



When GIS come to life.

Linking landscape- and population ecology for large population management modelling: the case of Ibex (*Capra ibex*) in Switzerland.

Thèse de doctorat

présentée à la

Faculté des Sciences de
L'Université de Lausanne

par

Alexandre Hirzel

Biologiste diplômé
Université de Lausanne

Jury

Prof. Aurelio Bay, Président
Prof. Nicolas Perrin, Directeur de thèse
Prof. Jérôme Goudet, Expert
Prof. Antoine Guisan, Expert
Dr. Cornelis Neet, Expert

Lausanne
2001



FACULTE DES SCIENCES

Institut d'Ecologie

When GIS come to life.

Linking landscape- and population ecology for large population management modelling: the case of Ibex (*Capra ibex*) in Switzerland.

Thèse de doctorat

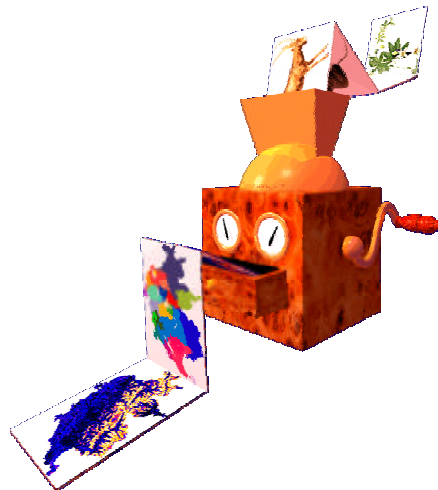
présentée à la

Faculté des Sciences de
L'Université de Lausanne

par

Alexandre Hirzel

Biologiste diplômé
Université de Lausanne



Jury

Prof. Aurelio Bay, Président
Prof. Nicolas Perrin, Directeur de thèse
Prof. Jérôme Goudet, Expert
Prof. Antoine Guisan, Expert
Dr. Cornelis Neet, Expert

Lausanne
2001

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président	Monsieur Prof.	Aurélio Bay
Directeur de thèse	Monsieur Prof.	Nicolas Perrin
Rapporteur		
Experts	Monsieur Prof.	Jérôme Goudet
	Monsieur Dr	Cornélis Neet
	Monsieur Dr	Antoine Guisan

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Alexandre Hirzel

Biologiste diplômé de l'Université de Lausanne

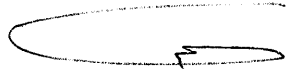
intitulée

When GIS come to life.

**Linking landscape-and population ecology for large population
management modelling: the case of Ibex (*Capra ibex*) in Switzerland**

Lausanne, le 7 septembre 2001

pour Le Doyen de la Faculté des Sciences


Prof. Aurélio Bay

Abstract

Conservation biology is commonly associated to small and endangered population protection. Nevertheless, large or potentially large populations may also need human management to prevent negative effects of overpopulation. As there are both qualitative and quantitative differences between small population protection and large population controlling, distinct methods and models are needed. The aim of this work was to develop theoretical models to predict large population dynamics, as well as computer tools to assess the parameters of these models and to test management scenarios. The alpine Ibex (*Capra ibex ibex*) – which experienced a spectacular increase since its reintroduction in Switzerland at the beginning of the 20th century – was used as paradigm species. This task was achieved in three steps:

A local population dynamics model was first developed specifically for Ibex: the underlying age- and sex-structured model is based on a Leslie matrix approach with addition of density-dependence, environmental stochasticity and culling. This model was implemented into a management-support software – named *SIM-Ibex* – allowing census data maintenance, parameter automated assessment and culling strategies tuning and simulating.

However population dynamics is driven not only by demographic factors, but also by dispersal and colonisation of new areas. Habitat suitability and obstacles modelling had therefore to be addressed. Thus, a software package – named *Biomapper* – was developed. Its central module is based on the Ecological Niche Factor Analysis (ENFA) whose principle is to compute niche marginality and specialisation factors from a set of environmental predictors and species presence data. All *Biomapper* modules are linked to Geographic Information Systems (GIS); they cover all operations of data importation, predictor preparation, ENFA and habitat suitability map computation, results validation and further processing; a module also allows mapping of dispersal barriers and corridors.

ENFA application domain was then explored by means of a simulated species distribution. It was compared to a common habitat suitability assessing method, the Generalised Linear Model (GLM), and was proven better suited for spreading or cryptic species.

Demography and landscape informations were finally merged into a global model. To cope with landscape realism and technical constraints of large population modelling, a cellular automaton approach was chosen: the study area is modelled by a lattice of hexagonal cells, each one characterised by a few fixed properties – a carrying capacity and six impermeability rates quantifying exchanges between adjacent cells – and one variable, population density. The later varies according to local reproduction/survival and dispersal dynamics, modified by density-dependence and stochasticity. A software – named *HexaSpace* – was developed, which achieves two functions: 1° Calibrating the automaton on the base of local population dynamics models (e.g., computed by *SIM-Ibex*) and a habitat suitability map (e.g. computed by *Biomapper*). 2° Running simulations. It allows studying the spreading of an invading species across a complex landscape made of variously suitable areas and dispersal barriers. This model was applied to the history of Ibex reintroduction in Bernese Alps (Switzerland).

SIM-Ibex is now used by governmental wildlife managers to prepare and verify culling plans. *Biomapper* has been applied to several species (both plants and animals) all around the World. In the same way, whilst *HexaSpace* was originally designed for terrestrial animal species, it could be easily extended to model plant propagation or flying animals dispersal. As these softwares were designed to proceed from low-level data to build a complex realistic model and as they benefit from an intuitive user-interface, they may have many conservation applications. Moreover, theoretical questions in the fields of population and landscape ecology might also be addressed by these approaches.

Résumé

La biologie de la conservation est communément associée à la protection de petites populations menacées d'extinction. Pourtant, il peut également être nécessaire de soumettre à gestion des populations surabondantes ou susceptibles d'une trop grande expansion, dans le but de prévenir les effets néfastes de la surpopulation. Du fait des différences tant quantitatives que qualitatives entre protection des petites populations et contrôle des grandes, il est nécessaire de disposer de modèles et de méthodes distinctes. L'objectif de ce travail a été de développer des modèles prédictifs de la dynamique des grandes populations, ainsi que des logiciels permettant de calculer les paramètres de ces modèles et de tester des scénarios de gestion. Le cas du Bouquetin des Alpes (*Capra ibex ibex*) – en forte expansion en Suisse depuis sa réintroduction au début du XX^{ème} siècle – sert d'exemple. Cette tâche fut accomplie en trois étapes :

En premier lieu, un modèle de dynamique locale, spécifique au Bouquetin, fut développé : le modèle sous-jacent – structuré en classes d'âge et de sexe – est basé sur une matrice de Leslie à laquelle ont été ajoutées la densité-dépendance, la stochasticité environnementale et la chasse de régulation. Ce modèle fut implémenté dans un logiciel d'aide à la gestion – nommé *SIM-Ibex* – permettant la maintenance de données de recensements, l'estimation automatisée des paramètres, ainsi que l'ajustement et la simulation de stratégies de régulation.

Mais la dynamique d'une population est influencée non seulement par des facteurs démographiques, mais aussi par la dispersion et la colonisation de nouveaux espaces. Il est donc nécessaire de pouvoir modéliser tant la qualité de l'habitat que les obstacles à la dispersion. Une collection de logiciels – nommée *Biomapper* – fut donc développée. Son module central est basé sur l'Analyse Factorielle de la Niche Ecologique (ENFA) dont le principe est de calculer des facteurs de marginalité et de spécialisation de la niche écologique à partir de prédicteurs environnementaux et de données d'observation de l'espèce. Tous les modules de *Biomapper* sont liés aux Systèmes d'Information Géographiques (SIG) ; ils couvrent toutes les opérations d'importation des données, préparation des prédicteurs, ENFA et calcul de la carte de qualité d'habitat, validation et traitement des résultats ; un module permet également de cartographier les barrières et les corridors de dispersion.

Le domaine d'application de l'ENFA fut exploré par le biais d'une distribution d'espèce virtuelle. La comparaison à une méthode couramment utilisée pour construire des cartes de qualité d'habitat, le Modèle Linéaire Généralisé (GLM), montra qu'elle était particulièrement adaptée pour les espèces cryptiques ou en cours d'expansion.

Les informations sur la démographie et le paysage furent finalement fusionnées en un modèle global. Une approche basée sur un automate cellulaire fut choisie, tant pour satisfaire aux contraintes du réalisme de la modélisation du paysage qu'à celles imposées par les grandes populations : la zone d'étude est modélisée par un pavage de cellules hexagonales, chacune caractérisée par des propriétés – une capacité de soutien et six taux d'imperméabilité quantifiant les échanges entre cellules adjacentes – et une variable, la densité de la population. Cette dernière varie en fonction de la reproduction et de la survie locale, ainsi que de la dispersion, sous l'influence de la densité-dépendance et de la stochasticité. Un logiciel – nommé *HexaSpace* – fut développé pour accomplir deux fonctions : 1° Calibrer l'automate sur la base de modèles de dynamique (par ex. calculés par *SIM-Ibex*) et d'une carte de qualité d'habitat (par ex. calculée par *Biomapper*). 2° Faire tourner des simulations. Il permet d'étudier l'expansion d'une espèce envahisseuse dans un paysage complexe composé de zones de qualité diverses et comportant des obstacles à la dispersion. Ce modèle fut appliqué à l'histoire de la réintroduction du Bouquetin dans les Alpes bernoises (Suisse).

SIM-Ibex est actuellement utilisé par les gestionnaires de la faune et par les inspecteurs du gouvernement pour préparer et contrôler les plans de tir. *Biomapper* a été appliqué à plusieurs

espèces (tant végétales qu'animales) à travers le Monde. De même, même si *HexaSpace* fut initialement conçu pour des espèces animales terrestres, il pourrait aisément être étendu à la propagation de plantes ou à la dispersion d'animaux volants. Ces logiciels étant conçus pour, à partir de données brutes, construire un modèle réaliste complexe, et du fait qu'ils sont dotés d'une interface d'utilisation intuitive, ils sont susceptibles de nombreuses applications en biologie de la conservation. En outre, ces approches peuvent également s'appliquer à des questions théoriques dans les domaines de l'écologie des populations et du paysage.

