimum level of interpretation accuracy in the identification of land-use and land-cover categories should be at least 85 per cent.

Overall classification accuracy is the ratio of the total number of correctly classified pixels to the total number of pixels in each class (Kalensky and Scherk 1975). Cohen (1960) and Bishop *et al.,* (1975:396-397) defined a measure of overall agreement between image data and the reference (ground

statistics called acceptance sampling. For the sake of consistency and in order to use conventional statistical terms, type I and type II errors will be used here.

[15] Hay (1979) noted the ideas expounded by Van Genderen and Lock (1977) were developed by himself.

truth) data called "Kappa" or "K". K ranges in value from 0 (no association, that is, any agreement between the two images equals chance agreement) through to 1 (full association, there is perfect agreement between the two images). K can also be negative, which signifies a less than chance agreement.

The methods discussed above for quantifying error in raster images are equally applicable to quantifying error in vector polygons. Instead of checking whether an image pixel is correctly classified, a point within the polygon is verified against the ground truth information. A specific problem encountered with vector images is ground truth samples which occur across boundary lines; in this case the class with the largest area within the sample area may be selected to represent the vector map image.

A method of assessing map accuracy based on line intersect sampling was described by Skidmore and Turner (1992). Line intersect sampling is used to estimate the length of cover class boundaries on a map[16] that coincide with the true boundaries of the cover classes on the ground. A ratio of coincident boundary to total boundary is proposed as a measure of map accuracy and this ratio is called the "boundary error". Though this technique has been developed for vector maps, it is equally applicable to raster maps.

Skidmore and Turner (1992) found that the true boundary lengths were not significantly different from the estimated boundary lengths sampled using line intersect sampling, with a' = 0.05. The estimated boundary accuracy (64%) was extremely close to the true boundary accuracy (65.1%), and there was no significant difference between the true and estimated boundary accuracy.

Reliability of the output

In addition to assessing the accuracy of the original data sets used to produce GIS models, the final products of the modelling effort must be validated.

As previously evidenced, the process of model building takes its move from a number of data maps of the same region (e.g. elevation, soils, ground cover, etc.) which are digitised and geographically rectified to a common projection in a GIS. Such maps may be stored as a series of layers in a GIS. Each point within a polygon, or each cell in a raster layer (where each raster cell is assigned one value), takes the values of the layers directly above the point. The model is then built defining specific questions which may be asked about specific points or cells, such as what is the slope or aspect at the point. Eventually the biologist is interested in mapping areas which satisfy the known ecological requirements of a species (e.g. slope greater than 30 degrees, with erodible soils, occurring on

[16] The map may be generated from remotely sensed imagery, or by traditional cartographic methods such as aerial photograph interpretation.

a southerly aspect). By overlaying the map layers, such simple queries may be answered.

Nevertheless the simple process of integrating layers propagates the errors of the original data and sometimes amplifies them. If the sources of errors are identified and their accuracy is known, an estimate of the accuracy of the final output can be achieved through error propagation analysis.

Sources of errors range from those evidenced in the previous paragraphs to those inherent to non-spatial or text data which are also part of a GIS model. For instance, non-spatial data include knowledge or rules used by expert systems (Skidmore 1989b). All these errors are contributing factors to error accumulation when overlaying GIS data layers.

Composite overlaying is the simplest overlaying technique. Two (or more) layers are combined, and the raster location (or polygons formed) describes the union of the classes on the layers. The composite overlay is in effect a universal Boolean "AND" operation over the whole map. That is, for a two layer data set comprising layer X and layer Y, we note $X_{i=1,n}$ " $Y_{j=1,m}$ (i.e. the intersection of X and Y for the n classes in map X and the m classes in map Y) at all points over the map.

Arithmetic and mathematical operators that may be applied to two or more layers include addition, subtraction, multiplication, division, maximum, minimum, average and exponent.

The method(s) by which error is accumulated during the overlaying process is important for modelling error in the final map products. The first necessity for modelling map error accumulation is to quantify the error in the individual layers being overlaid. As discussed above, there has been a lot of work expended on this problem, with some tangible results.

Newcomer and Szajgin (1984) used probability theory to calculate error accumulation through two map layers. This study assumed that the two map layers were dependent; that is, if we select a cell which is in error in layer 1 then that act reduces the probability of selecting an erroneous cell from layer 2. If the data layers are independent, then an erroneous cell selected from layer 1 will not reduce the probability of selecting an incorrect cell from layer 2.

Using the statistics of Parratt (1961) with empirical data, Burrough (1986) concluded that with two layers of continuous data, the addition operation is relatively unimportant in terms of error accumulation. The amount of error accumulated by the division and multiplication operations is much larger. The largest error accumulation occurs during subtraction operations. Correlated

variables may have higher error accumulation rates than non-correlated data because erroneous regions will tend to coincide and concentrate error rates there.

The use of Bayesian logic for GIS overlaying is explained in Skidmore (1989b). As with Boolean, arithmetic and composite overlaying, there is inherent error in the individual data layers when overlaying using Bayesian logic. In addition, Bayesian overlaying uses rules to link the evidence to the hypotheses: the rules have an associated uncertainty and are an additional source of error.

A number of possible solutions are suggested for modelling accumulation of error in a GIS. These include reliability diagrams as well as a probabilistic approach.

Wright (1942) suggested that reliability diagrams should accompany all maps. He also emphasised that the sources used to generate different regions of the map have varying accuracy and these sources should be clearly stated on the map. For example, one region may have been mapped using low altitude aerial photography and controlled ground survey, and would therefore be more accurate than another region mapped using high altitude photography and only reconnaissance survey. This theme was taken up by Chrisman (1987) and MacEachren (1985), who suggested that such a reliability diagram showing map pedigree should be included as an additional layer accompanying each map layer in a GIS. However, for the purposes of error accumulation modelling, reliability diagrams do not provide a quantitative statement about the accuracy (or error) of the map.

The supervised non-parametric classifier described by Skidmore and Turner (1988) and Skidmore *et al.,* (1996) classified remotely sensed and GIS digital data. The classifier outputs for all cells the empirical probability of correct classification for each class according to the training area data, and thereby gives an indication of map accuracy.

The few methods proposed for modelling error accumulation are limited in their application. Working with ideal data, these methods do allow some conclusions to be drawn about error accumulation during GIS overlay operations. However, the methods break down when used with map layers created under different conditions than assumed by the methods.

Newcomer and Sjazgin (1984) used probability theory to model error accumulation. Heuvelink and Burrough (1993) modelled the accumulation of error in Boolean models, using surfaces interpolated by kriging as the estimated error source.

Sensitivity analysis

When no information is available either on the extent of the errors of the original data sets or on the type of error propagation function applicable to the model, a way of defining levels of reliability of the output is to analyse its variability subject to changes in the input parameters. In the previous paragraphs of this chapter we have seen possible sources of variability and uncertainties which can arise both in the deductive and in the inductive approaches to species distribution modelling. These sources of variability range from subjective errors introduced by the specialist, which define the species-environment relationship, to locational errors of species observation due to possible biases in the sampling scheme or to inaccuracy of the instrument used to locate the species (e.g. radiotelemetry, GPS). A through discussion on the sources of uncertainties in species distribution modelling can be found in Stoms et al., (1992).

Once the sources and ranges of variability are identified, different input data sets can be systematically produced by selecting the variates (sensu Sokal and Rohlf 1995) from the variability range of the original input variables. These alternative data sets are used to build alternative models which can be compared with the original one, identifying the variability induced in the output by the uncertainty of the input variables. The variability induced in the output is a measure of the overall performance of the model, and can be compared with a predetermined acceptable significance level. As a general case, when dealing with high uncertainty in the measures of the input variables, a greater inertia (less subject to changes in the results) of the model is generally preferable.

Sensitivity analysis does not replace validation but can be used at any stage of the model building process to identify those parameters that should be monitored more carefully to maximise the reliability and the accuracy of the results.

DISCUSSION

The use of the GIS tool in species distribution modelling should follow precise steps in which each one of these issues highlighted in the chapter is discussed and addressed. First unambiguous use of some key terms such as scale and habitat is recommended. The misuse of the latter is not restricted to GIS applications alone but spans the entire field of ecological study (Hall *et al.,* 1997). GIS can be a valid tool to overcome the current ambiguity between the species-related and the land-related concept of the term habitat. In fact, the latter was introduced as a way of dealing with problems related to environmental mapping using traditional tools. The enhancements introduced by GIS

overcome those problems. In the meantime, however, replacing use of the word "habitat" with a less problematic term such as "environment" is suggested.

The ambiguity of the term habitat and most of the work on habitat use and habitat selection have also given rise to the question of whether GIS-based models can be used to explain the causal event of a species-environment relationship. Use of GIS is not central to a better understanding of causal effects in a species-environment relationship, especially if the quality of data does not support both high-resolution and large extension analyses. Currently our analytical capabilities are limited by the lack of both high resolution and global coverage data sets. Nevertheless the use of GIS's spatial analysis tools within the framework of a controlled environment, in which all the key variables are monitored at an adequate resolution, can increase our capability to assess causal effects in species-environment relationships.

Apart from further generic considerations, a few important issues have been overlooked in these first years of GIS application in the field of ecological modelling and especially in the field of species distribution modelling. There has been inadequate discussion and consideration of the assumptions underlying the model building process and of the related issue of spatial and temporal scale, which are of paramount importance for sound scientific[17] use of GIS. Adequately discussed assumptions can be a justification *per se* of the development of a model. Whenever a hypothesis is stated, and a model is built to test its congruence, it should be regarded as a problem-solving tool. For instance we cannot know the outcome of alternative management options, nevertheless we can state different hypotheses, state the assumptions which must be met for each hypothesis to hold true and try to model the result of the different options. In such cases we do not have direct control over the results of the management action but we can control that the assumptions are met. This means that the output of the model will hold true if the assumptions are met and if the model is built upon the logical consequences of these assumptions. In such cases validation, meant as an independent estimate of the "truth", can to a certain extent be neglected (Starfield 1997), as "truth" will be the result of the implementation of the management action. Nevertheless, most of the time assumptions are not adequately discussed. This is particularly evident when dealing with the constraints of scale dependency of biological events. Probably the issues of scale still suffer from inadequate support from the available tools. For example, we still lack convenient ways of handling spatio-temporal data in GIS software packages, not to speak of analysing the two components together.

[17] Here scientific is meant in the sense of scientific method: the recognition and formulation of a problem, the collection of data through observation and experiment, and the formulation and testing of the hypotheses.

If validation can be neglected when dealing with hypothesis testing models like the ones illustrated above, it becomes a fundamental issue when building analytical models, which are built to assess species environment relationships and ecological processes. In such cases, validation steps must be included right from the beginning of the model building process, first assessing the quality and reliability of the raw data used, then evaluating the limits of the relationships that drive the process and finally analysing the correspondence of the output with the "truth".

Validation can be a costly exercise in the model building process, and efforts are being made to find a cost effective approach to this issue. Since the issue of validation is general to GIS modelling and especially to GIS ecological applications, it can benefit from the experience matured in other fields of science (e.g. Remote Sensing). Methods for summarising the accuracy of raster and vector maps using point samples and error matrices are now widely used in GIS, and are beginning to find their way into ecological applications as well. However, standard techniques have not had universal acceptance for a number of reasons. For example, a number of alternative sampling designs have been proposed for analysing the accuracy of imagery. The choice of a sampling design is often subject to the particular problems associated with the area to be ground truthed. However, a number of general trends are obvious from the GIS and RS literature. Random and stratified random sampling are acknowledged to be methods that maximise precision and accuracy (albeit at a higher cost than cluster sampling or systematic sampling). It should be noted that in highly heterogeneous landscapes (e.g. native forests, especially tropical forests), stratification is often too costly to consider. Cluster sampling offers reduced sampling costs, but in order to be effective is dependent upon low intracluster variance. Systematic sampling schemes may lead to a bias in parameter estimation if periodic errors align with the sampling frame (for example, as a result of image banding, or linear topographic features such as in the Allegheny Mountains of Pennsylvania).

GIS data layers contain numerous errors. And these pose a number of problems as errors accumulate during the process of analysis and model building. Though modelling the accumulation of error during GIS overlay analysis is still in its infancy, some methods for measuring error accumulation during GIS analysis have been discussed.

Any procedure to reduce mapping error in individual layers in a GIS will improve the mapping accuracy of an overlay generated from the GIS. Until better error modelling techniques are developed for GISs, descriptive statistics should ideally be calculated for each layer in a GIS, as well as for each layer produced by GIS modelling. The descriptive statistics should include overall mapping accuracies as well as class mapping accuracies. An alternative way of

defining the performance of a GIS model, thus assigning a level of reliability to its results, is sensitivity analysis which identifies crucial parameters. These parameters are those that, within their range of variation, determine the highest variation in the model output.

CONCLUSIONS

A common opinion among epistemologists is that we are facing a break between the development of advanced technologies and our needs and capabilities to use them. As a result of this break, sophisticated systems are frequently used for relatively simple operations and their use becomes a goal in itself. Toraldo di Francia (1978) names this paradox "the law of inversion between the goal and the instrument". This has also been the case with GIS in recent years in which the accent was on the use of the tool rather than on solving the problems to which it was applied. During the infancy of the tool's development, its enthusiastic and acritical application is clear evidence of the inadequate understanding of the tool's limitations.

However in recent years there has been growth both in the capabilities of the tool and in the awareness of its users. Sufficient case studies exist to define a logical framework in which the process of GIS modelling, and more specifically species distribution modelling, should be kept. To accomplish this, the issues that need to be addressed throughout the process have been identified.

In an era in which the need to acquire and analyse data at wider scales is increasing, and globalisation in environmental assessment applications is becoming urgent, we should not waste the opportunity that GIS offers to wildlife biologists to cope with these needs. This does not mean that GIS modelling is a substitute for fieldwork and direct observation. All management action should be based on direct site-specific studies, especially in the case of rare and/or endangered species. Nevertheless, sound GIS distribution models, which can be achieved by addressing the different issues described in this paper, can help to identify areas that require more detailed investigation.

.

Chapter 4:
Mobilising biodiversity data in practice: the example of the African Mammals Databank[18]

INTRODUCTION

A Databank for the Conservation and Management of the African Mammals (AMD), published in 1999, illustrates the potential value of existing data to decision-making, and provides the first example of using an integrated approach to mobilising information. The AMD built species distribution models through a process of data gathering, data formalisation, expert review, distribution modelling and validation. It integrated existing data, experts' knowledge and GIS to produce a series of species distribution models that reflect habitat suitability.

The aim of the AMD is to provide a broad picture of the species distributions at the continental scale. While policy is set globally, regionally and nationally, action takes place at the local level. This has been the traditional focus of ecologists and conservationists. However, they are increasingly recognising that factors affecting local communities are better understood in the context of broader spatial and temporal analyses (Pimm, 1991). Projects with broader spatial extent are rare for a variety of reasons (funding, long-term commitment of programs and researchers, academic pressures, etc.). Yet they are important to help us understand the big picture of ecological processes and to manage them effectively. A broad scale approach to conservation has also been advocated as a way of dealing with the current conservation crisis (May, 1994).

The comparability of the models to actual presence was assessed in four countries. The result showed a level of accuracy that supports the original assumption: that habitat suitability can point to a species' true Area of Occupancy. Area of Occupancy is a key piece of information needed to implement effective conservation measures. Extent of Occurrence, the classic presentation of distribution ranges of species, lacks the resolution needed to

[18]Based on: Boitani L., F.Corsi, A. De Biase, I. D'Inzillo Carranza, M. Ravagli, G. Reggiani, I. Sinibaldi, P. Trapanese. (1999). AMD African Mammals Databank - A Databank for the Conservation and Management of the African Mammals. Istituto di Ecologia Applicata - Rome, Italy ISBN 88-87736-00-6.

define the species' true Area of Occupancy, or area actually occupied by the species within the Extent of Occurrence (Gaston, 1991).

The AMD models offer an improvement on traditional Extent of Occurrence maps, at a level of resolution appropriate for decision-making and natural resource management. Few applications of these types of models have been attempted at a continental scale (e.g., Busby, 1991; Walker, 1990; Skidmore *et al.*, 1996) and none for Africa. Thus, this project was seen as a good opportunity to produce a first assessment of their potential.

METHODOLOGY

The methodology included six phases: 1) identifying the species to be included in the databank and choosing the appropriate taxonomic classification; 2) gathering and formalising existing published information on the selected species; 3) involving experts in refining and validating the baseline data; 4) developing environmental suitability models; 5) validating results; and 6) completing the species conservation management analysis.

The Species

A total of 281 species, belonging to 12 orders and 28 families were included in the data bank. Selection of the species for this project was based on various criteria, although the size of a rabbit was taken as the general lower threshold.

Inclusions:

- All species belonging to the orders of Primates, Carnivora, Perissodactyla (except rhinos, see below), Hyracoidea, Tubulidentata, Artiodactyla, Pholidota and Lagomorpha were included, irrespective of their size;
- Although small, Macroscelidea are considered of particular conservation concern, and were thus added to the list;
- Among Rodentia, 7 species (belonging to Dipodidae, Pedetidae and Hystricidae) were included, as they are of medium size and/or particular conservation or ecological interest. Following the same criteria, the three species of Tenrecidae not endemic to Madagascar and the entire family of Erinaceidae were included for Insectivora.

Exclusions:

- Possibly requiring a different environmental data set and separate analyses, all Chiroptera, Cetacea and Sirenia were excluded, except for the African manatee (*Trichechus senegalensis*);

- Three notable species were excluded, the rhinoceros (*Diceros bicornis* and *Ceratotherium simum*) and the elephant (*Loxodonta africana*). The two species of rhinos were excluded because data on the last few areas in which they are found are being kept from the public and it was decided not to interfere with this important decision. The elephant was excluded because an excellent and detailed database in a format very similar to the one presented here is kept in Nairobi by the IUCN/SSC African Elephant Specialist Group. For a similar reason, all Madagascar species were also excluded, as there is an extensive database produced by the World Bank;
- Newly described species (i.e., *Pseudopotto martini*) were excluded due to the almost total lack of information, especially on their distribution range;
- Finally, species thought to be extinct in recent years (in the last 20-50 years) were excluded, along with those that occur in Africa only as recently introduced populations (i.e., fallow deer, *Dama dama*) or those that are totally domesticated (i.e., dromedary, *Camelus dromedarius*).

Species nomenclature followed "Mammals species of the world" (Wilson & Reeder, 1993). Other classifications were applied when a different systematic arrangement was recommended by the relevant IUCN/Species Survival Commission (SSC) Specialist Group[19], or was supported by recent evidence. For instance, *Procolobus badius* follows the indication of the IUCN/SSC Primate Specialist Group, which suggests retaining the different forms of red colobus monkeys (recognised as different species in Wilson and Reeder, 1993) in a single species classification until a thorough re-analysis of their relationships is undertaken (Oates, 1996). The arrangement adopted is not intended to be a final solution. Given the continuous changes in and refinements to mammalian phylogenetic information, the related systematic relationships and naming process will most probably continue to evolve (see Chapter 7 for a discussion about managing taxonomic differences within biodiversity data management systems).

Data gathering

Data gathering comprised a literature review and compilation of existing species distribution data. A comprehensive review of the literature was carried out to collect and store all relevant and available information on the status,

[19] The IUCN SSC is organised into a series of Specialist Groups each of which focus on a species, group of species or specific conservation issue.

distribution, biology and ecological requirements of each species. The search included literature published in the past 10 years, although for some poorly known species it was extended to include most of the literature ever published. A keyword-based search covering taxonomy, ecology and geographical distribution within the geographical and administrative bounds of the study area was carried out.

General and specialised publications were used as introductory literature for collecting references: i.e., "Keyword Index of Wildlife Research" (1990-1995) "Zoological Records" (1990-1996) and "Current Advances in Ecological and Environmental Sciences" (Pergamon Press, 1990-1995). Computer based searches were carried out on the following databases: Medline, Cab Abstracts, Life Sciences, Nisc Disc (either through direct access to CD-Rom or through public library access).

References listed in the most recent and comprehensive monographs concerning African mammals were also examined (Estes, 1991; Skinner & Smithers, 1990; Macdonald, 1984; Mammalian Species of the American Society of Mammalogy, IUCN/SSC Specialist Group Series, C. Helm Mammal Series, IUCN Red List Series, etc.).

The quantity and quality of information available for each species varied significantly. Therefore, although the same analyses were carried out on all species, the reliability of each of the results is dependent on these factors.

Distribution maps were chosen as the best representation of current knowledge on each species' occurrence. Extent of Occurrence boundaries were obtained from published distribution maps. In a few cases point locations were considered and the limits of the Extent of Occurrence extended to include single known records of a species. This was generally restricted to those cases for which the existing distribution maps were not current or precise enough to include confirmed records of a species. In general, however, distribution maps were retained as the primary sources of information regarding the boundaries of species distribution.

As with the information available for each species, the quality of distribution data varied. Distribution maps summarise and graphically represent distribution data. Generally, they provide only an explanatory support to the description of one characteristic of a species (its occurrence in space). Furthermore, Extent of Occurrence boundaries, usually extrapolated from point locations, are a subjective interpretation of the species' distribution and other biological characteristics, such as habitat preference. Therefore the published maps are subject to a degree of approximation.

Data Formalisation

Data formalisation followed two parallel routes. The distribution data was used to draw Extent of Occurrence maps, while the data on the status, biology and ecology of each species was standardised. In particular the data on biology and ecology were standardised to match the available environmental variables and serve as input for the GIS environmental suitability models (see Annex 1 for a complete list of available GIS layers).

Distribution data
To standardise the distribution maps for analysis, they were converted into GIS (Arc/Info, ESRI, USA) polygon coverage. Three steps were taken: rasterisation, georeferencing and vectorialisation.

The maps were initially acquired as digital raster images by means of a scanner. The acquisition was performed at either 300 or 600 dpi, depending on the quality and dimension of the original figure. The images were then converted into a georeferenced raster, so that each point corresponded to real world coordinates. Most of the original distribution maps were based on explanatory figures of published papers, action plans and field guides. As these generally lack precise references to projection parameters, scale and control points, it was not possible to georeference them through simple Arc/Info procedures of transformation or projection. Georeferencing therefore had to be performed by means of "rubber-sheeting". In this process the raster is stretched by correlating the identifiable geographic features to the same features present in a template coverage. The main geographic features of the continent, derived from the available baseline geographical data set (DCW and ADS, see later on for a description), were merged into a single coverage to obtain the template.

For each map scanned, a minimum of 10 to a maximum of 40 control points were used for registration. The points were always selected to match easily identifiable geographic features such as rivers, administrative borders or other attributes.

The result of the georeferencing process was visually checked and repeated (adding and/or deleting control points) until the geographic features on the map corresponded to those on the template. This process resulted in the introduction of a varying amount of inaccuracy, especially in those cases in which the original maps were at a very small scale or when the geographic features were poorly or unreliably depicted. To reduce the overall error, up to 10 different distribution maps from different sources were acquired for each species.

Once georeferenced, the raster image was used to video-digitise Arc/Info polygon coverage. In this phase, an operator manually digitised the arcs that defined the Extent of Occurrence by following the distribution boundaries on the raster image. Each polygon in the coverage was assigned a code to discriminate between those portions of the range that are certain and those that are possible, and to identify where the species has been introduced/reintroduced. Another code provides easy retrieval of the map's source in literature from the bibliographic database.

Multiple distribution maps were available for most species. Maps to be used for the final Extent of Occurrence of the species were evaluated in terms of the reliability of the literature source from which they were taken, currency, resolution and performance in the georeferencing process. In several cases, a single map was used to define the Extent of Occurrence of a species. Sometimes, however, more than one map had to be merged to obtain the final coverage. In such cases, each map was first converted into a polygon coverage. Various maps were then merged into a single final coverage using the Arc/Info overlay and editing tools.

Taxonomic uncertainties regarding some of the species considered in the project were also taken into account in selecting the distribution maps to be used for the final Extent of Occurrence. Whenever possible, only the distribution maps reflecting the taxonomic status of the species as adopted in this project were used.

When no map following the taxonomic arrangement used in this project could be found (e.g., *Genetta rubiginosa* and *G. pardina*), the Extent of Occurrence was obtained from other available maps. In a few cases the maps included several subspecies, including the taxon in question. In these cases, ranges were included only when they could be discretely identified with the specific taxon.

Ecological data

The species' ecological requirements found in literature were interpreted and then scored to match environmental suitability in relation to two main data layers: White's vegetation map (White, 1983) and Seasonal Land Cover regions map (GLCC project).

White's vegetation map of Africa provided adequate representation of the vegetation types selected by mammals. Floristic, physiognomic and spatial information (zonation) is synthesised and organised in a two-level hierarchical classification (17 and 80 categories, at the first and second level, respectively). Also very useful in this exercise were the accompanying notes to White's vegetation map, which provide further details useful in identifying specific habitat features. However, White's vegetation map does not adequately depict

cultivated or degraded vegetation, and the nomenclature used is rarely (though increasingly) found in mammals' habitat descriptions.

The Seasonal Land Cover map brings additional information to White's vegetation map by providing a recent and detailed description of mainly physiognomic features (195 categories). It includes a detailed classification of cultivated and degraded vegetation, and the nomenclature used is easily related to mammals' habitat descriptions. Moreover, its detail was considered sufficient to represent habitat fragmentation with the resolution required by this study.

1	suitable	• frequently cited as main or preferred habitat • presence of major populations • higher population densities
2	moderately suitable	• frequently cited as secondary habitat • presence of minor populations • lower population densities • patchy distribution
3	unsuitable	• cited as unsuitable • not specifically cited, but unsuitability easily inferred from its ecology
9	undefined	• not cited, possibly suitable or moderately suitable as inferred from its ecology • data deficient, possibly suitable or moderately suitable

Table 4.1: Criteria used to assign the suitability scores to each combination of environmental variables in the CD model.

For each species, a separate suitability score was assigned to each legend category of the two maps. Three categories of suitability were considered, suitability 1 being the highest. In order to provide a method for scoring that was as objective as possible, a data file was compiled with the species' habitat descriptions from different literature sources. The criteria used are summarised in table 4.1, together with scores and definitions.

Scoring of the individual categories found in the mosaic		**FINAL SCORE**
Category 1	*Category 2*	
suitable	unsuitable	moderately suitable
suitable	moderately suitable	suitable
unsuitable	moderately suitable	moderately suitable
suitable	unknown	suitable
unsuitable	unknown	moderately suitable

Table 4.2: Criteria use to score mosaic categories based on the scoring of the classes that make up the mosaic.

Both White's vegetation map and the Seasonal Land Cover map include mosaic categories in their legend. For some species these mosaics constitute the preferred habitat type. For the other cases, the rules shown in Table 4.2, which

are based on the scoring of the individual categories making up the mosaic, were applied.

Water dependency was recorded according to the classification in table 4.3.

Species occurring in or near water	Includes species inhabiting rivers, lakes or adjacent areas or spending a considerable part of their time in or near water (e.g., otters, *Aonyx* spp. and *Lutra*; hippo, *Hippopotamus amphibius*; waterbuck, *Kobus ellipsiprymnus*).
Species occurring in riverine or gallery forests	Some species (many primates and some ungulates) show preference for gallery or riverine forests, or extend their forest range in more arid zones through gallery forest.
Water-dependent species	Includes all species (mainly ungulates) whose density decreases away from permanent water, at least during the dry season.

Table 4.3: Criteria used to define water dependency of the species.

Finally, the altitudinal range of occurrence of the species was also recorded.

Expert Review

Experts on the ecology of each species played a key role in the development of species distribution models. At this point in the process they made two important inputs. First Extent of Occurrence maps produced in the process described above were distributed to the experts. They reviewed the maps and corrected the boundaries according to their knowledge of the species distribution patterns. Experts were also asked to review and correct data on the biology and ecology of each species, both in terms of the species habitat description and of the suitability scores.

To ensure needed inputs from networks of experts, the project adopted an approach to minimise demand on experts' limited time and ease the burden of information requests. Experts were not asked to provide Extent of Occurrence maps or describe the species' biology and ecology. Rather they were provided with data sets compiled from the extensive literature search and standardisation process, allowing them to react to comprehensive and available information and to use their expertise to improve on it. This approach recognises competing priorities that the experts face, and enables them to make valuable inputs with minimal burden to their workloads (see Chapter 2, *Mobilising Dispersed Information for Biodiversity Conservation*).

To review the data gathered for the AMD, the standardised data were sent to the Chairs of the appropriate IUCN/SSC Specialist Groups for review. These experts revised the data either personally or forwarded them to other experts of their Specialist Group. Of 24 experts contacted, 12 replied in useful terms. The referees corrected or commented on the data, often sending their most recent publications.

Once the refereed information was returned from the experts, it was incorporated into the AMD and used for the analyses described in the following section.

Modelling

To evaluate the species' distributions, GIS was used to map the relationship between species' biology and ecology, and the available environmental factors. The basic assumption was that environmental suitability can describe the patterns of use of space within the species' Extent of Occurrence, and that these patterns represent the Area of Occupancy.

The AMD applied two techniques to model species' environmental suitability. One relied more heavily on the expert inputs concerning the species-environment relationship (deductive approach). The deductive approach uses the known ecological requirements to extrapolate suitable areas from the environmental variable layers available in the GIS database. Analysis of the species-environment relationship is synthesised from the expertise of one or more specialists who decide, to the best of their knowledge, which environmental conditions are the most favourable for the existence of the species.

Generally, a logical (e.g., Breininger *et al.*, 1991; Jensen *et al.*, 1992) or arithmetic map overlay operation (e.g., Donovan *et al.*, 1987; Congalton *et al.*, 1993) is used to merge the different GIS environmental layers to yield a combined effect. This approach was chosen as the central modelling effort of the project, and was used to produce the Categorical-Discrete (CD) distribution models.

The second technique provides a means of comparison. It relied more heavily on the Extent of Occurrence, here used as a proxy to more detailed data, which to a large extent remains unavailable (inductive approach). With this approach, species-environment relationships are not known *a priori* but are derived by comparing species observations with the available environmental variables. Given that extensive data sets on species observations were lacking, the Extent of Occurrence was assumed to be a good representation of the average

conditions suitable for each species. Using the information implicit in the Extent of Occurrence, the species' environmental preferences were derived. These were then used to drive the environmental layers overlay and extrapolate the environmental suitability model (e.g., Pereira & Itami, 1991; Aspinall & Matthews, 1994).

CD models

In the CD model approach, the Area of Occupancy is described by integrating the known Extent of Occurrence and the environmental suitability model derived using the species' known environmental preferences. Using the suitability score assigned to each environmental variable, the environmental data layers are combined to identify the areas of varying suitability for the species.

For White's vegetation map and for the Seasonal Land Cover map, the combined score was obtained from the overlay of the two layers according to the matrix rule shown in Table 4.4.

		White's vegetation map		
		1	**2**	**3**
Land Cover map	**1**	1	1	2
	2	2	2	3
	3	3	3	3

Table 4.4: Cross table used to combine the scores of White's vegetation map and the Land Cover Map.

The matrix gives priority to the Land Cover classification, which is more recent and detailed, but the score decreases when the Land Cover units fall in unsuitable vegetation contexts as indicated by White's map. As a result, suitable habitat patches within unsuitable vegetation zones are maintained and scored moderately suitable (e.g., a Land Cover forest unit within a White's woodland zone is scored moderately suitable for a forest species). Moderately suitable patches in the same zone are disqualified. Unsuitable habitat patches within suitable and moderately suitable vegetation zones are identified by means of the Land Cover map (e.g., cultivated areas within White's forest zone are scored unsuitable for a forest species).

Although matrix rules were applied to all species, the resulting scores were modified for some specific combinations, particularly when White's

classification better described the specific vegetation types (e.g., montane and desert vegetation).

The river network was used to emphasise the water dependency of some species. Buffer areas of 1, 5 and 10 km around permanent water courses or lakes were calculated. The suitability scores obtained from the intersection of White's vegetation map and the Land Cover map were altered outside and/or inside these buffer areas. The degree of modification of the scores was decided according to the specific requirements of each species, as was the dimension of the buffer areas chosen to represent the threshold.

For species occurring in or near the water, all vegetation types occurring outside the buffer area were considered unsuitable. The buffer size assigned was either 1 km or 10 km, depending on the species' mobility. Most species ascribed to this category concentrate their activity within 500 m of water (e.g., otters), but 1 km is the minimum resolution of the data set. The threshold of 10 km for hippos and waterbuck was used as a standard distance, as the maximum distance of these species from water is known to be in this range (Rowe-Rowe, 1994; Spinage, 1982).

Given that riverine and gallery forests are not identified by either White's vegetation map or the Seasonal Land Cover map , it was assumed these riparian habitat types occur within 1 km of permanent water in selected White's vegetation map classes. White's vegetation classes were selected based on the indications provided in the vegetation map's explanatory notes (White, 1983). For the species preferring these habitat types (species occurring in riverine and gallery forest, see above) the assigned preference scores were increased in the buffer area. Species such as the blue duiker (*Cephalophus monticola*), the Diana monkey (*Cercopithecus diana*), the black and white colobus monkey (*Colobus guereza*) and several others belong to this group. Only for *Cercocebus galeritus* was a different buffer size (5 km) used (Homewood, 1975).

For the species whose density decreases away from permanent water, the suitability scores obtained from the intersection between White's vegetation map and the Land Cover map were decreased outside a 10 km buffer area around permanent rivers. In this way, the suitability of areas that can support high numbers of a species on a year-round basis was enhanced, as compared to areas that are suitable only during the wet season. Support for choosing the 10 km threshold distance was found in Western (1975), who showed that in the Amboseli ecosystem the biomass density of large mammals more than 15 km from water was extremely low, as compared to areas within 10 km of water. The 15 km belt enclosed 99.5% of the biomass of water-dependent species during the dry season. This fixed threshold, however, must be considered with great caution, as stressed by Rod East (*in litteris*), who revised and commented

on the list of species included in this group: each species is likely to have its own threshold distance, and no particular threshold distance from permanent water can be applied uniformly across Africa. Interestingly enough, the described procedure improved the model's results for the selected species. It could easily be repeated as soon as new or more detailed information is available, or could be applied with more efficacy using regional or local data sets.

Elevation derived from the Digital Terrain model of Africa was used for some species that are restricted to well-defined elevation zones (e.g., Lepus starki, Theropithecus gelada). For these species, all vegetation types occurring outside known elevation limits were considered unsuitable.

The Area of Occupancy was defined as the sum of suitable and moderately suitable areas resulting from the models. Nevertheless, while an area can be environmentally suitable for a species the species may be absent because of other factors not taken into account in the modelling process. For the interpretation of the results of the models, these other factors are assumed to be summarised by the Extent of Occurrence. Thus presence was considered "expected" in areas showing levels of environmental suitability within the Extent of Occurrence; while it was only considered to be "potential" in suitable and moderately suitable areas outside the boundaries of the Extent of Occurrence.

Probabilistic models

The probabilistic-continuous models were developed to investigate the use of the limited data available for the AMD with a more objective modelling approach (Corsi et al. 2000). In general, probabilistic-continuous models are built using observation data to define the "species ecological profile". For the AMD project, these data sets were not available. To overcome this problem, it was assumed that the average environmental condition inside the Extent of Occurrence could be representative of the "species ecological profile".

Thus, the ecological suitability was mapped in terms of ecological distance from the average ecological conditions inside the Extent of Occurrence. To account for possible correlation between variables used in the model, the ecological distance was measured in terms of Mahalanobis distance.

The variables used for the inductive models are: the General variation in NDVI, regardless of season, obtained as the first Principal Component of the 12 monthly averages of the NDVI index (Eastman, 1992); a seasonality map derived from the same NDVI data set (Baird, 1996); the population densities raster; and the DEM.

As a consequence of the assumption that the average environmental condition inside the Extent of Occurrence can be representative of the species "ecological profile", the quality and reliability of the results depended on the homogeneity of distribution throughout the Extent of Occurrence of both the species and the environmental variables. Thus the overall reliability is quite low.

While it seems important to underline that these models should be read as speculative projections, their importance resides in the possibilities that the use of GIS opens in species distribution modelling for conservation planning. By introducing the geographic dimension into analysis of species viability, PC models can be used to address part of the stochasticity inherent in locating an individual of a species, as well as problems of corridor design and meta-population modelling (Akçakaya, 1993).

Validation

To assess usefulness of the models for further conservation management and planning, the results of the Continuous-Discrete models were validated with an independent dataset collected through fieldwork.

Four sample areas were chosen, each encompassing an entire country. One hundred points were identified in each, following a random stratified sampling design. The strata were obtained by overlapping: 1) the vegetation layers defined by White's vegetation map of Africa (1983); 2) the layers of human population density; 3) the administrative borders of the four selected countries. Where necessary, points were added inside protected areas to provide more data for analysis of the influence of this variable.

Work in the field involved verifying the relationship between the presence of the species and the different classes/values of the environmental variables considered. The 100 plots defined in each country were not meant to produce an independent estimate of the distribution of the species. Rather they provided a statistical sample of locations against which presence or absence of a species could be compared to the model's prediction of presence or absence.

The presence (or absence) of the species at each of the 100 predetermined points was verified by direct observation, *in loco* collection of publications and scientific reports and interviews with local authorities and residents. The latter were carried out according to previously agreed standard procedures, using special checklists and pictures to help with recognition of individual species. Each item of information was judged for quality/reliability.

While trying to keep the sampling design as close as possible to the one devised in the laboratory, various factors hindered access to a predetermined point. These included security considerations due to civil disorders, impassable roads or paths, military restrictions, excessive distance from the road and prohibitive land conditions (e.g., desert, dense tropical forest). In such cases, points were shifted to more accessible areas within the same stratum.

The countries selected were Botswana, Cameroon, Uganda, and Morocco, based on the following criteria:

- high environmental diversity in each country;
- general representativeness of conditions in East, West, South and North Africa;
- medium-large size with respect to the rest of the African countries;
- feasibility of collaborating with local universities or ministerial institutes;
- interest on the part of local officials and/or local scientists in participating in the project or in receiving specific training in GIS;
- availability of detailed and quality data on natural resources in each country.

Morocco was added to take account of the peculiar environmental and faunistic conditions found in North Africa, which would otherwise have been excluded. Even though Morocco is the most outlying of North African countries, it has the greatest biodiversity and is the most accessible for fieldwork. In each country an agreement was reached with a local institution for collaboration that included services and personnel, training, participation in fieldwork and analysis of results.

For each species the validation dataset includes all of the points within the species' Extent of Occurrence plus those points outside the Extent of Occurrence where the species was actually found. Selecting the validation points in this way ensures that negative points do not over influence the accuracy measure.

An overall accuracy was calculated for each species based on the result of the fieldwork. Together with the overall accuracy, a Kappa statistic (Cohen, 1960; Bishop et al. 1975) was calculated.

Conservation Analyses

To support conservation management of the species included in the AMD, two analyses were carried out on the results of the models: analyses of fragmentation, and comparison to protected areas. The aim of the analyses was

to provide insight into the status of the species and to support policy actions. The analysis of fragmentation sheds some light into the patterns of habitat destruction that the species are facing, possibly thanks to the high resolution achieved by the AMD models. The overlay with the existing network of protected areas helps to assess the extent to which they protect the species.

Fragmentation statistics were calculated on rasters aggregated to 5x5 km pixel size. This allowed averaging of errors due to uneven registration of the environmental layers used to produce the models. It also allowed cleaning of isolated pixels that would overestimate the overall fragmentation of the Area of Occupancy.

While many fragmentation indices are available, the AMD used those provided by the FRAGSTAT package (McGarigal & Marks, 1994) as they are simple and easy to interpret. For both the "suitable" and the "moderately suitable" classes and for their sum (the overall Area of Occupancy), the following indices were calculated:

- **Number of patches** (**NP**): the number of patches within the classes and the Area of Occupancy;
- **Mean patch size** (**MPS**): the average dimension of the patches;
- **Patch size standard deviation** (**PSSD**): the standard deviation of the previous index;
- **Largest patch index** (**LPI**): the percentage of the total area of the class that falls within the largest patch;
- **Mean shape index** (**MSI**): the average index over all the patches belonging to the class (the shape index of a patch varies between 1, when the patch is square, to infinity);
- **Area-weighted mean shape index** (**AWMSI**): same as above, with larger patches contributing more to the result.

It is important to note that none of the above indices alone can give a picture of the fragmentation of the species' Area of Occupancy. They should be read together, each one supporting the other. The number of patches of the two classes compared with that of the overall Area of Occupancy can provide a key for interpreting the connectivity of the different patches. If the NP of the two suitability classes is very high compared with that of the total Area of Occupancy, the Area of Occupancy can be expected to be substantially homogeneous even though highly interspersed with suitable and moderately suitable patches. Moreover, comparison of the two shape indices with the total number of patches can provide an insight into the real fragmentation of the Area of Occupancy at the scale in question.

The shape index varies from one (for a patch that is perfectly square) to infinity. An MSI near one indicates either that all of the patches (up to patches of a single pixel) are square or that most of them are one pixel size. Thus comparing the MSI to the AWMSI can give further insight into the patchiness of the Area of Occupancy. In fact, if the AWMSI is not equal to one, this means that there is at least one big patch (of varying complexity) of the given class within the Area of Occupancy.

This outcome can be strengthened by observing the value of the LPI for the same class; if the LPI is high, this further supports evidence of one single patch in the Area of Occupancy.

The efficacy of the protected areas network was tested by overlaying the GIS layer of the protected areas with the Area of Occupancy derived from the models. The percentage of area pertaining to each suitability class within the Extent of Occurrence, falling inside and outside existing protected areas, was tabulated for each species.

It must be stressed, however, that these percentages do not necessarily reflect the precise extension of areas truly protected. This is due in part to the fact that not all existing protected areas are mapped by the coverage used in this project. In some countries, protected areas are not mapped with polygons, but are only available as point locations. Protected areas represented with point localities were excluded from the analysis because no overlay was possible. Hence the result of this calculation is of limited use for species with distribution restricted to certain areas or countries (e.g., South Africa, for which protected areas were available only as point data), because the percentage of the Area of Occupancy included in protected areas could be severely underestimated.

RESULTS

The methodology described above produced a number of models. These were used both to assess the conservation status of and level of knowledge available for each individual species. For each one of the 281 species results are given in separate species' accounts. Examples are provided in Appendix 2; Appendix 1 illustrates the data structure of the digital data produced.

For each species account, the information is arranged under the following headings:

Nomenclature (Scientific name). Mostly following Wilson and Reeder (1993), although some changes have been made to accommodate more recent views of the taxonomy of the mammals.

Describer. The name of the species' author.

Vernacular names. Common names in English (Eng) and French (Fr).

Taxonomic notes. This section is meant to provide no more than a comment on special or significant taxonomic aspects. Closely related species, systematic uncertainties and well-defined forms, if any, are noted here.

IUCN threat category. This section gives the species' most recent threat assessment (at time of publication) based on the 1996 IUCN Red List of Threatened Animals (Baillie & Groombridge, 1996).

Available information. This section gives a brief report of the information available in the scientific and grey literature on the species. A major effort was made to include in the bibliography the most important sources of information about each species, excluding minor sources and papers dealing with topics not relevant to this project.

Known Extent of Occurrence. This section provides information on the source(s) of the map, the name of the specialist (if any) who revised it (and the date), and a list of other publications used to update the output. Comments are given on the distribution maps that were available for each species and the one thought to be most recent/reliable and retained for modelling.

Categorical-discrete (CD) distribution model. This section presents the ecological information on the species used for the model. Two tables show the quantitative results of the model in terms of percentage of the various suitability classes within the Extent of Occurrence and the 6 indices of fragmentation.

Probabilistic-continuous (PC) distribution model. The PC distribution model represents distribution in the form of an environmental suitability surface, where the ecological variability has been mapped as distance from the average ecological quality from the Extent of Occurrence.

Validation. This table (if present) shows the percent of the species' Extent of Occurrence within the sample areas, the number of valid plots and the overall accuracy. The validation was performed only for the species present in the countries in which the fieldwork was carried out. The overall accuracy is discussed for each species to consider possible explanations of excessively high or low figures. Values just below the threshold were found for species with habitat specificities that do not appear on the maps because they are probably too small for the mapping resolution of the data bank (e.g., species associated to rocky outcrops or gallery forest, etc.). High accordance was found for those species whose ecology is better known, allowing for better identification of the preferred habitat requirements.

Comments and conservation issues. This specific section of the species account contains comments on models, fragmentation analysis and protected areas overlap. These comments provide a basic interpretation of the output. While these comments are not intended as an indication of necessary species' conservation actions, they do provide a continental setting for decision-making.

References. The bibliography listed in this section includes the main publications used to compile each account.

Results of the validation indicate the extent to which expert inputs improved the accuracy of species distributions. Model validation was possible for 180 species. The overall mean and median of the overall accuracy for the 181 species considered is just above 60% (Fig. 4.1).

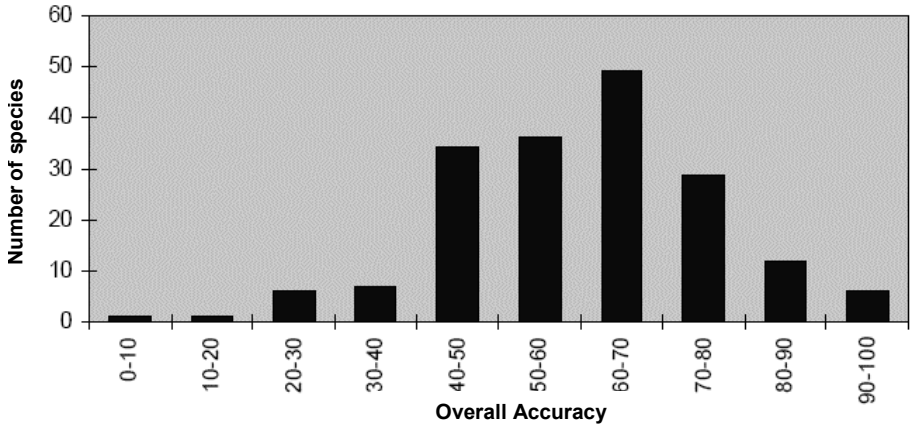


Fig. 4.1: Frequency distribution of the Overall Accuracy values.

To understand the extent to which species distribution accuracy was improved, the result of the modelling process must be compared to the original Extent of Occurrence. The average accuracy of the Extent of Occurrence is 38%. Thus the CD models have an average 22% increase in accuracy over the original distribution maps. While for 12 species the models did not improve the accuracy of the distribution, for 7 species the increased accuracy was more than 70% with the maximum increase of 86%.

A more detailed analysis was carried out to test the possible relationship between the Overall Accuracy of each species and the sampling effort. No significant correlation was found between the percentages of the Extent of Occurrence in the sample areas and the Overall Accuracy (Spearman R = -0.106642, p = 0.153) (Fig. 4.2); between the percentages of suitable areas of the

Extent of Occurrence and the Overall Accuracy (Spearman R = -0.078071; p = 0.296); between the percentages of total Area of Occupancy in the Extent of Occurrence and the Overall Accuracy (Spearman R = -0.051404; p = 0.4919). In summary, the Overall Accuracy is not related to the sampling effort and appears to be independent of variation in suitable areas and Area of Occupancy within the Extent of Occurrence.

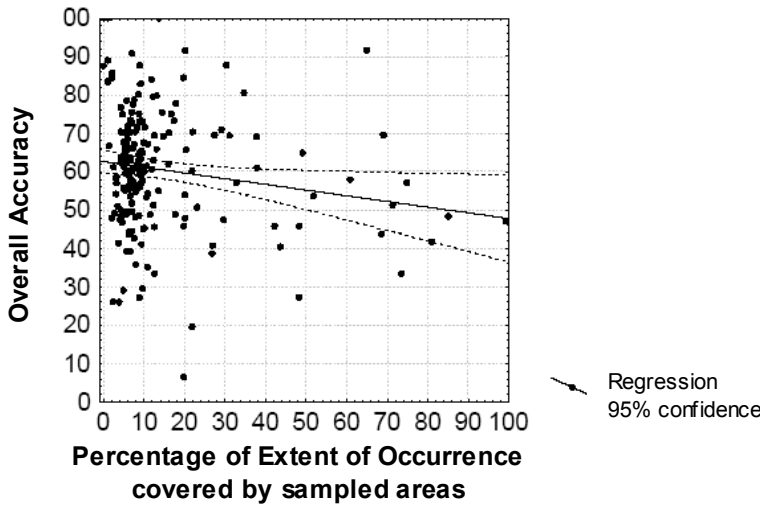


Fig. 4.2: Relation between the percentage of Extent of Occurrence in sample areas and the Overall Accuracy.

When considering the Kappa values, the average calculated for all the Extents of Occurrence is below 0 (K = -0.11). This indicates less than chance agreement; but also the models did not perform well, as the average of the Kappa is just 0.11. Nevertheless 168 models out of the 180 included in the validation analysis significantly (α=.01) enhanced the Kappa values of the corresponding Extent of Occurrence, and if we look at the increase of the Kappa values of the Continuous-Discrete models the average delta is 0.27.

DISCUSSION

The line marking the boundary of the geographic range of a species is usually drawn to separate the areas in which the species occur from those in which it does not occur (Corsi *et al.*, 2000).

We know that such a line is a human simplification and a representation artefact. Given that species distribute according to various parameters, which

include both environmental variables and individual selection, an accurate representation would only be achievable through a probability surface that assigns to each location the probability of finding there individuals of the species.

Obviously the accuracy of areas of presence delimited by the boundary lines of geographic ranges varies with scale. The boundary line drawn on a coarse scale map will include/exclude large areas that can be more accurately identified by a line drawn on a detailed map. In general, the finer the scale, the greater is the accuracy of the areas of presence delimited by the boundary.

Using GIS modelling techniques together with experts' knowledge, the AMD assesses the potential to refine areas of presence by analysing the distribution of the environmental suitability. The basic assumption is that mapping suitable areas within the boundaries of the geographic range (Extent of Occurrence) represents a better approximation of the Area of Occupancy of the species, because considering environmental suitability some of the parameters that drive the distribution of the species are taken into account (Heglund, 2002).

If the final results offer a distribution range whose accuracy is better than what was previously available, this is believed to be a significant improvement for the ecologists and the conservation planners. The Areas of Occupancy obtained through the modelling exercise were validated for 180 species (64%) out of the 281 analysed. With an overall average accuracy of 64% and an average increase in performance of 22% over the Extent of Occurrence, the overall results of the validation support the assumption that these models are a useful method for gaining insight into the potential areas of occupancy of the species within their Extent of Occurrence. Thus the use of expert knowledge provides a useful means of obtaining broad scale results that cannot be effectively gathered in any other way.

Validation values for the individual species range from 20% to 99%. Although difficult to quantify, in most cases the percentage seems to correlate to the extent and accuracy of information available for the species in question. In other words, poor results reflect mostly low quality (or insufficient) information, which in turn suggests that more effort should be put into understanding the ecology and biology of these species.

Nevertheless the reader is warned to appraise tables and figures correctly. The models use environmental parameters that have their own errors and approximations. Overlaying these data sets transfers their errors to the resulting model with an unknown error propagation function. Although the GIS can mask these errors from the ingenuous reader by displaying final outputs in the form of appealing maps, the errors are still there.

In particular, the AMD models identify various level of suitability within the range, and those levels must be interpreted as broad categories averaged over large areas. It is evident that a finer scale analysis would reveal much higher habitat complexity, which is probably evidenced by the low performance of the models when considering the Kappa values. Higher resolution analyses are indeed the natural next step after this initial global overlook.

The AMD offers the opportunity to refine the levels of approximation. Starting from the current output, the models can be loaded with other environmental parameters (covering the whole range or just a portion of it) and/or data at finer scale to narrow the definition of suitable areas, and the accuracy of their derived indices (e.g. Said *et al.*, 2003a and 2003b).

As a whole, the AMD process illustrates well the importance of mobilising existing information and putting it to effective use to support decision-making. With extensive inputs from species experts on both the Extent of Occurrence and status, biological and ecological characteristics of their respective species, the quality of the modelling process increased significantly and contributed to great extent to the level of accuracy found during the validation process. The method of extracting expert inputs, which eased the time burden to the extent possible, further increased the participation rate and quality of information.

Chapter 5: Modelling species distributions using an inductive approach: the example of the wolf in Italy[20]

INTRODUCTION

The wolf (*Canis lupus*) once ranged throughout most of Western Europe, but by the end of the last century was reduced to only a few small isolated populations in the Iberian peninsula, Italy and the Balkans (Promberger & Schröder 1993). In Italy, the population is believed to have reached its minimum in the early 1970s, when about 100 wolves were estimated, mostly in the central and southern portion of the peninsula (Zimen & Boitani 1975). After full legal protection was established in 1976, increased acceptance of wolves and a significant increase in wild ungulate populations favoured the numerical increase of the wolf and the re-colonisation of large areas of the former distribution range (Boitani 1992).

Currently, there are between 400 and 500 animals ranging along the Apennines from the French border to the southern tip of Italy, however, distribution is discontinuous and density varies (Boitani & Ciucci 1993). The natural re-colonisation of the Italian and French Maritime Alps started in 1992 and is likely to extend northward to the central Alps in the near future (Boitani & Ciucci 1993).

In spite of the expanding trend, population viability of the wolf is still threatened by small population size and significant adult mortality caused by illegal hunting, which is estimated to be 15-20% per year of the total population (Boitani & Ciucci 1993). The species has recently been confirmed as "endangered" (Pinchera et al. 1997). The re-colonisation of areas where the wolf had been absent for many years has increased conflicts between wolves and humans (predation on livestock) and has stressed the need for a national management plan (Boitani & Ciucci 1993). As pointed out by Noss (1992), a landscape approach in the range of 10,000-100,000 km^2 is likely to be the most adequate for integrating management of viable populations of wide-ranging animals and it is evident that an effective management plan for the wolf in Italy should consider all of Italy except the islands (about 250,000 km^2).

[20] Based on: F.Corsi, Duprè E., Boitani L., 1999 A large-scale model of wolf distribution in Italy for conservation planning Conservation Biology 13(1):150-159.

The most recent developments in population viability analysis have shown the usefulness of spatially explicit computer simulation and the integration of demographic and dispersal data with a detailed knowledge of the landscape geometry (Lamberson et al. 1992; McKelvey et al. 1992; but see Harrison et al. 1993; Harrison 1994). However, only a few studies have modelled spatial factors that determine wolf distribution. Mladenoff et al. (1995) built a multiple logistic regression model to assess the importance of landscape-wide factors in defining favourable wolf habitat in the Northern Great Lakes Region of the U.S. and found that road density and fractal dimension (land cover patch boundary complexity) were the most correlated variables. The model was also applied to the Northeastern U.S. to predict wolf favourable habitats (Mladenoff & Sickely 1998). A similar result on road density had previously been obtained through simple correlation analyses (Thiel 1985; Mech et al. 1988; Fuller et al. 1992). These studies all aim at defining the best habitat descriptor/predictor variables.

In this chapter a method is developed that uses multivariate analysis of GIS data to provide a spatially explicit model of wolf distribution. Using GIS applications, models are derived from a limited number of areas and locations of known presence of the wolf. From these locations the species-environment relationship is defined. The method shows potential for developing models useful for management planning when only limited information is available.

The aims of the model are to emphasise spatial patterns rather than habitat suitability and to contribute to the design of a country-wide conservation plan for the wolf by: 1) providing a basis for more advanced spatial and habitat analyses; 2) identifying the broad fragmentation patterns of the wolf distribution; and 3) providing insights into the most likely wolf re-colonisation of the Alps, an area where the wolf is expected to extend its range in the next few years.

METHODOLOGY

The methodology is based on two paradigms. First, given a set of environmental variables that potentially influence wolf distribution, a training set can be built using two groups, one of known wolf territories and the other of areas where the wolf is absent. A model can thus be built in the multivariate space defined by the variables that maximise the difference between the two groups. Second, given an adequate sample of areas where the wolf is found, it is possible to build a “signature” that best describes (and predicts) the areas where the wolf lives, based on the available environmental variables. The results can be used to: 1) identify all areas of the country where the environmental conditions are most similar to the known territories; 2) evaluate to what extent each portion of

the study area departs from the optimal conditions as defined by those of the territories.

The model, however, is not suitable for analysing habitat use, as no absolute value of the contribution of each environmental variable to the model is

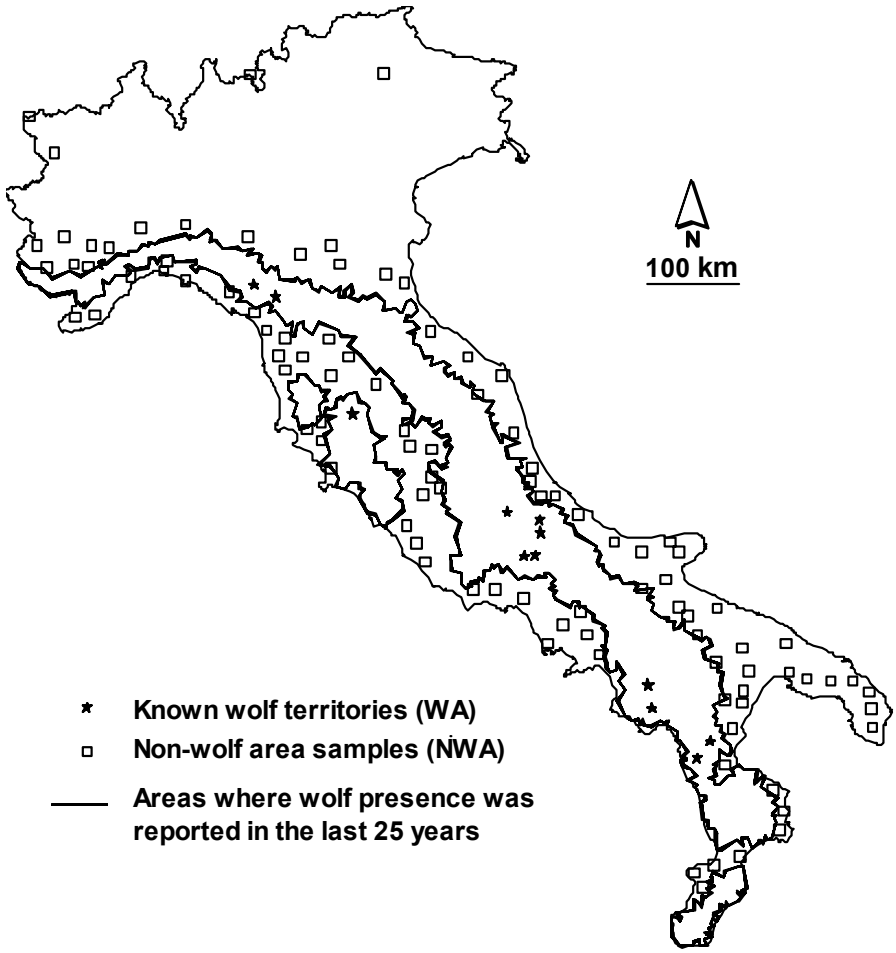


Figure 5.1: Area of current wolf presence (solid line); wolf territories (stars) and non-wolf sample areas (squares) used in the discriminant analysis.

obtained. In fact, changing the set of environmental data, and/or the training sets, the relative contribution of each variable is expected to change, while the model is expected to maintain overall stability in defining nation scale response (e.g., use of space).

The study area was all of continental Italy. The country is characterised by an extremely high variety of landscapes and ecological features. This is the consequence of the country's north-south extension, the mild coastal vs. the more continental climate of internal and northern regions, the altitude variation from sea level to 4800 meters and the intense habitat modification produced over thousands of years by human activity.

Data sets

Constrained by the limited country-wide information available, the data set was composed of three main sub-sets: 1) data on wolf presence/absence; 2) the environmental variables to be correlated to wolf presence; and 3) a list of 287 locations where dead wolves were collected during the past 25 years.

Wolf presence/absence

In order to define the training set, two groups of samples were used to describe the environmental features of the "wolf areas" (WAs) and the "non-wolf areas" (NWAs). The WAs were obtained using 12 wolf territories previously studied by radio-tracking (7) and/or intensive snow-tracking (5) in various parts of the wolf range (Zimen 1978; Boitani 1986; Ciucci 1994; unpublished data: Boitani, Ciucci, Francisci) (Fig. 5.1). A basic assumption is that the diversity of environmental conditions within these territories represents the average best conditions for a stable presence of the wolf in the Apennines, including human influence: all 12 territories were found within areas where wolves either have always been present or have recently (>10 years) and permanently colonised.

Using all available records (direct and indirect signs of wolf presence), NWA was identified as the area where no evidence of stable wolf presence had been gathered in the last 25 years (Fig. 5.1). Considering only the portion of Italy south of the Po river and given the size and shape of the Italian peninsula, it is reasonable to assume that any location within NWAs is within the reach of dispersing wolves (< 100 Km). As these areas have not been re-colonised in the last twenty years when wolves were expanding their range, it can be assumed that most of the NWAs' habitat is unsuitable. Therefore, random samples taken within the NWAs south of the Po river should provide samples of areas in which the values of the environmental variables are mostly unsuitable for the wolf. To minimise the risk that this group could include points of suitable wolf habitat, and to account for the diversity of habitat conditions, the NWA was oversampled and produced 100 non-overlaying circular areas (Fig. 1) by randomly sampling the centres of the circles within the NWA south of the Po river. The surface of each area (106 km^2) was equal to the average size of the 12 known wolf territories.

Of these 100 samples, 96 (8 times the number of available territories) were chosen randomly. The remaining four were selected at the site of the major icefields in the Alps in order to account for the different topographic conditions of the Apennines and the Alps, the latter exhibiting higher altitudes and presence of icefields which are absent in the Apennines. These differences can conceal the real relationship between an environmental variable and wolf presence.

Environmental variables

The second data set was used to describe the environmental characteristics of the training set and to extrapolate the result of the analysis to the entire study area. The thirteen variables used (Tab. 5.1) to define the multidimensional environmental space were selected not only to account for best knowledge of the basic wolf needs of space, food and cover, but also with respect to their availability in digital form and complete national coverage. Although clear insight into the real influence of each of the 13 variables on wolf distribution cannot be obtained, it was assumed that they describe fairly well the high diversity of ecological conditions to which the wolf is known to adapt (Mech 1970; Boitani & Ciucci 1993). Wolf distribution in Italy appears to be influenced primarily by human presence, food availability and, consequently, type of land use (Boitani & Fabbri 1983). The wolf in Italy has been reported to feed on wild ungulates, livestock, and garbage at dump sites (Boitani 1982; Ciucci 1994; Meriggi & Lovari 1996). Therefore, the selected set of variables included densities of sheep, number of ungulate species present (densities were not available) and density of dumping sites.

Cover was described in terms of percentage of land use classes (5 variables). Indices of diversity and dominance of land use were included to account for the overall landscape structure. Altitude was also included, and interpreted as an ancillary variable highly correlated both to human disturbance and cover availability. Human pressure is probably the most important factor affecting wolf distribution, especially in Italy where human impact on the environment is substantial (Boitani 1982). Additional variables such as human population and road densities were selected as habitat components to account for human disturbance.

All variables were obtained directly as digital thematic maps from various governmental sources and stored in a GIS (Arc/Info, ESRI, Ca.). Some of the data sets were used directly for analysis, such as the land use map in scale 1:200,000, while others were derived from the original digital thematic map by means of basic analyses (e.g., road density was computed from the original road network with a 10x10 km-cell grid). Human population and sheep densities (ISTAT 1990, 1991) were aggregated by *comune* (municipality), the median size of which is 21.5 km^2. A Digital Terrain Model with a square cell size of

250 meters was used to derive information on altitude, while the diversity and dominance indices were calculated using a 10x10 km-cell grid, following the Shannon-Weaver formula. The maps of dumping site density and number of ungulate species were available at smaller scale (about 1:1,000,000).

Variable	***Final model***	***Origin and resolution of data***
Farmland	√	Land Use maps (1962-1986). Scale 1:200.000
Forest	√	Land Use maps (1962-1986). Scale 1:200.000
Pasture		Land Use maps (1962-1986). Scale 1:200.000
Bare soil or water		Land Use maps (1962-1986). Scale 1:200.000
Urban settlement	√	Urban settlement contours (ENEL 1971). Scale 1:25.000
Altitude		Italy's DTM (Ministry of Environment). Resolution 250 m
Human density	√	13° National Census of the Population (ISTAT 1991). Aggregated by *comune*
Road density	√	Maps of the Italian Touring Club. Scale 1:200.000
Shannon diversity index	√	Land Use maps (1962-1986). Scale 1:250.000
Shannon dominance index	√	Land Use maps (1962-1986). Scale 1:250.000
Dump sites density	√	Census of the Ministery of Agricolture and Forests (1990). Aggregated by Region
Sheep density		4° National Census of Agricolture (ISTAT 1990). Aggregated by *comune*
Number of ungulate species	√	Species' distribution maps (Ministry of Environment 1993). Scale 1:1.250.000

Table 5.1. The thirteen environmental variables used in the analysis. The variables used in the final model are marked √.

Since data quality and homogeneity were a major concern, all original data sets underwent notable editing and all discrepancies were corrected. The final layers were then converted to raster format with a cell size of 250 m.

The 12 territories and the 100 NWAs were characterised by performing simple overlay with the 13 layers and calculating basic statistics (percentage of coverage, mean values, etc.) depending on the type of environmental variable. In order to extend the result of the modelling based on these training sets to the entire study area, map algebra focal functions (Tomlin 1990) were used to replicate the same statistics over the entire study area. Each raster of 13 variables was processed assuming each pixel to be the centre of a hypothetical wolf territory and assigning to that pixel the same statistics used to characterise

the training set, and calculated within a window of a 23 pixel radius. This radius gives an area of 103.8 km^2, the best approximation to the average dimension of the 12 wolf territories (106 km^2) obtainable with a cell size of 250m.

Dead wolf locations

The data set of 287 dead wolf locations was obtained by pooling all information collected by various Italian offices and scientists on the wolves found dead in the past 25 years. About half of the sample was collected by one of the authors (L.B.). Evidence of human related cause of death (e.g., poison, shotgun, car/train accidents, etc.) was available for at least 70 % of the sample, while, for the remaining 30 %, it was presumed through indirect ancillary information. Although only a portion of the total number of illegally killed wolves was recovered, and the collection was not organised through a pre-defined procedure, the sample can be assumed to reflect the gross spatial distribution of killed wolves. In Italy, a wolf illegally killed or (more rarely) found dead is still an event, and the news is immediately spread: local authorities usually recover the body and file a formal statement. The sample may not accurately represent regional variation in the recovery system and temporal distribution of killed wolves; however, this does not affect the sample utilisation in the analysis (see below for the tests on spatial distribution of the sample). This data set was used in exploring the correspondence of dead wolf locations to areas of marginal environmental quality, thus providing a tool to validate the model and to enhance its conservation interpretation.

Data analysis

The first step aimed to identify the most important areas of wolf presence (actual and potential). Discriminant function analysis (DFA) was used. This statistical method, even though constrained by its inherent limitations, has been widely applied to define a binary use of space (e.g., wolf/non wolf) (Verbyla & Litvaitis 1989; Livingston et al. 1990; Dubuc et al. 1990). Similar results could be obtained as well using a logistic regression and less rigorous statistical assumptions. However, it was preferred to normalise the variables through various transformations (see below) and use a DFA for its similarity to the methods adopted in the second step of the analysis (e.g. the Mahalanobis distance). A forward stepwise canonical discriminant analysis was performed on 13 variables, 2 groups and 112 observations (12 WAs and 100 NWAs). Dumping site density and number of ungulate species were normalised using logarithmic transformation, forest and pasture extension using the Freeman and Tukey transformation, and the remaining 9 variables using the Box-Cox transformation with different values (Sokal & Rohlf 1995).

The analysis was run with F = 0.6 (F being the probability value of the F statistics) to determine how significant the contribution of a variable to the regression had to be in order to be added to the discriminant function. The DFA results were used to classify the entire study area, i.e. all of Italy. The classification was calculated as the posterior probability of each pixel to belong to one of the groups

$$p(t|x) = \frac{\exp\left(-0.5\ D_t^2(x)\right)}{\sum_u \exp\left(-0.5\ D_t^2(x)\right)} \qquad [1]$$

where x is the vector containing the values of environmental characteristics for each pixel and D_t is the generalised squared distance of each pixel from the t group

$$D_t^2(x) = (x - m_t)' S_t^{-1} (x - m_t)$$

where S_t represents the within group covariance matrix and m_t the vector of the means of the variables of the t group. Equality of covariance matrices was tested by means of the Box M test (Davis 1986). The generalised squared distance (SAS Institute, Inc. 1985) was used instead of the simple Mahalanobis distance, as it accounts for differences in the variance-covariance matrix of the two groups. The *a priori* probabilities were considered to be equal, with the threshold set at 50% probability.

In the second step, independently calculated from the previous one, the aim was to describe potential interconnections between the areas of wolf presence. The Mahalanobis distance statistic was used as an index of environmental quality (distance from the wolf's best environmental conditions). The Mahalanobis distance statistic has been used as a multivariate index to rank habitat suitability in GIS raster maps (Clark et al. 1993; Knick & Dyer 1997). It avoids many difficult requirements of discriminant function and logistical regression, particularly problems involving incorrect classification of used versus non-used habitats (Clark et al. 1993). Wolf territories rather than a series of animal locations (radio-locations: Clark et al. 1993; sightings: Knick & Dyer 1997) were used: the smaller number of "observations" should be partly compensated by their higher ecological significance (large and stable areas).

A surface of actual/potential use of space for the entire study area was calculated. The environmental centroid of the WAs group represented the best description of the optimal environmental conditions for the wolf. An index of environmental quality was built based on the environmental Mahalanobis distance of any given location from the WAs centroid (the smaller the distance the more similar the environmental conditions of that location to the wolf's ecological profile). The environmental distance was calculated for a continuous raster covering the entire study area.

The third step was based on the overlay of the locations of dead wolves to the models produced by the above analyses. It served as validation of the previous ones and as support for their conservation interpretation.

The relationship between the distance from the WAs centroid and the probability of wolf occurrence was evaluated using the location of dead wolves. For each location the environmental distance from optimal wolf conditions was calculated through interpolation from the continuous surface. The resulting frequency distribution was fitted with different probability density functions to interpret changes in wolf distribution in response to variation in the environmental distances from WA's.

To analyse if an increase in population was related to expansion of wolf populations into areas of lower environmental quality, differences in the environmental quality levels of the areas where dead wolves were found in different time periods was searched. All casualties were ordered chronologically and then grouped in sets of 10, 20, 30, 40, 50, 60 and 70 time consecutive locations. Each set was tested for normality and analysed with ANOVA, the hypothesis being that if a density dependent pattern of re-colonisation is applicable to the wolf in Italy, then the distribution should show a shift toward an increasing number of dead wolves in lower quality environments.

Finally, a tentative model for wolf management was produced by extrapolating from the distance raster a new raster where the pixel values represented the expected percentage decrease of dead wolves (due to casualties) which would be achieved if all areas falling within the pixel's ecological distance were effectively protected. The index was obtained using the cumulative probability density function derived from the ecological distances of dead wolf locations. Plotting on the same graph the cumulative frequency distribution of areas of increasing ecological distance and the cumulative probability density function of dead wolf locations, the curves represent, for any given value of ecological distance, the percentage of territory that should be totally protected to achieve an expected percentage reduction of wolf casualties. This model was applied only to the portion of Italy south of the river Po (see below).

The stability of the environmental distance model was assessed using jackknife procedures (Sokal & Rohlf, 1995; Cressie, 1993). In analysing the results, the study area was divided into a portion south of the Po river including the peninsular part of continental Italy (Apennines), where most of the current wolf range is located, and a portion north of the Po river including the Alps where the wolf is currently expanding its range.

RESULTS

Areas of importance for wolf presence

Nine of the 13 variables were selected by the stepwise discriminant analysis (Wilks' lambda=0.567, F=8.64, p=0.0001). The 100 NWAs appeared to be well separated from the 12 WAs on the first canonical variates, with only a few NWAs within the pertinence of the WAs (Fig. 5.2). The 12 wolf territories were correctly classified, as overall probability of belonging to the wolf group was > 90%. Only 3 of the 96 random NWAs were assigned to the wolf distribution, while the 4 non-random NWAs were classified, as expected, into the NWAs group.

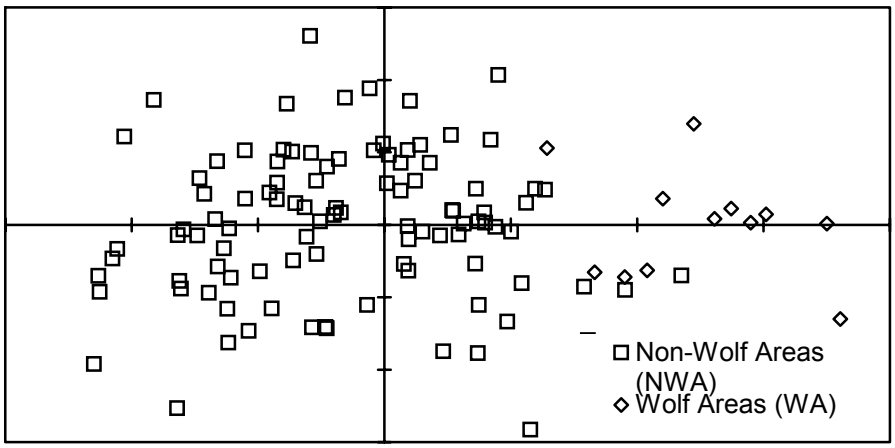


Figure 5.2 Discriminant analysis model. The second axis is plotted only to enhance readability as the first axis accounts for almost 100% of the total variability.

By applying the classification criterion to the entire study area, a map of the areas which are most important for wolf presence was obtained (Fig. 5.3). About 14,200 km^2 (about 5.7 % of continental Italy) with a posteriori probability of more than 50% were found to pertain to this category. Of these, 11,300 km^2 are in the peninsular portion of the country (Apennines), while 2,900 km^2 are located in the Alpine region. These areas, offering optimal environmental conditions, could be considered as the core of the wolf (actual and potential) distribution and should be expected to act as a source for less suitable areas.

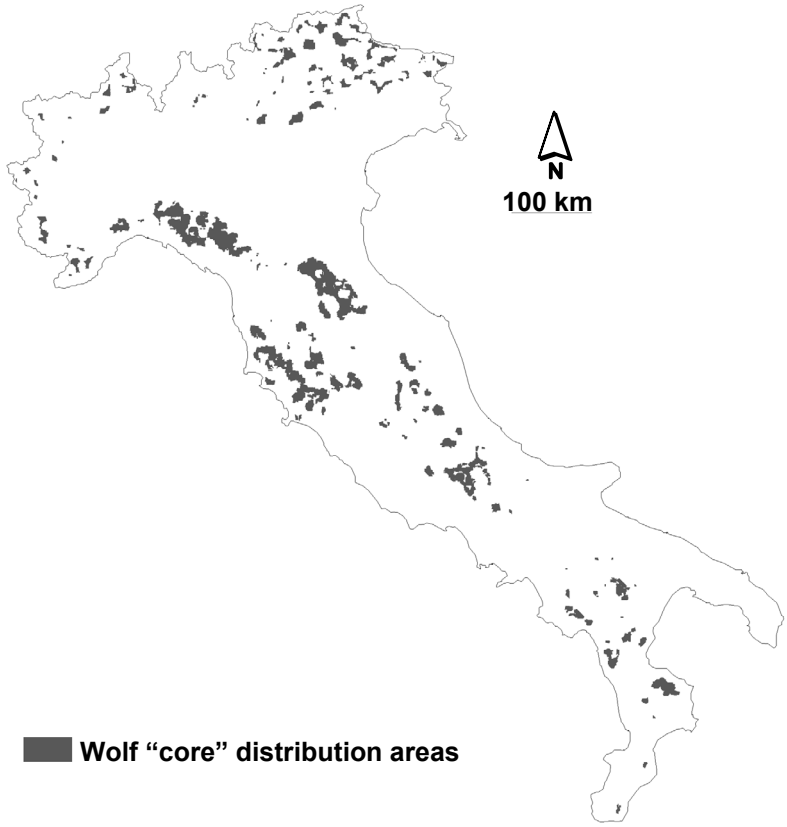


Figure 5.3 Wolf "core" distribution areas as obtained from the discriminant analysis model, i.e. with a posteriori probability >50%.

Index of environmental quality and the surface of actual/potential use of space

The values of the raster of the distances range from 0-2,933 (mean=297, s.d.=301). These absolute values have no specific meaning *per se*, and are of interest only when considered in relation to another variable such as the dead wolves' locations. The frequency distribution of these locations fitted a log-normal density function (mean=3.9068367, s.d.=0.8644463, Kolmogorov-Smirnov d=0.0124403, p=n.s.) (Fig. 5.4). The right side of the distribution (the decreasing part) indicates density dependence, that is, as we move away from the areas of high environmental quality, wolf numbers and consequently deaths tend to decrease. As for the left side, taking into account that all deaths are due to human related causes, there are various possible explanations: 1) interactions between humans and wolves tend to be less frequent in areas of high

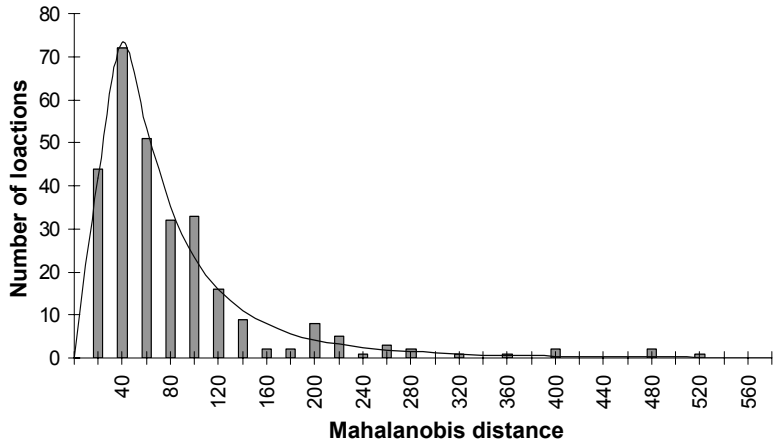


Figure 5.4 Log-normal distribution fitted to the environmental distances of dead wolf locations. The histograms show the observed distribution; the line shows the fitted log-normal distribution (mean = 3.9068, s.d. = 0.8644).

environmental quality (e.g. lower human population density, higher availability of wild prey); 2) the interactions do not cause any casualty (e.g. better/more cover availability); and/or 3) the casualties that result from these interactions are not included in the sample.

% of study area	*Expected % reduction of casualties*	*km²*
2.4	< 20	6060
8.9	20-35	16157
18.0	35-50	22661
30.6	50-65	31618
50.2	65-80	49173
100.0	>80	124516

Table 5.2. Expected percent reduction in wolf casualties that would be achieved through full protection of the areas pertaining to different classes of environmental quality. Percentages of study area are given as cumulative figures to the upper limit of the probability class.

The first two hypotheses are similar as both reduce the number of casualties. As for the third, it should be rejected, as the patterns of relatively high human presence throughout the country make it unrealistic to postulate lower efficiency in recovering dead wolves in the best wolf areas.

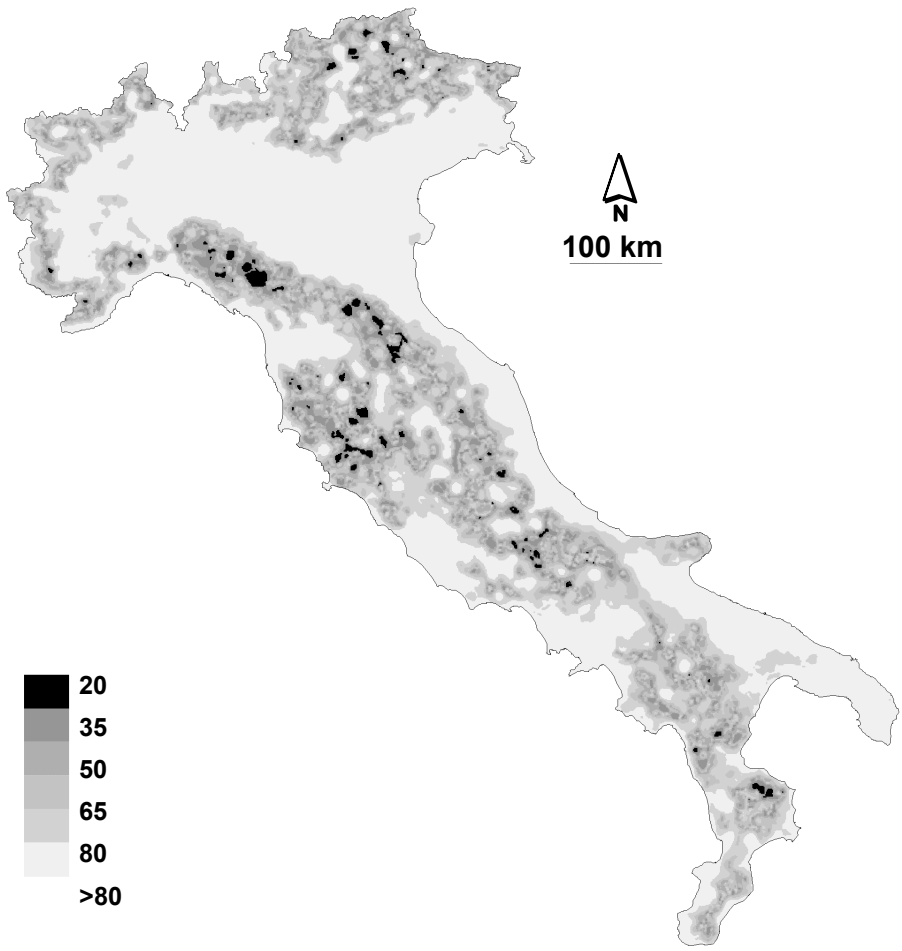


Figure 5.5 Probability map describing the expected decrease in the number of dead wolves due to casualties. It was obtained by calculating the probability associated with each distance based on the probability density function derived from the locations of dead wolves.

The map obtained from the conversion of the ecological distances based on the probability density function of the ecological distances of the dead wolf locations is shown in Figure 5.5. For management purposes, it can best be read as a map of the percent reduction in wolf casualties that would be achieved through full protection of the areas with different levels of environmental quality (Tab. 5.2).

Based on the cumulative frequency distribution of distance classes throughout continental Italy (Fig. 5.6), different conservation scenarios can be analysed by means of a cost/benefit approach. For example, the point of maximum "gain" occurs at an environmental distance of 109, where the number of casualties would decrease by more than 80%, while requiring that less than 40% of the study area south of the Po river would need to be protected.

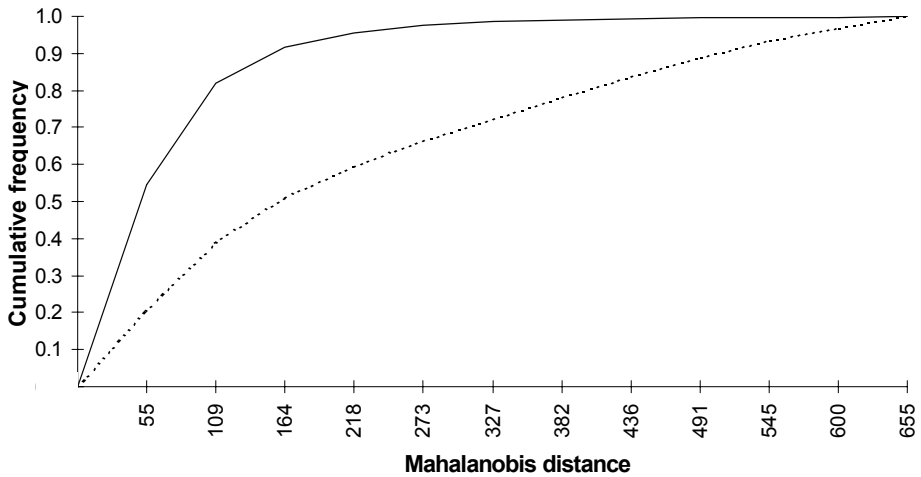


Figure 5.6 Model of expected percent decrease of wolf casualties with increasing size of protected areas. The solid line shows the cumulative log-normal distribution of dead wolves. The dashes line shows the cumulative distribution of the environmental distance classes in the study area.

Model validation and stability

If the pattern of space use in relation to the environmental distances has remained constant in the past 25 years then wolf densities in the areas included within any given distance level should not have changed in time, and the wolf population should have increased mainly through an increase in its area of occupancy. The results of the ANOVA applied to groups of sets of 10, 20, 30, 40, 50, 60 and 70 consecutive locations supported this hypothesis: all groupings but the first (with sets of 10 locations each) showed no significant differences among their sets. None of the tests gave significant results ($p > 0.5$). The low significance of the grouping with 10 locations ($p < 0.1$) may be explained as the result of local random effects (e.g. the temporary occupancy of areas of lower environmental quality during the dispersal phase) due to the excessive subdivision of the sample.

As for the overall stability of the models, the jackknife process showed an expected high variability (CVs of over 1000%) in the pixel values of the raster of environmental distances. This is not surprising as the relative contribution of the environmental variables varies according to the sub-set of samples used for the analysis. With the same 12 rasters obtained from the jackknife, the probability density function of the dead wolves' locations was also calculated, and from these a calculation of a probability raster for each of the 12 runs of the jackknife was made. In this last case, the coefficient of variation of each pixel dropped drastically (CV<=13%) in accordance with an expected stability of the distance values when considered as a relative measure of environmental quality.

DISCUSSION

There are at least two reasons for adopting a national scale approach to conservation of the Italian wolf metapopulation (*sensu* Harrison 1994, i.e., spatially structured). First, a national scale approach is needed to manage fragmentation of suitable habitats and the inevitable metapopulation structure of the resulting population (May 1994), and hence to manage conflicts with human economies and illegal hunting. Second, the future of the wolf in Italy, as well as that of most large carnivores elsewhere, will ultimately depend on our ability to designate a zoning system of areas and connecting "corridors" where the wolf will be managed in ways appropriate to local ecological and economic conditions (Boitani 1982; Mech 1995; Weaver et al. 1996; Noss et al 1996). An "integrated landscape management" (Saunders et al. 1991; Turner et al. 1995; Wiens 1996) appears to be the only rational approach to ensure the survival of a mobile and adaptable species like the wolf, particularly in a highly fragmented landscape mosaic such as Italy (and Europe).

The study explored a method for obtaining a spatial model of wolf distribution as contribution to the preparation of a conservation plan. Model building is a deductive-inductive process, with model formulation and validation occurring iteratively (Stormer & Johnson 1986; Clark *et al.* 1993) and developing through a feedback process with field studies (Price & Gilpin 1996). Good models are the key to good conservation management (Gilpin 1996), yet in the real world data are rarely adequate for complex and robust simulations (Dunning *et al.* 1995). This method is an example of integration of inductive expert knowledge and deductive modelling approaches (Stoms *et al.* 1992) to maximise the utility of limited data.

The model's predictions for the Alps may not be fully justified due to substantial ecological differences and should be taken only as a first indication of potential wolf distribution. Nevertheless, the model is based on current best

knowledge and it provides a first insight into the likely evolution of wolf presence in that area. The method is more suited to identifying spatial patterns than critical habitat factors for wolf distribution, since a variety of habitat combinations can produce identical distance values.

Even though human attitudes toward wolves are probably one of the most important factors determining wolf distribution (Boitani & Ciucci 1992; Mladenoff *et al.* 1995), it is neither a simple variable nor can its distribution be mapped. The method assumes that human attitude is hidden in the other variables (e.g., road density and land use), as suggested by Mladenoff *et al.* (1995), Thiel (1985), Mech *et al.* (1988), and Mech (1989). This approach implies that human disturbance is a density dependent variable (e.g. it increases linearly with human density). This is obviously a weak assumption as the human attitude toward wolves can greatly modify this relationship.

The "core" wolf areas as obtained from the discriminant analysis and the 12 wolf territories were characterised using the environmental distance surface, showing a conservative effect of the results of discriminant analysis. The average distance from the wolf optimal areas was 16.78 (s.d. = 6.08) for the "core" wolf areas and 31.09 (s.d. = 19.68) for the 12 territories. The high patchiness of these areas is expected in a highly fragmented landscape but their interpretation as a source for less suitable areas is constrained by the particular definition of "core" area being used.

A similar caution should be observed in interpreting the optimal areas as obtained from the means and the variance-covariance matrix of the 12 territories: the definition refers to the statistical method rather than to an analysis of biological factors. Although the jackknife process used to assess the variability of the environmental distance model justifies good confidence in its statistical stability, the model's best utilisation is as a conceptual guide for further insight into the biological and landscape reality of its results.

The output of the environmental distance model should never be interpreted as an absolute value. The high variability evidenced by the jackknife indicates that there is no direct functional relationship between these values and an absolute index of environmental quality. However, the jackknife also shows that the relative measure of these distances appears to remain constant, allowing their use as a relative index of environmental quality. The environmental distance raster can be interpreted as the relative expectation of wolves being at a given location (lower distances indicating higher expectation).

The general level of spatial fragmentation (i.e., fragments' size) appears within the order of magnitude of wolf territory size, allowing for future simulation of the effect of the territorial behaviour on interpatch dynamics (Gutierrez &

Harrison 1996). The fragmentation pattern in the Alps should be re-analysed when similar data is available also for neighbouring countries, as ecological continuity may be ensured by areas across the border.

The map of percentage expected decrease of dead wolves due to casualties allows for preliminary analysis of various conservation scenarios on a cost/benefit basis, albeit with caution due the use of a simplistic model. Assuming that the patterns shown by the frequency distribution of dead wolf locations in the Apennines can be extrapolated to the Alps, the percentage of area to be fully protected in order to expect a corresponding percentage decrease in the number of dead wolves may be inferred (Tab. 5.2). The distance raster, when analysed in conjunction with available GIS functions, can be used to address important conservation issues such as areas of occupancy, core areas, least conflict (with human activities) areas, conservation options between areas of different quality (source/sink) and corridors.

Within the limits of the practical utilisation of metapopulation conceptual models (Gutierrez & Harrison 1996), the model is currently being used to support the difficult technical and political process of preparing a conservation and zoning plan for the wolf in Italy. However, caution is called for in using the appealing predictions of computer models to make real-life decisions without a critical analysis of their inherent limitations (Price & Gilpin 1996).

Chapter 6: Expert-based species distribution maps for the assessment of the biodiversity conservation status: the Italian National Ecological Network[21]

INTRODUCTION

A number of approaches to prioritise conservation actions has been developed from the basic consideration that species are not distributed evenly across the Earth's surface (Nowicki, 1998). This uneven distribution derives from both natural and human-induced factors. Thus the distribution of species can be seen as a web of areas with different functions, linked together to form an ecological network (Bouwma *et al*., 2002).

From a practical point of view, ecological networks can be identified with different approaches. Through landscape ecology, for example, the network is derived from fragmentation and connectivity analysis of landscape patches (Forman, 1983; Foppen et al., 2000). Relationships between the landscape patches and the species, whose distributions are meant to be described by the network, are only marginal. Thus there is no real guarantee that the network will be of real use in the conservation of animal and plant species.

In conservation biology the approach to ecological networks evolves to include in the analysis the ecological perspective of those species that benefit from the ecological network (Battisti, 2002; Gustafson & Gardener, 1996). Under this approach the importance of a certain landscape unit, the presence of a barrier or of an ecotone (Manson et al., 1999), and the "permeability" of the environmental matrix are always defined as they relate to the species or group of species being analysed. In other words, the conservation biology approach to ecological networks aims to describe the joint environmental suitability of the species being analysed.

---

[21] Extract from: Boitani L., Corsi F., Falcucci A., Maiorano L., Marzetti I., Masi M., Montemaggiori A., Ottaviani D., Reggiani G., Rondinini C. 2002. Rete Ecologica Nazionale. Un approccio alla conservazione dei vertebrati italiani. Università di Roma "La Sapienza", Dipartimento di Biologia Animale e dell'Uomo; Ministero dell'Ambiente, Direzione per la Conservazione della Natura; Istituto di Ecologia Applicata.

To implement the conservation biology concept of ecological networks, detailed knowledge of species distributions is required. This allows both assessment of the areas that match the various requirements of a number of species, and an overall biodiversity assessment (Mittermeier *et al.*, 1998). This latter aspect is generally developed starting from the individual species' ranges by overlapping them to produce maps of species richness, or other biodiversity indices such as endemism, species abundances, phylogenetic and/or taxonomic local variation (Reid 1998). While species distributions are difficult to define, models developed based on the ecological requirements of the species can be a good approximation (Corsi *et al.* 2000).

While precise direct information on species distribution is rather limited, a number of techniques have been developed to derive species distribution models with the support of geoinformatics. These include expert-based deductive models, which use specialist's expertise to define the species ecological requirements to derive distribution information, and more objective inductive models, which derive species ecological requirements that drive the modelling process. In the latter case known localities where the species occurs are compared with a number of environmental GIS layers (Corsi *et al.* 2000).