When GIS come to life

SUMMARY

I. INTRODUCTION.....	5
A. ECOLOGICAL MODELLING.....	5
B. LARGE POPULATION MODELLING	8
C. AN EXAMPLE OF LARGE POPULATION: THE ALPINE IBEX.....	9
D. REFERENCES.....	12
II. POPULATION DYNAMICS AND MANAGEMENT.....	14
A. INTRODUCTION.....	14
B. MODEL AND PARAMETERS.....	15
C. SIMULATIONS	19
D. DISCUSSION.....	20
E. REFERENCES	22
III. HABITAT SUITABILITY: THE ECOLOGICAL NICHE FACTOR ANALYSIS.....	23
A. INTRODUCTION.....	23
B. MARGINALITY, SPECIALISATION, AND THE ECOLOGICAL NICHE.....	24
C. FACTOR EXTRACTION.....	26
D. HABITAT SUITABILITY MAP	29
E. AN APPLICATION TO THE ALPINE IBEX	31
DISCUSSION	33
G. REFERENCES.....	35
IV. ASSESSING HABITAT-SUITABILITY MODELS WITH A VIRTUAL SPECIES.....	37
A. INTRODUCTION.....	37
B. METHODS.....	39
C. RESULTS.....	45
D. DISCUSSION.....	45
E. CONCLUSION.....	48
F. REFERENCES	48
V. LANDSCAPE EXPLICIT POPULATION DYNAMICS: WHEN GIS COME TO LIFE!	50
A. INTRODUCTION.....	50
B. MODEL DESCRIPTION	51
C. MODEL PARAMETERISATION.....	54
D. APPLICATION TO IBEX.....	57
E. DISCUSSION	62
F. SEQUENTIAL GAUSSIAN ALGORITHM.....	64
G. REFERENCES.....	65
VI. GENERAL CONCLUSION.....	68
ACKNOWLEDGEMENTS.....	72
APPENDICES	74
APPENDIX A: IBEX OFFICIAL FORMS	75
APPENDIX B: BIOMAPPER MODULES.....	78
APPENDIX C: SIM-IBEX KEY PROCEDURES	81
APPENDIX D: BIOMAPPER KEY PROCEDURES.....	92
APPENDIX E: HEXASPACE KEY PROCEDURES	100

When GIS come to life

I. INTRODUCTION

A. Ecological modelling

1. What are models?

UNIVERSE is vast and complex. Its spatial extension is so huge, its temporal dimension so immense and the number of its components so gigantic that Human senses and brain can only apprehend a minuscule part of it. Our sense penetration power has been enormously multiplied by modern instruments and techniques. By means of radio-telescopes, we can study galaxies billion light-years away from us and billion years in the past; microscopes and particle accelerators extend our sight at the very components of life and matter. But what is the point to see if one does not understand what happens? By studying particular phenomena, by measuring entities and intensities, by varying experimentally parameters and observing how they affect the underlying processes, searchers are able to reconstruct how things work.

However, as studied phenomena become more complex and involved parameters more numerous, instruments are needed to amplify our understanding

power. Whilst measuring instruments and experiments are extensions to our senses, statistics, mathematics and numerical algorithmic are extensions to our brain. On the one hand, statistics allow to explore, extract and summarise useful information hidden in an overwhelming quantity of data, on the other hand, analytical and numerical models try to reproduce natural processes and test their sensitivity to parameter changes.

Most of the time, both approaches may be used concurrently but there are cases where the studied system is mostly or even completely out of reach and there is no methodological choice anymore. Small parts might possibly be isolated and studied directly, but the system remains globally intractable. It may be spatially out of reach, as e.g. in astrophysics where most studied objects are too far to be subjected to experimentation. It may be temporally out of reach, as e.g. in cosmology, geology or evolutionary biology where phenomena are too slow and/or too far in the past to be studied directly. Finally, even if the system of interest is at hand here and now, it may be too big or composed of too many interacting elements to be reachable, as e.g. in meteorology or climatology. Ecology, dealing with the innumerable interactions among living organisms and between them

and their abiotic environment (Haeckel, 1866), is therefore clearly belonging to this last class of systems.

Thus, as the real world is unreachable to experimentation, searchers try to duplicate it in the computer or on the blackboard where it may be easily manipulated: this operation is called “modelling”. The whole art consists then in selecting the few factors that will be included in order to reproduce adequately the phenomena of interest. The model mimics the real world by keeping only those elements that are relevant to achieve a particular purpose. This purpose may lie anywhere along a continuum going from “tactic” to “strategic” applications (May, 1973; Gillman and Hails, 1997). Strategic models aim at understanding general processes and patterns, often from a qualitative point of view; they tend to be very simplified, to take only a few factors into account and to be therefore difficult to apply to real situations. In contrast, tactical models are designed to address, often quantitatively, applied conservation problems; they must therefore have a higher level of predictive accuracy and thus to fit more tightly to real data. In both cases anyway, the very modelling task, through the summarising, abstracting, interconnecting and quantifying operations it entails, brings a significant support to the comprehension of the system mechanics. Thereafter, if the model proves sufficiently accurate for its purpose, it may be utilised to make predictions and to test various scenarios and hypotheses.

The real world may be modelled continuously or discretely. This dichotomy is present at all stages of the modelling work, from data structures to process representation. Both ways have pros and cons. Choosing among them is always the result of a trade-off. We will see however in the following chapters that they may be somewhat mixed.

As continuous (or analytical) models are based on mathematical parametric functions, all the power of analytical tools is in the service of comprehension and inter-

pretation. However, processes and interactions are constrained in the rigorous frame of simple relations (mostly linear (or linearisable) and quadratic).

Discrete (or numerical) models decompose space, time and elements into somewhat homogenous units, which are processed sequentially. Processes and relations may be either equation- or rule-based, which gives a great modelling liberty and allows one to fit more closely to reality. However, these numerical procedures are computer-time-consuming; moreover, in contrast with analytical models, mechanistic interpretation is not conspicuous and generally need heavy sensitivity analyses.

In this work, we have applied both approaches to various problems of population and landscape ecology; they will be presented in the next chapters.

2. Geographic Information Systems

Once the model has been precisely conceived, it must be implemented into a computer program. Basically, a program may be viewed as a black box that takes in an input data set, processes it following some algorithm sequence and user-set parameters, and finally produces an output data set. Data are therefore a crucial part of the modelling system and they must be organised efficiently. Here, efficiency is composed of two components: 1° Data must be easily and quickly processed and 2° they must be easy to manage and to interpret by the user. This problem is particularly crucial in landscape ecology, which deals with huge amount of spatially explicit data. Geographic Information Systems (GIS) were especially designed to achieve these goals, namely:

1. Storing and maintaining spatially explicit data
2. Displaying and analysing them
3. Performing complex spatial operations
4. Communicating results to managers, deciders and public.

Here is not the place to give a course about these tools and we refer the reader to

the existing literature (e.g. Eastman, 1993; Johnston, 1998; Eastman, 1999). However, as there will be made a wide use of these tools in this work, it might be useful to give here a few informations about how data are structured and how we made use of them.

Spatially explicit data are stored into *information layers*. This term emphasises the fact that all informations related to a particular area are distributed among several logical data sets that may be reassembled (or *overlaid*) to be compared and analysed. These layers may be organised following two modes:

Object-oriented mode

Data are attached to objects whose spatial coordinates are precisely defined. Objects may be zero-dimensional (points), one-dimensional (lines, curves, broken lines, “empty” polygons) or two-dimensional (regions, “filled” polygons). In ecology, these objects typically represent biological or landscape information as for example observation or capture location, roads, rivers, forests or lakes. Associated data are called *attribute*. There may be any number of attributes attached to an object and they can give any kind of information (e.g. name, date of capture, amount of observed individuals, river seg-

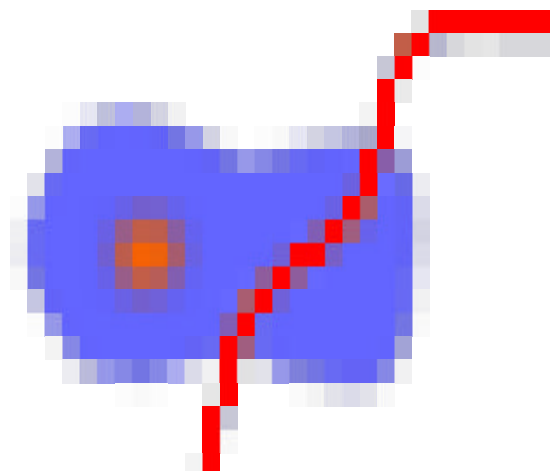


Figure 1.2: Example of image-oriented landscape modelling: the landscape is decomposed into a raster of isometric cells. A unique value is attached to each cell.

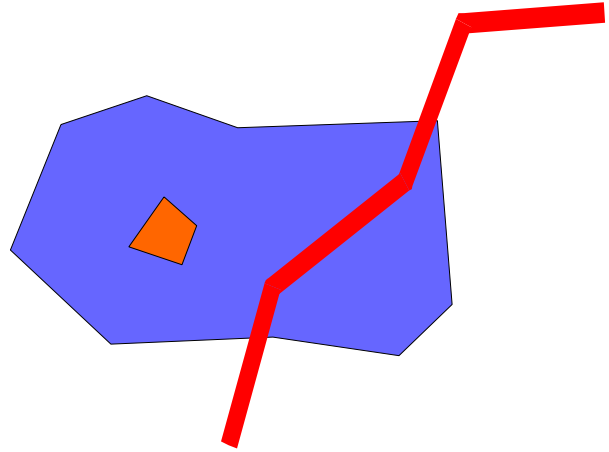


Figure 1.1: Example of object-oriented landscape modelling: The forest and the village are represented by polygonal objects whilst the road is represented by a broken line. Each object is defined by a sequence of spatially-located nodes. Data are attached to each object as attributes.

ment width, vegetation type, lake depth). Similar objects are grouped into information layers. Visually, a colour- or shape-code or superimposed information displays the value of their attributes. These data are said to be stored in a *vector structure* because all spatial information is contained into the vector-linked nodes defining object shapes (Fig.1.1).

This mode is perfectly suited to model qualitative information as e.g. land-cover, human structures, hydrographic network, administrative limits, etc. We mainly used this mode to store and to import qualitative data.

Image-oriented mode

The studied area is decomposed into a lattice (or raster) of isometric (generally square) cells. Each layer contains one type of qualitative or quantitative information, and this information is attached to each cell. As spatial information is contained into the lattice format, data are said to be stored into a *raster structure* (Fig.1.2).

This mode allows modelling spatially continuous information as, for example, elevation, temperature, depths, remote-

sensing imagery, etc. It is moreover the best mode to perform a wide palette of spatial analyses and operations. This data structure is also perfectly suited to represent landscape-processes and thus we used it extensively in our spatially explicit models.

The GIS software *Idrisi* (Eastman, 1993 and 1999) was chosen as the framework of our softwares for several practical reasons: 1° It is widely used by ecologists. 2° Its development environment is completely open and very easy to use 3° Its relatively low price makes it affordable to small ecology laboratories and private ecology companies.

3. Constraints on wildlife management models

Wildlife management may be defined as the management of wildlife populations (Caughley & Sinclair, 1994). This may apply in three coarse cases depending on the focal species situation: 1° it may have already disappeared from natural environment in the region of interest and suitable ecosystems have therefore to be *restored*; 2° it is on the way to become extinct and its last habitat patches must thus be *protected*; 3° it interferes with human activities and its populations have therefore to be *controlled*. Thus, conservation is implicitly bound to address fairly practical questions and this puts several constraints on conservation biologist works, namely: realism, accuracy, applicability and usability.

The realism constraint implies that conservation models tend to be more tactic than strategic; they must be applicable to real situations and address fairly concrete problems. The accuracy constraint stresses both model calibration and validation; As conservation models are liable to have actual applications that might have critical consequences, accuracy is of crucial importance both during parameter calibration and model validation and sensitivity analysis. Applicability and usability constraints intervene at a later time of the work but are not less important. If model parameters cannot be tuned or assessed, the model will

be useless; it belongs to the modeller to furnish the tools or methods allowing tuning his/her model. Finally, the model's implementation must be designed in order to facilitate the end-user work. When ported outside of scientific context, models tend to be either completely disregarded or taken for gospel; thus, because it is designed for non-scientific practitioners, care must be taken to prevent its misuses (e.g. used outside of its application domain or with poor data).

B. Large population modelling

Although conservation biology is classically associated with endangered species protection, large or potentially large populations may also need human management to prevent negative effects of overpopulation (Caughley & Sinclair, 1994). These negative effects may be inter- or intra-specific competition (e.g. Rushton et al., 1997), propagation of diseases (e.g. Mayer et al., 1997), environmental or agricultural damages (e.g. Ammer, 1996; Motta, 1996) or socio-economic conflicts (e.g. Le Lay et al., 2001).

This work was conducted on the impulsion of the Swiss Federal Office of Environment, Forests and Landscape, which is involved in several conservation projects (e.g. Capercaillie (*Tetrao urogallus*) restoration, Lynx (*Lynx lynx*) reintroduction, Wolf (*Canis lupus*) invasion). One of them is the sustainable management of Ungulates. Indeed, due to the absence of large predators, several species (e.g. Roe Deer (*Capreolus capreolus*), wild Boar (*Sus scrofa*) and alpine Ibex (*Capra ibex ibex*)) are growing and spreading fast, causing silvicultural, inter- and intra-specific competition problems.

Dealing with large populations entails particular problems, biological (e.g. density-dependence, spreading) as well as technical (large numbers of individuals and interactions). There was therefore a strong need for large population management models. It was the aim of this work to ad-

dress these questions. Several features of the alpine Ibex (developed in the next section) made it a good species for this kind of study and it was therefore the focal species of our work.

First, we approached the question at the local scale by building *SIM-Ibex*, an age- and sex-structured dynamics model specifically designed for Ibex management (chapter II); for this purpose, populations were assumed isolated from each other. Although this assumption is relevant in the case of this particular species, it may hardly be considered a generality. Thus, in a second part of our work we built a landscape-explicit population dynamics model.

To achieve this goal, information was needed about Ibex habitat and dispersal preferences. We therefore developed the GIS-toolkit *Biomapper*, based on the Ecological Niche Factor Analysis (ENFA), to answer to this kind of questions (chapter III). The validity of this method was assessed by means of a virtual species (chapter IV).

Informations provided by the dynamics and habitat models could now be merged into a spatial model. As we were dealing with potentially large populations, standard approaches (e.g. population-based or individual-based models) could hardly be applied; they are much more adapted to restoration and protection problems. Therefore, we developed in place a cellular automaton model (chapter V). It had the valuable property to merge easily the local population dynamics model developed in chapter II with the habitat suitability model of chapter III.

Whilst Ibex was the focal species of our studies, it was a permanent care of our work to make our models and softwares extendable to other species.

C. An example of large population: the alpine Ibex

Many features of alpine Ibex made it a good model species for our work: 1° Its

Swiss population is large, approaching carrying capacity (cf. fig. 1.4) 2° Accurate data are available, both temporally and spatially. 3° As a protected and popular species, its conservation interest Swiss government as well as public opinion. Here are a few major informations about its biology and history in Switzerland.

1. Systematic

Class:	<i>Mammalia</i>
Order:	<i>Ungulata</i>
Sub-order:	<i>Artiodactyla</i>
Family:	<i>Bovidae</i>
Sub-family:	<i>Caprinae</i> .
Genus:	<i>Capra</i>
Species:	<i>ibex</i>
Sub-species:	<i>ibex</i>

This alpine subspecies is by now present in the whole Alps. It may be hybridised with the domestic goat but though these individuals look very like pure Ibexes, their own offspring quickly come back to the goat type. Several other sub-species or related species are known though their genetic relationships are not well known: *C. pyrenaica* in Spain, *C. i. nubiana* in the Arabic peninsula and Red Sea shores, *C. i. siberica* in Himalayas and *C. i. caucasica* in the western part of Caucasian mountains.

2. Biology

The male Ibex weights typically more than 100 kg and the female from 40 to 50 kg. Body length, measured from muzzle to tail, is about 140 to 180 cm for the males and 120 to 140 cm for the females. Withers height is 100 cm for the males and 80 cm for the females. Females may live older than 20 years but males get rarely beyond 15 years old. The most characteristic feature of the Ibex is the sabre-like roll-lined horns bore by the adult males. Both sexes bear horns. By the males, they grow continuously during their whole life. This allows one to assess their age remotely by means of binoculars (cf. fig. 1.3).

The Ibex lives in populations of various sizes. It lives typically in rocky and steep areas, from 1600 to 3200 meters above sea

level, where south-facing steep meadows are its preferred habitat. Below the forest upper limit, it may be encountered in sunny open areas with rocky boulders. However, when suitable areas above the forest limit are rare or when colonies become overpopulated, it may be found even on wooded lands.

Ibex lives in different regions of its home range according to seasons. Winter quarters are the steepest south-facing areas, where snowslides are the most frequent and snow melts most quickly. From April to May, Ibexes go down to the areas where grass is already appearing. From end of May to the beginning of June, they wander upward and the females hide themselves into the most inaccessible places to give birth. From summer to the end of autumn, they live in highest parts of their home range (Rauch, 1941; Hainard, 1962; OFEFP, 1991).

In summer, males live in separate groups from females and young animals. From December to January – the rut period

– male and female groups come together and then separate again progressively. When about three years old, the young males quit their female group and try to join a male group. Most females give birth when reaching the age of four years (three years in favourable conditions). After a 23- to 25-week gestation, from end of May to middle of June, the Ibex gives birth to one single kid (OFEFP, 1991).

In autumn, Ibex alimentation is mainly composed of dry (39%) and fresh (45%) *Poaceae* complemented by other herbs (4.5%), small shrub (4.5%), coniferous shoots (4%) and leaves, mosses, lichens (3%). In winter, it feeds on low nutritive quality aliments (Klansek et al., 1995).

Ibex has few natural enemies. Kids may be killed by eagles; Lynx could also crop a few kids in the lower part of their home range, mainly in spring. However, Human is the main predator. Alimentary resources, climate, avalanches and diseases are the major causes of death (OFEFP, 1991).

The Ibex presents several characteristics

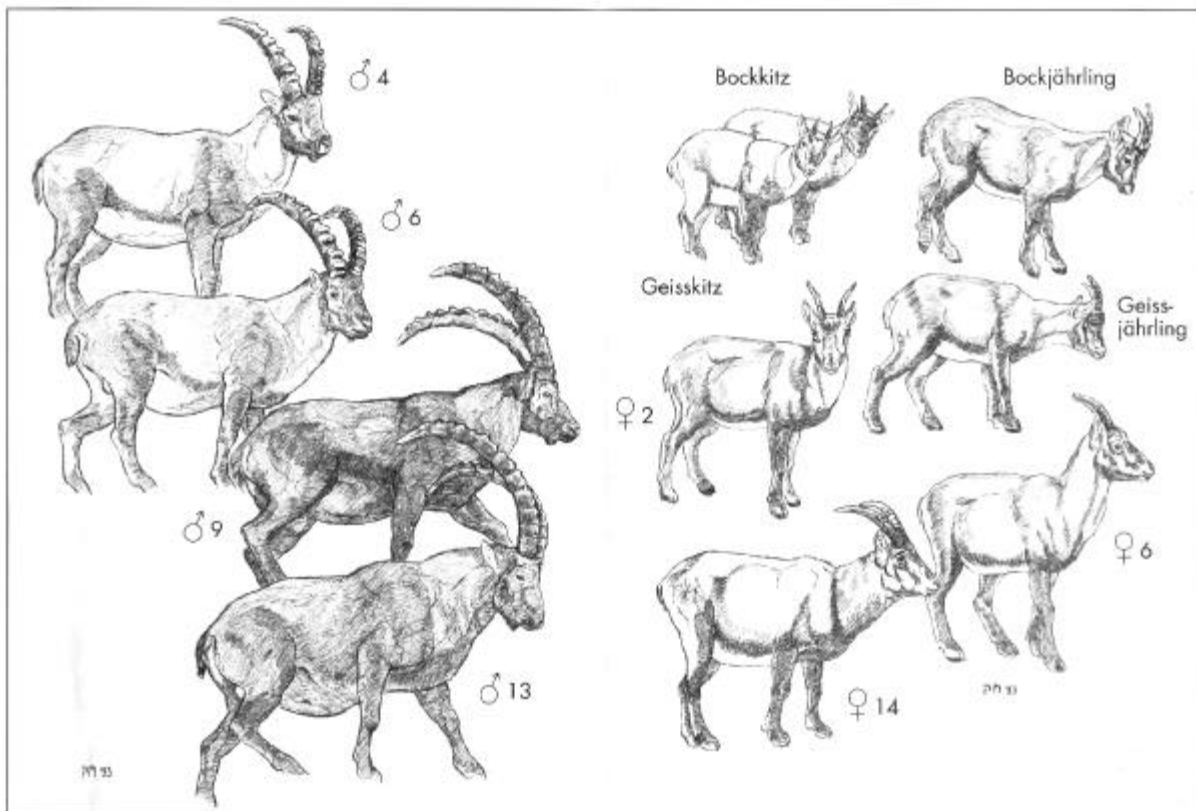


Figure 1.3: Ibexes. Sex and age (in years) are indicated beside the drawing. (Courtesy of Dr Peter Meile, Jagd- und Fischereiinspektorat Graubünden)

that make it particularly suitable for theoretical studies: Many ibex peculiarities make it a perfect modelling species: 1° It lives in high altitude, above forest limit, and therefore it may be accurately numbered; actually, every population in Switzerland is monitored since at least 1991, sometimes since its creation. 2° Its high elevation preferences make its habitat relatively patchy with only few exchange between populations. 3° Male age can be easily assessed with binocular (by counting horn marks). 4° As large carnivores have been exterminated from Switzerland, predation can be neglected.