Expert-based species distribution models are cost effective and efficiently address species conservation. Species distribution models are central to many of the conservation strategies proposed at the individual species level, for example individual species' conservation action plans (IUCN-SSC, 2003c) and CITES analyses (Rosser *et al.* 2001). The African Mammals Databank (AMD) was the first continent-wide high-resolution analysis of medium and large African mammal species' geographical ranges. The project was designed to mobilise and formalise distribution and ecological information for 281 species of large and medium size mammals in Africa. (see Chapter 4 of this thesis). The usefulness of the AMD has prompted a number of NGOs to support its continuation and the inclusion of the remnant vertebrate species in the system. The African Vertebrates Databank (AVD) is being conducted as a joint effort of the IUCN SSC, CI-CABS, BirdLife International, the Museum of Natural History of London, and the Instituto di Ecologia Applicata in Rome (Fonseca *et al.* 2000).

At the national level, the same methodology was applied to develop the Italian Ecological Network (Rete Ecologica Nazionale – REN), a system recently developed for the Italian Ministry of the Environment (Boitani *et al.* 2002). This chapter presents an approach to building understanding of the more general patterns of higher taxa, which were analysed at the end of the Italian ecological

Network project. The aim of the analyses was to illustrate the potential for applying the models' outputs as a contribution to understanding the patterns of biodiversity distribution, and to evaluate the efficiency of protected areas in conserving biodiversity. This latter aspect is particularly important, given that composite distribution areas of many species and their internal patterns are particularly useful for biodiversity assessment and conservation (Scott *et al.*, 1987; Reid and Miller, 1989; Miller, 1994).

METHODS

Combining a literature review and expert advice, a data bank of the ecological requirements for the 504 species of vertebrates of the Italian fauna was developed. The data bank includes scores of environmental suitability for environmental GIS data layers such as the Corine Land Cover map, elevation, distance from water and distance from roads. The data bank also stores the Extent of Occurrence (Gaston, 1994; IUCN 2000) of each species.
Using the environmental suitability scores and the environmental layers available in the geo-database (Table 1), deductive environmental suitability models (Corsi et al., 2000) were developed for each species. Models were developed with a nominal resolution of 1 ha. To accommodate different ecological requirements for the different species, nine model types were developed. Each model type corresponded to different rule-sets that were applied to score and combine the environmental layers (Boitani *et al.* 2002).
From the environmental suitability model a distribution map was derived. It assumed that the suitable areas within the Extent of Occurrence of the species are those actually occupied by individuals of the species (Corsi *et al.* 1999, Boitani *et al.* 1999).

Groups	Number of threatened species
Mammals	34
Birds	70
Reptiles	9
Amphibians	14
Fresh water fishes	22

Table 6.1, Taxonomic breakdown of the species included in the species richness map of the threatened vertebrates of Italy.

Species distribution maps were validated with independent datasets provided by the Italian Ministry of Environment. The Overall Accuracy was calculated for each model. Models were considered valid if the Overall Accuracy was above 50%. Individual species models with an Overall Accuracy above the 50%

threshold were integrated into species richness maps. Species richness is a simple way of addressing biodiversity patterns distribution (Williams, 1996), and has been successfully used for gap analysis (Caicco *et al.* 1995).

Species richness maps were produced (a) using all of the vertebrates, (b) by taxonomic group (e.g. mammals, birds, reptiles, amphibians and fishes) and (c) for the 149 threatened species (Table 6.1) listed in the Italian Red Book of Threatened species (Bulgarini *et al.* 1998).

Species richness maps were standardised by dividing the cell values by the maximum of each map.

The distribution of the standardised richness of threatened species was compared to the existing network of protected areas to assess its effectiveness to protect biodiversity. The effectiveness of the protected areas was analysed following the classification of the Italian Ministry of the Environment, which recognises 4 categories: National Parks, Natural Reserves of the State, Special Protection Areas (European Union. 1979) and Sites of Community Interest (European Union, 1992)).

For each class, the overall protection of Italy's threatened biodiversity was calculated as

$$P_a = \frac{B_a}{B}$$

where P_a is the protection index of the ath category of protected areas, B_a is the sum of all the biodiversity index values within the ath category of protected area and B is the total sum of biodiversity index values at the national level.

The effectiveness of the four categories of protected areas in conserving biodiversity was also analysed. The frequency distribution of the biodiversity index within the four categories of protected areas was compared with the frequency distribution of the biodiversity index of the map of threatened species. The Kolmogorov-Smirnov test was used to assess differences in the frequency distributions. To account for the multiple pair-wise comparisons that were made, significance levels were adjusted following Bonferroni's correction:

$$\alpha/n$$

where α is the significance level and n is the number of pair-wise comparisons.

Finally, gaps in the protection scheme were identified by selecting the areas with a biodiversity index higher than the 95th percentile of the frequency distribution. The distribution of these areas was then compared with the existing network of protected areas.

Based on this latter map, a management tool was developed. The tool aimed to identify a potential scheme for the creation of new protected areas. The areas of

higher biodiversity value were classified according to their distance from existing protected areas. The map indicates needs for new protected areas focussing on vertebrate conservation and/or possible extension of existing areas.

RESULTS

A comparison of the species richness map of threatened vertebrates with maps of the individual taxa, and with the overall vertebrate species richness, shows that the sub-sample of the threatened species is a good representation of species richness in Italy. In particular, the distribution of the species richness of the threatened species matches 66% of the distribution of the overall species richness (Boitani et al. 2002). Areas for which the two differ are areas of higher concentration of threatened species, therefore the map of species richness of threatened species can be considered an effective tool for biodiversity conservation planning (Boitani et al. 2002).

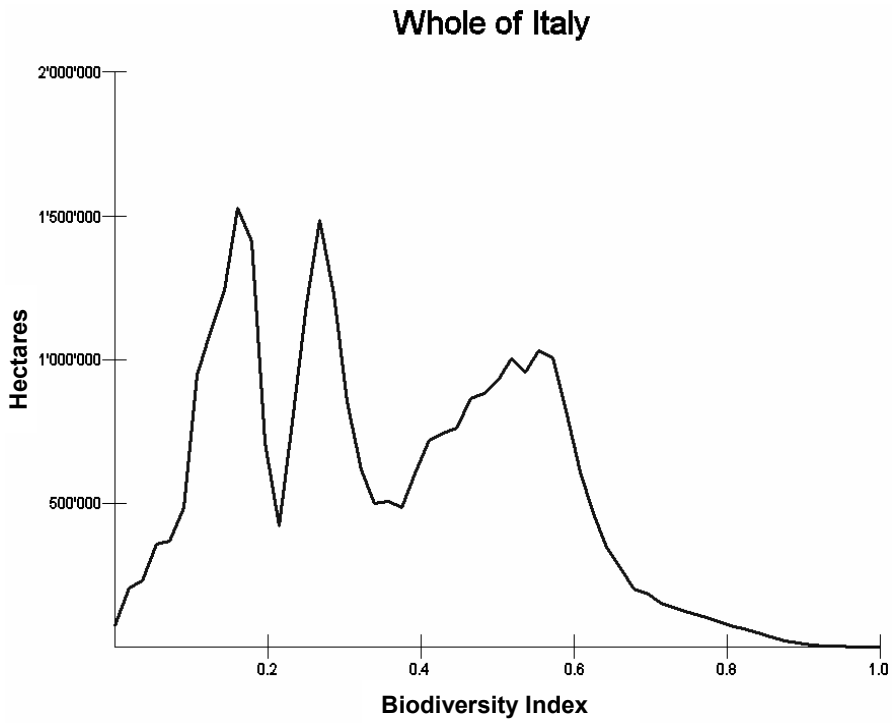


Figure 6.1 Frequency distribution of the biodiversity index for the threatened species over the whole of Italy.

For all of the 149 species included in the species richness map of the threatened vertebrates, the distribution model validation exceeded the 50% accuracy threshold.

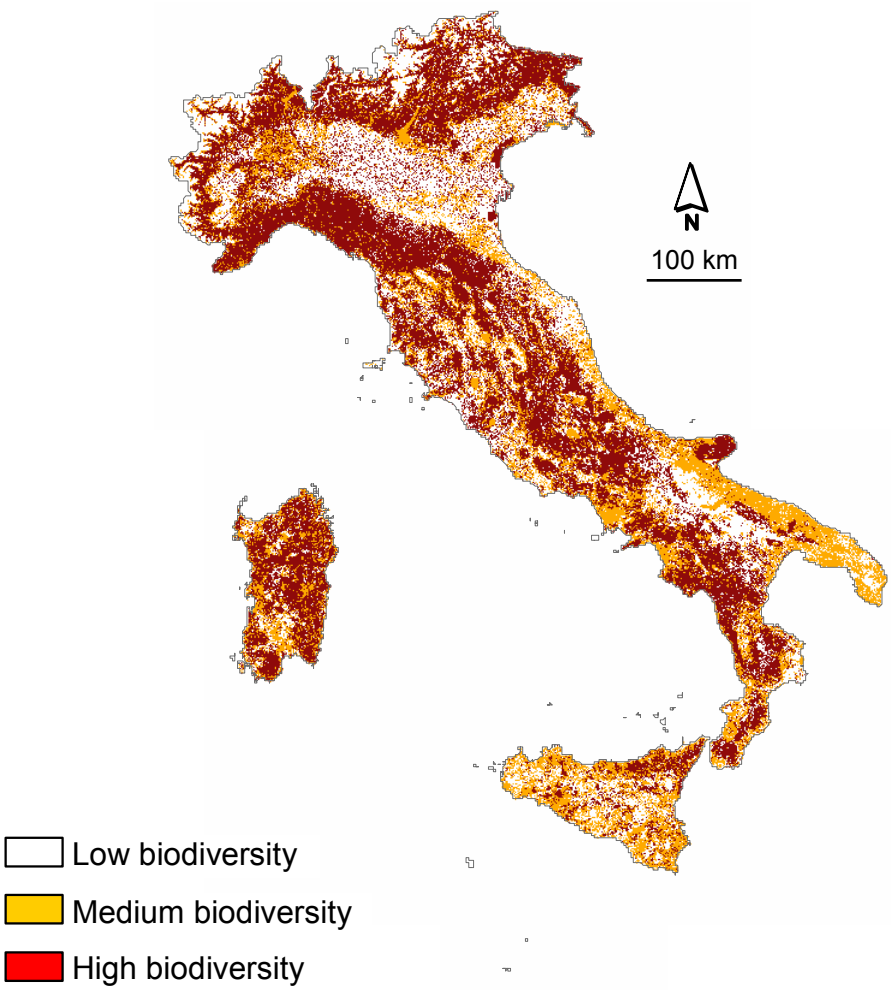


Figure 6.2 Map of the biodiversity index classified using the natural breaks of the frequency distribution.

The frequency distribution of the biodiversity index for threatened species (fig. 6.1) shows three peaks of frequency, which correspond to values of 0.17, 0.27 e 0.58 respectively. This allows identification of two natural breaks for mapping the biodiversity index according to high, medium and low biodiversity values (fig 6.2)

The distribution of the biodiversity index shows a pattern that follows the Alps range (excluding the higher altitudes) and the Apennines. Areas outside the major mountain ranges include known wilderness areas such as Sardinia and the Gargano peninsula. Given that the biodiversity index has been developed from the threatened species richness, which includes a number of water bird species, areas such as the Venice lagoon also fall within the high biodiversity class. Low biodiversity indices are found in the higher peaks and in the agriculture lands in the Po valley and along the coastal plains.

Type of Protected Area	Surface (km^2)	Surface (%)	Biodiversity Proportion (%)	Biodiversity Index (average ± s.d.)
NP	14243	4.7	5.5	0.407 ± 0.196
NRS	547	0.2	0.2	0.351 ± 0.173
SPA	11587	3.9	4.1	0.372 ± 0.188
SCI	24738	8.2	9.3	0.399 ± 0.192
Not protected	248780	83.0	81.0	0.345 ± 0.187
Total	**299895**	**100.0**	**100.0**	

Table 6.2 Values of the biodiversity index within the various categories of protected areas: NP: National Parks; NRS: Natural Reserves of the State; SPA: Special Protection Areas; SCI: Sites of Community Interest.

Examining the distribution of units within the high biodiversity class shows that, in accordance with the expected distribution, higher biodiversity index values are observed in the northern part of the peninsula, where the Italian fauna neighbours the Central European fauna. The lower values of the class (between 0.7 and 0.8) occur in the Central and Southern parts of the Apennine range.

Protected areas cover 17% of Italy's surface. An overlay of the protected areas network on the biodiversity distribution map shows that they protect 19% of the biodiversity index for threatened species. Sites of Community Interest (SCI) and National Parks (NP) are the two categories that protect the higher proportion of biodiversity (Table 6.2). This is even more evident when the average Biodiversity Index per hectare is taken into consideration.

The frequency distribution of the biodiversity index within the four categories of protected areas was compared with the overall frequency distribution of the index, and with the frequency distribution of the index outside of the protected areas. The number of pair-wise comparisons is 5, which brings the level of significance to 0.001 (Bonferroni adjustment for multiple comparison tests). All of the frequency distributions differed significantly ($p<0.001$). Figures 6.3-6.6 show the frequency distributions of the biodiversity index for the different types

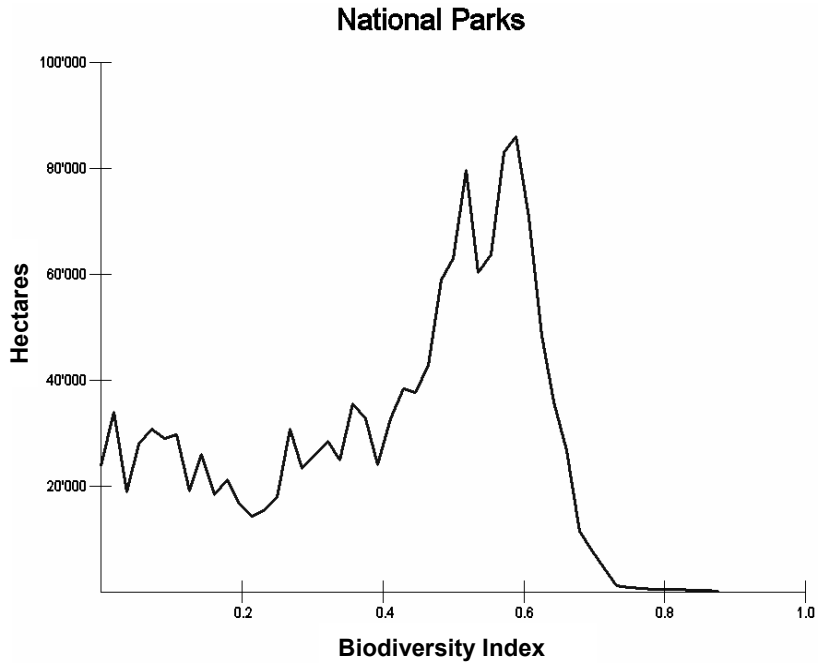


Figure 6.3 Frequency distribution of biodiversity index of threatened species in National Parks.

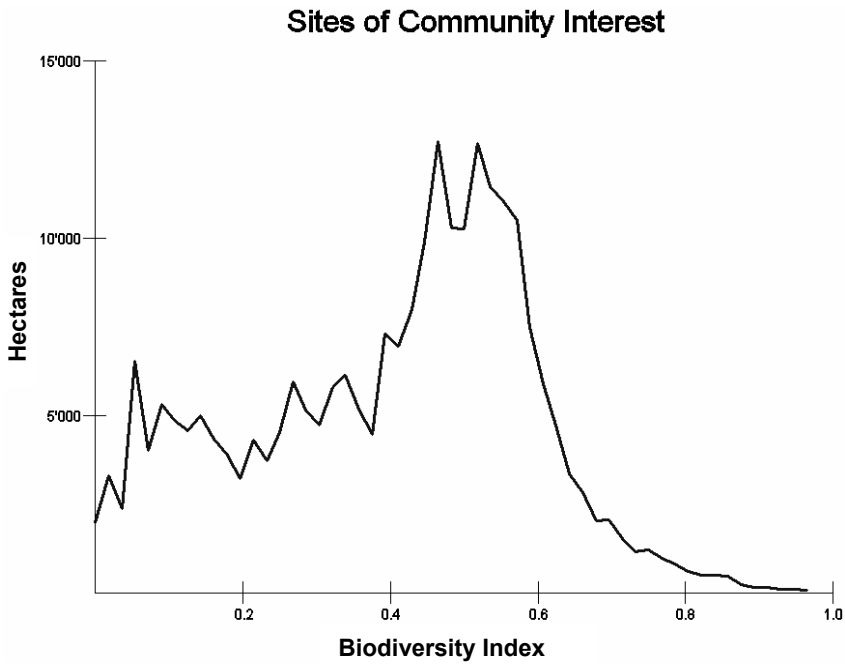


Figure 6.4 Frequency distribution of biodiversity index of threatened species in Sites of Community Interest.

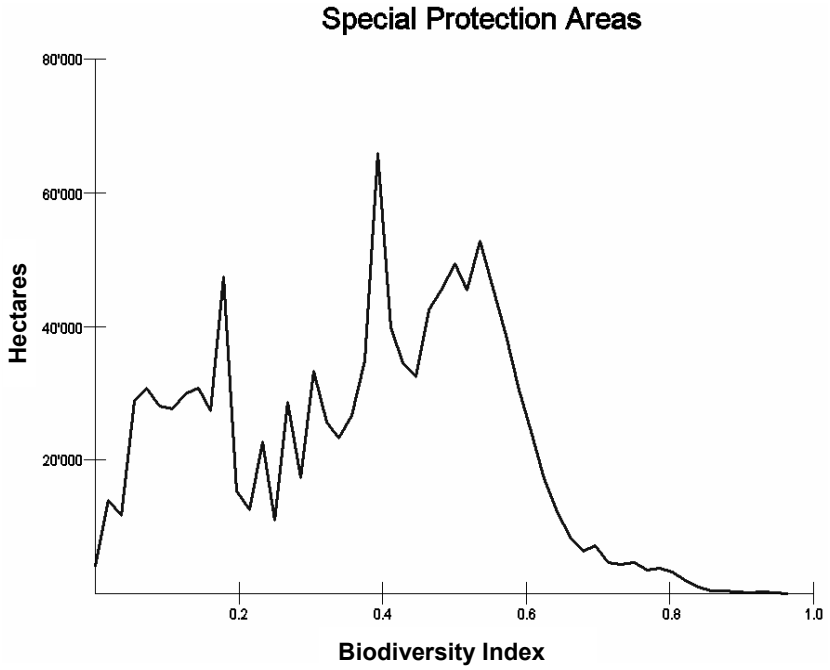


Figure 6.5 Frequency distribution of biodiversity index of threatened species in Special Protection Areas.

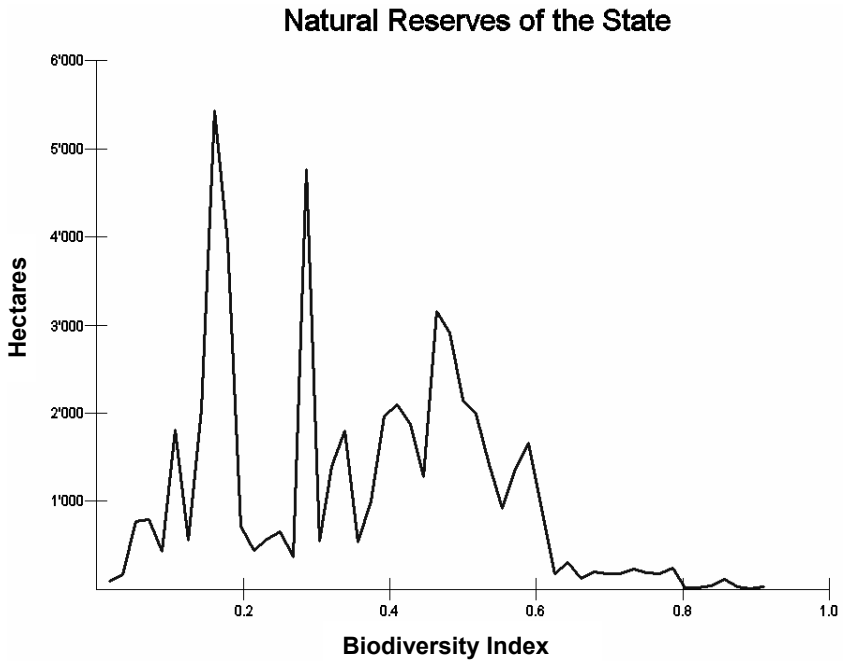


Figure 6.6 Frequency distribution of biodiversity index of threatened species in Natural Reserves of the State.

of protected areas. Note that the extent of the Y-axis is 1 order of magnitude lower for figures 6.3, 6.4 and 6.5 and 3 orders of magnitude lower for figure 6.6 when compared with the Y-axis in figure 6.1. While this change of scale allows for a better visual interpretation of the graphs, it should be kept in mind that the total area of the different types of protected areas is only a fraction of the overall surface of Italy.

The graphs show that both National Parks and Sites of Community Interest peak at a value of biodiversity index around 0.6. This value is close to the third peak in the overall distribution, and indicates an overall efficacy of these two types of protected areas, which seem to focus on areas of higher biodiversity value.

The frequency distribution of the Natural Reserves of the States is the one with the highest numbers of peaks, but overall the three basic peaks of the overall distribution can be observed. This seems to indicate a more limited efficacy of this type of protected area. Finally the Special Protection Areas show an intermediate situation. Though their frequency distribution shows the same three peaks of the overall distribution, the relative importance of the peaks is reversed, with a higher contribution of the higher biodiversity values.

The differences observed between the SPAs and the SCIs are due to the different aims of the two types of protected areas. SPAs are targeting the protection of bird species (European Union, 1979) while the SCI are established to protect mammal species.

The great majority of threatened mammal species live in areas of relatively pristine wilderness, which in general are also areas with high biodiversity value. On the other hand, some threatened bird species rely on agricultural areas, where the overall biodiversity is reduced. Thus some SPAs can be located in areas where they effectively achieve their conservation goal focused on birds, but at the same time they are less effective in the conservation of the overall biodiversity.

The distribution of the areas with biodiversity values within the upper 95th percentile (biodiversity index > 0.66) and that fall outside of the current protected areas network is shown in figure 6.7.

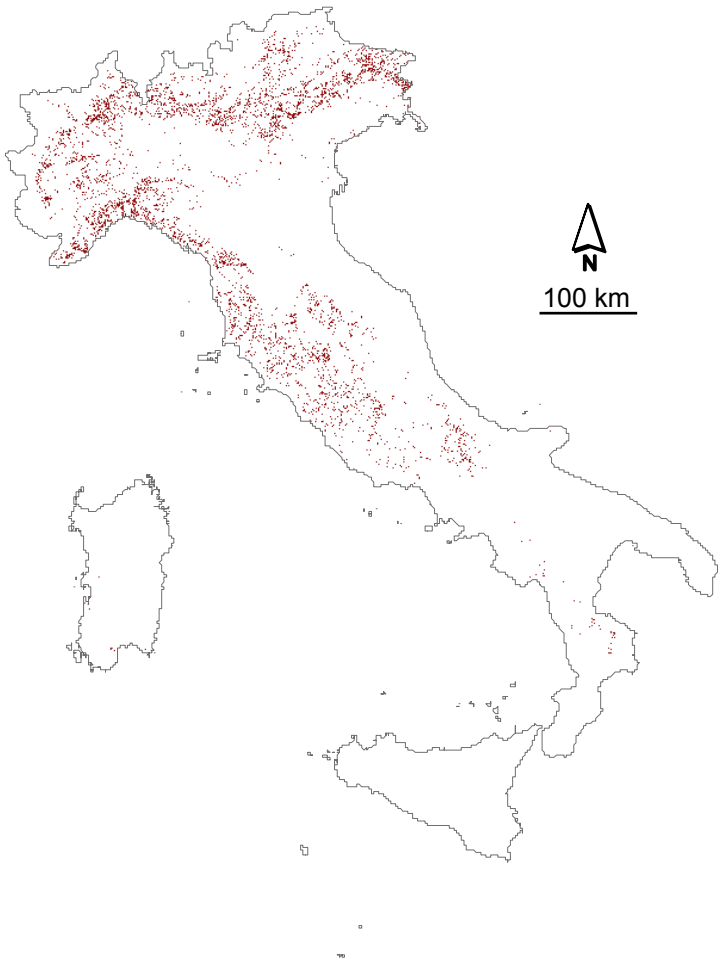


Fig. 6.7 Areas of high biodiversity value (biodiversity index in the 95th upper percentile) outside protected areas.

The high biodiversity value areas are distributed rather homogeneously in the central-northern part of the Italian peninsula within the higher biodiversity class of the map in figure 6.2. This is in accordance with the species richness gradient, which is known to decrease from North to South to the major islands (Sardinia and Sicily) (Massa, 1982; Stoch, 2000). While this fits well with the island biogeography theory (MacArthur and Wilson, 1967), it is important to note that the sampling effort follows a similar latitudinal gradient. Thus a lower biodiversity in the central and southern part of the Peninsula could indicate lack of knowledge (Stoch, 2003).

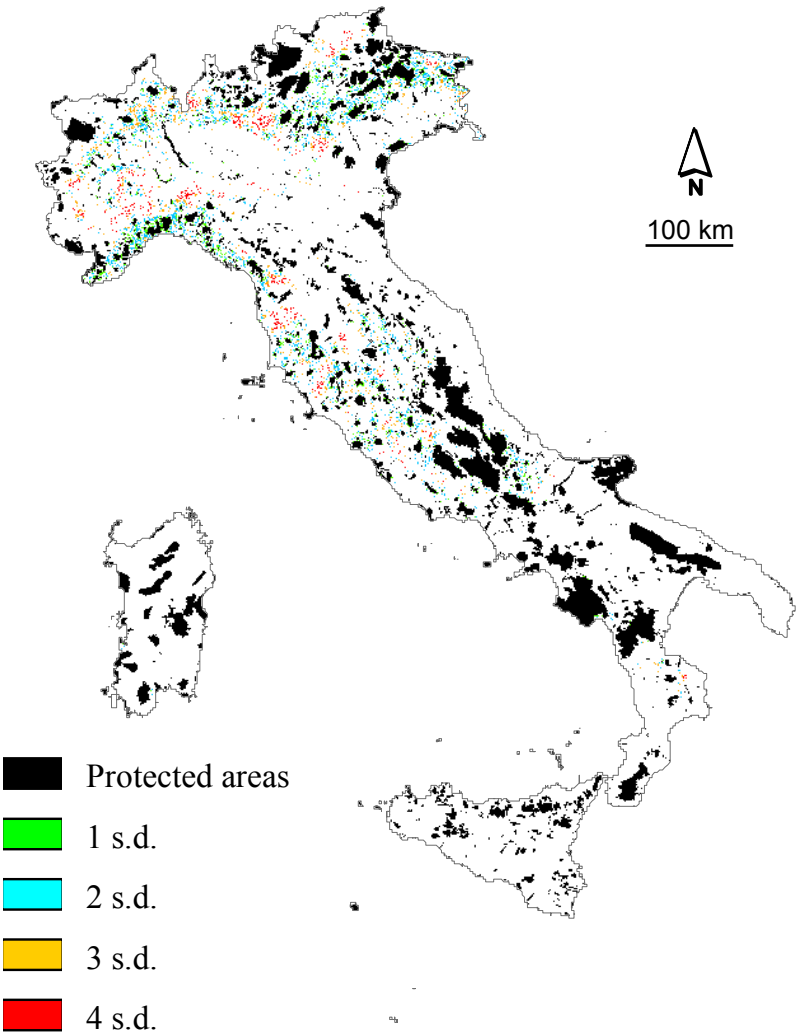


Figure 6.8 Distribution of the areas of higher biodiversity index values (higher then the 95th percentile of the frequency distribution), which fall outside of existing protected areas. The different colors indicate increasing distance classes from existing protected areas (shown in black in the figure). s.d. = standard deviation of the average patch distance.

The comparison of the patches of higher biodiversity value with the existing network of protected areas is shown in figure 6.8. To better appreciate the relationship of the various areas, high biodiversity value areas are colour-coded based on their distance from the nearest protected areas. Distances are grouped

in 4 classes of standard deviations. Assuming that closer areas will share similar biodiversity characteristics, higher distances are possible indicators of biodiversity underrepresented within the protected areas network.

The larger extent of the areas further from existing protected areas falls at the margins of the Po Valley, where the Alps and the Apennines begin to rise. This reflects the fact that most of the protected areas were created at higher altitudes, where the competition between nature conservation and human activities is lower.

CONCLUSIONS

In its basic simplicity, the analyses carried out for the Italian Ecological Network illustrate existing potential to enhance the extent and quality of information used to support decision-making (Fonseca *et al.*, 2000). The distribution models can be used to interpret species distributions (assuming that suitable habitats are the ones that are used), or environmental suitability models, in which case they contribute to defining the ecological network. Overlays of these models provide a predictive tool that advances us from single-species assessment to more comprehensive understanding of patterns of biodiversity. Relating these patterns to protected areas, as described above, allows us to assess the extent to which protected areas are effective in protecting biodiversity (Scott *et al.*, 1993). The potential utility, and power, of these overlays becomes clearer when one considers the possibility of integrating them, for example, with indices of threat (Root *et al.*, 2003) or areas proposed for development (Akçakaya *et al.*, 2003).

In general, the results of the overlap of models allow for more accurate identification of the centres of species diversity as compared with the gross results of Extent of Occurrence overlaps. It was beyond the aim of the study to analyse and discuss the details of the differences between the corresponding series of species overlaps (Extents of Occurrence and Areas of Occupancy). However, the patchiness of the present results suggests that the practice of using total distribution ranges or distribution grids to analyse the patterns of biodiversity may produce misleading conclusions. As many of the current biodiversity conservation efforts attempt to focus their actions on the areas of higher species concentration (Myers, 1988), accurate identification of these areas is of paramount importance.

In spite of the many limitations of the baseline data and the modelling technique, the order of magnitude of the results obtained provides a comforting picture of the conservation status of the Italian vertebrate species. As a whole the network of protected areas appears to adequately cover the distribution of biodiversity.

To this extent interpreting the overlays as joint suitability models for all of the species assessed (i.e. the areas where ecological conditions suitable for more than one species are met), adds another level to the management potential of expert-based species distribution models, such as the ones developed for the Italian Ecological Network. The patchiness and fragmentation of the summary models designs the backbone of a nation-wide ecological network for the vertebrate species. Using the conservation approach definition of ecological networks (Battisti, 2002; Gustafson & Gardener, 1996), these overlays can guide a more complete approach to the identification of core areas and linking corridors.

Chapter 7: Data Management for Biodiversity Conservation[22]

INTRODUCTION

Growing recognition of the extinction crisis and documented loss of biodiversity due to human action has led to a wide array of political and civil society responses (Bisby, 2000). Whether to initiate a sustainable community development project, monitor the status of biodiversity in a network of protected areas, assess the environmental impacts of a development project or implement the obligations of an environmental Convention, the scientific and political communities have great need for comprehensive, quality biodiversity information (Wilson, 2000).

Global Conventions, such as the Convention on Biological Diversity (CBD), require member States to integrate biodiversity considerations into decision-making processes inherent to the use of biological resources. Decisions concerning the use of biological resources within a country must take into consideration the impacts on sustainability for other countries. Global Conventions that address or are related to biodiversity conservation rely on qualitative information to implement and monitor stated objectives and obligate member states to comply with specific reporting requirements (Crain and Collins, 1999).

In response to these demands, a number of government initiatives have developed in recent years. For example, the CBD Clearing-house Mechanism (CHM) has as a primary objective to "develop a global mechanism for exchanging and integrating information on biodiversity". Others include large global generalist initiatives such as Eco-Portal, to more specialised ones like the Global Biodiversity Information Facility (GBIF), to national level initiatives and local e-communities that focus on specific biodiversity related topics or sites. Further evidence is found in a series of more localised networks such as the U.S. National Biological Information Infrastructure (NBII), the European

[22] Partially based on: F. Corsi, 2000. Presentation of the Species Information Service of the Species Survival Commission of IUCN The World Conservation Union. 2nd World Conservation Congress, October 2000. – Amman, Jordan.
and on: F.Corsi, G.Gadaleta, G.de Luise, 2002. IUCN - SSC Web Enabled Species Information Service: Scope, Objectives, Approach. Oracle Italia.

Network for Biodiversity Information (ENBI) and the ASEAN Regional Centre for Biodiversity Conservation (ARCBC).

Many of these initiatives "nest" within others, as illustrated by NBII, the U.S. segment within regional initiatives (the North American Biodiversity Information Network – NABIN, and the Inter-American Biodiversity Information Network – IABIN), that then build into the CHM and GBIF initiatives. These links, along with a growing number of NGO and University information networks, begin the process of forming an information "Commons" that connects a variety of types and sources of biodiversity information (Cotter and Bauldock, 2000). The networking approach recognises the need to manage data at the global level, in response to the natural circumstances (biodiversity and natural resources cross political borders) and political circumstances (policy is set and implemented at the national level).

Development of biodiversity data management infrastructures and the means for exchange and sharing amongst multiple systems face a number of challenges. These include inconsistent classification (taxonomic) systems, technical standards and historic data sets (Bowker, 2000).

This paper proposes a biodiversity data management infrastructure approach that addresses the challenges described above. The approach provides a system for managing historic, current and future biodiversity data with the consistency and flexibility required to serve as a basis for tracking changes over time. The system would support analysis of species status and distributions, comparison of these analyses to better understand species richness, and analysis of changes and trends over time. This spatial and temporal capacity, as shown in Chapters 4 to 6, provides a baseline to monitor status and trends of biodiversity, a need that has been clearly articulated in both the scientific and political communities.

The proposed infrastructure can support robust biodiversity analyses and monitoring when coupled with two other key elements:

1. Capacity to mobilise the wealth of existing but dispersed information in a systematic and quality-assured way (Chapter 2, *Mobilising Dispersed Information for Biodiversity Conservation*);
2. Application of proven and emerging analytical methodologies that maximise utility of limited data and produce relevant and timely biodiversity information products (see Corsi *et al.*, 1999, Boitani *et al.*, 1999, Corsi *et al.*, 2000).

Existing data, methods for mobilising that data and analytical techniques are sufficient to produce biodiversity analyses relevant to decision-making processes, as described in *Mobilising Dispersed Information for Biodiversity*

Conservation (Chapter 2), *Species distribution modelling with GIS* (Chapter 3), *Mobilising biodiversity data in practice: the example of the African Mammals Databank* (Chapter 4), *Modelling species distributions using inductive approach: the example of the wolf in Italy* (Chapter 5) and *Expert-based species distribution mpas for the assessment of the biodiversity conservation status: the Italian National Ecological Network* (Chapter 6). These examples illustrate the process of mobilising and analysing data for a single research purpose. When coupled with a data management infrastructure as described here, capacity to analyse biodiversity will be significantly enhanced in two ways. First, a source of consistent and current data could be made available for multiple analyses at a variety of scales, thus diminishing the high cost of data gathering and validation project by project. Second, a data infrastructure that ensures consistency across time and space can support the process of monitoring change over time, a capacity that is essential for current biodiversity management needs and obligations.

CHALLENGES OF BIODIVERSITY INFORMATION MANAGEMENT

Four key challenges in designing a suitable data structure, and solutions to address them, are discussed in detail below:

1. Taxonomic classification is the reference point for tracking individual pieces of information and, importantly, for linking information over time and space (Kim, 1993). Yet the science of taxonomy is imperfect: it harbours controversy over classification systems and changes significantly over time. *Therefore a biodiversity data management infrastructure must accommodate and manage the dynamic nature of taxonomy, guaranteeing at the same time data consistency for all of the information connected to each individual taxon.*

2. The spectrum of needs and expectations when dealing with biodiversity information are as varied as the number of people collecting or needing to use the information. There is no single recipe or accepted standard of what needs to be known to effectively support biodiversity decision-making. It is impossible to predict the types of information that will be required in the future. Furthermore, biodiversity decision making needs constant reference to the historical context of its evolution. Bowker (2000) describes a three-pronged challenge to the historical aspects of biodiversity information. The first relates to knowledge about biodiversity; its study demands knowledge about evolution. Second, accurate records of publication that go as far back as the mid-eighteenth century are essential especially in the field of taxonomy. Finally, biodiversity information is, by nature, reliant on data sets derived from a number of sources whose history and purpose must be understood to determine appropriate uses. *Thus a biodiversity information system should allow a flexible*

data structure which can accommodate new data requirements as they emerge, and define standards that support consistency over time and space.

3. Requiring inputs from a variety of experts also requires a clear basis for establishing collaboration agreements that support data exchange and sharing (BCIS, 2000c). These must be supported by rules for access to various aspects of the system, in terms of adding or changing information and in terms of viewing it. *Data exchange and sharing must be grounded in clear guidelines that articulate allowable data uses as well as limitations, and that clearly articulate ownership.*

4. Data quality is fundamental if an information system is to serve as an authoritative resource and produce analytical products in which scientists and decision-makers have confidence (BCIS, 2000d). *Information flow must support a system for validating data to ensure its quality and foster confidence in the system.*

The Dynamic Nature of Taxonomy

By nature information about biodiversity is a moving target. There is little agreement about the number of species that have been described and classified to date, or on the estimated number of species existing in the biosphere. For those organisms that have been described, there is also disagreement on the nature of the relationships between them, on where to place the "border" between different organisms, and even on the correct name to be used for each one. Further, names change amongst different taxonomic systems and as systems evolve, thus eroding the potential to consistently track information about individual taxa over time.

Biodiversity information systems should aim not only at managing current data, but also at providing the baseline for future monitoring and analyses of success stories and failures. The capacity to consistently relate information about a single taxon is therefore critical. However, while the current, most widely accepted systematic and taxonomic interpretation is what drives the scientific process as it relates to biodiversity conservation, that interpretation may change from day to day.

There are significant implications to accepting or rejecting specific systems of taxonomy. A species may not be threatened, while concurrently one or more of its sub-populations are on the verge of extinction. Different systematic interpretations may classify the sub-population as part of different species. Collection of biodiversity information is thus hampered by the lack of a clear reference point to which the information can be linked.

Controversy surrounding taxonomic classifications, and the resulting ambiguity of reference points, presents a significant challenge in the management of biodiversity information. This challenge is exacerbated by the inevitable inconsistencies arising from multiple entry points and inputs over time: scientific names are inconsistent due to taxonomic disagreements; misspellings are frequent and increase the number of names to be managed by an order of magnitude. This challenge is not simply a matter of developing new data structures to create consistency amongst thousands of related pieces. It requires capacity to preserve information as it becomes historic (the basis for monitoring change over time) while concurrently responding to taxonomic system changes and ensuring that historic and current information can be clearly linked. Failing to meet this particular challenge has serious implications in particular for the problem of data deterioration (see Chapter 2, *Mobilising Dispersed Information for Biodiversity Conservation*).

The challenge has been accepted at the political level through the recognition of the "taxonomic impediment" and the establishment of the Global Taxonomic Initiative (GTI) within the framework of the Convention on Biological Diversity (CBD, 2002a). The GTI aims to facilitate the establishment of adequate knowledge on taxonomy and foster data standards and exchange, providing guidance and economic means. A number of initiatives that aim to arrive at a globally accepted taxonomic standard have been initiated in recent years. Examples include the Integrated Taxonomic Information System (ITIS, 2003), Species 2000 (2003), the International Plant Name Index (IPNI, 1999) and the Index to Organism Names (ION) (BIOSIS, 2003). The main focus of these initiatives is complementary, yet they overlap and their formal channels of communication do not accommodate data exchange (Blum, 2000). The real shortcoming of these initiatives is, however, that they do not integrate the data management aspect with an adequate process to capture the dynamic nature of taxonomy.

Results of a Google search (URL: http://ww.google.com/ July 2003) illustrates the problem. A search was made on "mammals taxonomic databases". It returned 6,830 hits. By refining the search to databases on present day species (the "-palaeontology" option was used), the number of hits was reduced to 6,540. To assess the number of individual hits (e.g., the actual number of taxonomic databases which include mammals) a random sample of ten pages, each with ten links, was selected from the 683 pages returned by Google. For each one of these ten pages, the number of individual hits was counted (new databases not listed in pages already checked). After repeating the exercise ten times, an estimate of 1% to 5% independent databases was obtained. This leaves 65 to 325 databases that deal with the taxonomy of mammals.

While mammals are one of the two taxonomic groups for which the most knowledge is available (the other is birds), there is still a high number of unrelated databases on mammals (65 to 325). A rapid search of the first few hits reveals that the data structures, and the information itself, are also different. Comparing the genera under a very well known family such as "Canidae" (Table 7.1) in two of these databases, Mammals Species of the World-MSW at the National Museum of Natural History (NMNH) (Wilson and Reeder 1993) and the Integrated Taxonomic Information System (ITIS, 2003) the result is that the first has 14 genera while the second only 13 (the second one does not list *Alopex*, which it considers to be a synonym of *Vulpes*). Adding a third database (INNVISTA, 2003) gives yet another result, 14 genera like the NMNH, but listing *Fennecus* as an accepted genus (while the MSW lists it as a synonym of *Vulpes*) and not listing *Pseudalopex*. On the other hand, the Walker's Mammals of the World online (Novak, 1995) lists all of the above names, totalling 16 genera under the family Canidae.

NMNH-MSW	ITIS	INNVISTA	Walker's
Alopex		Alopex	Alopex
Atelocynus	Atelocynus	Atelocynus	Atelocynus
Canis	Canis	Canis	Canis
Cerdocyon	Cerdocyon	Cerdocyon	Cerdocyon
Chrysocyon	Chrysocyon	Chrysocyon	Chrysocyon
Cuon	Cuon	Cuon	Cuon
Dusicyon	Dusicyon	Dusicyon	Dusicyon
		Fennecus	Fennecus
			Lycalopex
Lycaon	Lycaon	Lycaon	Lycaon
Nyctereutes	Nyctereutes	Nyctereutes	Nyctereutes
Otocyon	Otocyon	Otocyon	Otocyon
Pseudalopex	Pseudalopex		Pseudalopex
Speothos	Speothos	Speothos	Speothos
Urocyon	Urocyon	Urocyon	Urocyon
Vulpes	Vulpes	Vulpes	Vulpes

Table 7.1. List of genera under the family Canidae in four taxonomic databases available on the Internet.

While the scientific debate on the systematics of the family Canidae will continue, an information system storing biodiversity data based on any one of the above taxonomies would encounter potential problems in matching data and exchanging it with other systems. Canidae is a very well known family. The problem increases exponentially when dealing with other groups of organisms for which we have less knowledge.

Taxonomic revisions, which "officially" change the systematics of groups of organisms, are not the only source of confusion. There is also significant confusion in the integration of taxonomies from different regions, and amongst countries that have contrasting views on the names and systematics of certain organisms. Databases without strict rules for spell checking and other means of information storage, for example plain paper data sets, add to the complexity of the system as they introduce misspelled entries. In a taxonomic data management system, data integrity can be guaranteed by treating misspelled entries as synonyms. Thus for each accepted entry in the taxonomic tree there will be a number of synonyms and misspelled entries that need to be mapped directly to it.

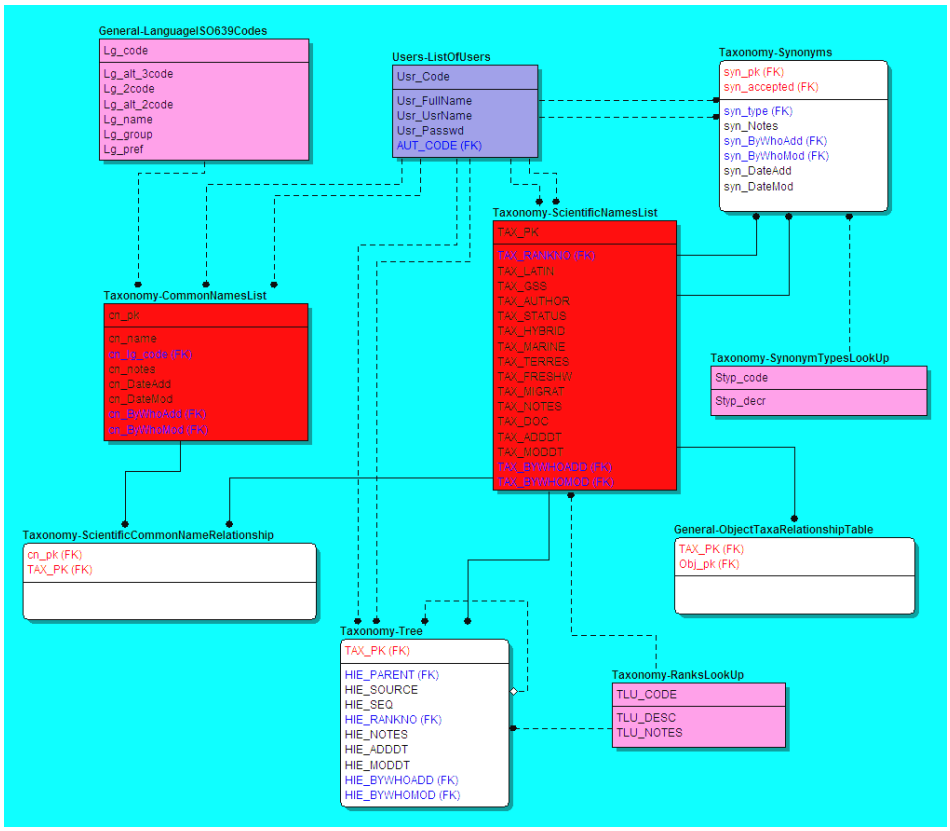


Figure 7.1 Entity-Relationship diagram of the proposed structure to manage taxonomy. The central role of the "taxonomic entities", here represented by record of the Taxonomy-ScientificNamesList table, is evidenced through the relationships with the other tables dealing with the systematics and data ownership.

The data structure proposed in figure 7.1 tackles the problem by allowing both storage of all terms used in the classification of known organisms, and their dependencies and relationships. Although it uses the Entity-Relational data model (as with the rest of the data structures illustrated in this paper), it could easily be translated into other models such as XML (Chen, 1999; Shuyun *et al.*, 2003).

The basic concept around which the data structure rotates is that of the "taxonomic entity". Taxonomic entities represent any term used for the classification of a living organism. To this extent, taxonomic entities are not only the specific epithet of a species but also the names of the families, of the orders and of any other taxonomic rank used with any branch of systematics. For instance, if we refer to the example of presented above, the family "Canidae" would be a taxonomic entity as well as the genera "*Alopex*", "*Atelocynus*", "*Canis*" etc. Furthermore, taxonomic entities are any synonyms or misspellings that have been used to identify a given organism and thus identify some data or pieces of information for that organism. As an example, if we were to follow the NMNH-MSW classification of the Canidae family, genera such as "*Fennecus*" and "*Lycalopex*" would not be listed as accepted scientific names (*sensu* ICZN, 2000). But they would still be part of the system as taxonomic entities that are synonyms of the accepted ones. Similarly "*Staphylococus*", "*Staphilococcus*" and *"Staphilococus"* are all misspellings of the accepted genus "*Staphylococcus*". All must then be included in the system as taxonomic entities that are misspellings of the accepted one.