3. History

Paleontological studies attest the presence of the Ibex in Switzerland from pre-historic times. Between the 16th and 18th century, almost all Alps populations were exterminated. A small population survived in the Gran Paradiso (Italy) only, because it was there actively protected in the royal hunting reserve (Hainard, 1962).

After several attempts, Ibex was finally successfully reintroduced in Switzerland in 1911 (Rauch, 1941; Hainard, 1962; OFEFP, 1991). It was then given a protected-species status. Since then, its numbers have grown rapidly (fig. 1.3). The Ibex is now distributed in most of the Swiss Alps and tends to be overcrowding (OFEFP, 1991). Overpopulation entails several negative effects: heavy damages to meadows and trees, health problems caused by lack of food and stress, diseases (e.g. pneumonia, parasitism, keratoconjunctivitis, scab) (Tataruch and Ondersheka, 1996; Mayer et al., 1997), competition with other ungulates (Chamois (*Rupicapra rupicapra*), red deer (*Cervus elaphus*)) (OFEFP, 1991).

In a first attempt to reduce local densities, since 1938, Ibexes were captured to be reintroduced in empty potentially suitable areas. But as this method proved insufficient to stabilise the populations at an acceptable level, in 1977 the Swiss government authorised a regulation culling (ORB,

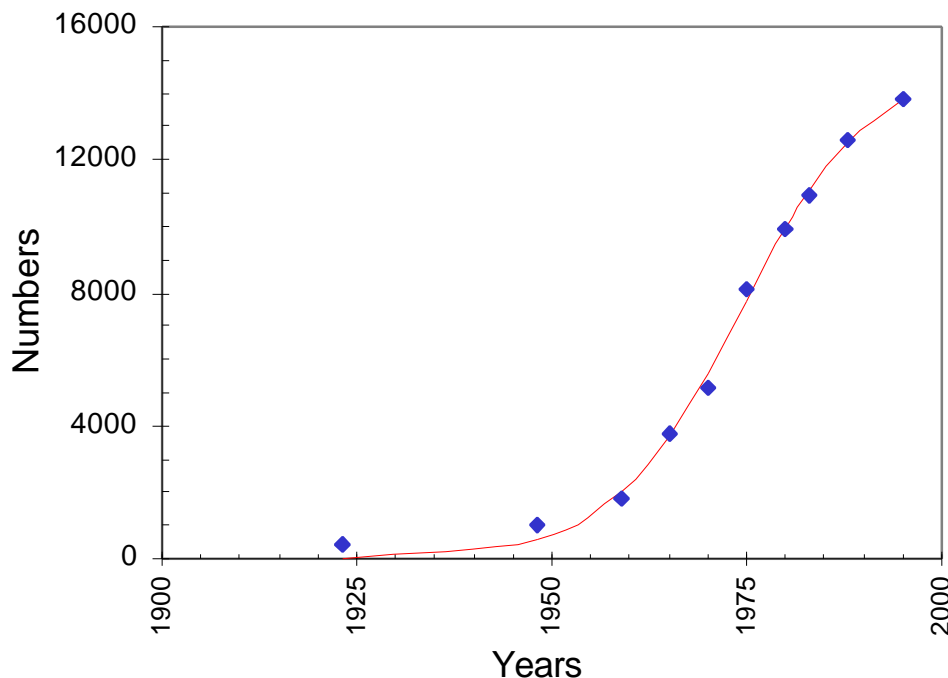


Figure 1.4: Since its reintroduction in Switzerland, Ibex population density has grown rapidly, following a logistic curve. ($r = 0.12$, $K = 15'000$, $R^2 > 0.99$)

1988; OFEFP, 1991; Ratti, 1994; Abderhalden and Buchli, 1997).

4. Monitoring

All Ibex populations are subject to a yearly monitoring since 1990. A few of them are even surveyed since their reintroduction. Thanks to the morphological particularities of Ibex, the censuses are fairly detailed, distinguishing the sexes and four male age-classes. These classes are chosen for binocular discrimination ease: young of both sexes (1 to 2 years old), adult females (more than 2 years old), males from 3 to 5, 6 to 10 and more than 10 years old.

Monitoring are conducted as follows: the whole area – possibly covering several administrative regions – is numbered simultaneously to minimise wandering individual inaccuracies. The massif is partitioned into several about-5-by-5-km homogeneously oriented sectors; each sector is assigned to a group of two or three persons comprising at least one experienced professional. Each group begins the survey simultaneously early in the morning. They climb up the mountain, exploring their assigned sector by means of binoculars. Ibex sex and age (male only) are determined and numbered. Observation time is also noted as well as the movement direction if any, in particular, when they are wandering toward an adjacent sector. Afterward, groups are debriefed and their observations pooled to compute the global census. Ibex living above forest limit, they are relatively easy to observe and therefore, numbering attains a good accuracy. Monitoring campaigns take place either in April or in early June. June censuses are less accurate because females tend to hide to give birth, but bad weather conditions are more frequent in April.

Census results (form I, appendix A) and a culling plan proposal (form II, appendix A) are sent by the local ranger to the canton wildlife manager to be controlled. They are further transmitted to the federal office where they must be approved. This complex administrative procedure is due to the

protected species status of the Ibex (ORB, 1988).

5. Culling

Culling takes place in early September. Whilst details may differ from one canton to another, there are many similarities. The culling plan must be precisely followed; each hunter is assigned a particular sex and age class. Hunting period lasts a few days and the fauna manager accompanies the hunters on the field. Each culled animal is immediately examined by the manager who takes a few measures including weight, age and health status.

When the culling plan has been completed, a report detailing the precise age of each culled individual is sent to the canton fauna inspector (form III, appendix A).

D. References

- Abderhalden, W., Buchli, C., 1997. Steinbockprojekt Albris/SNP, Schlussbericht. Zerne, ARINAS & FORNAT.
- Ammer, C., 1996. Impact of Ungulates On Structure and Dynamics of Natural Regeneration of Mixed Mountain Forests in the Bavarian Alps. *Forest Ecology & Management* 88, 43-53.
- Caughley, G., Sinclair, A.R.E., 1994. *Wildlife ecology and management*. Cambridge, Massachusetts, Blackwell Science.
- Eastman, J.R., 1993. *Idrisi : Un SIG en mode image*. Worcester, USA, Clark University.
- Eastman, J.R., 1999. *Idrisi32, Guide to GIS and image processing*. Worcester, Clark University.
- Gillman, M., Hails, R., 1997. *An introduction to ecological modelling. Putting practice into theory*. Oxford, Blackwell Science.
- Haeckel, E., 1866. *Generelle Morphologie der Organismen*, G. Reimer.
- Hainard, R., 1962. *Mammifères sauvages d'Europe*. Paris, Delachaux et Niestlé.
- Johnston, C.A., 1998. *Geographic Information Systems in ecology*. Oxford, Blackwell Science.
- Klansek, E., Vavra, I., Onderscheka, K., 1995. *Die Äsungszusammensetzung des Alpen-*

- steinwildes (*Capra i. ibex* L.) in Abhängigkeit von Jahreszeit, Alter und Äsungsangebot in Graubünden. *Zeitschrift für Jagdwissenschaft* 41, 171-181.
- Le Lay, G., Clergeau, P., Hubert-Moy, L., 2001. Computerized Map of Risk to Manage Wildlife Species in Urban Areas. *Environmental Management* 27, 451-461.
- May, R.M., 1973. Stability and complexity in model ecosystems. Princeton, New Jersey, Princeton University Press.
- Mayer, D., Degiorgis, M.P., Meier, W., Nicolet, J., Giacometti, M., 1997. Lesions Associated With Infectious Keratoconjunctivitis in Alpine Ibex. *Journal of Wildlife Diseases* 33, 413-419.
- Motta, R., 1996. Impact of Wild Ungulates On Forest Regeneration and Tree Composition of Mountain Forests in the Western Italian Alps. *Forest Ecology & Management* 88, 93-98.
- OFEFP, 1991. Commentaires relatifs à l'ordonnance sur la régulation des Bouquetins en Suisse (ORB), OFEFP.
- ORB, 1988. Ordonnance sur la Régulation des populations de Bouquetins en Suisse.
- Ratti, P., 1994. Stand von Hege und Erforschung des Steinwildes im Kanton Graubünden (Schweiz). *Zeitschrift für Jagdwissenschaft* 40, 223-331.
- Rauch, A., 1941. Le Bouquetin dans les Alpes. Paris.
- Rushton, S.P., Lurz, P.W.W., Fuller, R., Garson, P.J., 1997. Modelling the distribution of the red and grey squirrel at the landscape scale : a combined GIS and population dynamics approach. *J. Appl. Ecol.* 34, 1137-1154.
- Tataruch, F., Onderscheka, K., 1996. Untersuchungen zur Kondition des Steinwildes (*Capra i. ibex* L.) in Graubünden. *Zeitschrift für Jagdwissenschaft* 42, 97-103.

II. POPULATION DYNAMICS AND MANAGEMENT

A. Introduction

The alpine Ibex (*Capra ibex ibex*) became almost extinct in the Alps between the 16th and the 18th century. It was reintroduced in Switzerland in 1911 and given a protected species status. Since then, its numbers have grown rapidly. The Ibex is now distributed in most of the Swiss Alps and tends to be overcrowding. This is having damageable effects on its own health (e.g. pneumonia, parasitism, keratoconjunctivitis, scab) (Tataruch and Ondercheka, 1996; Mayer et al., 1997) as well as on vegetation, and is causing competition with other ungulates (Chamois (*Rupicapra rupicapra*), red deer (*Cervus elaphus*)) (OFEFP, 1991). Since 1938, in a first attempt to reduce local densities, Ibexes were captured from overpopulated colonies and reintroduced into low-density suitable areas. But this method proved insufficient to stabilise the numbers at an acceptable level, and therefore, in 1977, the Swiss government authorised a regulation culling (OFEFP, 1991; Ratti, 1994; Abderhalden and Buchli, 1997). In 1996, the Federal Office of the Forest, Environment and Landscape bade a study of Ibex population dynamics and realisation of a man-

agement-support software. This program would be used by regional gamekeepers for population management and by federal authorities for culling strategy verification. This paper covers some results of this study.

Many ibex peculiarities make it a perfect model species: 1° Its high elevation preferences make its habitat relatively patchy with extremely restricted exchanges between colonies. 2° It lives in high altitude, above forest limit, and therefore it may be accurately numbered; actually, every population in Switzerland is monitored at least since 1991, sometimes since its creation. 3° Male age can be easily assessed with binocular by looking at horn nodosity pattern. 4° As large carnivores have been exterminated from Switzerland, predation may be neglected.

The software had to fulfil several constraints: 1° It had to be built on an accurate population dynamics model, which had nevertheless to be sufficiently simple to allow easy and robust automated parameter assessment. 2° As the software was intended for non-biologists, it had to be easy to use whilst displaying all relevant information. 3° It had to implement various culling strategies, which the user might evaluate by mean of simulations. Basically, the aim was to give the gamekeeper a bet-

ter understanding of his population and help him/her to maintain it at an optimal density and an equilibrated age structure.

The software – named *SIM-Ibex* – was written in Borland Delphi 4. Two versions were produced (German and French) and are now used by gamekeepers to tune their culling strategy and by governmental authorities to verify them. It may be downloaded for free on Internet at <http://www.unil.ch/izea/software/simibex.html>.

B. Model and parameters

Model conceptualisation was guided by management needs and data availability and reliability. A few assumptions were made on the basis of field knowledge (P. Ratti, J.-C. Roch, C. Neet, H. Brunner, comm. pers.) and previous studies (Nievergelt, 1966; Nievergelt and Zingg, 1986; Giacometti, 1988; OFEFP, 1991; Abderhalden et al., 1997): 1° Ibex populations were considered as totally isolated from each other. 2° Survival was affected by age, sex and population density (linear density-dependence) 3° Fecundity being generally impossible to assess from available data, it was considered as fixed. 4° Both fecundity and survival were subjected to environmental stochasticity. 5° For pa-

rameter assessment, populations were assumed at stable age equilibrium.

1. Model

Population dynamics was modelled by a Leslie's matrix approach (Begon et al., 1996) with a one-year time unit. The individual numbers of age x , sex s at time t ($N_{s,x,t}$) are function of the numbers of the age $x-1$, sex s at time $t-1$ and an age- and density-dependent survival rate $S_{x,t}$ (Eq. 2.1). Numbers of kids ($x=0$) is a function of adult female numbers $N_{f,t}$ and fecundity B (Eq. 2.2).

As survival is age- and sex-dependent, three stages in both sexes were distinguished according to their homogeneous survival rates (OFEFP, 1991): 1° a “young” stage with relatively low survival (≈ 0.5), 2° an “adult” stage with higher survival (≈ 0.7) and 3° an “old” stage with low survival (≈ 0.5) (fig. 1); stage boundaries differ between sexes, in particular females' longevity (20 years) is greater than males' (15 years). The model has therefore three survival parameters (S_y , S_a , S_o) and one female fecundity parameter (B) (fig. 2.2).

Finally, survival rates are linearly decreasing with density and affected by uniformly distributed stochasticity. Each survival parameter is thus composed of a slope and an intersect (Eq. 2.4a,b,c).

Table 2.1: Mathematical symbols used in this chapter.

Symbol	Description
a_y, a_a, a_o	Intercept of the linear function giving the survival rate respectively for the young, adult and old animals.
b_y, b_a, b_o	Slope of the linear function giving the survival rate respectively for the young, adult and old animals.
B	Fecundity rate of the females older than 2 years. Fixed to 0.5
C_t	Numbers of animals culled at time t
e	Environmental stochasticity coefficient.
N_t	Numbers of the whole population, i.e. sum of all age and sex classes at time t
$N_{s,x,t}$	Numbers of individuals in the sex class s , age class x at time t
r_t	Per capita increase rate at time t . (See eq.2.12)
$S_{y,t}$	Survival rate of the young animals at time t (see text)
$S_{a,t}$	Survival rate of the adult animals at time t (see text)
$S_{o,t}$	Survival rate of the old animals at time t (see text)

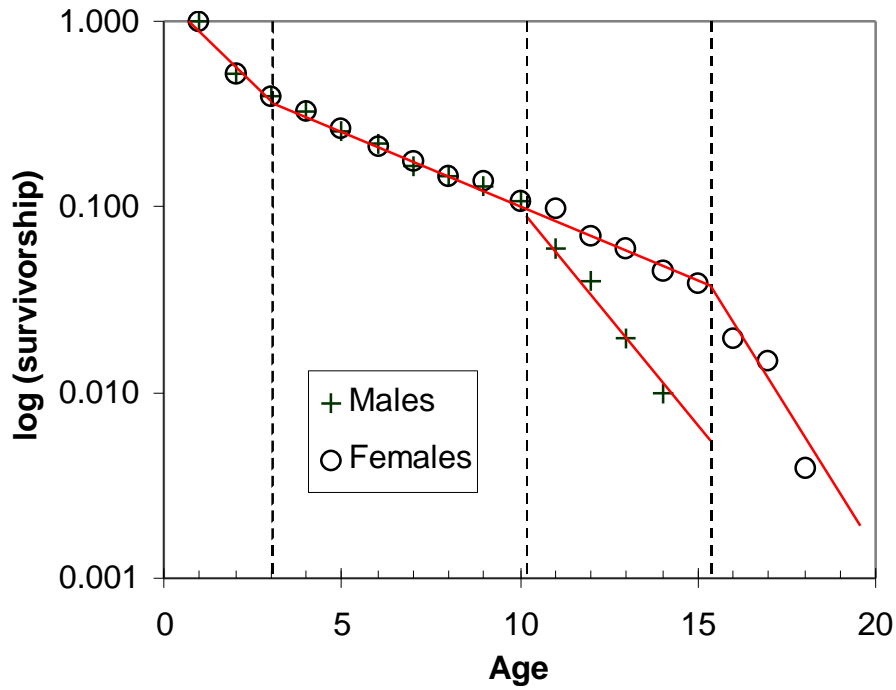


Figure 2.1: Survivorship curve of males and females. According to the survival rates, one can distinct three stages in the life of each sex: the survival rate is lower in the young age (~0.5), it is greater for the adults (~0.7) and then again lower for the oldest animals (~0.5). The eye-drawn lines show the survival rate of the three age classes. Note that there are differences between sexes. (Derived from OFEFP(1991), fig. 2.4)

Transitions between classes can thus be written:

$$\text{If age class } x > 0 \text{ then } N_{s,x,t+1} = N_{s,x,t} S_{x,t} \quad (2.1)$$

$$\text{If age class } x = 0 \text{ then } N_{0,t} = N_{f,3-20,t} B \quad (2.2)$$

$$\text{with } S_{.,t} \in \{S_{y,t}, S_{a,t}, S_{o,t}\} \quad (2.3)$$

$$S_{y,t} = \mathbf{a}_y + \mathbf{b}_y N_t \pm \mathbf{e} \quad (2.4a)$$

$$S_{a,t} = \mathbf{a}_a + \mathbf{b}_a N_t \pm \mathbf{e} \quad (2.4b)$$

$$S_{o,t} = \mathbf{a}_o + \mathbf{b}_o N_t \pm \mathbf{e} \quad (2.4c)$$

Poor reliability of kid numbers did not allow computing density-dependent fecundity parameters. By the way, density seems to affect more kid winter survival than female fecundity (P. Ratti, comm. pers.).

All the symbols used in this paper are summarised in table 2.1:

2. Available data (2.1)

Since 1991, every Swiss ibex population is subject to a yearly survey. Five age and sex classes – chosen for binocular discrimination ease – are monitored (table 2.2) as well as culled numbers.

Moreover, for several populations, much longer time series are available but only total numbers are counted.

3. Parameter estimation

For each population, there are thus seven parameters to estimate: slope and intercept for the three age dependent survival rates and the fecundity rate. The latter cannot be estimated for each population individually because of lack of reliability of kid numbers (the populations are

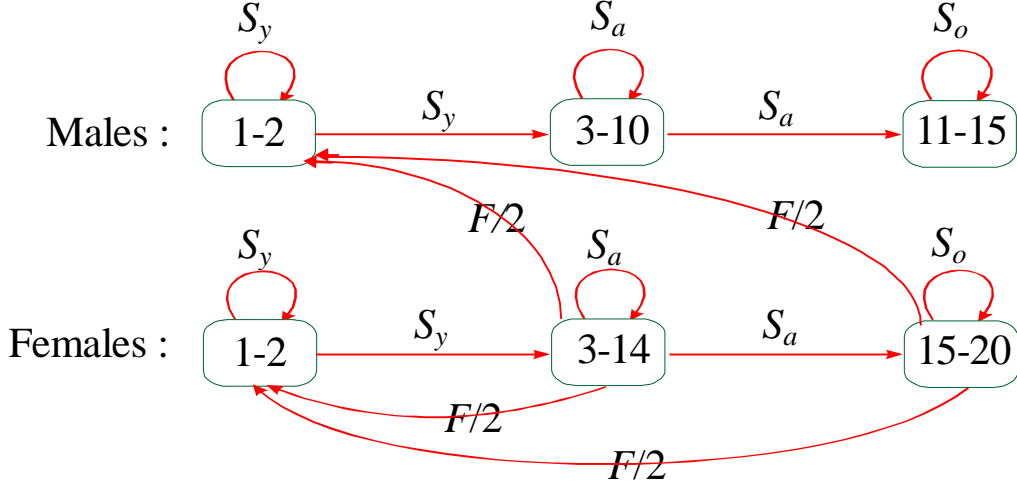


Figure 2.2: The model is based on three age classes (young, adult and old) distinguished according to survival (S_{xy}) and fecundity (F): Female fecundity is zero up to the age of three years, then constant for all ages. Birth sex ratio is assumed to be 1:1.

monitored either before or during the birth season) and the fecundity B is therefore set to 0.5 for adult females (three years old or more) and 0 for the young, with a 1:1 birth sex-ratio (Nievergelt, 1966).

The six density-dependent survival parameters (\mathbf{a} , \mathbf{b}) are computed as follows:

First, the number of kids $N_{0,t}$ at time t is computed, assuming a mean fecundity coefficient B for all females:

$$N_{0,t} = BN_{f,3-20,t} \quad (2.5)$$

where $N_{f,t}$ is the total number of females at time t .