Ideally, once a taxonomic entity is used in any form of publication, it should find its way within the system in the Taxonomy – ScientificNamesList table. This guarantees that a wealth of information published under the name (be it an accepted name, a synonym or a misspelling) is referenced in the system. For instance searching on Google with the term "*Staphilococcus*" returns 1980 hits, while searching on "*Staphylococus*" returned 1530 and only 235 searching on "*Staphilococus*". All this information would be lost to the system if the appropriate taxonomic entities were not referenced in the system.

The Taxonomy-ScientificNamesList is simply a list of names, with the basic attributes common to a taxonomic list such as Authority (which, in this system, also includes the year in which the taxonomic entity has been described) and the status of acceptance of the name. Although simple in its design, the structure is open to further atomisation of the data, which could include more complete descriptions of the taxonomic entities as proposed by the Task Group on Access to Biological Collection Data (ABCD, 2003). ABCD is a joint initiative of the Committee on Data for Science and Technology (CODATA) of the International Council of Science, and the International Working Group on Taxonomic Databases (TDWG). It is supported by both the Global Biodiversity

Information Facility (GBIF) and the project BioCASE: A Biodiversity Collection Access Service for Europe.

The structure allows management of the dynamic aspects of taxonomy through two distinct tables: Taxonomy–Tree and Taxonomy–Synonyms. These tables define respectively the access and visibility rules to the table of the taxonomic entities. The first one structures the hierarchy of the taxonomy, establishing the parent-child relationship (the systematics) between the taxonomic entities. The second one handles many-to-many relationships between the accepted name and all of its synonyms. Here the concept of synonymy is extended to include not only the usual taxonomic synonymies, but also the misspellings. To this extent, synonyms are qualified with a "type code" (field syn_type).

An important aspect of the proposed structure is the persistence of the taxonomic entity in the system; once an entity enters in the system it remains in it. It may be flagged as a synonym of another entity, but it cannot be removed. The importance is that it is unrealistic to change all of the references to the synonyms and/or misspellings in the existing data. As an example, consider the 3000 plus hits of misspelled references to the *Staphylococcus*. Thus by making entries permanent that information will still be fully accessible in the future.

Sub-dividing the data structure into static and dynamic components ensures data consistency in the system while allowing "users' views" on the taxonomy. Users' views are here defined as an alternative interpretation of the systematics and of the acceptance of scientific names. Users' views can be implemented easily by allowing multiple instances of both the Taxonomy-Tree and the Taxonomy-Synonym tables, in which users can create their personal "view" of the taxonomy. The users' views would still use the same "taxonomic entities" and codes that are stored into the Taxonomy – ScientificNamesList table, thus maintaining data consistency throughout the system.

A final aspect of the proposed structure is the management of common names, which allow a more general access to the individual taxonomic entities. Common names are stored in the table Taxonomy-CommonNamesList. Both the common name and the tongue in which it is used are stored in the table. Languages should be coded following the ISO 639-2 three-letter codes, which currently lists 507 major languages and dialects. The interesting aspect of using the ISO code is that it is an international standard, which implements precise rules to add, change and update entries (ISO 639/JAC, 2003). Given that one common name may refer to a group of species and that the same species can be identified with more than one common name, the relationship between the two tables is managed through an external relational table (Taxonomy-ScientificCommonNameRelationship).

By assigning each taxonomic entity a unique identifier that associates it to other types of data, data consistency over time is guaranteed. The approach serves as an historic track of the dynamic evolution of the taxonomy itself.

Evolving Data Needs: Consistency Across Time and Space

A common problem to all information systems is lack of data standards, clear data requirements and consistency over space and time. Biodiversity information systems suffer the same problem. They therefore must apply clear standards and allow a flexible data structure to accommodate new data requirements as they emerge.

An information system should define standards that are understandable to all users while it remains consistent in time and space. Time consistency means that as standards evolve, compatibility is maintained with data already in the system, while space consistency means that the standards should be independent from a specific location (that is, they are distinguished from views, language or standards of a specific region or country).

Effective broad-scale planning integrates information from different scientific disciplines that spans both time and space. For example, the process of developing a management plan for a specific species within a national park must consider specific information about the species both within the park and in the species' broader context (the meta-population structure of the species). This information might include the threats the species faces throughout its distribution range, its overall demographic trends, the historical knowledge of dispersion routes and range extension, the specific uses that are made of the species in different places and its economic value both within the local markets of the various places in which it is used and as a service provider in its own ecosystem.

Population parameters will likely have been collected by different zoologists in different parts of a species' distribution range. In some areas the historical data set may be extensive while in others there may be only records for a limited period of time. The economic importance of the species may have been analysed by a multinational corporation that is exploiting the species. At the local level anthropologists may have information on traditional uses. Ecologists might contribute information on the ecological requirements of the species, at least in some parts of its distribution range. The conservation status may be known from global and/or national red lists. Conservation activists may be monitoring the evolution of threats while denouncing new ones as they arise. Trade experts may contribute information on international agreements restricting the species import/exports.

Even if all of this information is readily accessible, the process of linking the pieces is a challenge. Standards for documenting the pieces of information will vary from source to source. Each scientific discipline has unique technical jargon, and even within the same discipline two experts may refer to the same thing with different terms.

A biodiversity data management infrastructure must therefore be flexible enough to accommodate inconsistent terminology and formats, and a wide range of types of information.

The results of a second Google query (June 2003) illustrates the variety of ways that different sources of biodiversity information are managed. The first four results of a Google query on "biodiversity information system" (including the quotes) were the Belize Biodiversity Information System (BBIS, 2001), the Israel Biodiversity Information System (BioGIS, 2002), rECOrd LRC - The Biodiversity Information System for Biological Information in the Cheshire region (rECOrd, 2003), and the Chinese Biodiversity Information System (CBIS, 2000).

The first site provides extensive accounts on the animal species found in Belize, with data as detailed as sonograms for individual bat species.

The second site provides observation points for plants and snails in Israel. For each species the description of its niche is based on a number of climatic parameters. Species niches are crossed with GIS layers of the same climatic parameters, providing predictive model maps.

The third site focuses on an inclusive data collection approach, that is anyone (including amateurs) willing to share is encouraged to contribute. The initiative is linked to the National Biodiversity Network, a private trust aimed at sharing biodiversity data throughout the United Kingdom. A form is provided to upload data on species observations. No query facility is provided online, but the site offers assistance through Email.

The fourth site provides the outline of a system in planning. Although it has not been updated since the beginning of the year 2000, it serves the purpose of describing the variability of data requirements as it lists the types of data that are, or should be, part of a comprehensive biodiversity data management infrastructure. These include specimen data, off-site conservation data, seed and geneplasm bank data and key ecosystem data. Listed requirements also include models to predict and monitor key species and ecosystems, evaluation of the current status of biodiversity in China, and more. A list of expert systems for

endangered species conservation, PVA of endangered species, reserve planning and other tools are also included.

To accommodate the data emerging from all of the above systems, and of the many more that have not been analysed in detail, a biodiversity data management system must be flexible. This can be achieved by creating an abstraction layer in the data structure that manages the primary relationship between the taxonomy and the associated information. Figure 7.2 shows the proposed structure.

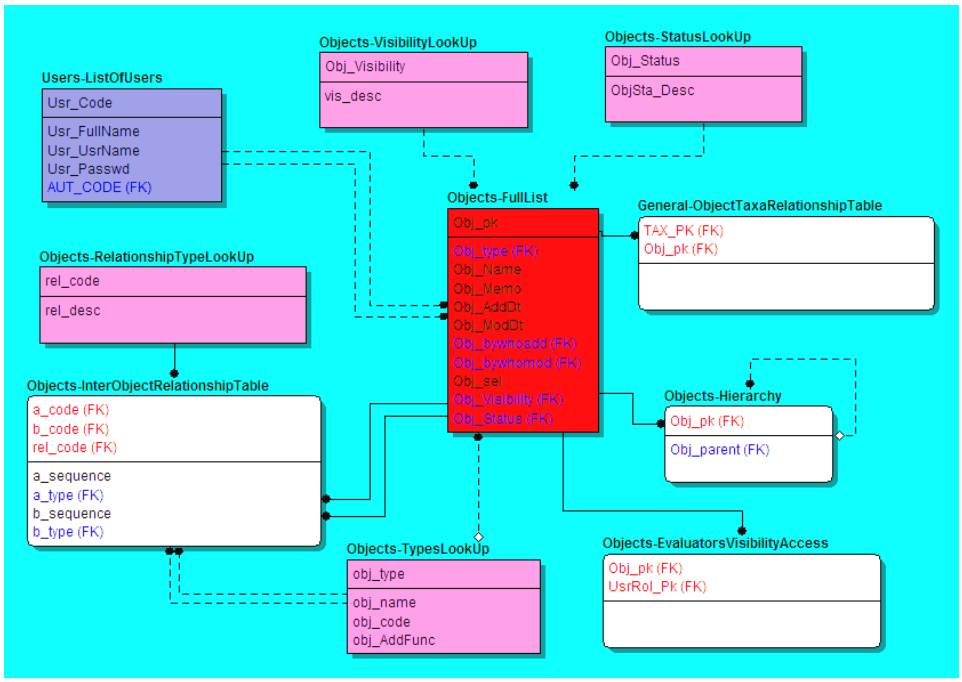


Figure 7.2 Entity-Relationship diagram of the proposed structure to manage the biodiversity information using the "information object" abstraction layer.

The central concept of this structure is the "information object". Information objects are defined as self-contained, homogeneous assemblages of data. To this extent, information objects can contain any type of data that is needed to support the conservation of a taxon.

Both the concept of information object and that of abstraction layer are derived from object-oriented database theory (Martin, 1993; Scheer and Smith, 1999), and have been proposed as the foundation of a biodiversity data management system (Srikanta and Haritsa, 2000). In this context they are used to explain not

only how the system manages different data types, but also how the system can accept new types of information that have not yet been considered.

Examples of information objects range from representation of the geographic ranges of species, to their habitat descriptions, to persons involved in the study and/or the management of the species. They can be in the form of structured database tables, word processor files, spreadsheets, images or GIS data. They can be composed of multiple atomised pieces of information or by one single figure. Information objects can even hold metadata describing other data sets (or data objects) (Gilliland-Swetland *et al.*, 2000) or define interoperability language to interface remote databases (Cleary *et al.*, 1998).

No matter what type of data the information object contains and how complicated it is to store and manage, the abstraction layer enables the user to create a relationship between it and the taxonomic entities.

The table Objects-FullList stores one entry for each information object in the system. Once an information object is created, it is assigned a unique identifier. The Objects-FullList table stores only the entry point to the actual data that are contained in the object. Each object type will then have a specific data structure to hold its relevant data. The link between the object entry point and its data is achieved through the propagation of the object identifier into the data structures devoted to the storage of the objects' data. This same identifier is also used as the link to the taxonomy through the General-ObjectTaxaRelationshipTable.

Maximum flexibility is achieved by allowing different objects to be related to each other through the Objects-InterObjectsRelationshipTable. For instance a geographic range of a species (an object within the system) might have been acquired from a publication; that same publication may also be stored in the system as a reference object. To fully document the origin of the geographic range object, this could be linked to the reference object of the publication from which it was derived. A relation code type (rel_code) enables the system to differentiate between types of relationships between objects. For example, the relationships between any object and a "person-object" could define an authorship relationship, or an evaluator relationship (a person who has contributed to the peer-commentary of the object).

Finally, objects are tagged with status and visibility codes. The status code defines the "processing" steps of the object, from the time it is input into the system by an expert, through to the peer review process and its release to the general public. The visibility code matches to the different type of users' roles of the system, and indicates which types of users have access to the information of the object. See the section that follows for further details.

Data Sharing Agreements

As seen in the previous section, building information about biodiversity requires input from a variety of experts. Participation in the common goal of creating the information required to support decision-making is fostered by clear data sharing agreements that guarantee the ownership of the original data. There are two main aspects to the issue: 1) identification and classification of the users who access the system; and 2) access policies.

The first aspect is a matter of implementing the data structures needed to build an authentication process. Users are provided with a user name and password, which uniquely identify each one of them to the system.

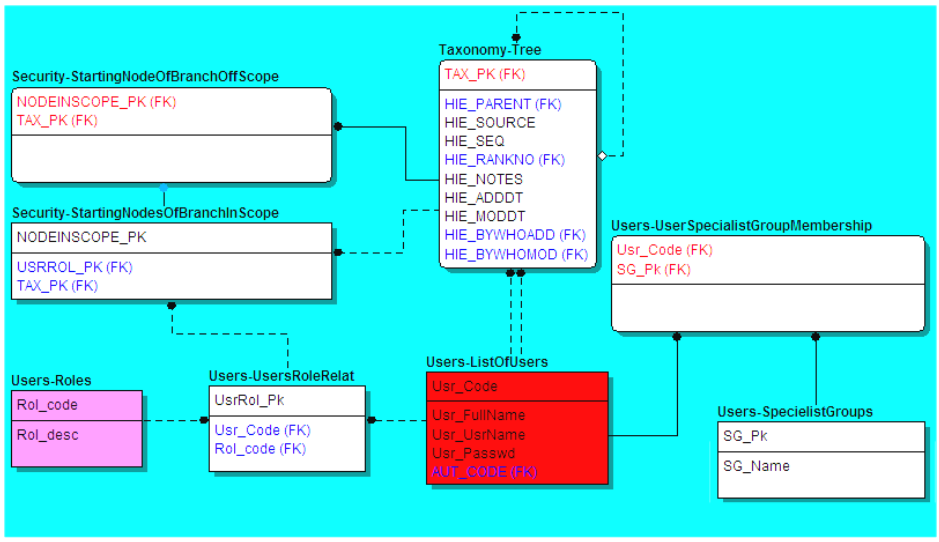


Figure 7.3 Entity Relationship diagram of the security layer that allows the implementation of the data sharing agreements. Central to the architecture are the two security tables (Security-StartingNodesOfBranchInScope and Security-StartingNodesOfBranchOffScope), which establish the relationship between the users and the authorised branches of the taxonomic tree.

From a practical point of view, this issue can be tackled by implementing a security layer that defines users' rights to access the data. This list of users is maintained in the Users – ListOfUsers table (Figure 7.3). To simplify system administration tasks, the authentication structure assigns users to groups created either on a taxonomic basis (groups of users that share the expertise on a given species or group of species) or on a thematic basis (groups of users that share the same interest on a specific topic related to biodiversity). The simplification of the administration tasks arises from the fact that once users are assigned to a

certain group they share the group's access rules. Groups are defined within the Users-SpecialistGroups table

Users have specific roles within the system. Roles define the rights that a user has in the segment of the system to which he or she has access. The following is a list of possible roles:

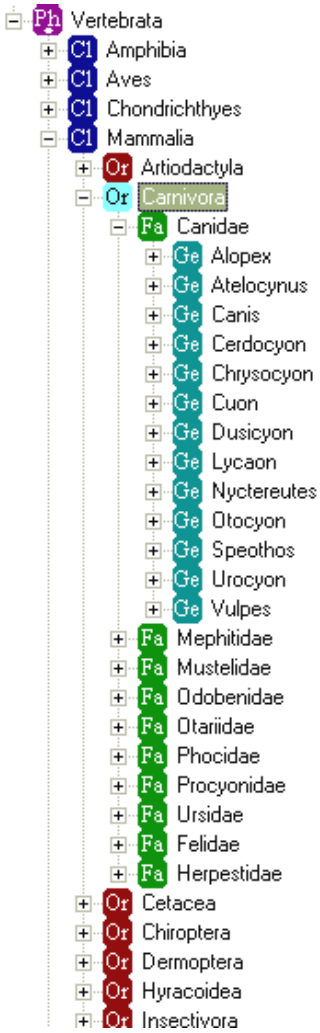
- **Administrator** – one or more. Ensure maintenance functions on the entire database. Does not need to have any specific authority on the scientific content of the system.
- **Scientific Coordinator** – one or more. Supervise the inclusion of the information into the system, coordinating the flow of the information from the experts through the peer-commentary process and to final publication. Can supervise portions of the data (as defined, for instance, by the grouping of the users) or the entire system (for example for the creation of information products that require cross-group data).
- **Expert** – any number. Contribute information into the system. Has full access rights to her/his contributions and starts the peer-commentary process by submitting the data to the relevant scientific coordinator.
- **Evaluator** – any number. Access individual information objects for peer-review.

Users can clearly have more than one role depending on the type of data they access. For example, a user may be a scientific coordinator for one group of species (one branch of the taxonomy), for which she/he has recognised standing; while at the same time serves as an expert contributing information for another group of species.

The structure accommodates this possibility by defining a unique user/role combination within the Users-UserRoleRelat table. The code assigned to the unique combination (UsrRol_PK) is then used to define the access rules and the rights of the user to segments of the system.

The data structure shown implements a way to define the taxonomic scope (species or group of species) for which the user has access and has a given role. The structure relies on the combined use of two tables: Security-StartingNodesOfBranchInScope; and Security-StartingNodesOfBranchOffScope. The first table stores the user/role code and the code of the starting node of a taxonomic branch to which the user/role combination is granted access. Given that within an authorised branch some segments may not be accessible to the specific user/role combination, an entry in the Security-StartingNodesOf-

BranchOffScope table indicates the starting node of a branch for which the user/role combination does not have access.



Users-UserRoleRelat		
User Code	Role Code	User\Role
32	3	345

Entry in the Users-UserRoleRelat table that identifies user 32 in the role of expert (role code 3). The user\role is assigned a unique code (345) by the system.

Security-StartingNodesOfBranchInScope		
User\Role	Taxonimic Entity	NodeInScope_PK
345	3423 (Carnivora)	654

Entry in the Security-StartingNodesOfBranchInScope table authorising user 32 to access, with the role code 3 (Expert), to all nodes in the branch that starts with node 3423 (Carnivora). The authorisation record is given a unique code (654) by the system.

Security-StartingNodesOfBranchOffScope	
Taxonimic Entity	NodeInScope_PK
3467 (Felidae)	654

Entry in the Security-StartingNodesOfBranchOffScope table excluding from the authorised nodes all the nodes of the branch that start at node 3467 (Felidae).

Figure 7.4 Example of authorisation scheme based on access to one branch of the taxonomy, with concurrent limitation of access to a sub-portion of it.

For example, in figure 7.4, a user having user code 32 could have the role of expert (role code 3) for the taxa in order Carnivora. The user/role combination is assigned a code by the system (in this case 345). The record in the Security-StartingNodesOfBranchInScope table would store 345 to identify the user/role combination and the code of the node Carnivora (3423). The system would take care of defining a unique code for this authorisation (NodeInScope_PK = 654). If the same user does not have expert rights on the taxa in the family Felidae (which would have been authorised based on the previous entry in the Security-

StartingNodeofBranchInScope table), a record is added in the Security-StartingNodesOfBranchOffScope table with the code NodeInScope_PK of the authorising record (654) and the code of the node Felidae (3467). The combination denies that user access to Felidae.

The access layer is completed with the visibility code introduced in the previous section. The visibility code is matched to the different type of users/roles of the system, and indicates which types of users have access to view the information of the object. Levels envisaged are the author, the scientific coordinator, the group of peers (members of the same group) and the general public.

Data Flow And Quality Assurance

A key aspect of the management of data is the way information flows from the data providers to the data consumers. Information flow must support the security layers described in the previous paragraphs and assure that users are confident in the veracity of sharing agreements. Information flow must also build into the process the data validation that ensures scientific quality of the information emerging from the system.

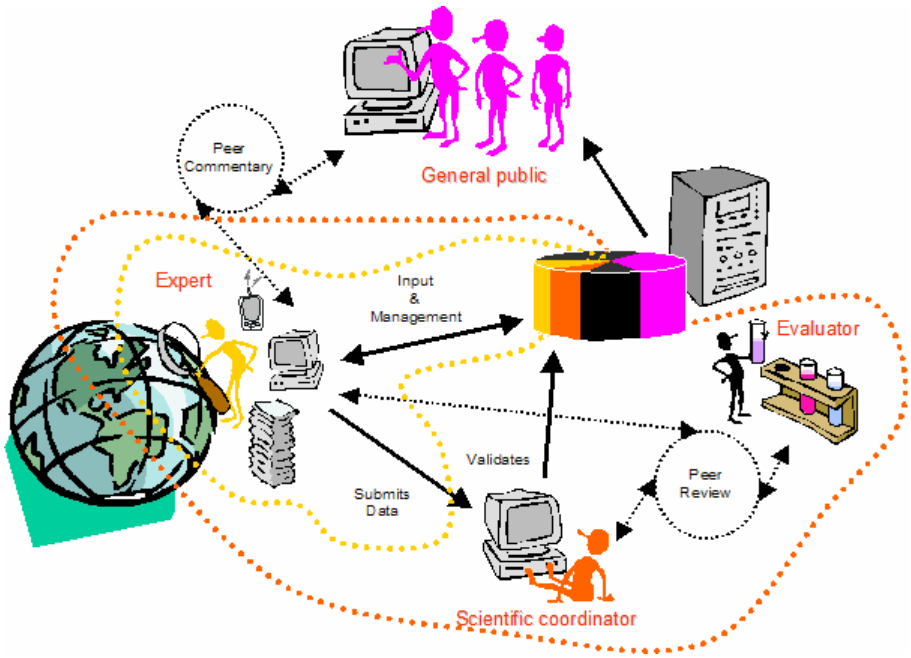


Figure 7.5 Data flow and data validation process. The dotted lines define the domains. The yellow one is the personal domain of the individual expert, while the orange one is the domain of the interest group. Once the data is validated it is made accessible to the general public.

The experts create the information objects that populate the system. Until they forward the objects for validation (i.e. initiate the peer-review process), the object is visible only to them. At that point, they still retain full management rights to the object. These include the right to delete the object as well as to change its contents. The area in the picture (Figure 7.5) outlined with the broken yellow line indicates the domain of the experts within the general system.

Once the object is ready to be made public, the expert forwards it to the scientific coordinator of the group to which she or he belongs. This starts the first level of validation within the group of experts. At this level of review, evaluators invited by the scientific coordinator of the group assess the object and suggest its publication or rejection within the system. Comments to the original object contribute to the definition of a final object that is ear-marked for publication.

To keep track of the peer-review process, all comments to the original object are held in the system. Similarly, objects for which the internal review process does not find an agreement for publication are also kept in the system and ear-marked as rejected. Both comments and/or rejected objects will be accessible only within the group of experts.

Publication, with the caveat of possible restriction due to the sensitive nature of data, makes the object available to the general public. At this level a second round of review occurs through the peer-commentary. Users of the data can send back into the system their comments and establish direct links to the original group or expert that published the object. Comments become an integral part of the original object. At any stage, the expert or the group can consolidate the comments into a new information object, which will update the original one.

Even once an object is updated, it is not deleted from the system. The most recent object will be the "current" one, but the others contribute to build the historic baseline for biodiversity monitoring, both in terms of our progress in conservation and in accumulating knowledge.

CONCLUSIONS

To date, biodiversity data management infrastructures have not accommodated inconsistencies and changes in taxonomy, data types and information needs over time. Those that aim to support current decision-making processes, for the most part still under development, are trying to address the need to accommodate and manage the dynamic nature of taxonomy and the broad array

of types of information and information formats that must be accommodated to create a fully inclusive and comprehensive system.

The data management infrastructure design described above responds to the challenges of managing information about a dynamic entity: biodiversity. It provides a key step forward for responding to the needs and demands for comprehensive, quality assured, spatially explicit and current biodiversity information. New approaches to solving the described challenged are introduced, along with technological solutions that support a system capable of addressing the major issues related to biodiversity data management.

The system introduces a new concept, the taxonomic entity, creating a consistent reference point in the taxonomy and guaranteeing a base reference point from which biodiversity information stays constant over time. At the same time, the system accommodates the dynamic aspects of systematics. By introducing the "information object" as an abstraction layer between the species and the actual biodiversity data, it allows for unlimited growth in the type and nature of the information collected.

The data management infrastructure also accommodates change over time, tracking pieces of data as they become historic. Data consistency in the historic data is guaranteed by their link to the static part of the taxonomy. The infrastructure presented also addresses the problem of access and security, which is fundamental for the correct functioning of an information network. Finally, the proposed data flow scheme addresses the problem of maintaining high standards of quality through a combination of peer review and peer commentary.

This chapter explains one approach to a system that can support an element fundamental to biodiversity monitoring: the capacity to track a sequence of events, or changes, as they relate to a place, a species or any element of biodiversity. Along with a comprehensive and active information network of contributing experts, the described data management infrastructure can underpin sophisticated biodiversity analyses and monitoring relevant not only to today's biodiversity conservation planning and management initiatives and policies, but also to the needs and circumstances that may emerge tomorrow.

An aspect that has not been covered in this paper is the actual link between the experts' network and the data management scheme: the user interface. This aspect can only be covered with an exhaustive user requirements review, to make sure that the users are comfortable with the way data needs to be entered and the way data can be retrieved. No pre-defined software tool will accommodate and satisfy all of the users' requirements, and it is likely that even user interfaces derived with the users' advice will still have a number of

functionalities that are only the most agreeable compromise between different needs and points of view. Nevertheless, failing to involve the users and delivering a user interface that frustrates them (either because it is too complicated or because it does not provide the functionalities needed) is a sure recipe for rapid failure of even the most powerful data management system.

Chapter 8: Applications of existing biodiversity information: capacity to support decision-making. Synthesis

INTRODUCTION

The demands for authoritative and current information about biological diversity and its status are varied and growing. Lack of access to and understanding about that information leads to lack of sound, scientific information in decision-making related to biodiversity conservation planning and management. While the gaps in data and understanding of biodiversity are significant, information and analytical methodologies available today are adequate to support decision-making:

- The information is dispersed but can be accessed with a well structured and managed information network supported through the electronic media;
- The true value of the information can be realised with application of analytical methodologies, in particular ecological modelling, which exploit existing technologies (e.g. GIS);
- Potential for information uses, including spatial and temporal analyses, can be maximised with a data management infrastructure that accommodates taxonomic and data inconsistencies, and evolving data needs and demands.

This thesis has shown that these three factors can be efficiently addressed with our current knowledge and the available technologies. The chapters present a series of approaches and examples that support an affirmative response to the question posed: *can we support biodiversity information needs of the scientific and political communities right now and in the immediate future?* Following is a review of the findings presented in Chapters 2-7, and an assessment of the relevance of the information and analytical capacities we have to the biodiversity information needs of today and into the future.

MOBILISING BIODIVERSITY INFORMATION: THE INFORMATION WE HAVE

The need for sound scientific guidance for biodiversity conservation and management actions has given rise to initiatives that focus on the improvement of knowledge, adding scientific data to the limited information available on biodiversity. Some try to fill the gap between current knowledge and the need for action, targeting two main aspects: improve information on biodiversity through an objective interpretation of the available data within simple frameworks, and clear documentation of all data used during the process.

In Chapter 2, *Mobilising Dispersed Information for Biodiversity Conservation,* the process for accessing dispersed information needed to support biodiversity conservation decision-making was described. When focusing on information about species, both in terms of pursuing species conservation and in terms of using species as a measure of the larger concept of biodiversity, there are several initiatives that are applying, at least partially, the process described. They are moving in the direction of linking networks of experts to collect and validate pieces of information that can increase our potential to manage and conserve biodiversity.

Initiatives similar to the approach highlighted in Chapter 2 rely on volunteer experts and consequently on the best environmental conditions to engage them. One such initiative is building its capacity to apply existing methodologies and technologies to maximise utility of limited biodiversity information. The IUCN Species Survival Commission (SSC) is a network of 7,000 species conservation experts organised into sub-networks focusing on specific groups of species. It is in the process of expanding its capacity to assess and monitor biodiversity through its Species Information Service (SIS). SIS aims to integrate into the human network the potentialities offered by modern IT, in line with what has been described in Chapter 7. The information mobilisation capabilities of SIS build around three major components: 1) the network of experts; 2) the Specialist Group sub-networks, which will serve as the primary data managers; 3) a Central Service Unit (CSU) to support network data gathering, management, and analysis of integrated data. One aim of SIS is to increase the SSC's capacity to gather dispersed biodiversity information. Two immediate steps are underway:

- Formalisation of partnerships to reduce duplication of effort and widen substantially the base of information sources: An example is the IUCN Red List Programme which has become a consortium of organisations that, historically, combined the networks of only BirdLife International and the SSC. That partnership has now grown to include NatureServe, Conservation International-Center

for Applied Biodiversity Science (CI-CABS) and The Ocean Conservancy.

- Systemisation of its assessment approach to focus on full groups of species: While all known mammal and bird species were assessed prior to publication of the 1996 IUCN Red List of Threatened Species, the current target is to complete Global Assessments of all amphibian species by the end of 2003[23] and reptile, freshwater fishe, shark, ray and chimera, and freshwater mollusc species by the end of 2004. Also in 2004, the re-assessment of all bird and mammal species will be completed. Similar targets exist for plant species in line with the Global Strategy for Plant Conservation approved at the 6th Conference of the Parties of the CBD (The Hague, 7-19 April 2002). The consortium aims to build understanding about the extent of the extinction crisis, while concurrently increasing knowledge about the patterns of change to biodiversity (IUCN-SSC, 2003a).

Another initiative that builds on an expert network to access and mobilise existing information is BioNET, the Global Network for Taxonomy (BioNet International, 2003). The aim of BioNET is to support the study of taxonomy and create the grounds for taxonomy as a viable profession. This is achieved through sensitising governments on the importance of taxonomy as a reference point of all our knowledge on biodiversity and by supporting capacity building in less developed countries. The network operates through sub-regional LOOPs (Locally Organised and Operated Partnerships; see, for example, http://safrinet.ecoport.org or http://eafrinet.ecoport.org or http://carinet.ecoport.org) of institutions in both developed and developing countries that are dedicated to making regions self-sufficient in their taxonomic needs. BioNET is actively working on overcoming the "taxonomic impediment" as defined by the Global Taxonomic Initiative of the CBD. It supports dialogue between the different LOOPS, and promotes definitions of standards that overcome local differences in taxonomic nomenclature. While this initiative does not tackle the issue of delivering standardised taxonomy to the conservation community (see Chapter 7 for a complete description of the issue), it does create a human network that can support taxonomic information needs that could then be made available through the technological innovations outlined in Chapter 7.

Other initiatives have taken a completely different approach from the mobilisation of information through networks of volunteer experts. In the year 2000 the ALL Species Foundation (2002) was launched. ALL Species aimed to

[23] Publication of these findings is scheduled for the fall of 2004.

develop an inventory of all species on Earth. Forty scientists and other professionals from around the world estimated that it would take 25 years and cost between $1 billion and $3 billion to complete the task. The scepticism of the scientific community has led to failure of the grand scheme of ALL Species; at the end of 2002 the Foundation office closed and currently it is hosted in the facilities of the California Academy of Sciences in San Francisco. However, the Foundation is still active and could certainly be revived should it be able to raise the necessary funds.

The ALL Species initiative is the first to move the challenge of biodiversity information gathering from the voluntary to the professional realm. Had it succeeded, this could have been an epochal change, as experts would have been paid to provide the information. This in turn would have guaranteed more general access to the information and reduced the proprietary issues that come with volunteer experts. Competing priorities would be reduced, and the timing of information flow would be more certain. Compensation for information is probably the single most important factor in stimulating its flow.

To a certain extent the Global Assessment approach, which is embedded in the SSC strategy for mobilising biodiversity information, takes this direction. Each global initiative is run as an individual project, and pays experts to compile the information required to complete the IUCN Red List assessment for each group of species.

Modelling species distribution: conveying the information

Mobilising dispersed biodiversity information would not be effective without providing tools to synthesise that information in ways that are concise and easy to understand. While models are simplified and a partial picture of reality, they are meant to provide interpretation keys that can help in the process of understanding natural phenomena.

Models presented in Chapter 5, *Modelling species distributions using an inductive approach: the example of the wolf in Italy*, as well as those derived within the framework of the African Mammals Databank (Chapter 4, *Mobilising biodiversity data in practice: the example of the African Mammals Databank*) provide a broad picture of the possible distribution of the species involved. They are meant to guide the decision making process in considering both local priorities and the context in which the local conditions are developing.

A measure of the usefulness of the approach is the extensive distribution of results of the African Mammals Databank, and the extent to which applications have been developed based on the methodologies outlined. The results were

published in a book distributed to more than 400 individuals and institutions throughout the world. More than 15,000 visitors accessed the Web site in the five years subsequent to publication. More than 300 links to the AMD web site exist through the World Wide Web.

The success of the AMD has prompted a number of NGOs to support its continuation and the inclusion of the remnant vertebrate species in the system. The African Vertebrates Databank (AVD) is being conducted as a joint effort of the IUCN SSC, CI-CABS, BirdLife International, the Museum of Natural History of London, and the Instituto di Ecologia Applicata in Rome. The same methodology also supports the Italian Ecological Network, a system recently developed for the Italian Ministry of the Environment (Boitani *et al.* 2002).

The model of the wolf in Italy (see Chapter 5 - *Modelling species distributions using an inductive approach: the example of the wolf in Italy*) has been used to assess areas of potential conflict between wildlife and human activities. Tuscany is an area recently re-colonised by the Italian wolf population. Without the presence of a major predator, livestock has been bred free-ranging, without protection from attacks. With the return of the wolves to the region, attacks and damage to livestock have become more frequent. In 1994, the Administration of Regione Toscana (Italy) approved a law on the payment of damages to livestock due to wolf attacks (Consiglio Regione Toscana, 1994). The law encourages breeders to adopt defence measures. Article 6 of the law introduces the "wolf map" (*carta del lupo*), which identifies areas for which wolf presence is confirmed. Within the areas identified, damages are paid only to the breeders who have adopted protection measures. The official version of the "carta del lupo" was largely derived from the analyses in Ciucci and Boitani (1996), which was to a significant extent based on the preliminary results of the research that was later published in Corsi *et al.* (1999).

The same methodology applied for the inductive analysis of the wolf in Italy has been also successfully applied to the other two species of large carnivores in the Alpine region: Lynx (*Lynx lynx*) and European brown bear (*Ursus arctos arctos*) (Corsi *et al.* 1998), and to the large herbivores: Ibex (*Capra ibex*), Alpine chamois (*Rupicapra rupicapra*), Apennine chamois (*Rupicapra pyrenaica ornate*), Red Deer (*Cervus elaphus*), Corsican Red Deer (*Cervus elaphus corsicanus*), roe deer (*Capreolus capreolus*) and the Mouflon (*Ovis gmelini musimon*) (Maiorano, 1998). At the request of the Council of Europe, the models for the wolf, lynx and bear were later integrated into a combined model to support the conservation strategy for large European carnivores (Corsi *et al.* 2002).

While the assumptions and limitations of GIS modelling approaches are discussed in Chapter 3, the overall use of GIS species distribution models is

well accepted in the scientific community (Scott *et al.* 2002). Its use to build informative and relevant syntheses from limited datasets is illustrated in the above examples.

To improve their performance, enhancements can be made to the models. Some of these are discussed later in the chapter. An important aspect to stress here is the need to advance the precision of the models, moving them beyond contextual information into the realm of verified information. This can be achieved through "ground-truthing" exercises, which can build confidence in the models and define their appropriate role within the decision-making process. See the African Mammals Databank (See Chapter 4 - *Mobilising biodiversity data in practice: the example of the African Mammals Databank*) for a description of the sampling approach, and Chapter 3 - *Species distribution modelling with GIS* for a general overview of the validation procedures.

Multi-species perspective: using limited information to assess biodiversity

Although single species management is important (IUCN-SSC, 2003c), it does not respond to the more holistic demand for sustainable management of natural resources. That demand calls for integration of the conservation requirements of multiple species, within the context of other natural resources. An example is the integrated analysis of the 149 species included in Chapter 6 - *Expert-based species distribution maps for the assessment of biodiversity conservation status: the Italian National Ecological Network*. Although based on simple overlays of the modelled Areas of Occupancy, which is interpreted as species richness, they provide grounds to prioritise interventions in certain areas (CABS, 1999; WWF, 2002). It is certainly true that species richness *per se* only represents one, although the most simple to understand, measure of biodiversity and that adequate conservation planning should take into account other measures such as the evenness of the population and the similarity of the species (Purvis and Hector, 2000). However it is also true that species richness is a concept easily understood by decision-makers and the public, and it has been the basis of successful prioritisation exercises for conservation (e.g. Myers *et al.* 2000).

The example of Chapter 6 also indicates that species richness can be successfully coupled with measures of conservations status (e.g. species threat category) to increase its use as a management tool (Lawler *et al.*, 2003). In Chapter 6 taking into account the threat category of the species allows for a reduction of the number of species needed to effectively describe species richness patterns. Reducing the number of species taken into account decreases the cost of acquiring information. At the same time, focusing on threatened species increases the chances of preserving that portion of biodiversity more directly threatened with extinction.

Integration of threat categories into species richness maps adds another level of information into biodiversity distribution models. Concepts such as the Multispecies Conservation Value (MCV) (Root, 2003) use the environmental suitability models coupled with the risk of extinction of each species to derive a unique model. MCV assigns higher values to areas that have high environmental suitability for highly threatened species. MCV has been successfully used in the California Gap Analysis Project (Root *et al.*, 2003).

Species distribution models, such as those presented in Chapters 4 to 6, also represent a useful input for complementarity analysis (Vane-Wright *et al.* 1991). Using species distribution models to quantify the "targeted natural features" (Pressey et al. 1996) or "conservation features" (Stewart *et al.*, 2003) in the land units analysed, the complementarity algorithms devise the most effective network of protected areas to maximise the protection of all biodiversity.

DATA MANAGEMENT INFRASTRUCTURES: MANAGING DISPERSED INFORMATION TO SUPPORT BIODIVERSITY ANALYSES

Comprehensive access to biodiversity information must be coupled with a data management infrastructure capable of managing historic, current and future biodiversity data with the consistency and flexibility required to serve as a basis for tracking changes over time. Such a system would support analysis of species status and distributions and allow comparisons of these analyses to better understand species richness. Examples of such analyses are found in Chapters 4-6. In addition to such analyses, the data management infrastructure described in Chapter 7, *Data Management for Biodiversity Conservation*, would provide the data management system needed to support analysis of changes and trends over time.

A number of government and civil society initiatives provide frameworks through which various NGO and government initiatives can be linked. These include the Convention on Biological Diversity Clearing-house Mechanism (CBD-CHM), the Global Biodiversity Information Facility (GBIF), a global government initiative, the Inter-American Biodiversity Information Network (IABIN), a regional government initiative, and the Biodiversity Conservation Information System (BCIS), a 10-member NGO initiative that includes IUCN Commissions, BirdLife International, Conservation International, NatureServe and others.

These initiatives build on top of information networks that aim to mobilise information; they provide a level of structure that aims to maximise the movement of data and information by capitalising on technology. While they

represent a giant leap in the mobilisation and accessibility of biodiversity information, there is still a gap between the "production" and the "distribution" of this information. By collating bits and pieces of the information provided by their constituency made of research organisation, museums, environmental agencies and NGOs, these initiatives do not attempt to either standardise the data and/or to create a framework for peer-review/peer-commentary quality assessments. These fundamental aspects of global biodiversity data sharing are left to the participating organisations, which seldom exploit current IT capabilities to their potential.

An example is the number of taxonomic initiatives that aim to provide standard access to biodiversity nomenclature, but continue to focus on providing checklists (e.g. the species of a certain country, the species of a certain family etc.) rather than aiming for standardisation and data consistency. This is clearly shown in the recent citation in Science (2004), where GBIF's recently launched taxonomic portal is depicted as the dream of all experts. While linking some 30+ repositories of taxonomic information, GBIF is still far from delivering the "taxonomic name service" which would allow standardisation of existing sources of biodiversity information and guarantee data consistency among biodiversity information systems. In a recent publication, Patterson (2003) argues that "what is missing is a unifying device to draw all these [taxonomic initiatives] together on the Internet, and the tools to use the indexing and organisational power of taxonomy". He then correctly focuses the attention of the scientific community on the need for a "taxonomic name server" as opposed to lose linkages of existing taxonomic databases.

The IUCN SSC again provides us with an example of a developing biodiversity data management infrastructure that meets many of the requirements described in Chapter 7. The SIS infrastructure, now in final stages of development, builds on taxonomic entities and information objects, aiming to achieve maximum flexibility and use of inconsistent data. Data is managed in a virtually distributed mode. The Specialist Groups manage their data establishing one or more data authorities that supervise the data flow, and manage the internal peer-review process for data quality assurance. Although retaining full ownership rights on their data, the Specialist Groups contribute to a centralised data pool that forms the core of the SSC information infrastructure. The Central Service Unit (CSU) coordinates the data gathering process and produces the analytical products, which are based on integration of part or all of the full data set.

Analytical products include those traditional to the SSC, such as the IUCN Red List, species conservation action plans and analyses of amendments to the CITES appendices, with a number of value-added aspects made possible by SIS. For example, a set of biodiversity indicators will be developed from the IUCN Red List that includes monitoring of the number of species threatened

with extinction, threats to biodiversity, and efficacy of conservation actions taken. In addition, new products aimed at monitoring biodiversity conservation, and performance in managing it, are envisioned. These will include, for example, spatial patterns of biodiversity, threats to biodiversity and overviews of trends and threatening processes. Once published, all of the information will be monitored for quality through a peer-commentary process. IUCN SSC is moving toward monitoring the status and trends in biodiversity over time.

From a technical point of view, a software tool supports Specialist Group management, allowing these sub-networks to manage data for their own focus on conserving the species about which they are experts. Concurrently Specialist Groups can share information to be included in broader biodiversity analyses. The system relies on Web information flow, thus ensuring rapid and efficient movement of data between Specialist Groups and the CSU.

Clearly, there are aspects that SIS cannot tackle by itself. One is the management of a global taxonomic service. The SSC could certainly have a catalytic function in such a service, but it is not positioned to lead the initiative. Although the SIS system is designed to accommodate the static and dynamic nature of taxonomy as described in Chapter 7, the development and coordination of such an effort is the task of a larger, dedicated agency. The most appropriate candidate is GBIF.

CURRENT APPLICATIONS: IS WHAT WE HAVE RELEVANT TO WHAT IS NEEDED?

From a scientific perspective, management and monitoring of biodiversity requires a baseline from which present circumstances can be assessed and future changes can be monitored. The baseline can be defined in a variety of ways and, as long as the information management system supports consistency across time, a variety of robust baselines might provide a starting point from which to measure change.

Existing information, methodologies and technologies support creation of one type of baseline: a series of models that predict species distribution patterns and provide a context within which conservation decisions can be made. As more species are analysed and their distributions integrated, the distribution patterns become increasingly representative of overall patterns of biodiversity *and* can provide increasingly robust analyses of those locations that support high levels of diversity.

The examples cited above illustrate potential for application of species distribution models to natural resource management. The SSC SIS (IUCN-

SSC, 2003b) provides an example of potential for moving this capacity to another level. SIS will provide three types of biodiversity information products:

- Baseline species data sets including species distribution maps (both expert-generated Extents of Occurrence maps and models based on species environmental suitability), population trends, ecological requirements (e.g., habitat preferences, altitudinal ranges), degree of threat (conservation status according to IUCN Red List Categories and Criteria), types of threat, conservation actions (taken and proposed) and key information on use;

- Biodiversity analyses including spatial patterns of biodiversity, threats to biodiversity and overviews of trends and threatening processes; and,

- Customised products, based on requests for analyses from governments, the private sector and NGOs.

SSC SIS is therefore aiming to provide a resource for information needs now and in the future, and will set one baseline from which to measure change over time.

From the political perspective, the benefit of increasing our understanding of species distribution patterns, and its concurrent benefit of building our knowledge about diversity and threat patterns, is to understand more clearly the potential impacts of human activities. More specifically, the political perspective can be explored in the context of the global Conventions. The information requirements of the Convention on Biological Diversity (CBD), the Convention on International Trade in Wild Species of Fauna and Flora (CITES) and the Ramsar Convention on Wetlands are examined here.

The CBD defines, in Article 6, the General Measures for Conservation and Sustainable Use:

"Each Contracting Party shall, in accordance with its particular conditions and capabilities:

(a) Develop national strategies, plans or programmes for the conservation and sustainable use of biological diversity or adapt for this purpose existing strategies, plans or programmes which shall reflect, inter alia, the measures set out in this Convention relevant to the Contracting Party concerned; and

(b) Integrate, as far as possible and as appropriate, the conservation and sustainable use of biological diversity into relevant sectoral or cross-sectoral plans, programmes and policies."

Article 7 further defines the elements of identifying and monitoring components of biodiversity:

"Each Contracting Party shall, as far as possible and as appropriate, in particular for the purposes of Articles 8 to 10:

(a) Identify components of biological diversity important for its conservation and sustainable use having regard to the indicative list of categories set down in Annex I;
(b) Monitor, through sampling and other techniques, the components of biological diversity identified pursuant to subparagraph (a) above, paying particular attention to those requiring urgent conservation measures and those which offer the greatest potential for sustainable use;
(c) Identify processes and categories of activities which have or are likely to have significant adverse impacts on the conservation and sustainable use of biological diversity, and monitor their effects through sampling and other techniques; and
(d) Maintain and organize, by any mechanism data, derived from identification and monitoring activities pursuant to subparagraphs (a), (b) and (c) above."

The description of national reporting provided on the CBD web site (www.biodiv.org) describes the requirements of Article 26 as follows: "The objective of national reporting, as specified in Article 26 of the Convention, is to provide information on measures taken for the implementation of the Convention and the effectiveness of these measures. The reporting process is therefore not intended to elicit information on the status and trends of biological diversity as such in the country concerned, except in so far as such information is relevant to the account of the implementation measures."

The CBD clarifies that States do not need to report about biodiversity, *per se*, but rather about the impacts of the measures they are taking to implement the CBD. *Impacts cannot be measured without a baseline and capacity to measure change over time.*

In ***the CITES Convention,*** Article III, paragraph 7 "requires each Party to submit an annual report on its CITES trade containing a summary of information on, *inter alia*, the number and type of permits and certificates granted, the States with which such trade occurred, the quantities and types of specimens and the names of species as included in Appendices I, II and III"

(CITES, 2003). This requirement does not stipulate reporting related to the status of species regulated within the framework of the CITES appendices.

However, to propose amendments to the Appendices, the Parties do need knowledge of a species' status and potential impacts of trade on that status. The IUCN SSC Wildlife Trade Programme publishes prior to each CITES Convention of the Parties (COP) *CITES: A Conservation Tool* (IUCN-SSC, 2003d). This manual describes the process for amending the Appendices and, in its most recent edition, explains the history of information requirements as they relate to this process. It states:

> "Although the Convention recognises both trade and biological reasons for determining whether a species should be listed on the Appendices, it provides little guidance on how to decide which species to list on the Appendices."
>
> "In 1994, to counteract these problems, the Parties adopted Resolution Conf. 9.24 which contains more objective, transparent criteria for categorising and listing species according to their risk of extinction….The Resolution Conf. 9.24 criteria rely on the fact that the risk of extinction can be gauged from information on the status and trends in species population size and their distributions. The measures in Resolution Conf. 9.24 criteria include estimates of population size, area of distribution, observed or predicted rates of population decline or habitat loss and a variety of combinations of these." (section 1.3 New Approaches to the Classification of Species into the CITES Appendices)

While national reporting does not call for measures or monitoring of species subject to trade, the fundamental workings of the Convention do. *To achieve their individual aims for the Convention, it is in the best interest of the Parties to have a clear understanding of the status of species in question.*

The ***Ramsar Convention on Wetlands*** focuses its information requirements on the status of wetlands. At issue in the first instance are the criteria for naming Ramsar sites. The text of the Convention (Article 2.2) states that: "Wetlands should be selected for the List [of Wetlands of International Importance] on account of their international significance in terms of ecology, botany, zoology, limnology or hydrology" and indicates that "in the first instance, wetlands of international importance to waterfowl at any season should be included".

The criteria require understanding of a number of ecological factors, with focus on "sites containing representative, rare or unique wetland types and sites of international importance for conserving biological diversity".

There are eight criteria, including for example, if the wetland supports: vulnerable, endangered, or critically endangered species or threatened ecological communities; 20,000 or more waterbirds; 1% of the individuals in a population of one species or subspecies of waterbird; a significant proportion of indigenous fish subspecies, species or families, life-history stages, species interactions and/or populations that are representative of wetland benefits and/or values and thereby contributes to global biological diversity. See Ramsar Convention (2001a) for a full list of the criteria.

The Framework for Wetland Inventory outlines the elements and approaches to carrying out wetlands inventories. The introductory text states: "The *Global Review of Wetland Resources and Priorities for Wetland Inventory (GRoWI)*, prepared in 1999 for the Ramsar Convention by Wetlands International and the Environmental Research Institute of the Supervising Scientist, Australia, indicated that few countries have comprehensive national inventories of their wetland resources, and lack this essential baseline information on their wetlands. In addition, the National Reports submitted to Ramsar COP8 indicated that insufficient progress has been made in wetland inventory" (Ramsar Convention, 2002).

The Framework proceeds to describe the purposes of wetland inventories. These include: listing wetlands of local, national and/or international importance; describing the occurrence and distribution of wetland taxa; establishing a baseline for measuring change in the ecological character of wetlands; assessing the extent and rate of wetland loss or degradation. *To fulfil these information requirements, the Parties to Ramsar must have at their disposal information about the conditions and taxa contained within wetlands, and the capacity to monitor any changes to those conditions.*

For these three Conventions, the species distribution models and assessment of patterns of diversity/threat described are a valuable and relevant beginning for member States to understand the location and status of biodiversity. They provide, for example, critical baseline information for CBD Parties to "*identify components of biological diversity important for its conservation and sustainable use*", for CITES Parties to apply the criterion for Resolution Conf. 9.24, and for Ramsar Parties to nominate and inventory wetlands of international importance.

CBD (2002b) lists 146 biodiversity indicators ranging from ecosystems to genes. About half of the indicators have a reference to the spatial (either

explicitly or implicitly) dimension; of this 7 explicitly reference species distribution and/or ranges, while another 2 reference "species richness". The CITES list criteria amended at the 12th COP (Santiago del Chile, 2002), explicitly reference fragmentation and range size as criteria for the inclusion of species in Appendix I and II. In the 4th European Regional Meeting on the Ramsar Convention (Ramsar Convention, 2001b) there is explicit reference to the use of GIS to investigate distribution ranges of species to drive the selection process of Ramsar sites.

As tools for priority setting exercises, the species distribution models have proven their use (Fonseca et al,. 2000). For example, the AMD models and analyses have been used in the Upper Guinea Forest Ecosystem of West Africa priority setting workshop (Elmina - Ghana, December 1999) (CABS, 1999) and in the workshop "Biological Priorities for Conservation in the Guinean-Congolian Forest and Freshwater Region" (Libreville - Gabon, 2000) (WWF, 2002), to support the decision making process of the experts involved.

The models also offer a chance to analyse the local situation, generally driven by political opportunity, within the broader ecological context of the species. They can be used as a reference point to ensure sustainable use of the environment, without encroaching on the neighbours' capacity to use their share.

Building the models serves also to highlight knowledge gaps and inconsistencies and to focus the attention of decision-makers on those gaps. Low performance models could well result from problems of the analysis, but can also suggest that a species is poorly known. The result of the models thus represents an extrapolation of our current knowledge to the entire study area.

Building a model, however, is never a final result. Models allow us to integrate new knowledge as it becomes available and by doing so they can provide control points not only on the status of biodiversity but also on how much more we know about it.

Is what we have relevant to what is needed? The answer is yes, especially in the context of the environmental Conventions, which call for spatial and temporal information about biodiversity to implement objectives and measure effectiveness.

CONCLUSIONS

This thesis promotes the premise that sufficient biodiversity information exists to be mobilised, quality assured and analysed in support of the needs and

demands posed by the scientific and political communities. As a direct consequence of mobilising the available knowledge, it also becomes clearer where the knowledge gaps are found. While we continue to accumulate knowledge on the lesser known aspects of biodiversity, the information that is already available can drive decision-making related to biodiversity planning and management.

Sufficient data and expertise exist, although it is dispersed and can be difficult to access. Structured with a supporting environment and incentives, information networks supported by electronic technologies can provide the access needed to existing and new information on a real-time basis. Access to data and expertise can enhance the value of technologies such as GIS, from which species distribution models can be derived. The models provide us with a context needed to develop management scenarios for threatened species and analyse distribution patterns of biodiversity. While there are limitations to the application of GIS (see Chapter 3 - *Species distribution modelling with GIS*), modelling projects described in chapters 4-6 illustrate a variety of uses relevant to current demand for information. The African Mammals Databank offers a broad perspective into vertebrate distribution patterns across an entire continent while A Large-Scale Model of Wolf Distribution in Italy for Conservation Planning illustrates the utility to management planning of single-species distribution modelling.

The Italian Ecological Network shows how the individual species distribution models can be integrated to assess the conservation status of the overall biodiversity. This type of analysis answers the question posed at the beginning of this synthesis: *can we support biodiversity information needs of the scientific and political communities right now and in the immediate future?* Integration of species distribution models, which already consider a number of environmental variables, and the possibility of including in them patterns of threat, suggest that we do. We have the capacity to draw on "real-time" information, develop a variety of models based on this information, explore their veracity, integrate them to better understand biodiversity patterns and establish a cooperative system that links scientific and political communities more directly to the benefit of both.

Finally, data management infrastructures that accommodate inconsistent and evolving data needs, and integrate a robust quality control element, provide us with the means to manage, synthesise and analyse information with consistency across space and time. Application of such infrastructures is now taking form through several global and regional initiatives. It is this step that could set in motion true monitoring capability, providing a baseline to monitor status and trends of biodiversity. Responding to the need that has been clearly articulated

in both the scientific and political communities may be the key to unlocking the resources and prioritisation needed to realise the potential.

References

ABCD, 2003 Task Group on Access to Biological Collection Data (ABCD). URL: http://www.bgbm.fu-berlin.de/TDWG/CODATA/default.htm (accessed 24 June 2003)

Adomeit, E. M., D. L. B. Jupp, C. I. Margules, and K. K. Mayo, 1981. The separation of traditionally mapped land cover classes by Landsat data. In: Gillison, A. N., and D. J. Anderson (Eds.), Vegetation classification in Australia: Canberra, ANU Press, p. 150-165.

Agee, J. K., S.C.F. Stitt, M. Nyquist and R. Root, 1989. A geographic analysis of historical grizzly bear sightings in the North Cascades. Photogrammetric Engineering & Remote Sensing 55(11): 1637-1642.

Akçakaya, H. R., 1993. RAMAS/GIS: linking landscape data with population viability analysis. Applied Biomathemathics, New York, NY.

Akçakaya, H. R., M. A. McCarthy, J. L. Pearce, 1995. Linking landscape data with population viability analysis: management options for the helmeted honeyeater *Lichenostomus melanops cassidix*. Biological Conservation, 73:169-176.

Akçakaya, H. R., V. C. Radeloff, D. J. Mladenoff, and H. S. He. 2003. Integrating landscape and metapopulation modeling approaches: viability of the sharp-tailed grouse in a dynamic landscape. Conservation Biology [in press].

Aldred, B. K., 1972. Point-in-polygon algorithms. IBM Peterlee UK.

Allen, A. W., P. A. Jordan, and J. W. Terrell, 1987. Habitat suitability index models: moose, Lake Superior region. US Fish and Wildlife Service Biological Reports 82(10.155) 47pp

ALLSpecies Fundation, 2002. Home Site. URL: http://www.all-species.org/ (accessed: 30 August 2003)

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer, 1976. A land use and land cover classification system for use with remote sensor data, U.S.G.S. Washington.

Aspinall, R., 1992. An inductive modelling procedure based on Bayes' theorem for analysis of pattern in spatial data. International Journal of Geographical Information Systems. 6(2):105-121

Aspinall, R., K. Matthews, 1994. Climate change impact on distribution and abundance of wildlife species: an analytical approach using GIS. Environmental Pollution 86: 217-223.

Aspinall, R., N. Veitch, 1993 Habitat mapping from satellite imagery and wildlife survey data using a bayesian modelling procedure in a GIS. Photogrammetric Engineering & Remote Sensing 59(4):537-543

Atkinson, S. F., 1985. Habitat based methods for biological impact assessment. The Environmental Professional 7: 265-282.

Bailey, R. G., 1988. Problems with using overlay mapping for planning and their implications for Geographic Information Systems. Environmental Management 12(1):11-17.

Baird, D. R., 1996. Predicting East African ungulate diversity from satellite sensor imagery. Master of Research in the Natural Environment. University of Edinburgh, Edinburgh, U.K.

Barnes, R. D., 1989. Diversity of Organisms: How much do we know? American Zoologist, 29, 1075-1084.

Battisti, C., 2002. Reti ecologiche e fauna: approcci strutturali, funzionali, gestionali – un contributo metodologico. Parchi - in press.

BBIS, 2001. Belize Biodiversity Information System. URL: http://fwie.fw.vt.edu/wcs/ (accessed on the 5 June 2003)

BCIS, 2000a. Framework for Information Sharing: Principals. Busby, J.R. (Series Editor). Available from Program Manager, BCIS [contact details on http://www.biodiversity.org]

BCIS, 2000b. Framework for Information Sharing: Data Custodianship. Busby, J.R. (Series Editor). Available from Program Manager, BCIS [contact details on http://www.biodiversity.org]

BCIS, 2000c. Framework for Information Sharing: Data Access. Busby, J.R. (Series Editor). Available from Program Manager, BCIS [contact details on http://www.biodiversity.org]

BCIS, 2000d. Framework for Information Sharing: Standards & Quality Assurance. Busby, J.R. (Series Editor). Available from Program Manager, BCIS [contact details on http://www.biodiversity.org]

Begon, M., J. L. Harper, and C.R. Townsend, 1990. Ecology: individuals, populations and communities. Blackwell, Cambridge, 945 p.

Bell, A., 1986. Satellite mapping of the Great Barrier Reef. Ecos, p. 12-15.

Bennett, M. D., 1998. Plant genome values: How much do we know? Proc. Natl. Acad. Sci. USA. 95:2011–2016

Berry, B. J. L., and A. M. Baker, 1968. Geographic sampling, in Berry, B.J.L., and Marble, D.F., eds., Spatial analysis: a reader in statistical geography: New Jersey, Prentice-Hall.

BioGIS, 2002. Israel Biodiversity Information System. URL: http://www.biogis.huji.ac.il/ (accessed on the 5 June 2003)

BioNet International, 2003. BioNet International, The Global Network for Taxonomy. Home page. URL: http://www.bionet-intl.org/ (accessed: 30 August 2003)

BIOSIS, 2003 Index to Organism Names. URL: http://www.biosis.org/free_resources/ion.html (24 June 2003)

Bisby, F.A., 2000 The Quiet Revolution: Biodiversity Informatics and the Internet. Science 289: 2309-2312.

Bishop, Y. M., S. E. Fienburg, and P. W. Holland, 1975. Discrete Multivariate Analysis: Theory and Practice: Cambridge Massachucetts, MIT Press.

Blum S. D., 2000. An Overview of Biodiversity Informatics URL: http://www.calacademy.org/research/informatics/sblum/pub/biodiv_informatics.html (accessed 13 February 2003)

Boggs, C. L. and J. Roughgarden, 2000. Biodiversity: Result of speciation and extinction: In W.G. Ernst (Ed.), Earth Systems: Processes and Issues: Cambridge University Press, Cambridge, UK, p.255-266.

Boitani L., F. Corsi, A. Falcucci, L. Maiorano, I. Marzetti, M. Masi, A. Montemaggiori, D. Ottaviani, G. Reggiani, C. Rondinini, 2002. Rete Ecologica Nazionale. Un approccio alla conservazione dei vertebrati italiani. Università di Roma "La Sapienza", Dipartimento di Biologia Animale e dell'Uomo; Ministero dell'Ambiente, Direzione per la Conservazione della Natura; Istituto di Ecologia Applicata.

Boitani L., F.Corsi, A. De Biase, I. D'Inzillo Carranza, M. Ravagli, G. Reggiani, I. Sinibaldi, P. Trapanese, 1999. AMD African Mammals Databank - A Databank for the Conservation and Management of the African Mammals. Istituto di Ecologia Applicata - Rome, Italy ISBN 88-87736-00-6.

Boitani, L., 1982. Wolf management in intensively used areas of Italy. Pages 158-172 in F. H. Harrington and P. C. Paquet, editors. Wolves of the world: perspectives of behaviour, ecology and conservation. Noyes Publishing Co., New Jersey.

Boitani, L., 1986. Dalla parte del lupo. Giorgio Mondadori, Milano, Italy.

Boitani, L., 1992. Wolf research and conservation in Italy. Biological Conservation, 61: 125-132.

Boitani, L., and M. L. Fabbri., 1983. Strategia nazionale di conservazione per il lupo (*Canis lupus*). Ricerche Biologia Selvaggina (Bologna, Italy) 72:1-31.

Boitani, L., and P. Ciucci., 1993. Wolves in Italy: critical issues for their conservation. Pages 74-90 in C. Promberger and W. Schröder, editors. Wolves in Europe: Status and perspectives. Munich Wildlife Society, Ettal, Germany.

Bouwma, I. M., R. H. G Jongman & A. Pickaver, 2002. Theoretical background and experiences with designing ecological networks. In: I. M.Bouwma, R. H. G. Jongman & R. O. Butovsky (Eds.) Indicative Map of the Pan-European Ecological Network for Central and Eastern Europe. ECNC, Tilburg, the Netherlands. 21-33

Bowker, G. C., 2000. Work and Information Practices in the Sciences of Biodiversity. VLDB 2000: 693-696

Bozek, M. A. and F. J. Rahel, 1992. Generality of microhabitat suitability models for young Colorado River Trout (*Oncorhynchus clarki pleuriticus*) across site and among years in Wyoming streams. Canadian Journal of Aquatic Science 49:552-564

Breininger, D. R., M. J. Provancha, and R. B. Smith, 1991. Mapping Florida Scrub Jay habitat for purposes of land-use management. Photogrammetric Engineering & Remote Sensing 57(11): 1467-1474.

Brown, J. H., G. C. Stevens, D. M. Kaufman, 1996. The Geographic Range: Size, Shape, Boundaries, and Internal Structure. Annual Review of Ecology and Systematics 27: 597-623.

Buckland, S. T. and D. A. Elston, 1993. Empirical models for the spatial distribution of wildlife. Journal of Applied Ecology 30: 478-495

Bulgarini, F., E. Calvario, F. Fraticelli, F. Petretti, and S. Sarrocco (Eds) 1998. Libro Rosso degli Animali d'Italia – Vertebrati. WWF Italia. Rome, Italy.

Bunnell, F. L., 1989. Alchemy and uncertainty: what good are models?. Gentile. Tech. Rep. PNW-GTR-232. Portland, OR: U.S: Dept. of Agriculture, Forest Service, Pacific Northwest Research Station. 27p.

Burrough, G. H., 1986. Principles of Geographical Information Systems for land resources assessment. Clarendon Press

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

Butler, M. J. A., C. LeBlanc, J. A. Belbin, J. L. MacNeill, 1986. Marine resources mapping: an introductory manual. FAO Fisheries Technical Paper (274): pp 256

CABS, 1999. Conservation Priority-Setting Workshop. From the Forest to the Sea. Biodiversity Connections form Guinea to Togo. URL: http://www.biodiversityscience.org/priority_outcomes/west_africa/index.html (accessed 31 August 2003)

CABS, 2003. The Conservation Priorities Setting Program. URL: http://www.biodiversityscience.org/xp/CABS/research/global_planning/priority_setting/priorities.xml (accessed 20 April 2003).

Caicco, S. L., J. M. Scott, B. Butterfield and B.Csuti, 1995. A gap analysis of the management status of the vegetation of Idaho (USA). Conservation Biology 9:498-511.

CBD, 2001. Report Of The Brainstorming Meeting On Scientific Assessment. Development of methodologies and identification of pilot studies. SBSTTA OSLO, 17-19 NOVEMBER 1999

CBD, 2002a. Global Taxonomy Initiative – Background. URL: http://www.biodiv.org/programmes/cross-cutting/taxonomy/ (accessed 24 June 2003)

CBD, 2002b. Questionnaire on Indicators for Biological Diversity. Secretariat. URL: http://www.biodiversite-sbstta.org/IMG/pdf/doc-28.pdf (accessed 31 August 2003)

CBIS, 2000. Chinese Biodiversity Information System. URL: http://cbis.brim.ac.cn/cbise/ (accessed 5 June 2003)

CEPF, 2003. Frequently Asked Questions - Regional Strategy Development. URL: http://www.cepf.net/xp/cepf/home/faq/faq_strategy.xml (accessed 20 April 2003)

Chen P. 1999 ER Model, XML and the Web. ER '99, 18th International Conference on Conceptual Modeling, Paris, France, November, 15-18, 1999, pp. 538.

Chrisman, N., 1987. Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. International Journal Of GIS 1(3):265-278

CITES, 2003. Convention on International Trade in endangered Species of Wild Fauna and flora. URL: http://www.cites.org/ (accessed 30 August 2003)

Ciucci, P., L. Boitani. 1996. Ricerca e Gestione del Lupo in Toscana: Final Report. Regione Toscana, Dipartimento Agricoltura e Foreste. pp 215 + 7 maps

Ciucci, P., 1994. Movimenti, attività e risorse del lupo (*Canis lupus*) in due aree dell'Appennino centro-settentrionale. PhD Thesis, Univ. of Rome, Italy. Pagg. 117.

Ciucci, P., L. Boitani. 1991. Viability assessment of the Italian Wolf and guidelines for the management of the wild and a captive population. Ricerche Biologia Selvaggina (Bologna, Italy) 89:1-58.

Clark, J. D., J. E. Dunn, and K. G. Smith., 1993. A multivariate model of female black bear habitat use for a geographic information system. Journal of Wildlife Management 57:519-526.

Cleary, A., S. Kohn, S .G. Smith, and B. Smolinski, 1998. Language Interoperability Mechanisms for High Performance Scientific Applications.Proceedings of the SIAM Workshop on Object-Oriented Methods for Interoperable Scientific and Engineering Computing, Yorktown Heights, NY, October 21-23 1998

Cochran, W. G., 1977. Sampling Techniques: New York, Wiley.

Cohen, J., 1960. A coefficient of agreement for nominal scales: Educational and Psychological Measurement, v. 20, p. 37-46.

Colwell, R. N., 1983. Manual of remote sensing: Falls Church Virginia, ASPRS.

Congalton, R. G., 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data: Photogrammetric Engineering and Remote Sensing, v. 54, p. 593-600.

Congalton, R. G., J. M. Stenback, R. H. Barrett, 1993. Mapping Deer habitat suitability using Remote Sensing and Geographic Information Systems. Geocarto International 3: 23-33.

Consiglio Regione Toscana, 1994. Danni causati al patrimonio zootecnico da animali predatori o da eventi meteorici. Delega di funzioni e finanziamenti regionali. Legge Regionale 31 Agosto 1994, n. 72. Bollettino Ufficiale della Regione Toscana - n. 60

Corsi, F., E. Duprè and L. Boitani, 1999. A large-scale model of wolf distribution in Italy for conservation planning. Conservation Biology 13(1):150-159.

Corsi, F., J. De Leeuw, A. K. Skidmore, 2000. Modeling species distribution with GIS. In: Research Techniques in Animal Ecology. Boitani L. and T.K. Fuller (Eds.) Columbia University Press, New York: 389-434.

Corsi, F., I.Sinibaldi, L.Boitani, 1998 Large carnivores conservation areas in Europe.Discussion Paper for the Large Carnivore Initiative. Istituto di Ecologia Applicata and WWF International.

Corsi, F., L. Boitani, I. Sinibaldi, 2002. Ecological Corridors And Species: Large Carnivores In The Alpine Region. Nature and environment, No. 127. Council of Europe Publishing. ISBN 92-871-4989-6

Corsi, F., G. Gadaleta, G. de Luise, 2002. IUCN - SSC Web Enabled Species Information Service: Scope, Objectives, Approach. Oracle Italia.

Cotter, G. A., B. T. Bauldock, 2000. Biodiversity Informatics Infrastructure: An Information Commons for the Biodiversity Community. VLDB 2000: 701-704

Council of Europe, UNEP & ECNC, 1996. The Pan-European Biological and Landscape Diversity Strategy, a vision for Europe's natural heritage. Mart Spruijt bv, Amsterdam.

Crain, I. K. and N. M. Collins, 1999. Bioinformatics Implications of the International Biodiversity Conventions URL: http://www.iupac.org/symposia/proceedings/phuket97/crain.html (accessed 11 February 2003)

Crance, J.H., 1987. Guidelines for using the Delphi technique to develop habitat suitability index curves. US Fish and Wildlife Service Biological Reports 82(10.134) 21pp

Crane D., 1972. Invisible colleges; diffusion of knowledge in scientific communities. Chicago: University of Chicago Press.

Cressie, N. A. C:, 1993. Statistics for spatial data revised edition. John Wiley & Sons, New York. 900 pp.

Curran, P. J., and H. D. Williamson, 1986. Sample size for ground and remotely sensed data: Remote Sensing of Environment, v. 20, p. 31-41.

Dansereau, P., 1957. Biogeography, an ecological perspective. Ronald Press Comp. New York, 394 pp.

Daubenmire, R., 1976. The use of vegetation in assessing the productivity of forest lands. Botanical review 42: 115-143.

Daugherty, C. H. and F. W. Allendorf, 2002. The Numbers That Really Matter. Conservation Biology 16(2):283-284.

Davis, J. C., 1986. Statistics and data analysis in geology. Second Edition. John Wiley & Sons, New York. 646 pp.

Donovan, M. L., D. L. Rabe, C.E. Olson, 1987. Use of Geographic Information Systems to develop habitat suitability models. Wildlife Society Bullettin. 15:574-579.

Dubuc, L. J., W. B. Krohn, and R. B. Owen, Jr., 1990. Predicting occurrence of river otters by habitat on Mount Desert Island, Maine. Journal of Wildlife Management 54:594-599.

Duncan, B. W., D. R. Breininger, P. A. Schmalze, V. L. Larson, 1995. Validating a Florida scrub jay habitat suitability model, using demography data on the Kennedy Space Center. Photogrammetric Engineering & Remote Sensing, 61(11): 1361-1370.

Dunning, J. B., Jr., D. J. Stewart, B. J. Danielson, B. R. Noon, T. L Root, R. H. Lamberson, and E. E. Stevens., 1995. Spatially explicit population models: current forms and future uses. Ecololgical Applications 5:3-11.

Eastman, J. R., 1992. Time series map analysis using standardised principal components. Proceedings, ASPRS/ACSM/RT '92 Convention: Mapping and Monitoring Global Change. ASPRS/ACSM, Bethesda, MA.

Eco-Portal, 2003. Eco-Portal - The Environmental Sustainability.Info Source. URL: http://www.eco-portal.com/ (Accessed 15 May 2003)

Edwards, P. J., R. M. May and N. R. Webb (Eds.), 1994. Large-scale Ecology and Conservation Biology. Blackwell Scientific Publications, Oxford.

Ehlers, M., and R. Welch, 1987. Stereocorrelation of Landsat TM images: Photogrammetric Engineering and Remote Sensing, v. 53, p. 1231-1237.

Encyclopaedia Britannica, 1994. The new encyclopaedia britannica. Encyclopaedia Britannica Inc., Chicago.

ESRI, 1993. Digital Chart of the World for use with Arc/Info software. Environmental Systems Research Institute, Inc. Redlands, CA.

ESRI. 1992. ArcInfo User's Manuals, version 6.1. Environmental System Research Institute, Inc. Redlands, California.

Estes, J. E. and D.W. Mooneyhan, 1994. Of Maps and Myths. Photogrammetric Engineering and Remote Sensing, 60(5):517-524.

Estes, R. D., 1991. The Behaviour Guide to African Mammals. Including Hoofed Mammals, Carnivores, Primates. The University of California Press.

European Union, 1992. "Habitat" Directive. Annex I, Directive 92/43/CEE.

European Union. 1979. Birds Directive. Directive 79/409/EEC

Flather, C. H., and R. M. King, 1992. Evaluating preformance of regional wildlife habitat models: Implication to resource planning. Journal of Environmental Management. 34: 31-46.

Flather, C. H., S. J. Brady and D.B. Inkley, 1992. Regional habitat appraisal of wildlife communities: a landscape-level evaluation of a resource planning model using avian distribution data. Landscape Ecology. 7(2):137-147

Fonseca, G. A. B., A. Balmford, C. Bibby, L. Boitani, T. Brooks, N. Burgess, F. Corsi, E. Dinerstein, C. Gascon, L. Hannah, J. Lovett, R. A. Mittermeier, D. Moyer, S. Olivieri, D. Olson, C. Rahbek, S. Stuart, P. Williams., 2000. …Following Africa's lead in setting priorities. Nature. Vol. 405, 393-394.

Foppen, R. P. B., I. M. Bouwma, J. T. R. Kalkhoven, J. Dirksen & S. van Opstal, 2000. Corridors in the Pan-European Ecological Network, concepts and examples for terrestrial and freshwater vertebrates.ECNC, Tilburg.

Ford, G. E., and C. I. Zanelli, 1985. Analysis and quantification of errors in the geometric correction of satellite images: Photogrammetric Engineering and Remote Sensing, v. 51, p. 1725-1734.

Forman, R. T. T., 1983. Corridors in a landscape their ecological structure and function. Ekologia 2:375-387.

Fritts, S. H., and L. N. Carbyn, 1995. Population viability, nature reserves, and the outlook for gray wolf conservation in North America. Restoration Ecology 3:26-38

Fuller, T. K., W. E. Berg, G. L Radde, M. S. Lenarz and G. B. Joselyn, 1992. A history and current estimate of wolf distribution and numbers in Minnesota. Wildlife Society Bullettin 20: 42-55.

Fung, T., and E. LeDrew, 1988. The determination of optimal threshold levels for change detection using various accuracy indices. Photogrammetric Engineering and Remote Sensing 54(10):1449-1454.

Gaston, K. J., 1994. Rarity. Chapman and Hall, London.

Gaston, K. J., 1991. How large is a species' geographic range? Oikos 61: 434-438

Gaston, K. J., 1994. Measuring geographic range size. Ecography 17(2):198-205

Gaston, K. J., 2000. Global patterns in biodiversity. Nature 405:220-227

GBIF, 2003 Global Biodiversity Information Facility Communications Portal. URL: http://www.gbif.org/ (Accessed 15 May 2003)

Genard, M. and F. Lescourret, 1992. Modelling wetlands habitats for species management: the case of Teal (*Anas crecca crecca*) in the Bassin d'Arachon (French Atlantic Coast). Journal of Environmental Management 34: 179-195.

Gilliland-Swetland, A., M. Bacca, and T. Gill, 2000. Introduction to Metadata: Pathaways to Digital Information. Getty Information Institute.

Gilpin, M., 1996. Metapopulations and Wildlife Conservation: Approaches to Modeling Spatial Structure. Pages 11-28 in D. R. McCullough, editor. Metapopulations and wildlife conservation. Island Press, Washington DC.

Ginevan, M. E., 1979. Testing land-use map accuracy - another look: Photogrammetric Engineering and Remote Sensing, v. 45, p. 1371-1377.

Ginsparg, P., 2003. Can Peer Review be better Focused? URL: http://www.arxiv.org/blurb/pg02pr.html (accessed 14 May 2003)

Goudie, A., 1985. The encyclopaedic dictionary of physical geography. Blackwell, Oxford, 528 pp.

Grumbine, R. E., M. E. Soule, 1993. Ghost Bears: Exploring the Biodiversity Crisis. Island Press

Gustafson, E. J. & R. H. Gardener, 1996. The effect of land heterogeneity on the probability of patch colonization. Ecology 77: 94-107.

Gutierrez, R. J., and S. Harrison, 1996. Applying Metapopulation Theory to Spotted Owl Management: A History and Critique. Pages 167-186 in D. R. McCullough, editor. Metapopulations and wildlife conservation. Island Press, Washington DC.

Haber, W., 1994. System ecological concepts for environmental planning. In: Klijn, F., (Ed.), Ecosystem classification for environmental management. Kluwer Academic Publishers, Dordrecht, pp. 49-67.

Hall, L.S., P.R. Krausman, and M.L. Morrison, 1997. The habitat concept and a plea for standard terminology. Wildlife Society Bulletin, 25(1): 173-182.

Harnad, S., 1996. Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals. In: Peek, R. & Newby, G. (Eds.) Scholarly Publication: The Electronic Frontier. Cambridge MA: MIT Press. Pp. 103-108.

Harnad, S., 2000. The Invisible Hand of Peer Review. Exploit Interactive, issue 5. URL: http://www.exploit-lib.org/issue5/peer-review/ (accessed 14 May 2003)

Harrison, S., 1994. Metapopulations and conservation. Pages 111-128 in . P. J. Edwards, R. M. May and N. R. Webb, editors. Large-scale Ecology and Conservation Biology. Blackwell Scientific Publications, Oxford.

Harrison, S., A. M. Stahl, and D. Doak, 1993. Spotted owl update: US judge rejects Forest Service plan. Conservation Biology 7:1-4

Haslett, J. R., 1990. Geographic Information Systems: a new approach to habitat definition and study of distributions. TREE 5(7): 214-218.

Hausser, J., 1995. Säugetiere der Schweiz – Mammifères de la Suisse – Mammiferi della Svizzera. Commission des Mémoires de l'Académie Suisse des Sciences Naturelles.

Haworth, P. F. and D. B. A. Thompson, 1990. Factors associated with the breeding distribution of upland birds in the South Pennines, England. Journal of Applied Ecology. 27:562-577.

Hay, A. M., 1979. Sampling design to test land use accuracy: Photogrammetric Engineering and Remote Sensing, v. 45, p. 529-533.

Heglund, P. J., 2002. Foundations of Species-Environment Relations. In: Scott J. M., P. J. Heglund, M. L. Morrison *et al.* (Eds.) 2002. Predicting Species Occurrences. Issues of Accuracy and Scale. Island Press. Washington.

Herr, A. M. and L. P. Queen, 1993. Crane habitat evaluation using GIS and Remote Sensing. Photogrammetric Engineering and Remote Sensing (59)10: 1531-1538

Heuvelink, G. B. M., and P. A. Burrough, 1993. Error propagation in cartographic modelling using Boolean logic and continuous classification. International Journal of Geographical Information Systems, v. 7, p. 231-246.

Hodgson, M. E., J. R. Jensen, H. E. Mackey Jr., and M. C. Coulter, 1987. Remote sensing of wetland habitat: a wood stork example. Photogrammetric Egineering and Remote Sensing 53: 1075-1080.

Hoffer, R. M., 1975. Computer-aided analysis of SKYLAB multispectral scanner data in mountainous terrain for land use, forestry, water resource, and geological applications, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47907.

Holling, C. S., (Ed), 1978. Adaptive environmental assessment and management. John Wiley, New York, New York, USA

Homewood, K. M., 1975. Can Tana River Mangabey survive?. Oryx 13: 53-59.

Hopkins, L., 1977. Methods for generating land suitability maps: a comparative evaluation. Journal of the American Institute of Planning, October 1977: 386-400.

ICZN, 2000. Fourth Edition of the International Code of Zoological Nomenclature. International Trust for Zoological Nomenclature. London.

INNVISTA, 2003 URL: http://www.innvista.com/science/zoology/mammals/canidae.htm (accessed 24 June 2003)

IPNI, 1999. International Plant Names Index. Home Page. URL: http://www.ipni.org/index.html (accessed 24 June 2003)

ISO 639/JAC, 2003 Codes for the Representation of Names of Languages Part 2: Alpha-3 Code. http://www.loc.gov/standards/iso639-2/langhome.html (accessed 15 December 2003)

ISTAT (Istituto Centrale Statistica), 1990. 4° Censimento generale dell'agricoltura. Roma.

ISTAT (Istituto Centrale Statistica), 1991. 13° Censimento generale della popolazione. Roma.

ITIS, 2003 Integrated Taxonomic Information System URL: http://www.itis.usda.gov/index.html (accessed 24 June 2003)

IUCN, 2000. IUCN Red List Categories. Version 3.1 Prepared by IUCN Species Survival Commission. IUCN

IUCN, 2001. IUCN Intersessional Programme. URL: http://www.iucn.org/ourwork/ppet/files/steppinginto.pdf (Accessed 15 June 2003)

IUCN-SSC, 2001. A Study into Voluntarism in the Species Survival Commission. URL: http://www.iucn.org/themes/ssc/memonly/voluntarism/intro.htm (accessed 2 February 2003)

IUCN-SSC, 2003a. Red List web site. URL: http://www.redlist.org/. (accessed 3 March 2003)

IUCN-SSC, 2003b. Species Information Service (SIS). URL: http://www.iucn.org/themes/ssc/programs/sisindex.htm. (accessed 3 March 2003)

IUCN-SSC, 2003c. SSC Action Plans. http://www.iucn.org/themes/ssc/pubs/sscaps.htm. (accessed 3 March 2003)

IUCN-SSC, 2003d. CITES: A Conservation Tool. Available from the following URL: www.iucn.org/themes/ssc/programs/cites/cites.htm (accessed 23 august 2003)

Jensen, J. R., S. Narumalani, O. Weatherbee, and K. S. Morris Jr., 1992 Predictive modelling of Cattail and Waterlily distribution in a South Carolina reservoir using GIS. Photogrammetric Engineering and Remote Sensing 58(11):1561-1568.

Jones, K. B., 1986. Data types. In: Cooperrider, A. Y., J. B. Raymond, and H. R. Stuart, (Eds.): Inventory and monitoring of wildlife habitat. U.S. Dept. Interior, Bureau of Lands, Denver, 858 pp.

Jongman, R. H. G., ter Braak C. J. F. and van Tongeren O. F. R., (Eds.), 1995. Data analysis in community and landscape ecology. Cambridge University Press, 299 pp.

Kalensky, Z. D., and L. R. Scherk, 1975. Accuracy of forest mapping from Landsat CCTs, Proceedings of the 10th International Symposium on Remote Sensing of Environment, Volume 2, ERIM: Ann Arbor, Michigan, p. 1159-1163.

Kerr, R. M., 1986. Habitat mapping. In: Cooperrider, A. Y., Raymond, J. B. and Stuart, H. R. (Eds.): Inventory and monitoring of wildlife habitat. U.S. Dept. Interior, Bureau of Lands, Denver, 858 pp.

Kim, K. C., 1993. Biodiversity conservation: a global perspective. Manus. Col. ISBC. Kei, pp.1-19.

Klopatek, J. M., J. R. Krummel, J. B. Mankin and R. V. O'Neil, 1983. A theoretical approach to regional environmental conflicts. Journal of Environmental Management 16: 1-15

Knick, S.T., and D. L. Dyer, 1997. Distribution of black-tailed jackrabbit habitat determined by GIS in Southwestern Idaho. Journal of Wildlife Management. 61(1):75-85.

Kolasa, J., 1989. Ecological systems in hierarchical perspective: breaks in community structure and other consequences. Ecology 70(1):36-47

Krebs, C. J., 1985. Ecology, the experimental analysis of distribution and abundance. Harper & Row, New York, 800 p.

Lam, N. S., 1983. Spatial interpolation methods: a review: The American Cartographer, v. 10, p. 129-149.

Lamberson, R. H., K. McKelvey, B. R. Noon, and C. Voss, 1992. A dynamic analysis of northern spotted owl viability in a fragmented forest landscape. Conservation Biology, 6: 505-512.

Laurin, R. and D. Thompson, 1992. Fundamental of spatial information systems. The APIC series. Accademic Press, London. 680 pp.

Lawler, J. J., D. White, J. C. Sifneos and L. L. Master. 2003. Rare Species and the Use of Indicator Groups for Conservation Planning. Conservation Biology 17(3):875–882.

Lehmkuhl, J. F. and M. G. Raphael, 1993. Habitat pattern around Northern Spotted Owl locations in the Olympic Peninsula, Washington. Journal of Wildlife Management. 57(2):302-315.

Leopold, A., 1933. Game management. C. Scribner Sons, New York. 481 pp.

Levin S. A., 1992. The problem of pattern and scale in Ecology. Ecology 73(6):1943-1967.

Livingston, S. A., C. S. Todd, W. B. Krohn, and R. B. Owen, 1990. Habitat models for nesting bald eagle in Maine. Journal of Wildlife Management 54(4): 644-653.

Lyon, J. G., 1983. Landsat-derived land-cover classification for locating potential Kestrel nesting habitat. Photogrammetric Engineering and Remote Sensing 49(2): 245-250

Macdonald, D. W. (Ed.), 1984. The Encyclopaedia of Mammals. Facts on File, New York, NY.

Mace, G. M., 1994. An investigation into methods for categorizing the conservation status of species. In: Edwards, P. J., May, R. M. and Webb, N. R. (Eds.) Large scale ecology and conservation biology. Blackwell Scientific Publications.

MacEachren, A. M., 1985. Accuracy of thematic maps - implications of choropleth symbolization: Cartographica, v. 22, p. 38-58.

Maiorano L. 1998 Modelli di idoneità ambientali di alcuni ungulate in Italia e nell'arco Alpino. MSc. Thesis, University of Rome "La Sapienza".

Manson, R. H., R. S. Ostfeld and C.D. Canham, 1999. Responses of small mammal community to heterogeneity along forest old field edge. Landscape ecology 14: 355-367.

Martin, J., 1993. Principles of Object-Oriented Analysis and Design. Englewood Cliffs, NJ: Prentice Hall.

Massa, B., 1982. Il gradiente faunistico nella penisola italiana e nelle isole. Atti Società Italiana Scienze Naturali, Museo Civico di Storia naturale di Milano, 123 (2-3): 353-374.

Maurer, B. A., 1994. Geographical Population Analysis: Tools for the Analysis of Biodiversity. Methods in Ecology, Blackwell Scientific Publications, Oxford. 130 p.

Maurer, B. A., S. G. Heywood, 1993. Geographic range fragmentation and abundance in neotropical migratory birds. Conservation Biology 7: 501-509.

May, R. M., J. H. Lawton, N. E. Stork, 1995. Assessing extinction rates. In: Lawton J. H., May R. M., (Eds.) Extinction Rates. Oxford: Oxford University Press. p 1-24.

May, R. M., 1994. The effects of spatial scale on ecological questions and answers. Pages 1-17 in P. J. Edwards, R. M. May and N. R. Webb, (Eds.) Large-scale Ecology and Conservation Biology. Blackwell Scientific Publications, Oxford.

Mayhew, S., and A. Penny, 1992. The concise Oxford dictionary of geography. Oxford university press, Oxford.

McCullagh, P., and J. A. Nelder, 1988. Generalized Linear Models. Chapman and Hall, London. 261pp.

McGarigal, K., B. J. Marks, 1994. FRAGSTATS Spatial Pattern Analysis Program for Quantifying Landscape Structure. Version 2.0. Oregon State University, Corvallis, OR.

McKelvey, K., B. R. Noon, and R. H. Lamberson, 1992. Conservation planning for species occupying fragmented landscapes: the case of the northern spotted owl. Pages 424-450 in P. M. Kareiva, J. G. Kingsolver and R. B. Huey, (Eds). Biotic Interactions and Global Change. Sinauer, Sunderland, MA.

McKinney, M. L., 1998. Branching models predict loss of many bird and mammal orders within centuries. Animal Conservation 1:159-164.

Mech, L. D., 1970. The wolf. The ecology and behavior of an endangered species. Natural History Press, Doubleday, New York. Pagg. 384.

Mech, L. D., 1995. The challenge and opportunity of recovering wolf populations. Conservation Biology 9:270-278.

Mech, L. D., S. H. Fritts, G. L. Radde, and W. J. Paul., 1988. Wolf distribution and road density in Minnesota. Wildlife Society Bulletin 16:85-87.

Meffe G. K., 2001 Crisis in a Crisis Discipline. Conservation Biology 15(2):303-304.

Meriggi, A., and S. Lovari., 1996. A review of wolf predation in southern Europe: does the wolf prefer wild prey to livestock ? Journal of Applied Ecology 33:1561-1571

Michener, W. K., J. W. Brunt, J. J. Helly, T. B. Kirchner, & S. G. Stafford, 1997. Nongeospatial metadata for the ecological sciences. The Ecological Applications, 7(1), 330–342.

Miller, R. I., (Ed.), 1994. Mapping the diversity of nature. Chapman & Hall, London.

Mittermeier, R., N. Myers, J. B. Thomsen, and G. A. B. da Fonseca., 1998. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. Conservation Biology 12:516–532.

Mladenoff, D. J., and T. A. Sickely., 1998. Assessing potetial gray wolf restoration in the Northeastern United States: a spatial prediction of favorable habitat and potential population levels. Journal of Wildlife Management 62 (1): 1-10.

Mladenoff, D. J., T. A. Sickley, R. G. Haight, and A. P. Wydeven., 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. Conservation Biology 9: 279-294.

Moore, W. G., 1967. A Dictionary of geography. A. & C. Black, London.

Moran, K., 2000. Bioprospecting: lessons from benefit-sharing experiences. International Journal of Biotechnology, Vol. 2, Nos. 1/2/3

Morrison, D. F. 1976 Metodi di analisi multivariata. Casa Editrice Ambrosiana 343 pp

Morrison, M. L., B. G. Marcot, and R. W. Mannan, 1992. Wildlife-habitat relationships: concepts and applications. Univ. Wisconsin Press, 343 p.

Myers, N., 1988. Threatened biotas: hotspots in tropical forests. The Environmentalist 8:178–208.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, J. Kent, 2000. Biodiversity hotspots for conservation priorities. Nature 403:853-858.

Newcomer, J. A., and J. Szajgin, 1984. Accumulation of thematic map errors in digital overlay analysis: The American Cartographer, v. 11, p. 58-62.

Norton, T. W., and H. P. Possingham, 1993. Wildlife modelling for biodiversity conservation. In: Jakeman A. J., Beck, M. B., McAleer, M. J. (Eds.) Modelling Change in Environmental Systems. John Wiley & Sons Ltd.

Noss, R. F., 1990. Indicators for Monitoring Biodiversity: A Hierarchial Approach. Conservation Biology 4(4):355-364

Noss, R. F., 1992. Issues of scale in conservation biology. Pages 240-250 in P. L. Fiedler and S. K. Jain, (Eds.). Conservation biology the theory and practice of nature conservation, preservation and management. Chapman and Hall, New York & London.

Noss, R. F., H. B. Quigley, M. G. Hornocker, T. Merril, and P.C. Paquet, 1996. Conservation Biology and Carnivore Conservation in the Rocky Mountains. Conservation Biology, 10: 949-963.

Novak, R. M., 1995 Walker's Mammals of the World Online 5.1. Johns Hopkins University Press URL: http://www.press.jhu.edu/books/walker/toc.html (accessed 24 June 2003)

Nowicki, P., (Ed.), 1998 The Green Backbone of Central and Eastern Europe. Tilburg: European Centre for Nature Conservation. ECNC Publication Series on Man and Nature, Vol. 3. ISBN: 90-802482-5-8

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen, 1986. A hierarchical concept of ecosystems. Princeton, N.J.: Princeton University Press.

Oates, J. F., 1996. African Primates Status Survey and Conservation Action plan. IUCN/SSC Primate Specialist Group, IUCN – World Conservation Union, Gland, Switzerland.

Odum, E. P., 1971. Fundamentals of Ecology. W.B. Saunders Company, Philadelphia - London - Toronto. 584 p.

Olson, D., and E. Dinerstein, 1998. The Global 200: a representation approach to conserving the Earth's most biologically valuable regions. Conservation Biology 12:502–515.

Palmeirin, J. M., 1988. Automatic mapping of avian species habitat using satellite imagery. Oikos 52: 59-68.

Parratt, L. G., 1961. Probability and experimental errors in science: New York, Wiley.

Pausas, J. G., L. W. Braithwaite, M. P. Austin, 1995. Modelling habitat quality for arboreal marsupials in the South Coastal forest of New South Wales, Australia. Forest Ecology and Management. 78:39-49

Pearce, J. L., M. A. Burgman and D. C. Franklin, 1994. Habitat selection by Helmeted Honeyeaters. Wildlife Research. 21:53-63

Pereira, J. M. C. and L. Duckstein, 1993. A multiple criteria decision-making approach to GIS-based land suitability evaluation. International Journal of Geographical Information Systems, 7(5): 407-424.

Pereira, J. M. C. and R. M. Itami, 1991. GIS-base habitat modeling using logistic multiple regression: a study of the Mt. Graham Red Squirrel. Photogrammetric Engineering & Remote Sensing. 57(11): 1475-1486.

Peuquet, D. J., 1984. A conceptual framework and comparison of spatial data models. In: Marble, D. F., Calkins, H. W., and Peuquet, D. J., (Eds.), Basic readings in geographic information systems: Williamsville, New York, SPAD Systems.

Peuquet, D. J., and A. R. Boyle, 1984. Raster scanning, processing and plotting of cartographic documents: Williamsville, New York, Spad Systems Ltd.

Picozzi, N., D. C. Catt and R. Moss, 1992. Evaluation of Capercaillie habitat. Journal of Applied Ecology. 29:751-762

Piersma, T., R. Hoekstra, A. Dekinga, A. Koolhaas, P. Wolf, P. Battley, and P. Wiersma, 1993. Scale and intensity of intertidal habitat use by knots Calidirs canutus in the Western Wadden Sea in relation to food, friends and foes. Netherlands Journal of Sea Research. 31(4):331-357

Pimm, S. L., 1991. The Balance of Nature? University of Chicago Press, Chicago, IL.