Knowing that young numbers $N_{1+2,t}$ is $N_{1+2,t+1} = S_y N_{0,t} \cdot S_y^2 N_{0,t-1}$, the young survival rate S_y may be extracted by solving:

$$S_y = \frac{-N_{0,t} + \sqrt{N_{0,t}^2 + 4N_{0,t-1}N_{1+2,t+1}}}{2N_{0,t-1}} \quad (2.6)$$

The reproductive rate I_t for this year t taking culling into account is then computed by:

$$I_t = \frac{N_{t+1} + C_t}{N_t} \quad (2.7)$$

The numbers of the two-year old ibexes $N_{2,t}$ may be estimated as a proportion of

young individuals (assuming stable age distribution):

$$N_{2,t} = \frac{S_y N_{1+2,t}}{S_y + I_t} \quad (2.8)$$

That allows finding the adult survival rate S_a by solving:

$$N_{2,t} = \frac{N_{m,3-10,t}}{\sum_{i=1}^8 \left(\frac{S_a}{I} \right)^i} \quad (2.9)$$

For the last step, the numbers of the 10-year old males $N_{m,10,t}$ are computed by:

$$N_{m,10,t} = \frac{N_{m,6-10,t}}{\sum_{i=0}^4 \left(\frac{S_a}{I} \right)^i} \quad (2.10)$$

And finally, the old survival rate is found by solving:

$$N_{m,10,t} = \frac{N_{m,11+,t}}{\sum_{i=1}^5 \left(\frac{S_o}{I} \right)^i} \quad (2.11)$$

To summarise, for every year t whose previous and next year are also monitored, survival rates are computed for young ($S_{y,t}$), adult ($S_{a,t}$) and old ($S_{o,t}$) age classes

Table 2.2: List of data yearly collected on each population, with associated symbol

Symbol	Collected data
C_t	Numbers of culled animal at time t
$N_{1+2,t}$	Numbers of youngs (1 to 2 years old) at time t
$N_{f,t}$	Numbers of adult females (older than 2 years) at time t
$N_{m,3-5,t}$	Numbers of males from 3 and 5 years old at time t
$N_{m,6-10,t}$	Numbers of males from 6 and 10 years old at time t
$N_{m,11+,t}$	Numbers of old males (older than 10 years)
N_t	Total numbers (sum of all the above classes)

(equations 2.6, 2.9 and 2.11). The density-dependence can then be assessed by computing the linear regression of these yearly survival rates on the total numbers, which gives slope (**a**) and intercept (**b**) parameters.

The carrying capacity K is needed to plan the management strategy. Empirically, we found the following method to robustly assess this value:

A *per capita* rate of increase r_t is computed for every year by:

$$r_t = \frac{N_{t+1} - N_t}{N_t} = \frac{\Delta N}{N} \quad (2.12)$$

On the scatter plot of r_t as function of N_t

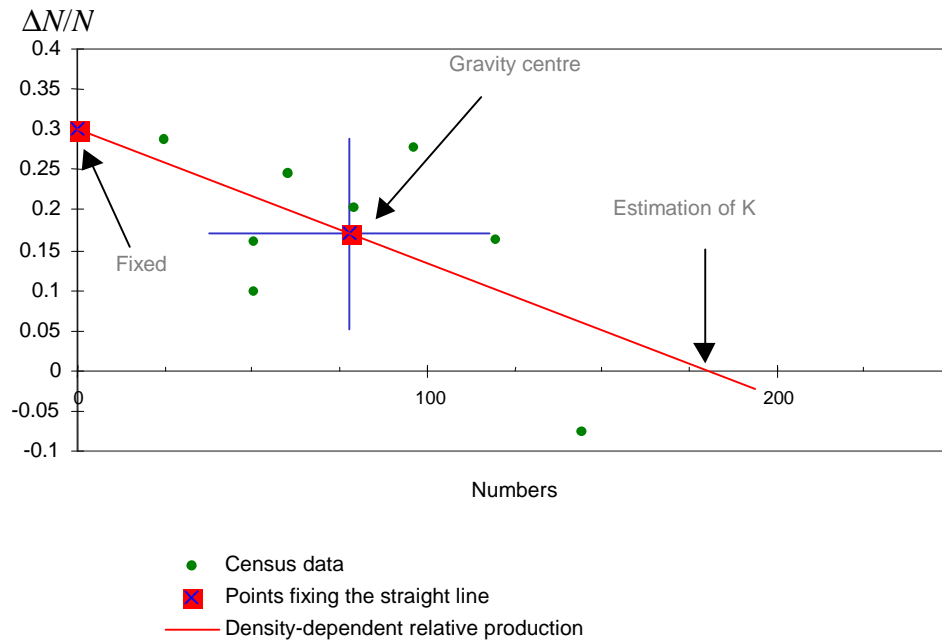


Figure 2.3: Method used to robustly assess the carrying capacity (K) of a population. The per capita rates of increase for each year are plotted against the total numbers (dots). The gravity centre of this scatter plot is computed. A straight line is drawn, passing by this point and by the $(N=0, \Delta N/N=0.3)$ point. This line intercepts the horizontal axis in K . See text for further explanation.

, the straight line passing by the centroid and by the point $(N=0, r_{max}=0.3)$ intersects the N -axis in K (see figure 2.3); this value of $r_{max}=0.3$ was computed as the maximum *per capita* increase possible for the ibex, assuming no mortality until the age of 20 and one kid by adult female and by year. This maximal rate of increase is in good agreement with data collected on young populations in favourable conditions (Gauthier et al.(1991) found 0.3 to 0.35), as well as with intrinsic rate of increase measured on well known populations (e.g. Brienzer Rothorn: 0.27)

4. Culling strategies

The aim of Ibex management is to maintain population densities at an “ideal” level according to health, extinction risk, forestry and interspecific competition. This is done by subjecting colonies to an adequate culling. Empirically, it has been found that this “ideal” target density was two thirds of the carrying capacity; culls have then to be distributed among sex and age classes. Three strategies were implemented into *SIM-Ibex* to achieve this management:

1° The governmental strategy consists in first deciding whether the population must be reduced, stabilised or let to grow. This decision is taken by comparing the numbers to the target density (set by default to $2/3 K$); if the numbers are 20% lower, the population must be let to grow; if the numbers are 20% higher, it must be reduced; otherwise, it must be stabilised. To reduce a population, 17% of the total numbers are culled; to maintain it, this rate is of 13% and to let it to grow, there is no cull at all. The culls are distributed among age and sex classes proportionally to their numbers. As this strategy was applied onto a few real populations, its implementation could therefore be used to validate our model (cf. below).

2° The custom strategy allows the user to distribute culls at will among age and sex classes. The amount of culls is computed in order to set the numbers to the target density, taking into account the population natural rate of increase.

3° The *SIM-Ibex* strategy is designed to distribute the culls among age and sex classes in order to both maintain the target density and guarantee a natural age pyramid, i.e. ages pyramid that would be theoretically attained by the population without any extrinsic effects (environment, management) (Begon et al., 1996). This theoretical pyramid is compared to the observed one and each class exceeding the natural pyramid will be culled.

Moreover, all these strategies are controlled by a “safety mechanism” that prevents from culling more than 20% of the total numbers. It makes extinction by over-culling very improbable.

Impacts of these strategies may be evaluated by mean of simulations.

C. Simulations

1. Validation

The model was validated on an ibex population (*Val Bever*, GR, Switzerland) whose dynamics was accurately known but was not used during model calibration phase. From 1981 to 1987, the population was let alone, then culling was introduced, following the governmental strategy. In the program, time was played back to 1981 and numbers initialised to their values at this date. Culling was enabled only from 1988. One thousand simulations were then run with these parameters to check how well the theoretical curves were fitting to the real one.

Fifteen years in 17, the real dynamics was framed inside interquartile range of the simulated data (see figure 2.4).

2. Culling strategy comparisons

The model was then applied to the Brienzer Rothorn population (Bern, Switzerland). This population has a carrying capacity of $K=157$ individuals. Three culling strategies were tried to explore model sensitivity. In all scenarios, the aim was to maintain the density around $2/3$ of K , i.e. about 110 individuals. Thousand replicates were simulated over a 20 years period. Extinction risk, i.e. probability to get lower than 20 individuals, and over-population risk, i.e. probability to get higher than 300 individuals ($\approx 2 K$), were computed, as well as the mean final density and the mean yearly culling.

Table 2.3: Governmental culling strategy. This strategy distributes culls assuming an equilibrium age pyramid, i.e. both sexes are equally culled and younger classes are more culled than older.

Culling strategy		Results	
Age classes	Culling proportion	Statistics	Value
Adult females:	32.5 %	Mean and SD final density:	108 ± 25
Males >11 years old:	7.5 %	Yearly culling:	12
Males 6-10 years old:	7.5 %	Extinction risk:	0.0 %
Males 3-5 years old:	17.5 %	Over-population risk:	0.0 %
Youngs 1-2 years old:	35.0 %		

Table 2.4: Female culling. This strategy concentrates on adult females. Fewer culls are needed to attain a lower than aimed final density.

Culling strategy		Results	
Age classes	Culling proportion	Statistics	Value
Adult females:	100.0 %	Mean and SD final density:	88 ± 34
Males >11 years old:	0.0 %	Yearly culling:	7
Males 6-10 years old:	0.0 %	Extinction risk:	3.9 %
Males 3-5 years old:	0.0 %	Over-population risk:	0.0 %
Youngs 1-2 years old:	0.0 %		

Table 2.5: Male culling. This strategy concentrates on adult males. Even a high culling pressure cannot maintain the population at the aimed density; in fact, the mean final density is even higher than population carrying capacity.

Culling strategy		Results	
Age classes	Culling proportion	Statistics	Value
Adult females:	0.0 %	Mean and SD final density:	183 ± 44
Males >11 years old:	15.0 %	Yearly culling:	33.5
Males 6-10 years old:	15.0 %	Extinction risk:	33.4 %
Males 3-5 years old:	70.0 %	Over-population risk:	0.0 %
Youngs 1-2 years old:	0.0 %		

First, an equilibrated strategy was applied: this strategy makes the assumption that the population follows a stable age pyramid. Culls are therefore proportionally higher on lower age classes and equally distributed between sexes (table 2.3). Then, two highly biased strategies were tried, either culling only females (table 2.4) or only males (table 2.5).