Pimm, S. L., G. J. Russell, J. L. Gittleman, T. M. Brooks, 1995. The future of biodiversity. Science 269:347-350

Pinchera, F., L. Boitani and F. Corsi, 1997. Application to terrestrial vertebrates of Italy of a system proposed by IUCN for a new classification of national Red List Categories. Biodiversity and Conservation 6: 959-978.

Pressey, R. L., H. P. Possingham and C. R. Margules. 1996. Optimality in reserve selection algorithms: When does it matter and how much? Biological Conservation 76:259-267.

Price, M. V. and M. Gilpin, 1996. Modelers, Mammologists, and Metapopulations: Designing Stephens' Kangaroo Rat Reserves. Pages 217-

240. In: D. R. McCullough, (Ed.). Metapopulations and wildlife conservation. Island Press, Washington DC.

Promberger, C., and W. Schröder, (Eds.), 1993. Wolves in Europe: Status and perspectives. Munich Wildlife Society, Ettal, Germany. 100 pp.

Purvis, A., and A. Hector 2000. Getting the measure of biodiversity. Nature 405:212-219

QED, 2002. Community support strategy for the Murray Darling Basin SA Region, SA Murray-Darling Basin Integrated Natural Resource Management Group, Adelaide SA, 78.

Ramsar Convention 2001a The Criteria for Identifying Wetlands of International Importance. URL: http://www.ramsar.org/about_infopack_5e (accessed 25 august 2003)

Ramsar Convention 2001b Practical Steps for Applying the Vision for the Ramsar List of Wetlands of International Importance. 4th European Regional Meeting on the Ramsar Convention, 16 October 2001. URL: http://www.ramsar.org/mtg_reg_europe2001_4yeo_e.doc (accessed 25 august 2003)

Ramsar Convention 2002 A Framework for Wetland Inventory. 8th Meeting of the Conference of the Contracting Parties to the Convention on Wetlands (Ramsar, Iran, 1971) Valencia, Spain, 18-26 November 2002. URL: http://www.ramsar.org/key_guide_inventory_e (accessed 25 august 2003)

Rapoport, E., 1982. Areography. Oxford: Pergamon Press.

rECOrd, 2003 A Biodiversity Information System for all your ecological & natural history information requirements in Cheshire, Halton, Warrington & Wirral. URL: http://www.consult-eco.ndirect.co.uk/lrc/index.htm (accessed 5 June 2003)

Reid, W. V, and K. Miller, 1989. Keeping Options Alive: The Scientific Basis for Conserving Biodiversity. World Resources Institute, Washington, D.C.

Reid, W. V., 1998. Biodiversity hotspots. Trends in Ecology and Evolution 13:275-280.

Rexstad, E. A., D. D. Miller, C. H. Flather, E. M. Anderson, J. W. Hupp, D. R. Anderson, 1988. Questionable multivariate statistical inference in wildlife habitat and community studies. Journal of Wildlife Management. 52(4):794-798

Richards, J. A., 1986. Remote sensing - digital analysis: Berlin, Springer-Verlag.

Root, K. V., 2003. Prioritizing conservation of biodiversity using a multispecies approach. Gap Analysis Bulletin 11:18-23.

Root, K. V., H. R. Akçakaya and L. R. Ginzburg, 2003. A multi-species approach to ecological valuation and conservation. Conservation Biology 17:196-206

Rosenfield, G. H., K. Fitzpatrick-Lins, and H. S. Ling, 1982. Sampling for thematic map accuracy testing: Photogrammetric Engineering and Remote Sensing, v. 48, p. 131-137.

Ross, R. M., R. M. Bennet, T. W. H. Backman, 1993. Habitat use by spawning adult, egg, and larval American Shad in the Delaware River. Rivers 4(3):227-238

Rosser A., M. Haywood and D. Harris 2001. CITES: A CONSERVATION TOOL. Prepared for the Twelfth Meeting of the Conference of the Parties, November 2002. IUCN Species Survival Commission.

Rowe-Rowe, D. T., 1994. The ungulates of Natal. Natal Parks, Game & Fish Preservation Board, Pietermaritzburg, Natal, R.S.A.

Said, M. Y., A. K. Skidmore, J. de Leeuw, H. M. Aligula, L. Kumar, and H. H. T. Prins, 2003a. Analysis of the relation between ungulate species richness in East Africa and climatic and remotely sensed productivity indices. In: Said M. Y. Multiscale perspectives of species richness in East Africa. PhD thesis, Wageningen University.

Said, M.Y., P. W. Wargute, A. K. Skidmore, J. de Leeuw, and H. H. T. Prins, 2003b. Mammalian species richness in relation to rainfall in East Africa: do similar patterns exist among taxonomic groups. In: Said M. Y. Multiscale perspectives of species richness in East Africa. PhD thesis, Wageningen University.

Salafsky N., R. Margoluis, K. H. Redford, and J. G. Robinson, 2002. Improving the Practice of Conservation: a Conceptual Framework and Research Agenda for Conservation Science Conservation Biology 16(6):1469–1479

Sanderson, G. C., E. D. Ables, R. D. Sparrowe, J. R. Grieb, L. D. Harris and A. N. Moen, 1979. Research needs in wildlife. Transcations of the north american Wildlife and Natural Resources Conference 44, 166-175.

SAS Institute, Inc., 1985. SAS user's guide: statistics, version 6.04 edition. SAS Institute, Inc., North Carolina. pp. 1028.

Saunders, D. A., R. J. Hobbs, and C. Margules, 1991. Biological consequences of ecosystem fragmentation: A review. Conservation Biology 5:18-32.

Scheer, T., and T. Smith, 1999 Accelerating Your Object-Oriented Development. An Objectivity, Inc. White Paper. URL: http://www.objectivity.com/DevCentral/Products/TechDocs/pdfs/AcceleratingOODB.pdf (accessed 2 August 2003)

Science 2004. One-stop taxonomy. Science 303:1267

Scott, J. M., B. Csuti, J. D. Jacobi, and J. E. Estes, 1987. Species Richness. BioScience 37(11): p. 782-788.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, C. E. Thomas Jr., J. Ulliman, and R. G. Wright, 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife Monographs 123:1-41

Scott, J. M., P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, F.B. Samson. (Eds.) 2002. Predicting Species Occurrence. Issues of Accuracy and Scale. Island Press. Washington.

Shaw, D. M., and S. F. Atkinson, 1990. An indroduction to the use of Geographic Information Systems for ornithological research. The Condor 92:564-570.

Shuyun Xu, Yu Li and Shiyong Lu, 2003 ERDraw: An XML-based ER-diagram Drawing and Translation Tool. In: Proc. of the 18th International Conference on Computers and Their Applications (CATA'2003), pp. 143--146, Honolulu, Hawaii, USA, March, 2003.

Skidmore, A. K., 1989a. An expert system classifies eucalypt forest types using Landsat Thematic Mapper data and a digital terrain model.: Photogrammetric Engineering and Remote Sensing, v. 55, p. 1449-1464.

Skidmore, A. K., 1989b. Unsupervised training area selection in forests using a nonparametric distance measure and spatial information: Int J Remote Sensing, v. 10, p. 133-146.

Skidmore, A. K., A. Gauld, P. W. Walker, 1996. A comparison of GIS predictive models for mapping kangaroo habitat. International Journal of Geographical Information Systems, 10: 441-454.

Skidmore, A. K., and B. J. Turner, 1988. Forest mapping accuracies are improved using a supervised nonparametric classifier with SPOT data. Photogrammetric Engineering and Remote Sensing 54(10):1415-1421

Skidmore, A. K., and B. J. Turner, 1992. Assessing map accuracy using line intersect sampling: Photogrammteric Engineering and Remote Sensing, v. 58, p. 1453-1457.

Skinner, J. D., R. H. N. Smithers, 1990. The mammals of the Southern African subregion. University of Pretoria, Pretoria, RSA.

Smith, A. T., L. Boitani, C. Bibby, D. Brackett, F. Corsi, G. A. B. da Fonseca, C. Gascon, M. Gimenez Dixon, C. Hilton-Taylor, G. Mace, R. A. Mittermeier, J. Rabinovich, B. J. Richardson, A. Rylands, B. Stein, S. Stuart, J. Thomsen, C. Wilson, 2000. Databases Tailored for Biodiversity Conservation. Science. 290(5499):2073-2074.

Sokal, R. R., and F. J. Rohlf, 1995. Biometry. W. H. Freeman e Co., New York. pp. 887.

Southwood, T. R. E., 1977. Habitat, the template for ecological strategies? Journal of Animal Ecology 46: 337-365.

Species 2000, 2003. Species 2000 Home Page. URL: http://www.sp2000.org/ (accessed 24 June 2003)

Spinage, C. A., 1982. A territorial antelope: the Uganda waterbuck. Academic Press, London and New York, NY.

Srikanta, B. J., J. R. Haritsa 2000 Design and Implementation of a Biodiversity Information Management System. In: K. Ramamritham, T.M. Vijayaraman (Eds.) Advance in Data Management 2000. Tata McGraw-Hill Publishing Company Ltd.

Starfield, A. M., 1997. A pragmatic approach to modeling for wildlife management. Journal of Wildlife Management. 61(2):261-270.

Steffens, F. E., 1992. Geostatistical estimation of animal abundance in the Kruger National Park, South Africa. In: Soares A. (Ed.), Geostatistics Troia Vol.2 – Quantitative Geology and Geostatistics. Klumer Academic Publisher

Stelfox, H. A., and G. R. Ironside, 1982. Land/Wildlife integration No. 2. Proceedings of a technical workshop to discuss the incorporation of wildlife information into ecological land surveys. Ecological Land Classification Series No. 17, Lands Directorate, Environment, Canada.

Stewart, R. R., T. Noyce and H. P. Possingham, 2003. The opportunity cost of ad-hoc marine reserve design decisions - An example from South Australia. Marine Ecology Progress Series 253:25-38.

Stoch, F., 2000. How many endemic species? Species richness assessment and conservation priorities in Italy. Belgian Journal of Entomology, 2(1): 125-133.

Stoch, F., 2003. The project "Distribution of the Italian fauna: the invertebrates". Atti Convegno "La conoscenza botanica e zoologica in Italia: dagli inventari al monitoraggio", Università di Roma "La Sapienza", 14 dicembre 2001, Ministero dell'Ambiente e della Tutela del Territorio: 17-32.

Stoms, D. M., and J. E. Estes, 1993. A remote sensing research agenda for mapping and monitoring biodiversity. International Journal of Remote Sensing 14(10): 1839-1860.

Stoms, D. M., 1992. Effects of habitat map generalization in biodiversity assessment. Photogrammetric Engineering & Remote Sensing. 58(11): 1587-1591

Stoms, D. M., F. W. Davis and C. B. Cogan, 1992 Sensitivity of wildlife habitat models to uncertainties in GIS Data. Photogrammetric Engineering & Remote Sensing. 58(6): 843-850.

Stork, N. E., 1988. Insect diversity: facts, fiction and speculation. Biological Journal of the Linnean Society, 35, 321-337.

Stormer, F. A., and D. H. Johnson, 1986. Introduction: biometric approaches to modeling. Pages 159-160 in J. Verner, M. L. Morrison, and C. J. Ralph,

(Eds), Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin.

Teillet, P. M., 1986. Image correction for radiometric effects in remote sensing: International Journal of Remote Sensing, v. 7, p. 1637-1651.

Thiel, R. P., 1985. The relationship between road densities and wolf habitat suitability in Wisconsin. American Midland Naturalist, 113: 404-407.

Thomasma, L. E., T. D. Drummer, R. O. Peterson, 1991. Testing habitat suitability index model for the Fisher. Wildlife Society Bulletin 19:291-297

Todd, W. J., D. G. Gehring, and J. F. Haman, 1980. Landsat wildland mapping accuracy: Photogrammetric Engineering and Remote Sensing, v. 46, p. 509-520.

Toll, D. L., 1984. An evaluation of simulated Thematic Mapper data and Landsat MSS data for discriminating suburban and regional land use and land cover: Photogrammetric Engineering and Remote Sensing, v. 50, p. 1713-1724.

Tomlin, C. D., 1990. Geographic Information Systems and cartographic modelling. Prentice-Hall, London.

Toraldo di Francia, G., 1978. Il rifiuto. Nuovo Politecnico 99, Einaudi.

Turner, M. G., G. J. Arthaud, R. T. Engstrom, S. J. Hejl, J. Liu, S. Loeb, and K. McKelvey, 1995. Usefulness of spatially explicit population models in land management. Ecological Appllications 5:12-16.

Twery, M. J., G. A. Elmes and C. B. Yuill, 1991. Scientific exploration with an intelligent GIS: predicting species composition from topography. AI Applications, 5(2): 45-53

U.S. Congress, 1987. Technologies To Maintain Biological Diversity. Office of Technology Assessment (OTA-F-330) Washington, DC: U.S. Government Printing Office.

U.S. Fish and Wildlife Service, 1980a. Habitats as a basis for Environmental Assesment (HEP). Ecological Service Manual 101, U.S. Dept. of Interior, Fish and Wildlife Service, Division of Ecological Services, Government Printing Office, Washington, D.C., 26 p.

U.S. Fish and Wildlife Service, 1980b. Habitats Evaluation Procedures (HEP). Ecological Service Manual 102, U.S. Dept. of Interior, Fish and Wildlife Service, Division of Ecological Services, Government Printing Office, Washington, D.C., 84 p. + appendices.

U.S. Fish and Wildlife Service, 1981. Standards for the development of Habitat Suitability Index models. Ecological Service Manual 103, U.S. Dept. of Interior, Fish and Wildlife Service, Division of Ecological Services, Government Printing Office, Washington, D.C., 68 p. + appendices

Udvardy, M. D. F., 1959. Notes on the ecological concepts of habitat, biotope and niche. Ecology 40: 725-728.

Udvardy, M. D. F., 1969. Dynamic zoogeography. Van Nostrand Reinhold Company, New York, NY.

UNEP, 1999. Global Environmental Outlook 2000. UNEP GEO team Division of Environmental Information, Assessment and Early Warning (DEIA&EW). Nairobi, Kenya

Van Apeldoorn, R. C., C. Celada, and W. Nieuwenhuizen, 1994. Distribution and dynamics of the Red Squirrel (*Sciurus vulgaris* L.) in a landscape with fragmented habitat. Landscape Ecology 9(3):227-235

Van Genderen, J. L., and B. F. Lock, 1977. Testing land-use map accuracy. Photogrammetric Engineering and Remote Sensing, v. 43, p. 1135-1137.

Van Genderen, J. L., B. F. Lock, and P. A. Vass, 1978. Remote sensing: statistical testing of thematic map accuracy: Remote Sensing of Environment, v. 7, p. 3-14.

Vane-Wright, R. I., C. J. Humphries, and P. H. Williams. 1991. What to protect? Systematics and the agony of choice. Biological Conservation 55:235–254

Verbyla, D. L., and J. A. Litvaitis, 1989. Resampling methods for evaluationg classification accuracy of wildlife habitat models. Environmental Management 13(6):783-787.

Vincke, P., 1986. Analysis of multicriteria decision aid in Europe. European Journal of Operations Research, 25:160-168.

von Foerster, T., 2001. The Future (?) of Peer Review In: R. S. Berry and A. S. Moffat (Eds.) The Transition from Paper: Where Are We Going and How Will We Get There? Published online URL: http://www.amacad.org/publications/trans5.htm (accessed 19 May 2003)

Walker, J., Jupp D. L. B., Penridge L. K., and Tian G., 1986. Interpretation of vegetation structure in Landsat MSS imagery: A case study in disturbed semi-arid eucalypt woodlands. Part 1. Field Data Analysis. Environmental Management, vol. 23, pp. 19-33.

Walker, P. A., 1990. Modelling wildlife distributions using geographic information system: kangaroos in relation to climate. Journal of Biogeography. 17:279-289.

Walker, P. A., and Moore D.M., 1988. SIMPLE An inductive modelling and mapping tool for spatially-oriented data. International Jaurnal of Geographical Information Systems. 2(4):347-363

Wallin, D. O., C. C. H. Elliott, H. H. Shugart, C. J. Tucker and F. Wilhelmi, 1992. Satellite remote sensing of breeding habitat for an African Weaver-bird. Landscape Ecology 7(2):87-99

Walters, C., 1992. Trends in applied ecological modelling. In: Wildlife 2001: Populations. D.R.McCullough and R.H.Barrett Editors. Pp. 117 - 122

Walters, C., 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology [online]1(2):1. Available from the Internet. URL: http://www.consecol.org/vol1/iss2/art1 (accessed 6 August 2003).

WCMC, 1999. CBD Implementation Measures Pilot Project. URL: http://www.unep-wcmc.org/cbd/measures/project_planning.pdf (accessed 2 February 2003)

Weaver, J. L., P. C. Paquet, and L. F. Ruggier, 1996. Resilience and Conservation of Large Carnivores in the Rocky Mountains. Conservation Biology, 10: 964-976.

Webster's, 1981. Webster's third new international dictionary. Merriam Webster Inc., Chicago.

Western, D., 1975. Water availability and its influence on the structure and dynamics of a savannah large mammal community. East African Wildlife Journal, 13: 265-286.

White, F. J., 1983. The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. UNESCO, Paris, France.

Whittaker, R. H., S. A. Levin, and R. B. Root, 1973. Niche, habitat and ecotope. American Naturalist 107: 321-338.

Wiens, J. A., 1996. Wildlife in Patchy Environments: Metapopulations, Mosaics and Management. Pages 53-84. In: D. R. McCullough, (Ed.) Metapopulations and wildlife conservation. Island Press, Washington DC.

Williams, G. L., 1988. An assessment of HEP (Habitat Evaluation Procedures) applications to Bureau of Reclamation Projects. Wildlife Society Bulletin 16:437-447

Williams, P. H., 1996. Measuring biodiversity value. World Conservation [formerly IUCN Bulletin], 1: 12-14

Wilson, D. E., and D. M. Reeder (Eds.), 1993. Mammals species of the World: A Taxonomic and Geographic Reference. Second edition. Smithsonian Institution Press, Washington D.C. 1206 pp. (Database avialable online at URL: http://nmnhwww.si.edu/msw/)

Wilson, E. O., 1992. The diversity of life. Cambridge, Massachusetts, Harvard University Press.

Wilson, E. O., 1997. Introduction. Pp. 1-3. In: Reaka-Kudla, M. L., D. E. Wilson, and E. O. Wilson, (Eds.) Biodiversity II: Understanding and protecting our biological resources. Joseph Henry Press, Washington

Wilson, E. O., 2000. On the Future of Conservation Biology. Conservation Biology 14(1):1-3

WRI, IUCN, UNEP, FAO, UNESCO, 1992. Global Biodiversity Strategy: Guidelines for Action to Save, Study, and Use Earth's Biotic Wealth Sustainably and Equitably. ISBN 0-915825-74-0

Wright, J. K., 1942. Map makers are human: comments on the subjective in maps: The Geographical Review, v. 32, p. 527-544

WWF, 2002. Biological Priorities for Conservation in the Guinean-Congolian Forest and Freshwater Region. Report of the Guinean-Congolian Forest and Freshwater Region Workshop, Libreville, Gabon March 30–April 2, 2000. URL: http://www.worldwildlife.org/ecoregions/congo_report_pre.pdf (accessed 31 August 2003).

Yapp, R. H., 1922. The concept of habitat. Journal of Ecology 10: 1 - 17.

Zeleny, M., 1982. Multiple Criteria Decision Making, McGraw-Hill, New York

Zimen, E., 1978. Der Wolf, Mythos und Verhalten. Meyster Verlag. Wien.

Zimen, E., and L. Boitani, 1975. Number and distribution of wolves in Italy. Zeitschrift für Säugetierkunde, 40: 102-112.

Zonneveld, I. S., 1974. Aerial photography, remote sensing and ecology: ITC Journal, v. 4, p. 553-560.

Zonneveld, I. S., 1989. The land unit - a fundamental concept in landscape ecology, and its applications. Landscape Ecology 3: 67-86.

Zonneveld, I. S., 1995. Land ecology: an introduction to landscape ecology as a base for land evaluation, land management and conservation. SPB, Amsterdam, 199 pp.

Appendix A
African Mammals Databank project outputs

The AMD project's outputs aim to contribute to the understanding of status, trends, and threats of African mammals, both the large charismatic and some of the smaller, lesser-known, species. The project's output is expected to benefit all institutions dealing with research and conservation of animal species at national and international level.

Two complementary sets of outputs have been produced to maximise the diffusion of results. The printed volume (Boitani *et al.*, 1999) provides a way of distributing the project's results to offices and institutions that do not yet have the technical facilities to handle electronic data. Conversely, the magnetic data bank remains available largely as a reference body of data for other applications. The technical capacity to handle georeferenced data and GIS is spreading rapidly, including in African countries. Costs are reasonable today and the software is becoming increasingly user-friendly and accessible to lower levels of university instruction.

The magnetic data set is available on CD-ROM and is accessible and downloadable from the web at the following URL:

http://www.gisbau.uniroma1.it/amd.

THE PRINTED DATA BANK

Orders and families are presented in a phylogenetic sequence. Within each of these, species (or subfamilies, when present within the families) are arranged in alphabetical order, as in the main reference source (i.e., Wilson & Reeder, 1993).

Each of the 281 species is given a separate account, while subgroups, infra-groups and super-groups were purposely ignored to avoid unnecessary repetitions and to maintain the overall scheme. Each species was given an identification code (ID code) and a hierarchical number for identification of its family and order: the former is indicated at the top left of each sheet, while the latter is used to number tables and figures (i.e., Order 2: Primates; Family 1: Loridae; Species 2.1.1: *Arctocebus aureus*; Species 2.1.2: *Arctocebus calabarensis*, etc.).

The species accounts have neither biology nor ecology sections. Such information is easily found in the many sources listed in the bibliography and the related FoxPro file.

The information presented in the species account is arranged under the following headings:

Nomenclature (***Scientific name***). In order to keep an international standard format, species nomenclature follows the taxonomy set out in Wilson & Reeder (1993). Exceptions were made when important literature sources or IUCN/SSC specialists expressed a strong preference for a different nomenclature, suggesting a valid alternative name and justifying the choice (for example: *Parahyaena brunnea* sensu Wilson & Reeder = *Hyaena brunnea*; *Herpestes naso* sensu Wilson & Reeder = *Xenogale naso; Genetta maculata* sensu Wilson & Reeder = *Genetta pardina* + *Genetta rubiginosa*). In the case of discrepancies between the Wilson & Reeder system of classification and available distribution maps, the reference text was not necessarily always given priority (i.e., *Procolobus badius* sensu Wilson & Reeder + *Procolobus preussi* sensu Wilson & Reeder + *Procolobus pennanti* sensu Wilson & Reeder + *Procolobus rufomitratus* sensu Wilson & Reeder = *Procolobus badius* sensu IUCN).

Describer. The name of the species' describer, written just below the scientific name, was also taken from Wilson & Reeder (1993).

Vernacular names. On each sheet, species' names are given in English (Eng) and French (Fr). These vernacular names are supplied as a convenient and useful tool for species identification, even if, unlike the Latin names, they are neither precise nor univocal and there is no way of dictating a generally accepted standardisation. In fact, there may be many different names for the same species in general use and deeply entrenched in the language within even a very restricted area. For English common names, the main reference texts were the IUCN/SSC Action Plans, while for French names, the main source was Kingdon (1997). Wherever a species was found to have more than one colloquial name, at least two were reported to facilitate consultation.

Taxonomic notes. This section is meant to provide no more than a comment on special or significant taxonomic aspects, not a listing of the entire species' or genus' taxonomic history. Closely related species, systematic uncertainties and well-defined forms, if any, are noted here. Subspecies are generally ignored, as the project focuses on the species taxonomic level; moreover, the validity and precise distribution of many subspecies are still uncertain, debated or pending review. Only in a few cases are subspecies listed: when they are classified at a different threat level by the IUCN or are of special conservation interest; when the named subspecies is considered a full species by other zoology texts; when there is no general agreement on the taxon classification; or in special cases (i.e., the gorilla, *Gorilla gorilla*) when the species' subdivision is accepted worldwide. The decision to include or exclude one or more subspecies in this section was arbitrary. No morphological description is given, as that would be beyond the scope of this work. Information in this section is based mainly on Wilson & Reeder (1993) and the IUCN Action Plans available for different

taxa, but sometimes, and especially in case of discrepancy, reference is made to Meester & Setzer (1971) and the more recent data published by Kingdon (1997).

IUCN threat category. This section gives the species' recent threat assessment according to the 1996 IUCN Red List of Threatened Animals (Baillie & Groombridge, 1996). The status category is fully listed for each taxon (at the species and, when described, subspecies level) together with its symbol and criteria in parentheses. Countries of occurrence are often listed for species (or subspecies), together with some of the following symbols (in particular for the antelopes, as indicated by East, 1996): «?» indicates that the taxon may occur, but its presence (former or current) is not confirmed; «ex» indicates that it is now extinct; and «ex?» indicates that it is now probably extinct.

The categories are: Extinct (EX), Extinct in the Wild (EW), Critically Endangered (CR); Endangered (EN); Vulnerable (VU); Lower Risk (LR); Data Deficient (DD); Not Evaluated (NE). The Lower Risk category includes three subcategories: conservation dependent (LR: cd); near threatened (LR: nt); least concern (LR: lc) (see Baillie & Groombridge, 1996, for definitions of categories and criteria).

Available information. This section gives a brief report of the information available in the scientific and grey literature on the species. The amount of information included here varies from species to species, depending on the complexity of each species' data, how much it has been studied and the main focus of the research done. The longest accounts are obviously those referring to the most popular and/or well-studied mammals (such as impala, *Aepycerus melampus*, gorilla, *Gorilla gorilla*, cheetah, *Acinonyx jubatus*, and lion, *Panthera leo*). In these cases (and in any case wherever relevant) the description is broken down into North, South, East and West African subregions.

A major effort was made to include in the bibliography the most important sources of information about each species, excluding minor sources and papers dealing with topics not relevant to this project.

Known Extent of Occurrence. This section summarises the species' known distribution range. It also provides the source(s) of the map, the name of the specialist (if any) who revised it (and the date), and a list of other publications used to update the output.

Comments are given on the distribution maps that were available for each species and the one thought to be most recent/reliable and retained for modelling. The maps show the limits within which - to the best of available knowledge - the species occurs, and underline special geographic features such as "introduced or reintroduced" areas, and the distinction between "possible" and "certain" presence. Each species' geographical range has been described in a standard form to make the results as uniform and essential as possible: the text

either lists the countries, or parts of them, in which the species occurs**,** or makes reference to geographic coordinates. The most recent names have been used as much as possible (however, the Democratic Republic of Congo is still referred to as former Zaire).

Categorical-discrete (CD) distribution model. This section presents the ecological information on the species used for the model. It also includes a table with the preference scores for the vegetation types occurring inside the EO (1 for suitable; 2 for moderately suitable; 3 for unsuitable; 9 for unknown).

Two tables show the quantitative results of the model in terms of percentage of the various suitability classes within the EO and the 6 indexes of fragmentation.

Probabilistic-continuous (PC) distribution model. The PC distribution model represents distribution in the form of an environmental suitability surface, where the ecological variability has been mapped as distance from the average ecological quality from the EO.

Validation. The validation parameters of the categorical-discrete (CD) distribution model were calculated only for those species for which at least 1% of the known Extent of Occurrence (EO) is included in the sample countries. This table (if present) shows the percent of the species' Extent of Occurrence (EO) within the sample areas, the number of valid plots and the Index of Accordance. A plot was analysed if it was included in the species' EO or if the species was observed in it. The Index of Accordance is the percentage of valid plots which agree with the CD model's results: it is considered sufficient if it exceeds 50%.

Comments and conservation issues. This specific section of the species account contains comments on models, fragmentation analysis and protected areas overlap. These comments provide an initial interpretation of the output. First, the two models' output is compared with the known Extent of Occurrence to underline the inconsistency, on a large scale, between the blotch shapes and the models' suitability areas. The section also indicates differences in the output of the two models and tries to interpret these differences. The shape and size of the Extents of Occurrence are compared with the shape and size of Areas of Occupancy, also using the first fragmentation indexes such as the patch numbers, their average dimension and their Standard Deviation. Attention is drawn to small fragments distant from the core of the species' distribution. The other fragmentation indexes are used to confirm the connectivity between different suitability areas. The fragmentation indexes should be used only as support for the maps.

These are followed by a discussion of the EO and AO percentages included in the protected areas. This information was influenced by the completeness and accuracy of protected areas coverage, which is not very reliable on a detailed scale, but is nevertheless a valid general indication of the degree of protection

received. It is important to underline the differences between the EO percentages (in the tables) and the AO percentages (in the comment sections).

These comments are not intended as an indication of necessary species' conservation actions, as analysis in this project was on a global scale and was not concerned with local situations.

References. The bibliography listed in this section includes the main publications used to compile each account and tends to be selective rather than exhaustive. Sources are listed by author and year in the text and are fully cited at the end of each account. Unless otherwise specified, information found in the Taxonomic notes and IUCN threat categories sections were compiled directly from Wilson & Reeder (1993) and Baillie & Groombridge (1996), respectively.

Tables and Figures. Tables and Figures are numbered using the hierarchical number of the species to which they refer as prefix (e.g., 8.6.30 *Ourebia ourebi*: Fig. 8.6.30.a, Fig. 8.6.30.b, Tab.8.6.30.a, etc.).

Indexes. The index contains all scientific names of the species considered in taxonomic and alphabetical order with reference to the ID code, the complete taxonomic code and the species account. Each species is then inserted into the respective family and order, both of which are ordered numerically.

THE MAGNETIC DATA BANK

A meta-data bank was set up containing the information that allows users to retrieve all the data relative to each of the species included in the project.

FIELD N°	FIELD NAME	FIELD TYPE	FIELD WIDTH	DESCRIPTION
1	IDSYS	CHAR	8	Hierarchic systematic code
2	ORDER	CHAR	20	Order name
3	FAMILY	CHAR	20	Family name
4	GENUS	CHAR	20	Genus name
5	SPECIES	CHAR	20	Species name
6	AUTHOR	CHAR	40	Author's name and date of description
7	IDCOD	CHAR	6	Unique identification code
8	STATUS	CHAR	30	IUCN threat category
9	SUMMARY	CHAR	254	Full path and filename of the species' account sheet
10	EO	CHAR	254	Full path and filename of the species' distribution coverage
11	HABITAT	CHAR	254	Full path and filename of the species-habitat relationships
12	CDMODEL	CHAR	254	Full path and filename of the CD-model image
13	PCMODEL	CHAR	254	Full path and filename of the PC-model image
14	VDS	CHAR	254	Full path and filename of the validation data set

Table A.1: Structure of the magnetic data bank index file (mammals.dbf).

- **The data bank structure**

A DBF structure contains all information for each species. Table A.1 gives the structure of the DBF file (mammals.dbf) which contains 281 records, one for each species.

The user can browse the data bank file for available data on a taxon with any DBF compatible software (e.g., Dbase III, Dbase IV, FoxPro for Windows etc.). The data bank will provide, along with data regarding the species' taxonomy and the IUCN status (record fields : IDSYS, ORDER, FAMILY, GENUS, SPECIES, AUTHOR, IDCOD, STATUS), a list of files storing the species' data as reported in the printed data bank, some additional data used to model the areas of occupancy and the results of the validation analysis (record fields: SUMMARY, EO, HABITAT, CDMODEL, PCMODEL, VDS). In this way the user can retrieve the files of interest, copy them and use them with common software to manage texts, images and spreadsheets. The list below explains each field in more detail.

- **IDSYS** is the hierarchic code, in which the first, second and third numbers refer respectively to the order, the family and the species (genus and species). The arrangement of orders and families follows Wilson & Reeder (1993): within these, the species are ordered alphabetically.
- **ORDER**, **FAMILY**, **GENUS**, **SPECIES** and **AUTHOR** refer to the taxonomy of a given species.
- **IDCOD** is the unique sequential code used throughout to identify a species. It acts as a link between the various pieces of information stored in distinct files, i.e., the numeric part of the IDCOD is common to all filenames dealing with a given species. For instance, the filenames amd001.doc and blo001.cgm refer to the file storing the species account sheet and the file storing the same species' Extent of Occurrence, respectively.
- **STATUS** is the IUCN threat category, if applicable. The user will find the threat category along with the criteria adopted to assign it.
- **SUMMARY** contains the filename and the full path of the species account sheet. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\amd001\amd001.rtf).
- **EO** contains the filename and the full path of the species Extent of Occurrence coverage file. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\amd001 \blo001.e00).

- **HABITAT** contains the filename and the full path of the species-habitat relationships matrix. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\ amd001\hab001.dbf).
- **CDMODEL** contains the filename and the full path of the species CD-model image file. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\amd001\ cdm001.tif)[24].
- **PCMODEL** contains the filename and the full path of the species PC-model image file. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\amd001\ pcm001.tif)[1].
- **VDS** contains the filename and the full path of the fieldwork validation data set. The user will find the full path in the format «CdRomNumber\directory\filename» (e.g., CdRom01\amd001\vds001.txt). If the validation analysis was not performed for a given species, the field is left blank.

- **Storage supports**

All data were stored on CD-Rom. Each CD-Rom contains several directories, each of which pertains to a single species and is named with the species' unique ID code (see the field IDCOD; e.g., amd001, amd255, etc.). The directories contain files with the species' data as specified below:

- the species account sheet (see the field SUMMARY; e.g., amd001.rtf in Rich Text Format);
- the coverage of the species Extent of Occurrence (see the field EO; e.g., blo001.e00 in Arc/Info Ms- Dos export format[25]);
- the species-habitat relationships matrix (see the field HABITAT; e.g., hab001.dbf in Xbase compatible format);
- the CD-model for the given species (see the field CDMODEL; e.g., cdm001.tif in TIFF PackBits compressed format)[1];
- the PC-model for the given species (see the field PCMODEL; e.g., pcm001.tif in TIFF PackBits compressed format)[1];

24 The user will find in the same directory two files with extension tfw (e.g. wat001.tfw and pro001.tfw) which were generated by Arc/Info during the conversion from the grid data set to the TIFF file. These are text files which store the information required by Arc/Info to georeference the image by the conversion from the TIFF file to a new grid data set.

25 Require dos2unix conversion if used on Unix platform.

- the validation data set (see the field VDS; e.g., vds001.txt in text delimited format).

- **Information structure**

The information stored in the files cited above is structured as follows.

Species account sheet: The data stored in the species account sheet are mainly texts describing the taxonomy, the extent of occurrence, and the results of the analyses performed. For more information on the structure of the sheet, see section 6.1.

The Extent of Occurrence coverage: The boundaries of the EO of the given species were stored in a vectorial Arc/Info coverage and converted into the Arc/Info export Ms-Dos format. The Lambert geographical projection was used with the following parameters:

Projection: Lambert Azimuthal

Units: meters

Radius of the sphere of reference: 6370997.00 (Sphere ArcInfo datum)

Longitude of centre of projection: 20° 00' 0.00

Latitude of centre of projection: 5° 00' 0.00

The polygons in the coverage have a 4- or 5-digit code: its meaning is summarised in Table A.2. The codification of polygons enables the user to retrieve readily the source of the map from the bibliographic database and to ascertain the quality of the information on presence associated with the various polygons.

DIGIT	VALUE	DESCRIPTION
1-4	Nnnn	Bibliographic code of the source reference of the distribution map **certain presence**
1-5	9nnnn	Polygon enclosing areas for which the presence is **uncertain**
1-5	90000	Polygon enclosing areas of **certain presence** on the basis of expert support and without bibliographic support
1-5	99991	Polygon enclosing areas in which the species is reported to be **introduced** or **re-introduced**
1-5	99999	Polygon completely enclosed within the EO of the species but in which the species is **absent**

Table A.2: Codification of the polygons in the Extent of Occurrence coverage.

The species-habitat relationships matrix: For each species a DBF file containing the suitability scores for each Land Cover class, White's vegetation type and the combination of the two falling inside the EO is provided. The scores listed are those used to produce the CD model. For the species for which the distance from water or the elevation was taken into account, the suitability

scores for each combination of vegetation and land cover inside and outside the specified threshold are also listed, together with the specified thresholds considered. Each file contains at least the following fields:

- **COD**: Numeric code corresponding to the unique combination of White's vegetation type and Land Cover class;
- **VEG**: White's classification second rank (see legend);
- **CODE2**: White's classification first rank (see legend);
- **LAND_CODE:** Land Cover class (see legend);
- **SCOREW**: Suitability score assigned to the vegetation type;
- **SCOREL**: Suitability score assigned to the Land Cover class;
- **SCOREWLC**: Suitability score assigned to the White's vegetation type/ Land Cover class combination.

For the species for which the distance from water was taken into account, the following fields were added:

- **DISTR**: threshold distance from permanent water (in meters) used to build the buffer;
- **IN**: Suitability score assigned to the White's vegetation type / Land Cover class combination at distances below the specified threshold;
- **OUT**: Suitability score assigned to the White's vegetation type / Land Cover class combination at distances greater than the specified threshold.

For the species for which elevation was taken into account, two elevation thresholds were defined, and scores were assigned to each combination depending on whether they fall either below the lower threshold, between the two thresholds or above the upper threshold. The following fields are in the file:

- **ELEV1**: Lower elevation threshold in meters above sea level;
- **ELEV2**: Upper elevation threshold in meters above sea level;
- **SCOREBE**: Suitability score assigned to the White's vegetation type / Land Cover class combination at elevations below the specified lower threshold;
- **SCOREIN**: Suitability score assigned to the White's vegetation type / Land Cover class combination at elevations between the two thresholds;

- **SCOREOV**: Suitability score assigned to the White's vegetation type / Land Cover class combination at elevations over the specified upper threshold.

The Categorical-Discrete (CD) model image: The file stores the raster file of the CD model. The results of the analyses carried out with the software developed under the Arc/Info environment were stored in the Arc/Info grid data set format. These were then converted to a TIFF graphic file format (with PackBits compression) to make them readable with any image processing software package. The raster is a rectangular flat image in the same projection as the Extent of Occurrence coverage. Along with the TIFF file the user will find in the same directory two files with the extension tfw (e.g., cdm001.tfw and pcm001.tfw), which were generated by Arc/Info during conversion of the grid data set into the TIFF file. The files store the information required by Arc/Info for georeferencing the TIFF image when converting it into a new grid data set.

The image pixel values are the outcome of the intersect procedure between the species' EO and the GLCC Land Cover data set, the White's vegetation map and for selected species the river network or the elevation data. The values range from 1 to 8 as shown in Table A.3.

PIXEL VALUE	DESCRIPTION	MAP COLOR
1	Suitable	green
2	Moderately suitable	pale green
3	Unsuitable	red
4	Undefined	gold
5	Environmental classes not found inside the EO	pale yellow
8	Water	cyan

Table A.3: Codification of the values for the CD model image.

The Probabilistic-Continuous (PC) model image: The file stores the raster file of the PC model. The results of the analyses carried out with the software developed within the project (using Microsoft Visual C++ compiler under Windows NT operating system) were stored in the Arc/Info grid data set format. The Arc/Info grid was later converted to the TIFF graphic file format (with PackBits compression) to make them readable with any image processing software package. The raster is a rectangular flat image in the same projection as the Extent of Occurrence coverage. Along with the TIFF file the user will find, in the same directory, two files with extension tfw (e.g., cdm001.tfw and pcm001.tfw) which were generated by Arc/Info during the conversion of the grid data set into the TIFF file. The files store the information required by Arc/Info to georeference the TIFF image when displaying it or converting it back into a grid.
The pixel values are the outcome of the analysis procedure described in Chapter 4, stretched to obtain an integer image file whose values range

from 1 to 255. Before stretching the image, the original Mahalanobis distance values were multiplied by 10 and truncated to the lower integer. In the stretched image 255 corresponds to all values with a Mahalanobis distance above 150, which was set as the upper limit (infinity) of the ecological distance. The false colours used for the map range continuously from a darker to a paler shade of green, indicating a more suitable (less ecologically distant) or less suitable (more ecologically distant) area for the species, respectively.

The validation data set: Not all species have a validation data set. As explained above, only species that have at least 1% of their EO within any of the four sample areas were included in the validation process. Thus the validation data set is only available for those species.

When present, this information can be found in a text file stored under the species' directory and named vdsXXX.txt (where XXX stands for the species code used throughout the project). This text files are tab delimited and can easily be imported into the most popular spreadsheet programs, and if desired, DMBS packages.

The first line of the file contains the species scientific name while the rest of the file is organised according to the following list:

- **Point_code**: unique identification code.
- **CD_score**: score of the location according to the CD model.
- **Inside_EO**: set to " in" if the location falls inside EO
- **Found**: set to 1 if the specie was observed during field work.
- **Country**: country code of the location (B = Botswana, C = Cameroon, M = Morocco, U = Uganda).
- **White_Cod**: White's Vegetation Map class code observed at the location.
- **Latitude**: Latitude of location.
- **Longitude**: Longitude of location.

Bibliographic data: To allow for standard use of the bibliographic data, the literature database is stored in ASCII text format. The maximum length of a CHAR Xbase field is 254 byte, enabling storage of a 254 character string: yet several reference titles are longer than that. Use of the MEMO field type would not have solved the problem because management of these kinds of fields is not standard and changes according to the DBMS used. The ASCII text format enables the user to import the data easily into any DBMS structure. The structure of the literature database is organised as shown in Table A.4. Each field is quoted and comma delimited. (e.g., "1499","Wilson & Reeder

(1993)","Wilson D.E., Reeder D.M. (Eds.) (1993).","…","Smithsonian Institution Press, Washington D.C.").

FIELD	FIELD TYPE	FIELD WIDTH	DESCRIPTION
1	CHAR	free text length	Unique identification code
2	CHAR	free text length	Author/s and date in short form
3	CHAR	free text length	Author/s and date in full form
4	CHAR	free text length	Title
5	CHAR	free text length	Reference
6	CHAR	free text length	Keywords (comma delimited)

Table A.4: Structure of the literature database.

Appendix B
Examples of Species Accounts from the African Mammals Databank

Carnivora Id code: amd084

Canidae

Canis adustus

Sundevall, 1847

(Eng) Side-striped jackal

(Fre) Chacal à flancs rayés

Taxonomic notes

According to Ginsberg & Macdonald (1990) five subspecies are recognised: *C. a. adustus, C. a. bwela, C. a. centralis, C. a. keffensis, C. a. lateralis*.

IUCN threat category

Not listed.

Available information

Literature available on the ecology of the side-striped jackal is quite limited. Although widely distributed on the continent, most of the information has been collected in East and Southern Africa, and very little is known on the species' ecology and distribution in West Africa (Crawford-Cabral, 1989b; Happold, 1987; Rosevear, 1974).

East Africa: Fuller et al. (1989) analysed the ecology of the three species of jackals represented there, but the data collected by the authors on the side-striped jackals are quite scarce. Niche separation and coexistence among the three species of jackals were also studied by Lamprecht (1978), who focused mainly on diet and foraging behaviour, and by Moehlman (1978). Notes on its presence in Ethiopia and in Somalia are found in Funaioli (1971) and Yalden et al. (1980; 1996). Duckworth (1992) reports some information on its presence and density in the Nechisar National Park (Ethiopia). Notes on its presence in eastern former Zaire are found in Rahm (1966).

Southern Africa: The main reference for the ecology and distribution of this species is Skinner & Smithers (1990). Some information on the diet in agricultural areas is found in Smithers & Wilson (1979), while Dunham (1994) reports on its habitat in the Zambezi woodland (Zimbabwe). Some data on its

feeding habits and activity patterns are found in Rautenbach & Nel (1978), who briefly discuss the interspecific relationships among coexisting carnivores in the Transvaal (South Africa). Data on the species' presence are available for most of South Africa (Bruton, 1978; Pringle, 1977; Rowe-Rowe, 1978, 1992), and Angola (Crawford-Cabral, 1989a).

West Africa: General information on the ecology and distribution of the species is found in several authors (Clutton-Brock et al., 1976; Estes, 1991; Kingdon, 1997; Stuart & Stuart, 1997). Status and distribution are discussed in Ginsberg & Macdonald (1990).

Known extent of occurrence

Fig. 3.1.1.a Known extent of occurrence

The side-striped jackal occurs throughout a large portion of Africa, ranging from Senegal to Ethiopia, southwards to northern Namibia, northern Botswana, Zimbabwe, Mozambique and northern South Africa (Transvaal, Kwazulu and Natal provinces) (Wilson & Reeder, 1993) (Fig. 3.1.1.a). The West African population is largely unknown, and its range there was drawn on the basis of Dr. C. Sillero-Zubiri's notes (15 May '97). Ginsberg & Macdonald (1990) did not include this part of Africa in the species' range. In East Africa, *C. adustus* is

sympatric both with *C. aureus* and *C. mesomelas* (van Valkenburgh & Wayne, 1994).

Categorical-discrete (CD) distribution model

The species lives in a variety of moist savanna types, through woodlands and thickets to the edges of forest; common in cultivated and swampy areas (Estes, 1991; Ginsberg & Macdonald, 1990; Kingdon, 1971-77; Skinner & Smithers, 1990; Yalden et al., 1980).

Based on these environmental preferences, the following scores were assigned (Fig. 3.1.1.b) (3.1.1.a):

Score

1 Moist savannas and their mosaics; forest savanna mosaics.
2 Dry savannas; croplands.
3 Forests, deserts.

suitable		moderately suitable		unsuitable		Total	
km^2	%	km^2	%	km^2	%	km^2	%
6 396 233	62	2 972 892	29	913 384	9	10 282 509	100

Tab 3.1.1.a: Cumulative size (km^2) of areas pertaining to each environmental suitability class within the Extent of Occurrence.

	Number Patches (NP)	Mean Patch Size (MPS) km^2	Patch Size SD (PSSD) km^2	Largest Patch Index (LPI) %	Mean Shape Index (MSI)	Area-Weighted Mean Shape Index (AWMSI)
suitable	2 469	2 592	90 501	46.18	1.27	33.92
moderately suitable	6 987	425	13 472	9.19	1.27	17.38
Total AO	698	13 426	351 474	99.16	1.25	21.38

Tab 3.1.1.b: Area of Occupancy fragmentation indexes.

Probabilistic-continuous (PC) distribution model

The output of the probabilistic-continuous (PC) distribution model is shown in Fig. 3.1.1.c.

Validation

% of EO in sample areas	Number of valid plots	Overall Accuracy (%)
4.31	155	70.32

Tab 3.1.1.c: Categorical-discrete (CD) distribution model validation parameters.

Comments and conservation issues

Comparing the known EO with the two models shows two areas that would require further field work to find evidence of the species' presence or absence. The known EO includes a large area covering the forests of central-western Congo which is ranked as unsuitable by the models. Conversely most of southern Chad is excluded by the EO, while it is indicated as largely suitable by the models. The known EO should also be verified in its southern boundaries in West Africa where large suitable areas are found south of the known limits. The AO is very large (>90% of the total EO) but is relatively unfragmented, as shown by the high SD of patch size and the LPI (99.16%). In spite of the large AO, about 10% is included in the existing protected areas. The species' AO matches very nicely with those of the other two African jackals, confirming the good quality of the data for these species. This is also confirmed by the good performance of the CD model in the validation process (Overall Accuracy 70.32%).

	Inside	outside	Total
suitable	6.48	55.73	62.20
moderately suitable	2.92	25.99	28.91
unsuitable	0.63	8.26	8.88
Total	10.02	89.98	100

Tab 3.1.1.d: Percent of environmental suitability classes within EO (as obtained from the categorical-discrete distribution model) inside and outside the protected areas.

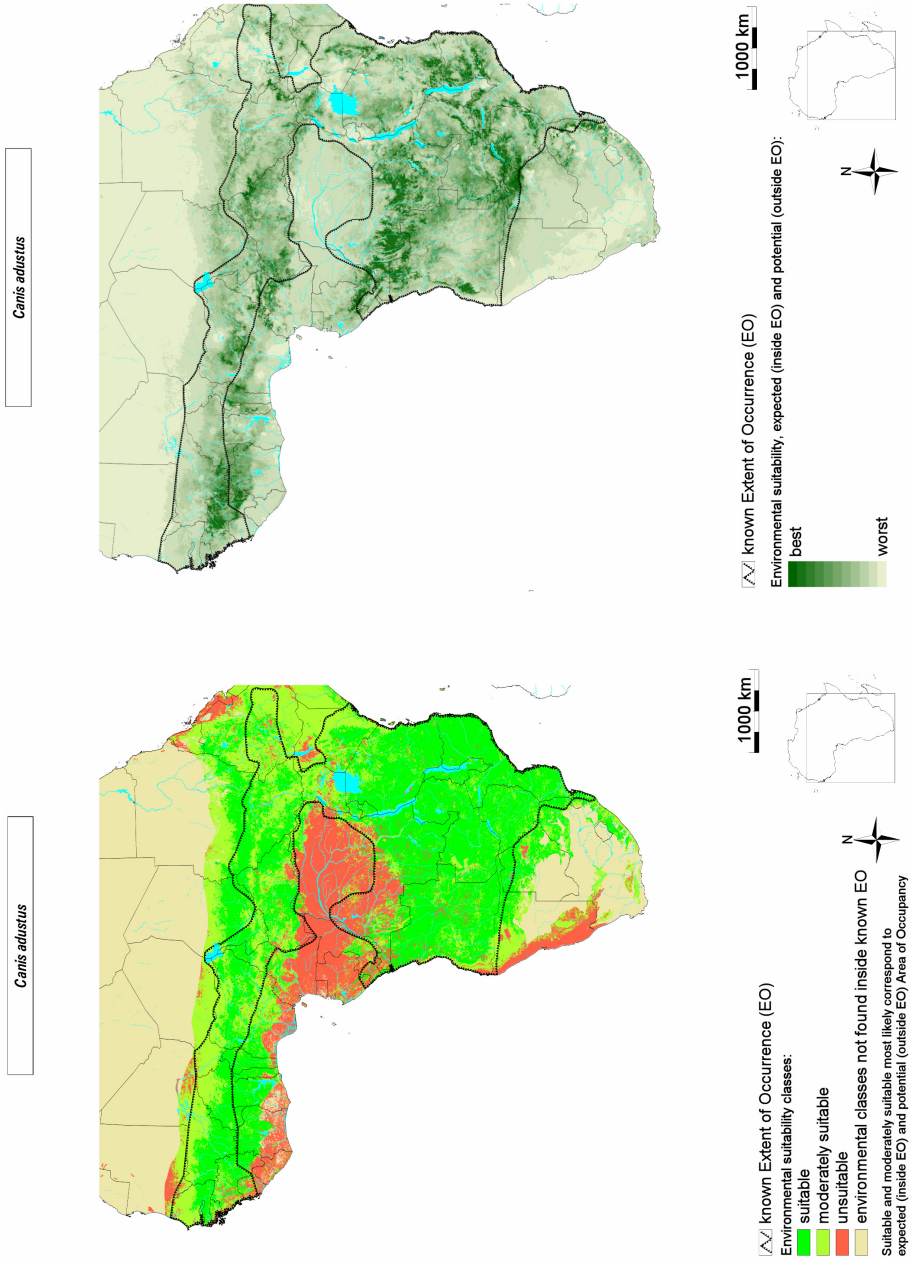


Fig. 3.1. 1 b (left) Categorical-discrete model; c (right) Probabilistic-continuous model

References

Bruton M.N. (1978). Recent mammal records from eastern Tongoland in Kwazulu, with notes on Hippopotamus in Lake Sibaka. Lammergeyer: 24, 19-27.

Clutton-Brock J., Corbet G.B., Hills M. (1976). A review of the family Canidae, with a classification by numerical methods. Bull. Br. Mus. (Nat. Hist.): 29, 117-199.

Crawford-Cabral J. (1989a). Distributional data and notes on Angolan carnivores (Mammalia: Carnivora). I - Small and medium-sized species. Garcia de Orta, Série de Zoologica: Lisboa 14(2), 3-27.

Crawford-Cabral J., (1989b): Noticia sobre a ocorrencia da Chacal-de-flancos-raiados (*Canis adustus* Sundevall) na Guine-Bissau. Garcia de Orta, Ser. Zool., Lisboa: 15(1), 25-28.

Duckworth J.W. (1992). Sighting frequencies of nocturnal mammals in an Ethiopian Rift Valley national park. Afr. J. Ecol.: 30, 90-97.

Dunham K.M. (1994). The effect of drought on the larger mammal populations of Zambezi riverine woodlands. J. Zool., Lond.: 234(3), 489-526.

Estes R.D. (1991). The Behavior Guide to African Mammals. Including Hoofed Mammals, Carnivores, Primates. The University of California Press.

Fuller T.K. et al., (1989): The ecology of three sympatric jackal species in the Rift Valley of Kenya. Afr. J. Ecol.: 27, 313-323.

Funaioli U. (1971). Guida breve dei mammiferi della Somalia. Istituto Agronomico per l'Oltremare. Biblioteca Agraria Tropicale.

Ginsberg J.R., Macdonald D. (1990). Foxes, Wolves, Jackals, and Dogs. An Action Plan for the Conservation of Canids. IUCN/SSC Canid Specialist Group.

Happold D.C.D., (1987): The mammals of Nigeria. Clarendon Press, Oxford.

Kingdon J. (1971-77). East African Mammals. VOL I: primates, hyraces, pangolins, protoungulates, sirenians. VOL IIA: Insectivores and bats. VOL IIB: hares and rodents. VOL IIIA: carnivores. VOL IIIB: large mammals. VOL IIIC: bovids. VOL IIID: bovids. Academic Press, London and New York: vol I - IIID.

Kingdon J. (1997). The Kingdon field guide to African Mammals. Academic Press, London and New York: Natural World.

Lamprecht J., (1978): On diet, foraging behaviour and interspecific food competition of jackals in the Serengeti National Park, East Africa. Z. Saug.: 43, 210-223.

Moehlman P.D. (1978). Jackals of the Serengeti. Wildlife News: 13(3), 2-6.

Pringle J.A. (1977). The distribution of mammals in Natal: 2. Carnivora. Ann. Natal Mus.: 23(1), 93-116.

Rahm U., Christiaensen A. (1966). Les mammiferès de l'Ile Idjwi (Lac Kivu, Congo). Les mammifères de la foret equatoriale de l'est du Congo. Ann. Mus. Roy. Afr. Cent.: n° 149.

Rautenbach I.L., Nel J.A.J. (1978). Coexistence in Transvaal Carnivora. Bull Car Mus Nat His: 6, 138-145.

Rosevear D.R., 1974: The carnivores of West Africa. British Museum (Nat. Hist.), London.

Rowe-Rowe D.T. (1978). The small carnivores of Natal. Lammergeyer: 25, 1-48.

Rowe-Rowe D.T. (1992). The carnivores of Natal. Natal Parks, Game & Fish Preservation Board, Pietermaritzburg, Natal, R.S.A.

Skinner J.D., Smithers R.H.N. (1990). The mammals of the Southern African subregion. Univ. of Pretoria, Pretoria.

Smithers R.H.N. & Wilson V.J., (1979): Check list and atlas of the mammals of Zimbabwe Rhodesia. Mus. Mem. Natl. Mus. Monum.Rhod.: 9, 1-147.

Stuart C., Stuart T. (1997). Field guide to the larger mammals of Africa. Struik Publishers.

van Valkenburgh B., Wayne R.K. (1994). Shape divergence associated with size convergence in sympatric East African jackals. Ecology: 75(6), 1567-1581.

Wilson D.E., Reeder D.M. (Eds) (1993). Mammal Species of the World: A Taxonomic and Geographic Reference. Second edition. Smithsonian Institution Press, Washington D.C.

Yalden D.W., Largen M.J., Kock D. (1980). Catalogue of the mammals of Ethiopia 4. Carnivora. Monitore Zoologico Italiano: suppl. XIII, 169-272.

Yalden D.W., Largen M.J., Kock D., Hillman J.C. (1996). Catalogue of the mammals of Ethiopia and Eritrea. 7. Revised checklist, zoogeography and conservation. Tropical Zoology: 9, 73-164.

Carnivora Id code: amd087

Canidae

Canis simensis

Rüppell, 1840

(Eng) Ethiopian wolf

(Fre) Loup d'Abyssinie

Taxonomic notes

Two subspecies have been described: *C. s. simensis* (north-west of the Rift) and *C. s . citernii* (south-east of the Rift) (Ginsberg & Macdonald, 1990).

IUCN threat category

Critically Endangered (CR: criteria A1b+2be, C1, E).

Available information

The ecology, distribution, and status of this wolf are described in Sillero-Zubiri & Macdonald (1997). Sillero-Zubiri & Gottelli (1994) reviewed literature and information available on this species. Information on the main aspects of the species' ecology are found in Gottelli & Sillero-Zubiri (1992), Morris & Malcolm (1978), and in Yalden & Largen (1992); its social behaviour, spacing pattern, and habits are also analysed by Gottelli & Sillero-Zubiri (1992) and Yalden & Largen (1992). Notes on its behaviour are found in Morris & Malcolm (1978). Habitat requirements are discussed in Morris & Malcolm (1978) and Sillero-Zubiri & Gottelli (1994). General information on the species' ecology and distribution are given by Kingdon (1997) and Stuart & Stuart (1997). Status and threats, together with notes on its ecology, particularly on its habitat requirements, are found in Ginsberg & Macdonald (1990) and in Gottelli & Sillero-Zubiri (1992).

Known extent of occurrence

Fig. 3.1.4.a Known extent of occurrence

C. simensis is endemic to the Ethiopian mountains, being present in the Arssi, Bale and Simen Mts., north-east Shoa, Gojjam and Mt. Guna (Ginsberg & Macdonald, 1990). Its distribution range was obtained by overlapping maps found in Sillero-Zubiri & Gottelli (1994), Ginsberg & Macdonald (1990) and Sillero-Zubiri & Macdonald (1997); the map was then revised and updated by Dr. C. Sillero-Zubiri (15 May '97; Nov. '98). The current range of the main subpopulation is rather well known, due to several research studies and monitoring projects carried out during the last 15 years.

Categorical-discrete (CD) distribution model

Optimal habitat is provided by Afro-alpine grasslands and heathlands. Also found in montane grasslands, sub-alpine heathlands and secondary *Helichrysum* dwarf-scrub (Sillero-Zubiri & Macdonald, 1997).

Based on these environmental preferences, the following scores were assigned (Fig. 3.1.4.b) (3.1.4.a):

Score

1 Altimontane grasslands and grassland mosaics.
2 Afro-montane bushlands, shrublands and grasslands.
3 Forests, woodlands, lowland bushlands & thickets, croplands.

suitable		Moderately suitable		unsuitable		Total	
km^2	%	Km^2	%	km^2	%	Km^2	%
2 864	25	3 520	31	2 864	25	3 520	100

Tab 3.1.4.a: Cumulative size (km^2) of areas pertaining to each environmental suitability class within the Extent of Occurrence.

	Number Patches (NP)	Mean Patch Size (MPS) km^2	Patch Size SD (PSSD) km^2	Largest Patch Index (LPI) %	Mean Shape Index (MSI)	Area-Weighted Mean Shape (AWMSI)
suitable	24	113	138	10.36	1.27	1.5
moderately suitable	25	143	223	16.33	1.28	2.03
Total AO	17	369	534	35.06	1.54	2.34

Tab 3.1.4.b: Area of Occupancy fragmentation indexes.

Probabilistic-continuous (PC) distribution model

The output of the probabilistic-continuous (PC) distribution model is shown in Fig. 3.1.4.c.

Validation

No occurrence of the species within the four sample areas.

Comments and conservation issues

The species is restricted to a very small and fragmented range, < 12 000 km^2, broken into at least 10 smaller sections. The two models indicate similar patterns of suitability: moreover, the PC model shows an interesting continuity of potentially suitable areas among some of the southern and central fragments. This feature should be further explored. AO is likely to be only 56% of the EO (< 6 500 km^2), and only one third pertains to the best suitable areas. The suitable patches within AO appear highly fragmented in small patches (mean size = 113 km^2); this seems to be confirmed by the small LPI (10.36%). The small patch

size is confirmed by the mean size of patches of the total AO (369 km^2) and the similarity of the two shape indexes also confirms a high fragmentation level. However, the LPI shows that at least one large patch which would include about 35% of the total AO is available: it is likely that this area supports the most viable portion of the population and it would deserve the highest conservation attention. Existing protected areas include an important portion of AO (>49%), but it appears urgent that these areas be effectively protected and that full protection be extended to the total AO. Given the small size of the population, the areas of occupancy, and the highly fragmented spatial patterns, the species is confirmed as Critically Endangered and requires the greatest conservation priority.

SUITABILITY CLASS	inside	outside	Total
suitable	9.51	15.66	25.16
moderately suitable	17.82	13.11	30.93
unsuitable	26.73	17.18	43.91
Total	54.06	45.94	100

Tab 3.1.4.c: Percent of environmental suitability classes within EO (as obtained from the categorical-discrete distribution model) inside and outside the protected areas.

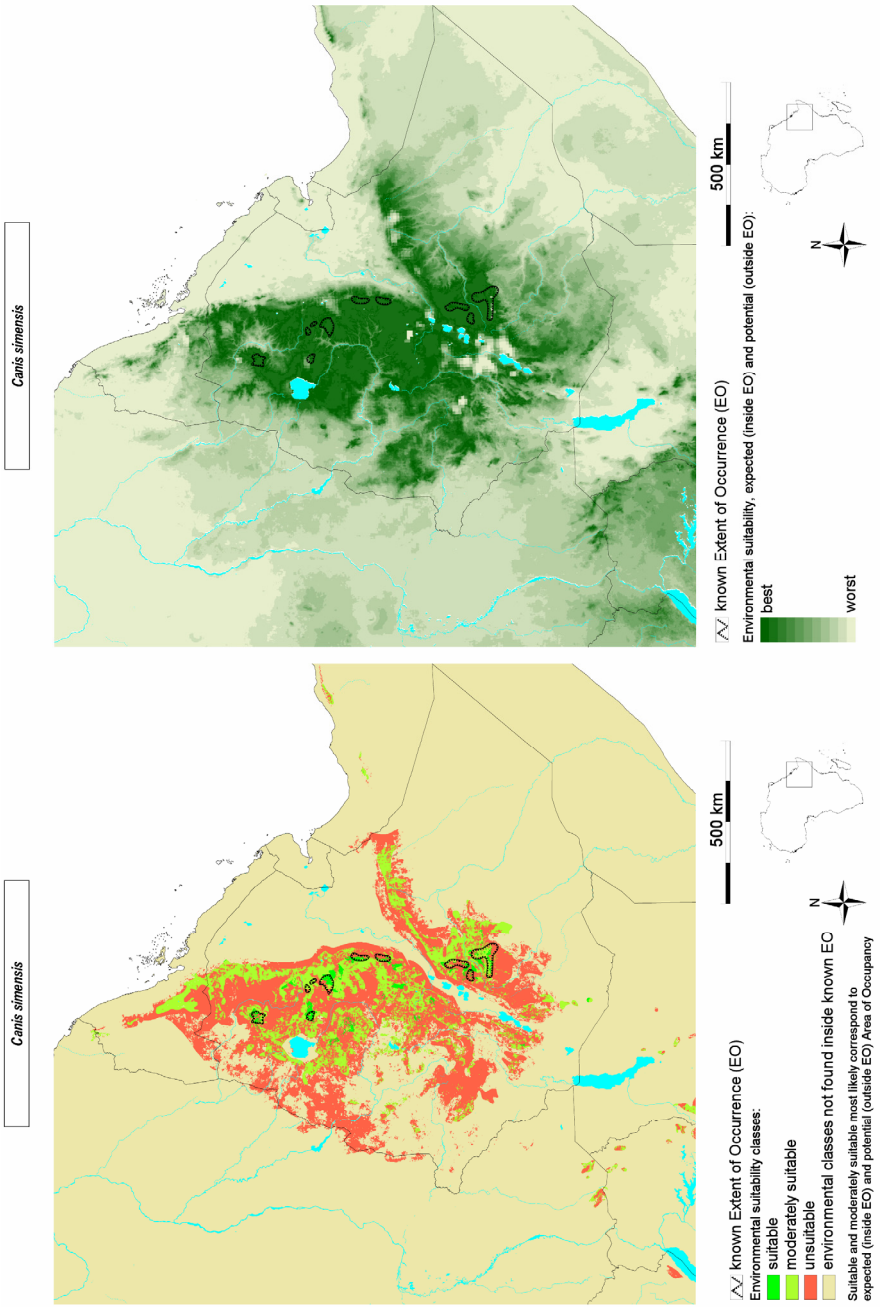


Fig. 3.1. 4 b (left) Categorical-discrete model; c (right) Probabilistic-continuous model

References

Ginsberg J.R., Macdonald D. (1990). Foxes, Wolves, Jackals, and Dogs. An Action Plan for the Conservation of Canids. IUCN/SSC Canid Specialist Group.
Gottelli D., Sillero-Zubiri C. (1992). The Ethiopian wolf - an endangered endemic canid. Oryx: 26(4), 205-214.
Kingdon. (1997). The Kingdon field guide to African Mammals. Academic Press
Morris P.A., Malcolm J.R. (1978). The Simen fox in the Balè Mountains. Oryx: 14, 151-160.
Sillero-Zubiri C., Gottelli D. (1994). *Canis simensis*. Mammalian Species: 485, 1-6.
Sillero-Zubiri C., Macdonald D. (Eds) (1997). The Ethiopian wolf. Status survey and conservation action plan. IUCN, Gland, Switzerland: Canid Spec. Group
Stuart C., Stuart T. (1997). Field guide to the larger mammals of Africa. Struik Publishers.
Yalden D.W., Largen M.J. (1992). The endemic mammals of Ethiopia. Mammal Rev.: 22 (3/4), 115-150.

Artiodactyla **ld code: amd175**

Bovidae

Kobus kob

(Erxleben, 1777)

(Eng) Kob

(Fre) Cobe de Buffon

Taxonomic notes

Ten subspecies have been described (Meester & Setzer, 1971), but only three are generally recognised as distinct (East, 1996).

IUCN threat category

Lower Risk, conservation dependent (LR: cd) as *Kobus kob,* as *K. k kob* (Buffon's kob) in Benin, Burkina Faso, Cameroon, C.A.R., Chad, Ivory Coast, Gambia (ex), Ghana, Guinea, Guinea-Bissau, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo (ex) and as *K. k. thomasi* (Uganda kob) in Kenya (ex), Sudan, Tanzania (ex), Uganda, former Zaire; while Lower Risk, near threatened (LR: nt) as *K. k. leucotis* (White-eared kob) in Ethiopia, Sudan, Uganda.

Available information

Literature available on the ecology of this antelope is quite rich, and the species has been studied in most of its range of distribution.

West and Central Africa: Several authors have reported data on its habitat use and requirements (Ajayi et al, 1981; De Bie, 1991; Geerling & Bokdam, 1973; Muhlenberg & Roth, 1985; Stark, 1986; Tchamba & Elkan, 1995). The authors mentioned above studied the species in protected areas, and also supply data on population numbers and density. Mertens (1985), Muhlenberg & Roth (1985), and Wanzie (1986) include information on population structure, mortality, and reproduction. A brief note on the ecology and presence of the species in former Zaire is found in Verschuren (1975).

East Africa: Most of the authors focus on behavioural aspects of the species' ecology, particularly its reproductive behaviour (Deutsch, 1994a; 1994b;

Deutsch & Ofezu , 1994; Fryxell, 1987). Deutsch (1994a) includes in his analysis a detailed description of its territorial behaviour and habitat selection in relation to its breeding behaviour. Information on habitat, diet, and habits are found in Fryxell (1985) and Fryxell & Sinclair (1988). The presence of the species in Ethiopia and Eritrea is discussed in Yalden et al. (1996), who also gives some information on its habitat.

General information on the species' ecology is found in Estes (1991), Kingdon (1997), Spinage (1982), and Stuart & Stuart (1997). East (1988, 1990) discusses its status and distribution in each country in its distribution range, and also gives information on its ecology, particularly its habitat requirements.

Known extent of occurrence

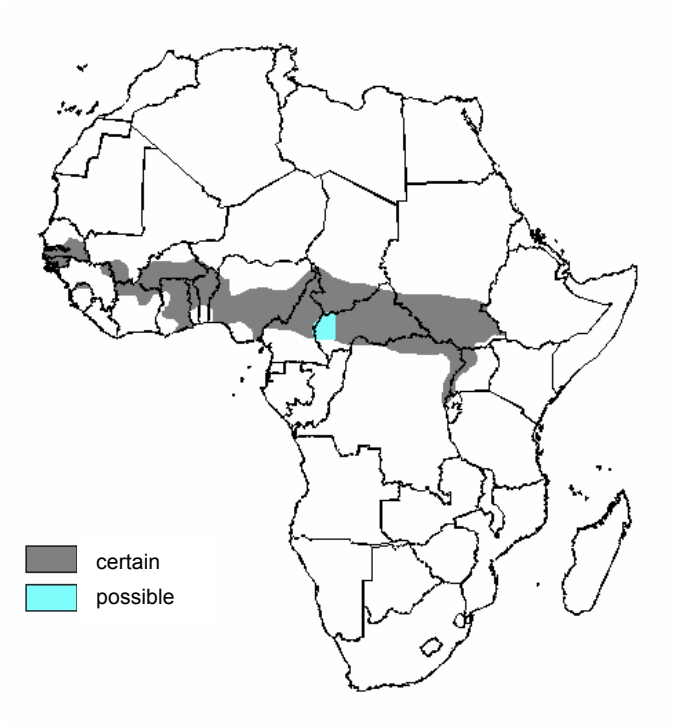


Fig. 8.6.73.a Known extent of occurrence

The kob occurs from Senegal to Sudan and Uganda (Wilson & Reeder, 1993) (Fig. 8.6.73.a). It is listed as extinct in Gambia, Kenya, Sierra Leone and Tanzania (East, 1996). Its distribution map was first obtained from Chardonnet (1995) and then corrected using the country maps in East (1988) and East (1990), as indicated by Dr. R. East (9 July '97).

Categorical-discrete (CD) distribution model

The species prefers open savanna woodland and grasslands close to water (East, 1988, 1990; Kingdon, 1997).

Based on these environmental preferences, the following scores were assigned (Fig. 8.6.73.b) (Tab. 8.6.73.a):

Score(*)

1 Grasslands, grassland mosaics and woodland mosaics and transitions.
2 Semi-arid vegetation, bushlands and thickets, savanna/forest mosaics.
3 Forests, deserts and croplands.

(*) Scores decreased for vegetation types occurring outside a 10-km buffer around permanent water

	suitable		moderately suitable		unsuitable		Total	
OCCURRENCE	km^2	%	km^2	%	km^2	%	km^2	%
certain	788 357	25	1 544 965	49	714 566	23	3 047 888	97
possible	3 629	0	64 746	2	12 132	0	80 507	3
Total	791 986	25	1 609 711	51	726 698	23	3 128 395	100

Tab 8.6.73.a: Cumulative size (km^2) of areas pertaining to each environmental suitability class within the Extent of Occurrence.

	Number Patches (NP)	Mean Patch Size (MPS) km^2	Patch Size SD (PSSD) km^2	Largest Patch Index (LPI) %	Mean Shape Index (MSI)	Area-Weighted Mean Shape Index (AWMSI)
suitable	1 180	672	10 606	13.96	1.31	13.51
moderately suitable	1 435	1 121	27 046	42.06	1.34	20.07
Total AO	487	4 930	103 275	94.99	1.29	20.61

Tab 8.6.73.b: Area of Occupancy fragmentation indexes.

Probabilistic-continuous (PC) distribution model

The output of the probabilistic-continuous (PC) distribution model is shown in Fig. 8.6.73.c.

Validation

% of EO in sample areas	Number of valid plots	Overall Accuracy (%)
8.91	98	59.18

Tab 8.6.73.c: Categorical-discrete (CD) distribution model validation parameters.

Comments and conservation issues

This species has a wide EO but only 25% appears to be suitable, while 49% is moderately suitable. These figures are supported by the results of the validation process (Overall Accuracy 59.81%). The area in western C.A.R. where the species' presence is possible is classified as moderately suitable. The most suitable areas are more frequent in the western part of the range and in southern Sudan. The small area in Sierra Leone does not appear to be of high quality. Both models indicate a rather unsuitable area along the border between former Zaire and Uganda. The AO is not significantly fragmented, but its shape might be very indented. About 8% of the total AO is included in existing protected areas. As the range is mostly in the Sahel region, where extensive poaching and competition with livestock occur, the amount and strategic distribution of protected AO is of particular importance for the conservation of this species.

OCCURRENCE	SUITABILITY CLASS	inside	outside	Total
certain	suitable	2	23.20	25.20
	moderately suitable	4.55	44.84	49.39
	unsuitable	1.84	21	22.84
possible	suitable	0	0.12	0.12
	moderately suitable	0	2.07	2.07
	unsuitable	0	0.39	0.39
Total		8.39	91.61	100

Tab 8.6.73.d: Percent of environmental suitability classes within EO (as obtained from the categorical-discrete distribution model) inside and outside the protected areas.

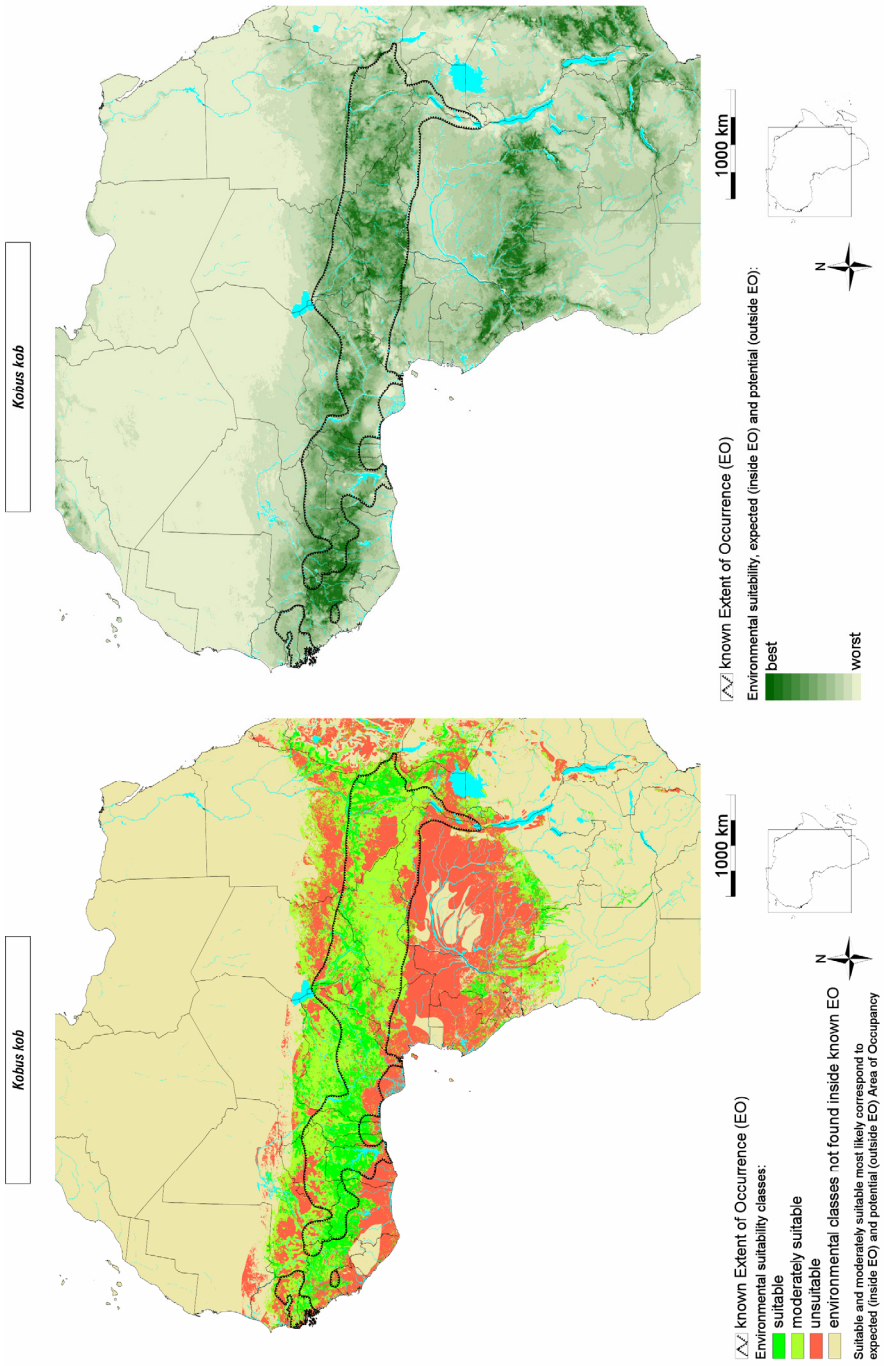


Fig. 3.6.6 b (left) Categorical-discrete model; c (right) Probabilistic-continuous model

References

Ajayi S.S., Afolayan T., Milligan K. (1981). A survey of wildlife in Kwiambana Game Reserve, Nigeria. Afr. J. Ecol.: 19, 295-298.

Chardonnet P. (1995). La ressource oubliée. Synthése. Faune Sauvage Africaine.Commission Européenne: vol . I.

De Bie S. (1991). Wildlife resources in the West African savanna. Wageningen Agricultural University Papers: 91 (2), 1-267.

Deutsch J. (1994a). Lekking by default: female habitat preferences and male strategies in Uganda Kob. J. Anim. Ecol.: 63, 101-115.

Deutsch J. (1994b). Uganda kob mating success does not increase on larger leks. Behav. Ecol. Sociobiol.: 34, 451-459.

Deutsch J.C., Ofezu A. A. (1994). Uganda Kob reproductive seasonality: optimal calving seasons or condition-dependent oestrus? Afr. J. Ecol.:32, 283-295.

East R. (Ed.) (1988). Antelope Global Survey and Regional Action plans, Part I: East and North East Africa. IUCN/SSC Antelope Specialist Group.

East R. (Ed.) (1990). Antelopes Global Survey and Regional Action Plans. Part 3: West and Central Africa. IUCN/SSC Antelope Specialist Group.

East R. (Ed.) (1996). Antelope survey update. IUCN/SSC Antelope Specialist Group: n° 2.

Estes R.D. (1991). The Behavior Guide to African Mammals. Including Hoofed Mammals, Carnivores, Primates. The University of California Press.

Fryxell J. (1985). Resource limitation and population ecology of the white-eared kob. Ph.D. Thesis: British Columbia.

Fryxell J. (1987). Seasonal reproduction of white-eared kob in Boma National Park, Sudan. Afr. J. Ecol.: 25, 117-124.

Fryxell J., Sinclair A.R.E. (1988). Seasonal migration by white-eared kob in relation to resources. Afr. J. Ecol.: 26, 17-31.

Geerling G., Bokdam J. (1973). Fauna of the Comoé National Park, Ivory Coast. Biological Conservation: 5(4), 251-257.

Kingdon J. (1997). The Kingdon field guide to African Mammals. Academic Press, London and New York: Natural World.

Meester J., Setzer H.W. (Eds) (1971). The Mammals of Africa, an Identification Manual. Smithsonian Institution Press, Washington D.C.

Mertens H. (1985). Structures de population et tables de survie des buffles, topis et cobs de buffon au parc National des Virungas, Zaire. Revue de Ecologie (La Terre et la Vie): 40, 33-51.

Muhlenberg M., Roth H.H. (1985). Comparative investigations into the ecology of the kob antelope *Kobus kob kob* (Erxleben, 1777) in the Comoe National Park, Ivory Coast. S. Afr. J. Wildl. Res.: 15, 25-31.

Spinage C.A. (1982). A territorial antelope: the Uganda waterbuck. Academic Press, London and New York.

Stark M.A. (1986). The numbers, seasonal distribution patterns, densities and biomass of the large herbivores, Benoue National Park, Cameroon. Mammalia: 50, 341-350.

Stuart C., Stuart T. (1997). Field guide to the larger mammals of Africa. Struik Publishers.

Tchamba M., Elkan P. (1995). Status and trends of some large mammals and ostriches in Waza National Park, Cameroon. Afr. J. Ecol.: 33(4), 366-376.

Verschuren J. (1975). Wildlife in Zaire. Oryx: 13(2), 149-163.

Wanzie C. (1986). Mortality factors of Buffon's kob *Kobus kob kob* (Erxleben) in Waza National Park, Cameroon. Mammalia: 50(3), 351-356.

Wilson D.E., Reeder D.M. (Eds) (1993). Mammal Species of the World: A Taxonomic and Geographic Reference. Second edition. Smithsonian Institution Press, Washington D.C.

Yalden D.W., Largen M.J., Kock D., Hillman J.C. (1996). Catalogue of the mammals of Ethiopia and Eritrea. 7. Revised checklist, zoogeography and conservation. Tropical Zoology: 9, 73-164.Meester J., Setzer H.W. (Eds) (1971). The Mammals of Africa, an Identification Manual. Smithsonian Institution Press, Washington D.C.

Appendix C
List of species included in the African Mammals Databank

INSECTIVORA	**INSECTIVORES**
Tenrecidae	Otter Shrews
Micropotamogale lamottei	Lesser Otter Shrew
Micropotamogale ruwenzorii	Ruwenzori Otter Shrew
Potamogale velox	Giant Otter Shrew
Erinaceidae	Hedgehogs
Atelerix albiventris	Four-toed Hedgehog
Atelerix algirus	Algerian Hedgehog
Atelerix frontalis	South African Hedgehog
Atelerix sclateri	Sclater's or Somali Hedgehog
Hemiechinus aethiopicus	Desert Hedgehog
Hemiechinus auritus	Long-eared Hedgehog

Primates	**PRIMATES**
Loridae	Lorisids
Arctocebus aureus	Golden Potto
Arctocebus calabarensis	Angwantibo
Perodicticus potto	Potto
Galagonidae	Galagos or Bushbabies
Eoticus elegantulus	Western Needle-clawed Bushbaby (South)
Eoticus pallidus	Western Needle-clawed Bushbaby (North)
Galago alleni	Allen's Bush Baby
Galago gallarum	Somali Galago or Bushbaby
Galago matschiei	Eastern Needle-clawed or Small Bushbaby
Galago moholi	S. Africa Lesser Bushbaby
Galago senegalensis	Lesser Bushbaby
Galagoides demidoff	Demidoff's Dwarf Bushbaby
Galagoides thomasi	Thomas's Dwarf Galago
Galagoides zanzibaricus	Zanzibar Galago
Otolemur crassicaudatus	Large-eared Greater Bushbaby
Otolemur garnetti	Garnett's Galago or Small-eared Greater Galago

Cercopithecidae	Cercophitecines and Colobines
Allenopithecus nigroviridis	Allen's Swamp Monkey
Cercocebus agilis	Agile or Golden-bellied Mangabey
Cercocebus galeritus	Crested or Tana River Mangabey
Cercocebus torquatus	White-collared or Sooty Mangabey
Cercopithecus ascanius	Redtail Guenon
Cercopithecus campbelli	Campbell's Guenon
Cercopithecus cephus	Moustached Guenon
Cercopithecus diana	Diana Monkey
Cercopithecus dryas	Dryas Guenon
Cercopithecus erythrogaster	Red-bellied Guenon
Cercopithecus erythrotis	Red-eared Guenon
Cercopithecus hamlyni	Owl-faced or Hamlyn's Guenon
Cercopithecus lhoesti	L'hoest's Guenon
Cercopithecus mitis	Blue, Sky or Diademed Monkey
Cercopithecus mona	Mona Guenon
Cercopithecus neglectus	De Brazza's Guenon
Cercopithecus nictitans	Spot-nosed Guenon
Cercopithecus petaurista	Lesser Spot-nosed Guenon
Cercopithecus pogonias	Crowned Guenon
Cercopithecus preussi	Preuss' Guenon
Cercopithecus sclateri	Sclatter's Guenon
Cercopithecus solatus	Sun-tailed Guenon
Cercopithecus wolfi	Wolf's Guenon
Chlorocebus aethiops	Green, Vervet, Grivet Monkey
Erythrocebus patas	Patas Monkey
Lophocebus albigena	Grey-cheeked Mangabey
Macaca sylvanus	Barbary Ape or Barbary Macaque
Mandrillus leucophaeus	Drill
Mandrillus sphinx	Mandrill
Miopithecus talapoin	Talapoin
Papio hamadryas	Hamadryas Baboon
Theropithecus gelada	Gelada Baboon
Colobus angolensis	Angolan Black-and-white Colobus
Colobus guereza	Guereza, Eastern Black-and-white Colobus
Colobus polykomos	Western Black-and-white Colobus
Colobus satanas	Black Colobus
Colobus vellerosus	White Thighted Black-and-white Colobus
Procolobus badius	Red Colobus

Procolobus verus	Olive Colobus
Hominidae	Apes
Gorilla gorilla	Gorilla
Pan paniscus	Pygmy Chimpanzee, Bonobo
Pan troglodytes	Chimpanzee
Carnivora	**carnivores**
Canidae	Dogs and allies
Canis adustus	Side-striped Jackal
Canis aureus	Common or Golden Jackal
Canis mesomelas	Black-backed Jackal
Canis simensis	Ethiopian Wolf
Lycaon pictus	African Wild Dog or Hunting Dog
Otocyon megalotis	Bat-eared Fox
Vulpes cana	Blanford's Fox
Vulpes chama	Cape Fox
Vulpes pallida	Sand or Pale Fox
Vulpes rueppelli	Ruppel's Sand Fox
Vulpes vulpes	Red Fox
Vulpes zerda	Fennec
Felidae	Cats
Acinonyx jubatus	Cheetah
Caracal caracal	Caracal
Felis chaus	Swamp Cat
Felis margarita	Sand Cat
Felis nigripes	Black-footed or Small-spotted Cat
Felis silvestris	African Wild Cat, Kaffir Cat
Leptailurus serval	Serval
Profelis aurata	Golden Cat
Panthera leo	Lion
Panthera pardus	Leopard
Herpestidae	Mongooses
Atilax paludinosus	Marsh Mongoose
Bdeogale crassicauda	Bushy-tailed Mongoose
Bdeogale jacksoni	Jackson's Mongoose
Bdeogale nigripes	Black-legged Mongoose

Crossarchus alexandri	Alexander's Mongoose
Crossarchus ansorgei	Angolan Ansorge's Cusimanse
Crossarchus obscurus	Kusimanse
Crossarchus platycephalus	Flat-headed Cusimanse
Cynictis penicillata	Yellow Mongoose
Dologale dybowskii	Pousargues' Mongoose
Galerella flavescens	Small Grey Mongoose
Galerella pulverulenta	Cape or Small Grey Mongoose
Galerella sanguinea	Slender Mongoose
Helogale hirtula	Somali Dwarf Mongoose
Helogale parvula	Dwarf Mongoose
Herpestes ichneumon	Ichneumon Egyptian Mongoose
Ichneumia albicauda	White-tailed Mongoose
Liberiictis kuhni	Liberian Mongoose
Mungos gambianus	Gambian Mongoose
Mungos mungo	Banded Mongoose
Paracynictis selousi	Selous' Mongoose
Rhynchogale melleri	Meller's Mongoose
Suricata suricatta	Suricate, Meerkat
Xenogale naso	Long-nosed Mongoose
Hyaenidae	Hyaenids
Crocuta crocuta	Spotted Hyaena
Hyaena brunnea	Brown Hyaena
Hyaena hyaena	Striped Hyaena
Proteles cristatus	Aardwolf
Mustelidae	Mustelids
Aonyx capensis	Clawless Otter
Aonyx congicus	Congo Clawless Otter
Lutra lutra	Common Otter
Lutra maculicollis	Spotted-necked Otter
Mellivora capensis	Ratel or Honey Badger
Ictonyx libyca	Libyan or North African Striped Weasel
Ictonyx striatus	Zorilla or Striped Polecat
Mustela nivalis	Weasel
Mustela putorius	European Polecat

Poecilogale albinucha	Striped or African Weasel
Viverridae	Genets and Civets
Nandinia binotata	African Palm or Tree Civet
Civettictis civetta	African Civet
Genetta abyssinica	Abyssinian Genet
Genetta angolensis	Angolan Genet
Genetta cristata	Crested Genet
Genetta genetta	European Common Genet
Genetta johnstoni	Johnston's Genet
Genetta pardina	Pardine Genet
Genetta rubiginosa	Large Spotted Genet
Genetta servalina	Servaline Genet
Genetta thierryi	False or Hausa Genet
Genetta tigrina	Tigrine Genet
Genetta victoriae	Giant Forest Genet
Osbornictis piscivora	Aquatic or Fishing Civet
Poiana richardsoni	African Linsang or Oyan
Sirenia	**sirenians**
Trichechidae	Manatees
Trichechus senegalensis	African Manatee
Perissodactyla	**odd-toed ungulates**
Equidae	Horses
Equus africanus	Wild Ass
Equus burchellii	Burchell's or Plain Zebra
Equus grevyi	Plain Zebra
Equus zebra	Mountain Zebra
Hyracoidea	**hyraxes**
Procaviidae	Hyraxes
Dendrohyrax arboreus	Southern Tree Hyrax
Dendrohyrax dorsalis	Western Tree Hyrax
Dendrohyrax validus	Eastern Tree Hyrax
Heterohyrax antinae	Ahaggar Hyrax

Heterohyrax brucei	Bruce's Hyrax or Yellow Spotted Rockdassie
Procavia capensis	Cape Hyrax
Tubulidentata	**aardvark**
Orycteropidae	Aardvark
Orycteropus afer	Aardvark, Antbear
Artiodactyla	**even-toed ungulates**
Suidae	Pigs
Phacochoerus aethiopicus	Desert Warthog
Phacochoerus africanus	Common Warthog
Hylochoerus meinertzhageni	Giant Forest Hog
Potamochoerus larvatus	Bushpig
Potamochoerus porcus	Red River Hog
Sus scrofa	Wild Boar
Hippopotamidae	Hippopotamuses
Hexaprotodon liberiensis	Pigmy Hippopotamus
Hippopotamus amphibius	Common Hippopotamus
Tragulidae	Chevrotains
Hyemoschus aquaticus	Water Chevrotain
Giraffidae	Giraffes
Giraffa camelopardalis	Giraffe
Okapia johnstoni	Okapi
Cervidae	Deer
Cervus elaphus	Barbary stag
Bovidae	Bovids, Horned Ungulates
Aepyceros melampus	Impala
Alcelaphus buselaphus	(Red) Hartebeest
Connochaetes gnou	White-tailed Gnu, Black Wildebeest
Connochaetes taurinus	Brindled Gnu, Blue Wildebeest
Damaliscus hunteri	Hunter's Antelope, Hirola
Damaliscus lunatus	Topi, Sassaby, Tsessebe

Damaliscus pygargus	Bontebok, Blesbok
Sigmoceros lichtensteinii	Lichtenstein 's Hartebeest, Konzi
Ammodorcas clarkei	Dibatag
Antidorcas marsupialis	Springbok
Dorcatragus megalotis	Beira
Gazella cuvieri	Edmi Gazelle, Atlas' Gazelle
Gazella dama	Addra, Dama Gazelle
Gazella dorcas	Dorcas Gazelle
Gazella granti	Grant's Gazelle
Gazella leptoceros	Rhim or Slender-horned Gazelle
Gazella rufifrons	Red-fronted Gazelle
Gazella soemmerringi	Soemmering's Gazelle
Gazella spekei	Speke's Gazelle
Gazella thomsonii	Thomson's Gazelle
Litocranius walleri	Gerenuk
Madoqua guentheri	Guenther's Dik-dik
Madoqua kirki	Kirk's or Damara Dik-dik
Madoqua piacentini	Piacentini's Dik-dik
Madoqua saltiana	Salt's Dik-dik
Neotragus batesi	Bates' Pigmy Antelope, Dwarf Antelope
Neotragus moschatus	Suni
Neotragus pygmaeus	Royal Antelope
Oreotragus oreotragus	Klipspringer
Ourebia ourebi	Oribi
Raphicerus campestris	Steinbok
Raphicerus melanotis	Cape Grysbok
Raphicerus sharpei	Sharp's Grysbok
Syncerus caffer	African Buffalo
Taurotragus derbianus	Derby, Giant Eland
Taurotragus oryx	Common Eland
Tragelaphus angasii	Nyala
Tragelaphus buxtoni	Mountain Nyala
Tragelaphus eurycerus	Bongo
Tragelaphus imberbis	Lesser Kudu
Tragelaphus scriptus	Bushbuck
Tragelaphus spekii	Sitatunga
Tragelaphus strepsiceros	Greater Kudu