D. Discussion

These three culling strategies were addressed principally to demonstrate how useful such management software could be

to evaluate the impact of given scenarios. Equilibrated culls strategy proved to be sound, maintaining the population near the target density with neither extinction nor over-population risk. In contrast, sex-biased strategies were shown of poor efficiency. Male culling, far from lowering population density, tended to make it higher, conducting to over-population; female culling had a too great impact, conducting density lower than expected, therefore causing extinctions. Lowering female density make the population highly prone to random accidents and thus to extinction; in contrast, lowering male density has very low impact on recruitment and, by lower-

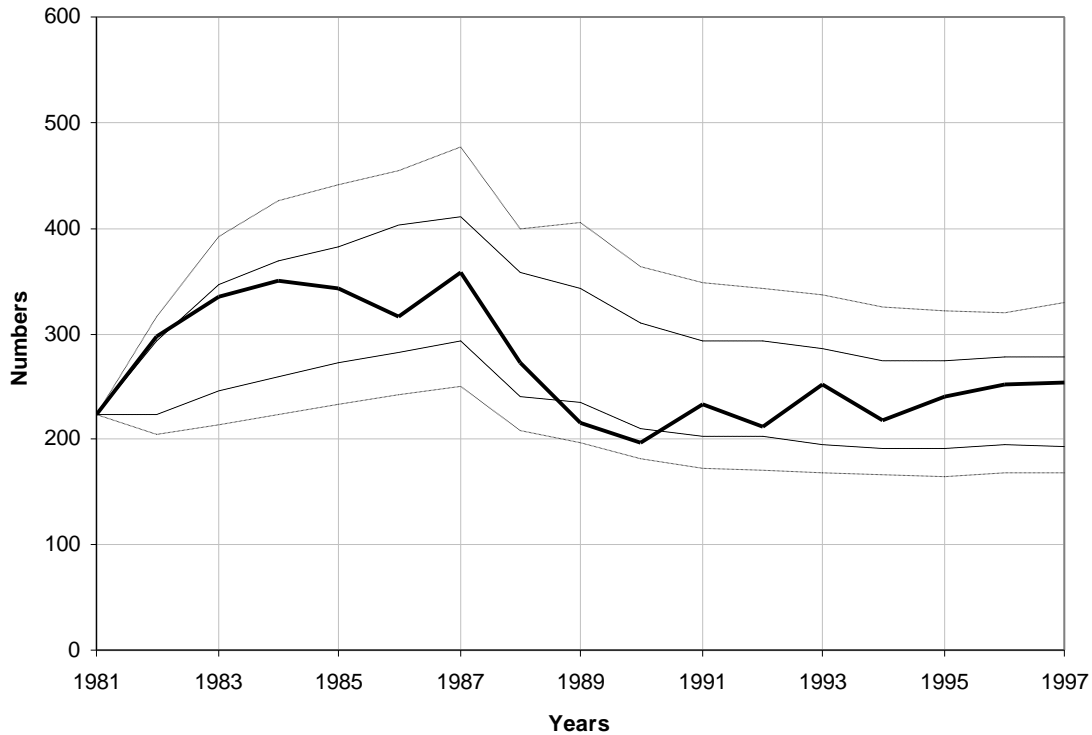


Figure 2.4: Validation of the model. The bold line shows the dynamics followed by the Val Bever population between 1981 and 1997; it was let alone until 1987 when a regulation culling was introduced, using the governmental strategy (see text). Starting with the initial conditions of 1981, 1000 simulations were run to see how well the model could reproduce the real dynamics (Hunt was also introduced in 1987). Dashed lines encompass the interquartile values of the 1000 replicates; solid lines encompass the space between the first and the ninth decile.

ing density-dependence negative effect, may entail stochastic outbursts. Both these extreme scenarios show the importance of females on population dynamics and the necessity to maintain equilibrated culling. Nevertheless, the former equilibrated strategy is fixed; it assumes the population to follow a stable age distribution and this may not be true in the case of a catastrophic event affecting only a part of it (e.g., an avalanche decimating a female group). In this perspective the *SIM-Ibex* strategy, which is based on the observed age pyramid, offers more security. These problems of biased culling are not purely academic: in the case of Ibex, trophy hunting might tend to bias cull toward older males and induce overpopulation problems.

The goal of this project was to produce a support software for fauna managers.

This implied several constraints on robustness and user-friendliness. Actually, available data, although sharing a common structure, are variable both in quality and quantity. Particularly, density-dependence parameters need several accurately-monitored years in order to get good results. As the age data are collected only since 1991, the small sample size does not allow a large manoeuvring room. As time passes, new data will be available and this problem will be reduced, but for now, a few bad years suffice to prevent density-dependence parameter computations. On the other hand, carrying capacity is based on total numbers only and is therefore much more robust.

The *SIM-Ibex* and custom management strategies rely highly on survival parameters and thus may only be used when these values are correctly estimated. When it is

not the case, governmental strategy must be preferred. The program includes checking procedures that raise an alert when model accuracy is dubious: all estimated survival parameters must lie inside a biological relevancy range ($0 < \text{intercept} \leq 1$, $-0.005 < \text{slope} \leq 0$). When this is not the case, the governmental strategy must be preferred.

Survival parameter reliability is mainly affected by three factors: 1° Monitoring accuracy has the greatest effect. In particular, age and sex classes must be equally detectable in the field; this might not be the case for example when the monitoring is performed during the birth season as pregnant females tend to hide in highly elevated areas. 2° Immigrating or emigrating individuals may also introduce biases into the model as they will be interpreted as increased or decreased survival. The model has been developed with isolated populations in mind; when this is not the case, the connected populations should be managed together. 3° Unpredictable catastrophic events like avalanches or large epidemics unrelated to density-dependence may also decrease the reliability of the data.

Although precisely studied Ibex populations have shown a concave quadratic response (J.-M. Gaillard, pers.comm.), we chose to model survival rates by a linear function of density. Indeed, available data were not numerous enough to fit accurately a non-linear behaviour. Doing so, a bias could be introduced into the model, tending to overestimate carrying capacity; this was a deliberate choice preventing from over-culling and diminishing extinction risk.

Model and culling strategies were implemented into a user-friendly software – *SIM-Ibex* – developed to assist fauna managers. It is now used by gamekeepers to tune their culling strategy as well as by governmental inspectors to control them. *SIM-Ibex* is available for free at:

<http://www.unil.ch/izea/software/simibex.html>.

E. References

- Abderhalden, W., Buchli, C., 1997. Steinbockprojekt Albris/SNP, Schlussbericht. Zerne, ARINAS & FORNAT.
- Begon, M., Harper, J.L., Townsend, C.R., 1996. Ecology. Oxford, Blackwell Science.
- Gauthier, D., Martinot, J.-P., Choisy, J.-P., Michallet, J., Villaret, J.-C., Faure, E., 1991. Le Bouquetin des Alpes. Rev. Ecol. (Terre Vie) 6, 233-272.
- Giacometti, M., 1988. Zur Bewirtschaftung der Steinbockbestände (Capra i. ibex L.). Mit einem geschichtlichen Abriss der Steinbockkolonien im Kanton Graubünden. Institut für Zuchthygiene. Zürich, 115.
- Mayer, D., Degiorgis, M.P., Meier, W., Nicolet, J., Giacometti, M., 1997. Lesions Associated With Infectious Keratoconjunctivitis in Alpine Ibex. Journal of Wildlife Diseases 33, 413-419.
- Nievergelt, B., 1966. Der Alpensteinbock (Capra ibex L.) in seinem Lebensraum : Ein ökologischer Vergleich verschiedener Kolonien.
- Nievergelt, B., Zingg, R. (1986). Capra ibex Linnaeus, 1758 - Steinbock. Handbuch der Säugetiere Europas : Paarhufer. J. Niethammer and F. Krapp. Wiesbaden, AULA-Verlag. 2/II, 384-404.
- OFEFP, 1991. Commentaires relatifs à l'ordonnance sur la régulation des Bouquetins en Suisse (ORB), OFEFP.
- Ratti, P., 1994. Stand von Hege und Erforschung des Steinwildes im Kanton Graubünden (Schweiz). Zeitschrift für Jagdwissenschaft 40, 223-331.
- Tataruch, F., Onderscheka, K., 1996. Untersuchungen zur Kondition des Steinwildes (Capra i. ibex L.) in Graubünden. Zeitschrift für Jagdwissenschaft 42, 97-103.

III. HABITAT SUITABILITY: THE ECOLOGICAL NICHE FACTOR ANALYSIS

A. Introduction

CONSERVATION ecology nowadays crucially relies on multivariate, spatially-explicit models in all research areas requiring some level of ecological realism. This includes population viability analyses (Akçakaya et al., 1995; Akçakaya and Atwood, 1997; Roloff and Haufler, 1997), biodiversity-loss risk assessment (Akçakaya and Raphael, 1998), landscape management for endangered species (Livingston et al., 1990; Sanchez-Zapata and Calvo, 1999), ecosystem restoration (Mladenoff et al., 1995; Mladenoff et al., 1997) and alien-invaders expansions (Higgins et al., 1999). Such studies often conjugate the power of Geographical Information Systems (GIS) with multivariate statistical tools to formalize the link between the species and their habitat, in particular to quantify the parameters of habitat-suitability models.

Most frequently used among multivariate analyses are logistic regressions (Jongman et al., 1987; Peeters and Gardieniers, 1998; Higgins et al., 1999; Manel

et al., 1999; Palma et al., 1999), Gaussian logistic regressions (ter Braak and Looman, 1987; Legendre and Legendre, 1998), discriminant analyses (Legendre and Legendre, 1984; Livingston et al., 1990; Corsi et al., 1998; Manel et al., 1999), Mahalanobis distances (Clark et al., 1993) and artificial neural networks (Manel et al., 1999; Özesmi and Özesmi, 1999; Spitz and Lek, 1999). All these methods share largely similar principles:

The study area is modelled as a raster map composed of N adjacent isometric cells.

The dependent variable is in the form of presence/absence data of the focal species in a set of sampled locations.

Independent, ecogeographical variables (EGV) describe quantitatively some characteristics for each cell. These may express topographical features (e.g. altitude, slope), ecological data (e.g. frequency of forests, nitrates concentration), or human superstructures (e.g. distance to the nearest town, roads density).

A function of the EGV is then calibrated so as to classify as correctly as possible the cells as suitable or unsuitable for the species. The details of the function and of its calibration depend on the analysis.