Ammotragus lervia	Barbary Sheep, Aoudad
Capra nubiana	Nubian Ibex
Capra walie	Walia or Abyssinian Ibex
Cephalophus adersi	Ader's Duiker
Cephalophus callipygus	Peter's Duiker
Cephalophus dorsalis	Bay Duiker
Cephalophus harvey	Harvey's Red Duiker
Cephalophus jentinki	Jentink's Duiker
Cephalophus leucogaster	White-bellied or Gabon Duiker
Cephalophus maxwellii	Maxwell's Duiker
Cephalophus monticola	Blue Duiker
Cephalophus natalensis	Red Duiker
Cephalophus niger	Black Duiker
Cephalophus nigrifrons	Black-fronted Duiker
Cephalophus ogilbyi	Ogilby's Duiker
Cephalophus rubidus	Ruwenzori Black-fronted Duiker
Cephalophus rufilatus	Red-flanked Duiker
Cephalophus silvicultor	Yellow-backed Duiker
Cephalophus spadix	Abbott's Duiker
Cephalophus weynsi	Weyns' Duiker
Cephalophus zebra	Zebra Duiker
Sylvicapra grimmia	Common, Grey or Bush Duiker
Addax nasomaculatus	Addax
Hippotragus equinus	Roan Antelope
Hippotragus niger	Sable Antelope
Oryx dammah	Scimitar-horned Oryx
Oryx gazella	Oryx or Gemsbok
Pelea capreolus	Vaal or Grey Rhebok
Kobus ellipsiprymnus	Common Waterbuck
Kobus kob	Kob
Kobus leche	Lechwe
Kobus megaceros	Nile Lechwe, Mrs Gray's Lechwe
Kobus vardonii	Puku
Redunca arundinum	Common or Southern Reedbuck
Redunca fulvorufula	Mountain Reedbuck
Redunca redunca	Bohor Reedbuck

Pholidota	**scaly ant-eaters**
Manidae	Pangolins
Manis gigantea	Giant Ground Pangolin
Manis temminkii	Cape or Temninck's Ground Pangolin
Manis tetradactyla	Long-tailed Pangolin
Manis tricuspis	Three-cusped or White-bellied Pangolin

Rodentia	**rodens**
Dipodidae	Jerboas
Allactaga tetradactyla	Four-toed Jerboa
Jaculus jaculus	Small Desert Jerboa, Jumping Mouse
Jaculus orientalis	Large Desert Jerboa
Pedetidae	Spring Hares
Pedetes capensis	Spring Hare
Hystricidae	Pocupines
Atherurus africanus	African Brush-tailed Porcupine
Hystrix africaeaustralis	South African Crested or Cape Porcupine
Hystrix cristata	North African Crested Porcupine

Lagomorpha	**hares**
Leporidae	Hares
Bunolagus monticularis	Bushman Hare, Riverine Rabbit
Lepus capensis	Cape Hare, Brown Hare
Lepus fagani	Ethiopian Hare
Lepus saxatilis	Scrub Hare
Lepus starcki	Ethiopian Highland Hare
Lepus victoriae	African Savanna or Whyte's Hare
Oryctolagus cuniculus	European Rabbit
Poelagus marjorita	Uganda Grass Hare or Bunyoro Rabbit
Pronolagus crassicaudatus	Natal Red Hare
Pronolagus randensis	Jameson's Red Rockhare
Pronolagus rupestris	Smith's Red Hare

Macroscelidea	**elephant shrews, or sengis**
Macroscelididae	Elephant Shrews
Elephantulus brachyrhynchus	Short-snouted Elephant Shrew
Elephantulus edwardii	Cape Elephant Shrew
Elephantulus fuscipes	Short-nosed or Dusky-footed Elephant Shrew
Elephantulus fuscus	Dusky or Peter's Short Snouted Elephant Shrew
Elephantulus intufi	Bushveld Elephant Shrew
Elephantulus myurus	Cliff or Eastern Rock Elephant Shrew
Elephantulus revoili	Somali Elephant Shrew
Elephantulus rozeti	North African Elephant Shrew
Elephantulus rufescens	Rufous, Spectacled, Long-eared Elephant Shrew
Elephantulus rupestris	Western Rock Elephant Shrew
Macroscelides proboscideus	Round- or Short-eared Elephant Shrew
Petrodromus tetradactylus	Four-toed or Forest Elephant Shrew
Rhynchocyon chrysopygus	Golden-rumped Elephant-shrew
Rhynchocyon cirnei	Giant or Chequered Elephant Shrew
Rhynchocyon petersi	Black-and-rufous Elephant Shrew

Appendix D
List of species included in the Italian Ecological Network

CLASS CEPHALASPIDOMORPHA	
Lampetra fluviatilis	Lampreda di fiume
Petromyzon marinus	Lampreda di mare
Lampetra planeri	Lampreda di ruscello
Lethenteron zanandreai	Lampreda padana
CLASS ACTINOPTERYGII	
Abramis brama	Abramide
Gymnocephalus cernuus	Acerina
Alburnus arborella	Alborella
Alburnus albidus	Alborella meridionale
Anguilla anguilla	Anguilla
Barbus tyberinus	Barbo appenninico
Barbus caninus	Barbo canino
Barbus barbus	Barbo europeo
Barbus plebejus	Barbo padano
Abramis bjoerkna	Blicca
Coregonus macrophthalmus	Bondella
Lota lota	Bottatrice
Salaria fluviatilis	Cagnetto
Carassius auratus	Carassio dorato
Cyprinus carpio	Carpa
Hypophthalmichthys molitrix	Carpa argentata
Ctenopharyngodon idellus	Carpa erbivora
Aristichthys nobilis	Carpa testa grossa
Salmo carpio	Carpione del Garda
Salmo fibreni	Carpione del Fibreno
Leuciscus cephalus	Cavedano
Leuciscus lucumonis	Cavedano di ruscello
Alosa agone	Cheppia-agone
Barbatula barbatula	Cobite barbatello
Sabanejewia larvata	Cobite mascherato
Cobitis bilineata	Cobite padano
Coregonus fera	Coregone
Gambusia holbrooki	Gambusia
Pomatoschistus canestrinii	Ghiozzetto cenerino

Knipowitschia panizzae	Ghiozzetto di laguna
Padogobius nigricans	Ghiozzo appenninico
Padogobius bonelli	Ghiozzo padano
Gobio gobio	Gobione europeo
Gobio benacensis	Gobione padano
Chondrostoma genei	Lasca
Atherina boyeri	Latterino
Esox lucius	Luccio
Sander lucioperca	Lucioperca
Barbus comizo	Messinobarbo
Misgurnus fossilis	Misgurno
Chondrostoma nasus	Naso
Aphanius fasciatus	Nono
Pachychilon pictum	Pachichilo
Knipowitschia punctatissima	Panzarolo
Perca fluviatilis	Persico reale
Lepomis gibbosus	Persico sole
Micropterus salmoides	Persico trota
Clarias sp.	Pesce gatto africano
Ameiurus nebulosus	Pesce gatto bruno
Ictalurus melas	Pesce gatto nero
Ictalurus punctatus	Pesce gatto punteggiato
Odonthestes bonariensis	Pesce re
Rutilus pigus	Pigo
Pseudorasbora parva	Pseudorasbora
Rhodeus amarus	Rodeo amaro
Rutilus rubilio	Rovella
Rutilus rutilus	Rutilo
Salvelinus alpinus	Salmerino
Salvelinus fontinalis	Salmerino di fonte
Phoxinus phoxinus	Sanguinerola
Chondrostoma soetta	Savetta
Scardinius scardafa	Scardola appenninica
Scardinius erythrophthalmus	Scardola comune
Cottus gobio	Scazzone
Silurus glanis	Siluro
Gasterosteus aculeatus	Spinarello
Acipenser naccarii	Storione cobice
Acipenser sturio	Storione comune
Huso huso	Storione ladano

Thymallus thymallus	Temolo
Tinca tinca	Tinca
Rutilus aula	Triotto
Oncorhynchus mykiss	Trota iridea
Salmo marmoratus	Trota marmorata
Salmo trutta	Trota mediterranea
Salmo cettii	Trota mediterranea
Telestes muticellus	Vairone
Telestes agassii	Vairone danubiano
CLASS AMPHIBIA	
Discoglossus pictus	Discoglosso dipinto
Discoglossus sardus	Discoglosso sardo
Speleomantes italicus	Geotritone italiano
Speleomantes ambrosii	Geotritone occidentale
Speleomantes flavus	Geotritone sardo giallo
Speleomantes genei	Geotritone sardo sud-occiden.
Speleomantes imperialis	Geotritone sardo sud-orient.
Speleomantes strinatii	Geotritone di Strinati
Speleomantes supramontis	Geotritone del Supramonte
Pelobates fuscus	Pelobate fosco italiano
Pelodytes punctatus	Pelodite punteggiato
Proteus anguis	Proteo
Hyla arborea + intermedia	Raganella comune e r. italiana
Hyla meridionalis	Raganella mediterranea
Hyla sarda	Raganella tirrenica
Rana dalmatina	Rana agile
Rana italica	Rana appenninica
Rana latastei	Rana di Lataste
Rana lessonae et esculenta COMPLEX	Rana di Lessona e Rana verde
Rana temporaria	Rana temporaria
Rana catesbeiana	Rana toro
Rana ridibunda	Rana verde maggiore
Bufo bufo	Rospo comune
Bufo viridis	Rospo smeraldino
Salamandra atra	Salamandra alpina, s. nera
Salamandra lanzai	Salamandra di Lanza
Salamandra salamandra	Salamandra pezzata
Salamandrina terdigitata	Salamandrina dagli occhiali
Triturus alpestris	Tritone alpino
Triturus carnifex	Tritone crestato italiano

Triturus italicus	Tritone italiano
Triturus vulgaris	Tritone punteggiato
Euproctus platycephalus	Tritone sardo (Euprotto sardo)
Bombina variegata	Ululone dal ventre giallo
CLASS REPTILIA	
Algyroides fitzingeri	Algiroide di Fitzinger
Algyroides nigropunctatus	Algiroide magnifico
Coluber viridiflavus	Biacco
Natrix tessellata	Biscia tessellata
Elaphe quatuorlineata	Cervone
Coluber gemonensis	Colubro dei Balcani
Macroprotodon cucullatus	Colubro dal cappuccio
Coluber hippocrepis	Colubro ferro di cavallo
Malpolon monspessulanus	Colubro lacertino
Elaphe situla	Colubro leopardino
Coronella austriaca	Colubro liscio
Coronella girondica	Colubro di Riccioli
Cyrtodactylus kotschyi	Geco di Kotschy
Hemidactylus turcicus	Geco verrucoso
Chalcides ocellatus	Gongilo
Podarcis melisellensis	Lucertola adriatica
Lacerta agilis	Lucertola agile
Archaeolacerta bedriagae	Lucertola di Bedriaga
Podarcis sicula	Lucertola campestre
Lacerta horvathi	Lucertola di Horvath
Podarcis filfolensis	Lucertola di Malta
Podarcis muralis	Lucertola muraiola
Timon lepidus	Lucertola ocellata
Podarcis wagleriana	Lucertola sicilian
Psammodromus algirus	Lucertola striata comune
Podarcis tiliguerta	Lucertola tirrenica
Zootoca vivipara	Lucertola vivipara
Chalcides chalcides	Luscengola
Vipera berus	Marasso
Natrix natrix	Natrice dal collare
Natrix maura	Natrice viperina
Anguis fragilis	Orbettino
Ophisaurus apodus	Pseudopo europeo, Ofisauro eu.
Lacerta viridis + bilineata	Ramarro occidentale + oriental
Elaphe longissima	Saettone, Colubro di Esculapio

Telescopus fallax	Serpente gatto, Colubro gatto
Tarentola mauritanica	Tarantola muraiola
Phyllodactylus europaeus	Tarantolino
Testudo hermanni	Testuggine comune
Emys orbicularis	Testuggine palustre europea
Vipera aspis	Vipera comune
Vipera ammodytes	Vipera dal corno
Vipera ursinii	Vipera dell'Orsini
CLASS AVES	
Egretta alba	Airone bianco maggiore
Ardea cinerea	Airone cenerino
Bubulcus ibis	Airone guardabuoi
Ardea purpurea	Airone rosso
Circus pygargus	Albanella minore
Strix aluco	Allocco
Strix uralensis	Allocco degli Urali
Alauda arvensis	Allodola
Anas crecca	Alzavola
Hieraaetus fasciatus	Aquila del Bonelli
Aquila chrysaetos	Aquila reale
Otus scops	Assiolo
Accipiter gentilis	Astore
Lanius senator	Averla capirossa
Lanius minor	Averla cenerina
Lanius collurio	Averla piccola
Recurvirostra avosetta	Avocetta
Delichon urbica	Balestruccio
Ficedula albicollis	Balia dal collare
Motacilla alba	Ballerina bianca
Motacilla cinerea	Ballerina gialla
Tyto alba	Barbagianni
Panurus biarmicus	Basettino
Scolopax rusticola	Beccaccia
Haematopus ostralegus	Beccaccia di mare
Gallinago gallinago	Beccaccino
Sylvia borin	Beccafico
Cisticola juncidis	Beccamoschino
Sterna sandvicensis	Beccapesci
Calonectris diomedea	Berta maggiore
Puffinus yelkouan	Berta minore

Circaetus gallicus	Biancone
Sylvia hortensis	Bigia grossa
Sylvia nisoria	Bigia padovana
Sylvia curruca	Bigiarella
Melanocorypha calandra	Calandra
Calandrella brachydactyla	Calandrella
Anthus campestris	Calandro
Anas strepera	Canapiglia
Hippolais polyglotta	Canapino
Acrocephalus scirpaceus	Cannaiola
Acrocephalus palustris	Cannaiola verdognola
Acrocephalus arundinaceus	Cannareccione
Sylvia atricapilla	Capinera
Neophron percnopterus	Capovaccaio
Galerida cristata	Cappellaccia
Carduelis carduelis	Cardellino
Himantopus himantopus	Cavaliere d'Italia
Turdus pilaris	Cesena
Ciconia ciconia	Cicogna bianca
Ciconia nigra	Cicogna nera
Cygnus olor	Cigno reale
Parus palustris	Cincia bigia
Parus montanus	Cincia bigia alpestre
Parus cristatus	Cincia dal ciuffo
Parus ater	Cincia mora
Parus major	Cinciallegra
Parus caeruleus	Cinciarella
Athene noctua	Civetta
Aegolius funereus	Civetta capogrosso
Glaucidium passerinum	Civetta nana
Pyrrhula phyrrula	Ciuffolotto
Aegithalos caudatus	Codibugnolo
Phoenicurus phoenicurus	Codirosso
Phoenicurus ochruros	Codirosso spazzacamino
Monticola saxatilis	Codirossone
Columba palumbus	Colombaccio
Columba oenas	Colombella
Phalacrocorax carbo	Cormorano
Corvus corone	Cornacchia
Charadrius dubius	Corriere piccolo

Corvus corax	Corvo imperiale
Alectoris graeca	Coturnice
Loxia curvirostra	Crociere
Cuculus canorus	Cuculo
Clamator glandarius	Cuculo dal ciuffo
Oenanthe oenanthe	Culbianco
Motacilla flava	Cutrettola
Phasianus colchicus	Fagiano comune
Tetrao tetrix	Fagiano di monte
Falco vespertinus	Falco cuculo
Circuì aeroginosus	Falco di palude
Pernis apivorus	Falco pecchiaiolo
Falco eleonorae	Falco della regina
Carduelis cannabina	Fanello
Phoenicopterus ruber	Fenicottero
Regulus ignicapillus	Fiorrancino
Netta rufina	Fistione turco
Fulica atra	Folaga
Acrocephalus schoenobaenus	Forapaglie
Acrocephalus melanopogon	Forapaglie castagnolo
Bonasa bonasia	Francolino di monte
Sterna albifrons	Fraticello
Charadrius alexandrinus	Fratino
Fringilla coelebs	Fringuello
Montifringilla nivalis	Fringuello alpino
Coccothraustes coccothraustes	Frosone
Larus ridibundus	Gabbiano comune
Larus melanocephalus	Gabbiano corallino
Larus audouinii	Gabbiano corso
Larus cachinnans	Gabbiano reale
Larus genei	Gabbiano roseo
Tetrax tetrax	Gallina prataiola
Gallinula chloropus	Gallinella d'acqua
Tetrao urogallus	Gallo cedrone
Egretta garzetta	Garzetta
Pica pica	Gazza
Anas platyrhynchos	Germano reale
Falco tinnunculus	Gheppio
Garrulus glandarius	Ghiandaia
Coracias garrulus	Ghiandaia marina

Pyrrhocorax graculus	Gracchio alpino
Pyrrhocorax pyrrhocorax	Gracchio corallino
Gyps fulvus	Grifone
Falco naumanni	Grillaio
Merops apiaster	Gruccione
Asio otus	Gufo comune
Bubo bubo	Gufo reale
Falco biarmicus	Lanario
Falco subbuteo	Lodolaio
Carduelis spinus	Lucarino
Phylloscopus bonelli	Lui bianco
Phylloscopus collybita	Lui piccolo
Phylloscopus sibilatrix	Lui verde
Sylvia undata	Magnanina
Sylvia sarda	Magnanina sarda
Phalacrocorax aristotelis	Marangone dal ciuffo
Phalacrocorax pygmeus	Marangone minore
Alcedo atthis	Martin pescatore
Anas querquedula	Marzaiola
Turdus merula	Merlo
Cinclus cinclus	Merlo acquaiolo
Turdus torquatus	Merlo dal collare
Anas clypeata	Mestolone
Emberiza schoeniclus	Migliarino di palude
Plegadis falcinellus	Mignattaio
Chlidonias niger	Mignattino
Chlidonias leucopterus	Mignattino alibianche
Chlidonias hybridus	Mignattino piombato
Oenanthe hispanica	Monachella
Aythya fuligula	Moretta
Aythya nyroca	Moretta tabaccata
Aythya ferina	Moriglione
Milvus migrans	Nibbio bruno
Milvus milvus	Nibbio reale
Nycticorax nycticorax	Nitticora
Nucifraga caryocatactes	Nocciolaia
Sylvia melanocephala	Occhiocotto
Burhinus oedicnemus	Occhione
Carduelis flammea	Organetto
Emberiza hortulana	Ortolano

Passer italiae	Passera d'Italia
Petronia petronia	Passera lagia
Passer montanus	Passera mattugia
Passer domesticus	Passera oltremontana
Passer hispaniolensis	Passera sarda
Prunella modularis	Passera scopaiola
Monticola solitarius	Passero solitario
Vanellus vanellus	Pavoncella
Falco peregrinus	Pellegrino
Remiz pendolinus	Pendolino
Lagopus mutus	Pernice bianca
Glareola pratincola	Pernice di mare
Alectoris rufa	Pernice rossa
Alectoris barbara	Pernice sarda
Tringa totanus	Pettegola
Erithacus rubecula	Pettirosso
Picus canus	Picchio cenerino
Picoides leucotos	Picchio dorsobianco
Tichodroma muraria	Picchio muraiolo
Sitta europaea	Picchio muratore
Muscicapa striata	Pigliamosche
Dryocopus martius	Picchio nero
Picoides major	Picchio rosso maggiore
Picoides medius	Picchio rosso mezzano
Picoides minor	Picchio rosso minore
Picoides tridactylus	Picchio tridattilo
Picus viridis	Picchio verde
Columba livia	Piccione selvatico
Actitis hypoleucos	Piro piro piccolo
Limosa limosa	Pittima reale
Charadrius morinellus	Piviere tortolino
Buteo buteo	Poiana
Porphyrio porphyrio	Pollo sultano
Rallus acquaticus	Porciglione
Anthus trivialis	Prispolone
Coturnix coturnix	Quaglia
Certhia brachydactyla	Rampichino
Certhia familiaris	Rampichino alpestre
Crex crex	Re di quaglie
Regulus regulus	Regolo

Oriolus oriolus	Rigogolo
Hirundo rustica	Rondine
Ptyonoprogne rupestris	Rondine montana
Hirundo daurica	Rondine rossiccia
Apus apus	Rondone
Apus melba	Rondone maggiore
Apus pallidus	Rondone pallido
Locustella luscinioides	Salciaiola
Saxicola torquata	Saltimpalo
Porzana parva	Schiribilla
Troglodytes troglodytes	Scricciolo
Ardeola ralloides	Sgarza ciuffetto
Prunella collaris	Sordone
Accipiter nisus	Sparviere
Platalea leucorodia	Spatola
Anthus spinoletta	Spioncello
Perdix perdix	Starna
Sterna hirundo	Sterna comune
Sterna bengalensis	Sterna del Ruppel
Gelochelidon nilotica	Sterna zampenere
Sylvia communis	Sterpazzola
Sylvia conspicillata	Sterpazzola di Sardegna
Sylvia cantillans	Sterpazzolina
Saxicola rubetra	Stiaccino
Sturnus vulgaris	Storno
Sturnus unicolor	Storno nero
Miliaria calandra	Strillozzo
Caprimulgus europaeus	Succiacapre
Podiceps cristatus	Svasso maggiore
Corvus monedula	Taccola
Ixobrychus minutus	Tarabusino
Botaurus stellaris	Tarabuso
Riparia riparia	Topino
Jynx torquilla	Torcicollo
Turdus viscivorus	Tordela
Turdus philomelos	Tordo bottaccio
Streptopelia turtur	Tortora
Streptopelia decaocto	Tortora dal collare orientale
Lullula arborea	Tottavilla
Tachybaptus ruficollis	Tuffetto

Hydrobates pelagicus	Uccello delle tempeste
Upupa epops	Upupa
Luscinia megarhynchos	Usignolo
Cettia cetti	Usignolo di fiume
Serinus citrinella	Venturone
Carduelis chloris	Verdone
Serinus serinus	Verzellino
Tadorna tadorna	Volpoca
Porzana porzana	Voltolino
Emberiza melanocephala	Zigolo capinero
Emberiza citrinella	Zigolo giallo
Emberiza cia	Zigolo muciatto
Emberiza cirlus	Zigolo nero
CLASS MAMMALIA	
Microtus agrestis	Arvicola agreste
Microtus arvalis	Arvicola campestre
Microtus multiplex	Arvicola di Fatio
Chionomys nivalis	Arvicola delle nevi
Clethrionomys glareolus	Arvicola rossastra
Microtus savii	Arvicola di Savi
Microtus subterraneus	Arvicola sotterranea
Arvicola terrestris	Arvicola terrestre
Barbastella barbastellus	Barbastello
Rupicapra rupicapra	Camoscio alpino
Rupicapra pyrenaica	Camoscio appenninico
Capra hircus	Capra selvatica
Capreolus capreolus	Capriolo
Cervus elaphus	Cervo nobile
Cervus elaphus	Cervo sardo
Sus scrofa	Cinghiale
Oryctolagus cuniculus	Coniglio selvatico
Crocidura suaveolens	Crocidura minore
Crocidura russula	Crocidura rossiccia
Crocidura sicula	Crocidura siciliana
Crocidura leucodon	Crocidura ventre bianco
Dama dama	Daino
Mustela nivalis	Donnola
Dryomys nitedula	Driomio
Mustela erminea	Ermellino
Martes foina	Faina

Rhinolophus blasii	Ferro di cavallo di Blasius
Rhinolophus euryale	Ferro di cavallo euriale
Rhinolophus ferrumequinum	Ferro di cavallo maggiore
Rhinolophus mehelyi	Ferro di cavallo di Mehely
Rhinolophus hipposideros	Ferro di cavallo minore
Felis silvestris	Gatto selvatico
Glis glis	Ghiro
Hystrix cristata	Istrice
Lepus corsicanus	Lepre appenninica
Lepus europaeus	Lepre europea
Lepus capensis	Lepre sarda
Lepus timidus	Lepre variabile
Lynx lynx	Lince euroasiatica
Lutra lutra	Lontra comune
Canis lupus	Lupo
Marmota marmota	Marmotta alpina
Martes martes	Martora
Sylvilagus floridanus	Minilepre
Miniopterus schreibersi	Miniottero
Tadarida teniotis	Molosso di Cestoni
Muscardinus avellanarius	Moscardino
Ovis orientalis	Muflone
Suncus etruscus	Mustiolo
Nyctalus noctula	Nottola comune
Nyctalus lasiopterus	Nottola gigante
Nyctalus leisleri	Nottola di Leisler
Myocastor coypus	Nutria
Plecotus auritus	Orecchione comune
Plecotus austriacus	Orecchione meridionale
Ursus arctos	Orso bruno
Pipistrellus kuhli	Pipistrello albolimbato
Pipistrellus pipistrellus	Pipistrello nano
Pipistrellus nathusii	Pipistrello di Nathusius
Hypsugo savii	Pipistrello di Savi
Mustela putorius	Puzzola europea
Eliomys quercinus	Quercino
Rattus norvegicus	Ratto delle chiaviche
Rattus rattus	Ratto nero
Erinaceus europaeus	Riccio europeo
Erinaceus concolor	Riccio orientale

Canis aureus	Sciacallo dorato
Sciurus vulgaris	Scoiattolo comune
Sciurus carolinensis	Scoiattolo grigio
Vespertilio murinus	Serotino bicolore
Eptesicus serotinus	Serotino comune
Amblyotus nilssonii	Serotino di Nilsson
Capra ibex	Stambecco delle alpi
Talpa caeca	Talpa cieca
Talpa europaea	Talpa europea
Talpa romana	Talpa romana
Meles meles	Tasso
Ondata zibethicus	Topo muschiato
Apodemus sylvaticus	Topo selvatico
Apodemus alpicola	Topo selvatico alpino
Apodemus flavicollis	Topo selvatico collo giallo
Apodemus agrarius	Topo selvatico dorso striato
Mus domesticus	Topolino domestico
Micromys minutus	Topolino delle risaie
Sorex alpinus	Toporagno alpino
Neomys fodiens	Toporagno d'acqua
Neomys anomalus	Toporagno acquatico di Miller
Sorex samniticus	Toporagno appenninico
Sorex araneus	Toporagno comune
Sorex minutus	Toporagno nano
Myotis bechsteini	Vespertilio di Bechstein
Myotis brandti	Vespertilio di Brandt
Myotis capaccinii	Vespertilio di Capaccini
Myotis dasycneme	Vespertilio dasicneme
Myotis daubentoni	Vespertilio di Daubenton
Myotis myotis	Vespertilio maggiore
Myotis blythi	Vespertilio minore
Myotis mystacinus	Vespertilio mustacchino
Myotis nattereri	Vespertilio di Natterer
Myotis emarginatus	Vespertilio smarginato
Mustela vison	Visone americano
Vulpes vulpes	Volpe comune

Curriculum Vitae

Over the past 20 years, Fabio Corsi has provided consulting services in GIS and ecological modelling to the UN Food and Agriculture Organisation, IUCN-The World Conservation Union, ESRI Italia, ITALECO, AGRICONSULTING, and other organisations. From 1990 to 2002 he served as a partner and Board member of Istituto di Ecologia Applicata (IEA), an Italian NGO focused on environmental conservation research. In 1999 and 2001 he was Professor of Information Systems Applied to the Management of Natural Resources at the University of Molise, Italy. He is currently Assistant Professor for Spatial Data Handling for Natural Resources Management, in the Department of Natural Resources of ITC (April 2003 to present).

Fabio Corsi's expertise includes eighteen years of teaching at the University-level, and through seminars, workshops, and corporate courses. He has nineteen years of experience in species distribution modelling, identification of wilderness areas using quantitative indexes, and assessment of fish stocks through spatially explicit models and application of the models to conservation biology. His work has been supported through partnerships with networks of experts, NGOs, and government agencies in pursuit of both funding and project execution.

Fabio Corsi earned a laureate with honours in Natural Science in 1984 at the University of Rome "La Sapienza", defending the thesis "Fisheries' Population Dynamics of Lake Monaci (Southern Latium)" and a complementary short thesis "Stratigraphy of some wells in the Kufra Basin (Southern Libya)".

He is author of more then 20 peer-reviewed scientific papers, 2 books, 1 multimedia CD-ROM and more then 80 other publications (technical reports, training manuals, software packages and congress presentations).

ITC Dissertation List

1. **Akinyede**, 1990, Highway cost modelling and route selection using a geotechnical information system
2. **Pan He Ping**, 1990, 90-9003757-8, Spatial structure theory in machine vision and applications to structural and textural analysis of remotely sensed images
3. **Bocco Verdinelli, G.**, 1990, Gully erosion analysis using remote sensing and geographic information systems: a case study in Central Mexico
4. **Sharif, M**, 1991, Composite sampling optimization for DTM in the context of GIS
5. **Drummond, J.**, 1991, Determining and processing quality parameters in geographic information systems
6. **Groten, S.**, 1991, Satellite monitoring of agro-ecosystems in the Sahel
7. **Sharifi, A.**, 1991, 90-6164-074-1, Development of an appropriate resource information system to support agricultural management at farm enterprise level
8. **Zee, D. van der**, 1991, 90-6164-075-X, Recreation studied from above: Air photo interpretation as input into land evaluation for recreation
9. **Mannaerts, C.**, 1991, 90-6164-085-7, Assessment of the transferability of laboratory rainfall-runoff and rainfall - soil loss relationships to field and catchment scales: a study in the Cape Verde Islands
10. **Ze Shen Wang**, 1991: 90-393-0333-9, An expert system for cartographic symbol design
11. **Zhou Yunxian**, 1991, 90-6164-081-4, Application of Radon transforms to the processing of airborne geophysical data
12. **Zuviria, M. de**, 1992, 90-6164-077-6, Mapping agro-topoclimates by integrating topographic, meteorological and land ecological data in a geographic information system: a case study of the Lom Sak area, North Central Thailand
13. **Westen, C. van**, 1993, 90-6164-078-4, Application of Geographic Information Systems to landslide hazard zonation
14. **Shi Wenzhong**, 1994, 90-6164-099-7, Modelling positional and thematic uncertainties in integration of remote sensing and geographic information systems
15. **Javelosa, R.**, 1994, 90-6164-086-5, Active Quaternary environments in the Philippine mobile belt

16. **Lo King-Chang**, 1994, 90-9006526-1, High Quality Automatic DEM, Digital Elevation Model Generation from Multiple Imagery
17. **Wokabi, S.**, 1994, 90-6164-102-0, Quantified land evaluation for maize yield gap analysis at three sites on the eastern slope of Mt. Kenya
18. **Rodriguez, O.**, 1995, Land Use conflicts and planning strategies in urban fringes: a case study of Western Caracas, Venezuela
19. **Meer, F. van der**, 1995, 90-5485-385-9, Imaging spectrometry & the Ronda peridotites
20. **Kufoniyi, O.**, 1995, 90-6164-105-5, Spatial coincidence: automated database updating and data consistency in vector GIS
21. **Zambezi, P.**, 1995, Geochemistry of the Nkombwa Hill carbonatite complex of Isoka District, north-east Zambia, with special emphasis on economic minerals
22. **Woldai, T.**, 1995, The application of remote sensing to the study of the geology and structure of the Carboniferous in the Calañas area, pyrite belt, SW Spain
23. **Verweij, P.**, 1995, 90-6164-109-8, Spatial and temporal modelling of vegetation patterns: burning and grazing in the Paramo of Los Nevados National Park, Colombia
24. **Pohl, C.**, 1996, 90-6164-121-7, Geometric Aspects of Multisensor Image Fusion for Topographic Map Updating in the Humid Tropics
25. **Jiang Bin**, 1996, 90-6266-128-9, Fuzzy overlay analysis and visualization in GIS
26. **Metternicht, G.**, 1996, 90-6164-118-7, Detecting and monitoring land degradation features and processes in the Cochabamba Valleys, Bolivia. A synergistic approach
27. **Hoanh Chu Thai**, 1996, 90-6164-120-9, Development of a Computerized Aid to Integrated Land Use Planning (CAILUP) at regional level in irrigated areas: a case study for the Quan Lo Phung Hiep region in the Mekong Delta, Vietnam
28. **Roshannejad, A.**, 1996, 90-9009284-6, The management of spatio-temporal data in a national geographic information system
29. **Terlien, M.**, 1996, 90-6164-115-2, Modelling Spatial and Temporal Variations in Rainfall-Triggered Landslides: the integration of hydrologic models, slope stability models and GIS for the hazard zonation of rainfall-triggered landslides with examples from Manizales, Colombia
30. **Mahavir, J.**, 1996, 90-6164-117-9, Modelling settlement patterns for metropolitan regions: inputs from remote sensing
31. **Al-Amir, S.**, 1996, 90-6164-116-0, Modern spatial planning practice as supported by the multi-applicable tools of remote sensing and GIS: the Syrian case
32. **Pilouk, M.**, 1996, 90-6164-122-5, Integrated modelling for 3D GIS

33. **Duan Zengshan**, 1996, 90-6164-123-3, Optimization modelling of a river-aquifer system with technical interventions: a case study for the Huangshui river and the coastal aquifer, Shandong, China
34. **Man, W.H. de**, 1996, 90-9009-775-9, Surveys: informatie als norm: een verkenning van de institutionalisering van dorp - surveys in Thailand en op de Filippijnen
35. **Vekerdy, Z.**, 1996, 90-6164-119-5, GIS-based hydrological modelling of alluvial regions: using the example of the Kisaföld, Hungary
36. **Pereira, Luisa**, 1996, 90-407-1385-5, A Robust and Adaptive Matching Procedure for Automatic Modelling of Terrain Relief
37. **Fandino Lozano, M.**, 1996, 90-6164-129-2, A Framework of Ecological Evaluation oriented at the Establishment and Management of Protected Areas: a case study of the Santuario de Iguaque, Colombia
38. **Toxopeus, B.**, 1996, 90-6164-126-8, ISM: an Interactive Spatial and temporal Modelling system as a tool in ecosystem management: with two case studies: Cibodas biosphere reserve, West Java Indonesia: Amboseli biosphere reserve, Kajiado district, Central Southern Kenya
39. **Wang Yiman**, 1997, 90-6164-131-4, Satellite SAR imagery for topographic mapping of tidal flat areas in the Dutch Wadden Sea
40. **Saldana-Lopez, Asunción**, 1997, 90-6164-133-0, Complexity of soils and Soilscape patterns on the southern slopes of the Ayllon Range, central Spain: a GIS assisted modelling approach
41. **Ceccarelli, T.**, 1997, 90-6164-135-7, Towards a planning support system for communal areas in the Zambezi valley, Zimbabwe; a multi-criteria evaluation linking farm household analysis, land evaluation and geographic information systems
42. **Peng Wanning**, 1997, 90-6164-134-9, Automated generalization in GIS
43. **Lawas, C.**, 1997, 90-6164-137-3, The Resource Users' Knowledge, the neglected input in Land resource management: the case of the Kankanaey farmers in Benguet, Philippines
44. **Bijker, W.**, 1997, 90-6164-139-X, Radar for rain forest: A monitoring system for land cover Change in the Colombian Amazon
45. **Farshad, A.**, 1997, 90-6164-142-X, Analysis of integrated land and water management practices within different agricultural systems under semi-arid conditions of Iran and evaluation of their sustainability
46. **Orlic, B.**, 1997, 90-6164-140-3, Predicting subsurface conditions for geotechnical modelling

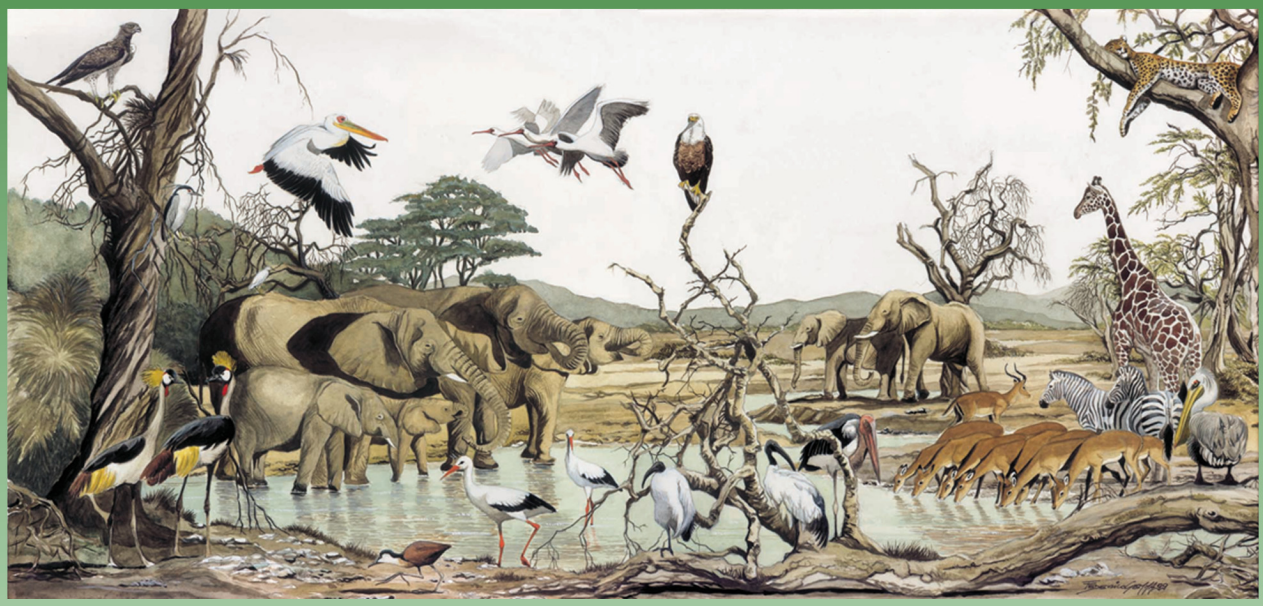
47. **Bishr, Y.**, 1997, 90-6164-141-1, Semantic Aspects of Interoperable GIS
48. **Zhang Xiangmin**, 1998, 90-6164-144-6, Coal fires in Northwest China: detection, monitoring and prediction using remote sensing data
49. **Gens, R.**, 1998, 90-6164-155-1, Quality assessment of SAR interferometric data
50. **Turkstra, J.**, 1998, 90-6164-147-0, Urban development and geographical information: spatial and temporal patterns of urban development and land values using integrated geo-data, Villaviciencia, Colombia
51. **Cassells, C.**, 1998, Thermal modelling of underground coal fires in northern China
52. **Naseri, M.**, 1998, 90-6164-195-0, Characterization of Salt-affected Soils for Modelling Sustainable Land Management in Semi-arid Environment: a case study in the Gorgan Region, Northeast, Iran
53. **Gorte B.G.H.**, 1998, 90-6164-157-8, Probabilistic Segmentation of Remotely Sensed Images
54. **Tegaye, Tenalem Ayenew**, 1998, 90-6164-158-6, The hydrological system of the lake district basin, central main Ethiopian rift
55. **Wang Donggen**, 1998, 90-6864-551-7, Conjoint approaches to developing activity-based models
56. **Bastidas de Calderon, M.**, 1998, 90-6164-193-4, Environmental fragility and vulnerability of Amazonian landscapes and ecosystems in the middle Orinoco river basin, Venezuela
57. **Moameni, A.**, 1999, Soil quality changes under long-term wheat cultivation in the Marvdasht plain, South-Central Iran
58. **Groenigen, J.W. van**, 1999, 90-6164-156-X, Constrained optimisation of spatial sampling: a geostatistical approach
59. **Cheng Tao**, 1999, 90-6164-164-0, A process-oriented data model for fuzzy spatial objects
60. **Wolski, Piotr**, 1999, 90-6164-165-9, Application of reservoir modelling to hydrotopes identified by remote sensing
61. **Acharya, B.**, 1999, 90-6164-168-3, Forest biodiversity assessment: A spatial analysis of tree species diversity in Nepal
62. **Akbar Abkar, Ali**, 1999, 90-6164-169-1, Likelihood-based segmentation and classification of remotely sensed images
63. **Yanuariadi, T.**, 1999, 90-5808-082-X, Sustainable Land Allocation: GIS-based decision support for industrial forest plantation development in Indonesia
64. **Abu Bakr**, **Mohamed**, 1999, 90-6164-170-5, An Integrated Agro-Economic and Agro-Ecological Framework for Land Use Planning and Policy Analysis

65. **Eleveld, M.**, 1999, 90-6461-166-7, Exploring coastal morphodynamics of Ameland (The Netherlands) with remote sensing monitoring techniques and dynamic modelling in GIS
66. **Yang Hong**, 1999, 90-6164-172-1, Imaging Spectrometry for Hydrocarbon Microseepage
67. **Mainam, Félix**, 1999, 90-6164-179-9, Modelling soil erodibility in the semiarid zone of Cameroon
68. **Bakr, Mahmoud**, 2000, 90-6164-176-4, A Stochastic Inverse-Management Approach to Groundwater Quality
69. **Zlatanova, Z.**, 2000, 90-6164-178-0, 3D GIS for Urban Development
70. **Ottichilo, Wilber K.**, 2000, 90-5808-197-4, Wildlife Dynamics: An Analysis of Change in the Masai Mara Ecosystem
71. **Kaymakci, Nuri**, 2000, 90-6164-181-0, Tectono-stratigraphical Evolution of the Cankori Basin (Central Anatolia, Turkey)
72. **Gonzalez, Rhodora**, 2000, 90-5808-246-6, Platforms and Terraces: Bridging participation and GIS in joint-learning for watershed management with the Ifugaos of the Philippines
73. **Schetselaar, Ernst**, 2000, 90-6164-180-2, Integrated analyses of granite-gneiss terrain from field and multisource remotely sensed data. A case study from the Canadian Shield
74. **Mesgari, Saadi**, 2000, 90-3651-511-4, Topological Cell-Tuple Structure for Three-Dimensional Spatial Data
75. **Bie, Cees A.J.M. de**, 2000, 90-5808-253-9, Comparative Performance Analysis of Agro-Ecosystems
76. **Khaemba, Wilson M**., 2000, 90-5808-280-6, Spatial Statistics for Natural Resource Management
77. **Shrestha, Dhruba**, 2000, 90-6164-189-6, Aspects of erosion and sedimentation in the Nepalese Himalaya: highland-lowland relations
78. **Asadi Haroni, Hooshang**, 2000, 90-6164-185-3, The Zarshuran Gold Deposit Model Applied in a Mineral Exploration GIS in Iran
79. **Raza, Ale**, 2001, 90-3651-540-8, Object-oriented Temporal GIS for Urban Applications
80. **Farah, Hussein**, 2001, 90-5808-331-4, Estimation of regional evaporation under different weather conditions from satellite and meteorological data. A case study in the Naivasha Basin, Kenya
81. **Zheng**, **Ding,** 2001, 90-6164-190-X, A Neural - Fuzzy Approach to Linguistic Knowledge Acquisition and Assessment in Spatial Decision Making
82. **Sahu, B.K.**, 2001, Aeromagnetics of continental areas flanking the Indian Ocean; with implications for geological correlation and reassembly of Central Gondwana

83. **Alfestawi, Y.,** 2001, 90-6164-198-5, The structural, paleogeographical and hydrocarbon systems analysis of the Ghadamis and Murzuq Basins, West Libya, with emphasis on their relation to the intervening Al Qarqaf Arch
84. **Liu, Xuehua,** 2001, 90-5808-496-5, Mapping and Modelling the Habitat of Giant Pandas in Foping Nature Reserve, China
85. **Oindo, Boniface Oluoch**, 2001, 90-5808-495-7, Spatial Patterns of Species Diversity in Kenya
86. **Carranza, Emmanuel John**, 2002, 90-6164-203-5, Geologically-constrained Mineral Potential Mapping
87. **Rugege, Denis**, 2002, 90-5808-584-8, Regional Analysis of Maize-Based Land Use Systems for Early Warning Applications
88. **Liu, Yaolin**, 2002, 90-5808-648-8, Categorical Database Generalization in GIS
89. **Ogao, Patrick**, 2002, 90-6164-206-X, Exploratory Visualization of Temporal Geospatial Data using Animation
90. **Abadi, Abdulbaset M.**, 2002, 90-6164-205-1, Tectonics of the Sirt Basin – Inferences from tectonic subsidence analysis, stress inversion and gravity modelling
91. **Geneletti, Davide**, 2002, 90-5383-831-7, Ecological Evaluation for Environmental Impact Assessment
92. **Sedogo, Laurent G.**, 2002, 90-5808-751-4, Integration of Participatory Local and Regional Planning for Resources Management using Remote Sensing and GIS
93. **Montoya, Lorena**, 2002, 90-6164-208-6, Urban Disaster Management: a case study of earthquake risk assessment in Carthago, Costa Rica
94. **Ahmad, Mobin-ud-Din**, 2002, 90-5808-761-1, Estimation of Net Groundwater Use in Irrigated River Basins using Geo-information Techniques: A case study in Rechna Doab, Pakistan
95. **Said, Mohammed Yahya**, 2003, 90-5808-794-8, Multiscale perspectives of species richness in East Africa
96. **Schmidt, Karin**, 2003, 90-5808-830-8, Hyperspectral Remote Sensing of Vegetation Species Distribution in a Saltmarsh
97. **Lopez Binnquist, Citlalli**, 2003, 90-3651-900-4, The Endurance of Mexican Amate Paper: Exploring Additional Dimensions to the Sustainable Development Concept
98. **Huang, Zhengdong**, 2003, 90-6164-211-6, Data Integration for Urban Transport Planning
99. **Cheng, Jianquan**, 2003, 90-6164-212-4, Modelling Spatial and Temporal Urban Growth
100. **Campos dos Santos, Jose Laurindo**, 2003, 90-6164-214-0, A Biodiversity Information System in an Open Data/Metadatabase Architecture

101. **Hengl, Tomislav**, 2003, 90-5808-896-0, PEDOMETRIC MAPPING, Bridging the gaps between conventional and pedometric approaches
102. **Barrera Bassols, Narciso,** 2003, 90-6164-217-5, Symbolism, Knowledge and management of Soil and Land Resources in Indigenous Communities: Ethnopedology at Global, Regional and Local Scales
103. **Zhan, Qingming**, 2003, 90-5808-917-7, A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data
104. **Daag, Arturo S.**, 2003, 90-6164-218-3, Modelling the Erosion of Pyroclastic Flow Deposits and the Occurrences of Lahars at Mt. Pinatubo, Philippines
105. **Bacic, Ivan**, 2003, 90-5808-902-9, Demand-driven Land Evaluation with case studies in Santa Catarina, Brazil
106. **Murwira, Amon**, 2003, 90-5808-951-7, Scale matters! A new approach to quantify spatial heterogeneity for predicting the distribution of wildlife
107. **Mazvimavi, Dominic**, 2003, 90-5808-950-9, Estimation of Flow Characteristics of Ungauged Catchments. A case study in Zimbabwe
108. **Tang, Xinming**, 2004, 90-6164-220-5, Spatial Object Modelling in Fuzzy Topological Spaces with Applications to Land Cover Change
109. **Kariuki, Patrick**, 2004, 90-6164-221-3, Spectroscopy and Swelling Soils; an integrated approach
110. **Morales, Javier**, 2004, 90-6164-222-1, Model Driven Methodology for the Design of Geo-information Services
111. **Mutanga, Onisimo**, 2004, 90-5808-981-9, Hyperspectral Remote Sensing of Tropical Grass Quality and Quantity
112. **Šliužas, Ričardas V.**, 2004, 90-6164-223-X, Managing Informal Settlements: a study using geo-information in Dar es Salaam, Tanzania
113. **Lucieer, Arko**, 2004, 90-6164-225-6, Uncertainties in Segmentation and their Visualisation

Front and back cover art created by
Federico Gemma



WAGENINGEN UNIVERSITY

WAGENINGEN UR

ITC INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

ITC Dissertation number: 114
ISBN: 90-8504-090-6