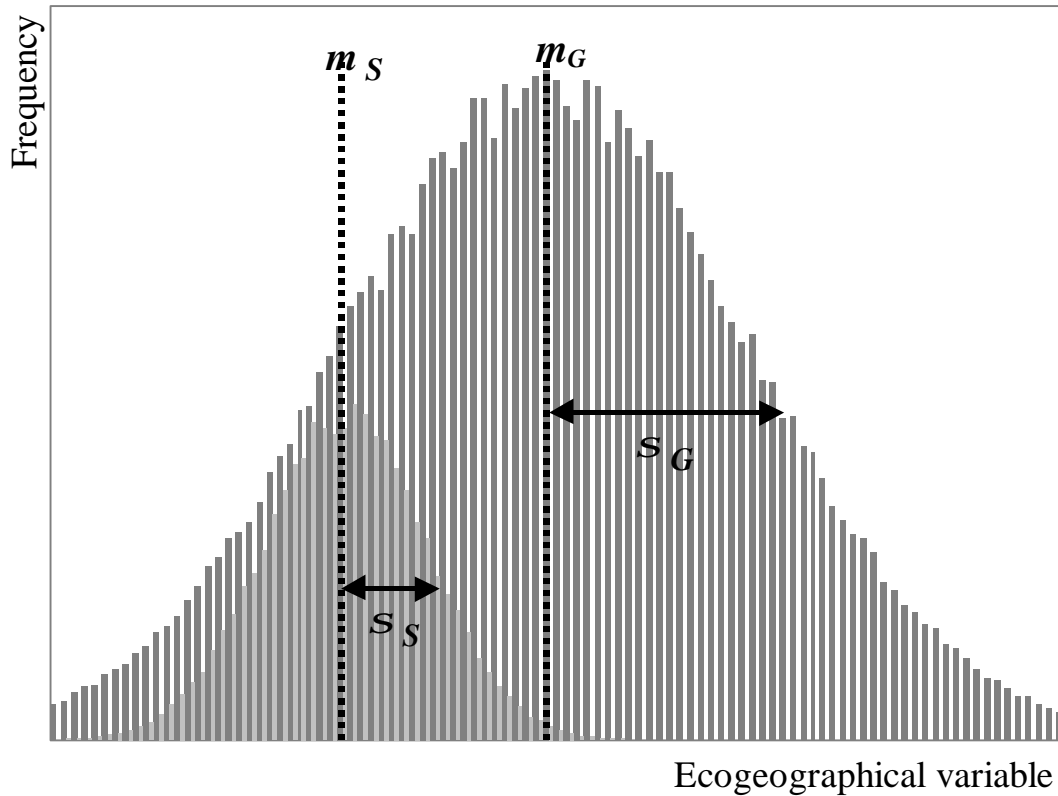


Fig.3.1: The distribution of the focal species on any eco-geographical variable (white bell curve) may differ from that of the whole set of cells (grey bell curve) with respect to its mean ($m_S \neq m_G$), thus allowing to define marginality; it may also differ with respect to standard deviations ($s_S \neq s_G$), thus allowing to define specialisation (See text for further explanations.). Note that these curves represent numbers and not probability densities.

Sampling the presence/absence data is a crucial part of the process. The sample must be unbiased to be representative of the whole population. Absence data in particular are often difficult to obtain accurately. A given location may be classified in the “absence” set because 1) the species could not be detected even though it was present (McArdle, 1990; Solow, 1993;). Kéry, (submitted) found that 34 unsuccessful visits were needed before one can assume with 95% confidence that the snake *Coronella austriaca* was absent from a given site), 2) for historical reasons the species is absent even though the habitat is suitable, or 3) the habitat is truly unsuitable for the species. Only the last cause is relevant for predictions, but occurrence of “false absences” may considerably bias analyses.

Here we propose a new approach specifically designed to circumvent this diffi-

culty. Requiring only presence data as input, the Ecological-Niche Factor Analysis (ENFA) computes suitability functions by comparing the species distribution in the EGV space with that of the whole set of cells. In the present paper, we expose the concepts behind the ENFA, we develop the mathematical procedures required (and implemented in the software *Biomapper*) and we illustrate this approach through an habitat-suitability analysis of the alpine Ibex (*Capra ibex*).

B. Marginality, specialisation, and the ecological niche

Species are expected to be non-randomly distributed regarding ecogeographical variables. A species with a temperature optimum, for instance, is expected to occur preferentially in cells lying

within its optimal range. This may be quantified by comparing the temperature distribution of the cells in which the species was observed with that of the whole set of cells. These distributions may differ with respect to their mean and their variances (Fig.3.1). The focal species may show some *marginality* (expressed by the fact that the species mean differs from the global mean) and some *specialisation* (expressed by the fact that the species variance is lower than the global variance).

Formally, we define the marginality (M) as the absolute difference between global mean (m_G) and species mean (m_S), divided by 1.96 standard deviations (s_G) of the global distribution,

$$M = \frac{|m_G - m_S|}{1.96s_G}. \quad (3.1)$$

Division by s_G is needed to remove any bias introduced by the variance of the global distribution: a cell randomly-chosen from a distribution is a priori expected to lie the further from the mean that the variance in distribution is large. The coefficient weighting s_G (1.96) ensures that marginality will be most often comprised between zero and one. Namely, if the global distribution is normal, the marginality of a randomly-chosen cell has only a 5% chance of exceeding unity. A large value (close to one) means that the species lives in a very particular habitat *relative to the reference set*. Note that equation (3.1) is given here mainly to explain the principle of the method; the operational definition of marginality implemented in our software is provided by equation (3.10), which is a multivariate extension of (3.1).

Similarly, we define the specialisation (S) as the ratio of the standard deviation of the global distribution (s_G) to that of the focal species (s_S).

$$S = \frac{s_G}{s_S} \quad (3.2)$$

A randomly-chosen set of cells is expected to have a specialisation of one, and any value exceeding unity indicates some form of specialisation. We reemphasise that specific values for these indexes are bound to depend on the global set chosen as reference, so that a species might appear extremely marginal or specialised on the scale of a whole country, but much less so on a subset of it.

Extending these statistics to a larger set of variables directly leads to Hutchinson's (1957) concept of the ecological niche, defined as an hyper-volume, in the multidimensional space of ecological variables, within which a species can maintain a viable population (Hutchinson, 1957; Begon et al., 1996). The concept is used here exactly in the same sense: by ecological niche we refer to the subset of cells in the ecogeographical space where the focal species has a reasonable probability to occur. This multivariate niche can be quantified on any of its axes by an index of marginality and specialisation.

Some of these axes are obviously more interesting than others, and this is why a factor analysis is introduced. The reasons are actually double. First, ecological variables are not independent. As more and more are introduced in the description, multicollinearity and redundancy arise. One aim of factor analyses is to transform V correlated variables into the same number of uncorrelated factors. As these factors explain the same amount of total variance, subsequent analyses may be restricted to the few important factors (e.g., those explaining the largest part of the variance) without losing too much information.

Second, specialisation is expected to depend on interactions among factors. For instance, the temperature one species prefers might vary with humidity. Species may thus specialise on a combination of variables, rather than on every variable independently. A factor analysis may allow

extracting those combinations of the original variables on which the focal species shows most of its marginality and specialisation. In Principal Component Analyses (Cooley and Lohnes, 1971; Legendre and Legendre, 1998), axes are chosen so as to maximise the variance of the distribution. In ENFA, by contrast, the first axis is chosen so as to account for all the marginality of the species, and the following axes so as to maximise specialisation, i.e. the ratio of the variance in the global distribution to that in the species distribution.

C. Factor extraction

1. Outline of the principles

We use raster maps, which are grids of N isometric cells covering the whole study area. Each cell of a map contains the value of one variable. Eco-geographical maps

contain continuous values, measured for each of the V descriptive variables. Species maps contain boolean values (0 or 1), a value of 1 meaning that the presence of the focal species was proved on this cell. A value of zero simply means absence of proof.

Each cell is thus associated to a vector whose components are the values of the EGV in the underlying area, and can be represented by a point in the multidimensional space of the EGVs. If distributions are multinormal, the scatterplot will have the shape of a hyper-ellipsoid (fig.3.2). The cells where the focal species was observed constitute a subset of the global distribution and are plotted as a smaller hyper-ellipsoid within the global one. The first factor, or *marginality factor*, is the straight line passing through the centroids of the two ellipsoids. The marginality of the species is the distance between these centroids, standardised as in equation (3.1).

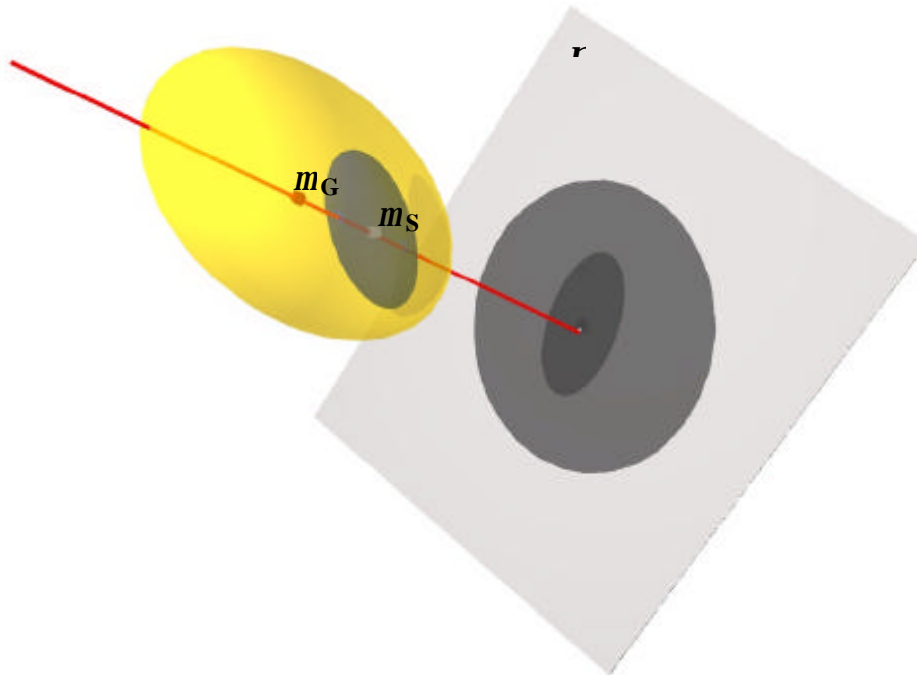


Fig 3.2: Geometrical interpretation of the Ecological Niche Factor Analysis. Square cells of the study area are represented in a three-EGVs space. The large grey balloon represents the global cloud of cells, while the smaller, dark balloon represents the subset of cells where the focal species was observed. The straight line passing through their centroids (μ_G and μ_S) is the marginality factor. In order to extract the variance associated with this factor, the cells coordinates are projected on a plane π perpendicular to it, thereby producing the light and dark grey ellipses. In reality, those operations are typically conducted with 20 to 30 EGVs.

Fig.3.2 plots this step for a three-dimensional initial set.

To obtain the *specialisation factors*, a change of reference system is then performed, in order to transform the species ellipsoid into a sphere, the variance of which equals unity in each direction. In this new system, the first *specialisation factor* is the one that maximises the variance of the global distribution (while being orthogonal to the marginality factor). The following specialisation factors are then extracted in turn, each step removing one dimension from the space, until all V factors are extracted. The amount of specialisation accounted for by a factor is necessarily lower than that due to the one previously extracted. All specialisation factors are furthermore orthogonal, in the sense that the distribution of the species subset on any factor is uncorrelated with its distribution on the others. A limited number (F) of the first factors will thus generally contain most of the relevant information. Their small number and independence make them easier to use than the original EGVs, so that all following operations will be restricted to them. In particular, the suitability of any cell for the focal species (be it classified as 0 or 1 for observation data) will be calculated according to its position in the F -dimensional space.

2. Mathematical procedures

Ecogeographical variables are first normalized as far as possible, e.g. through Box-Cox transformation (Sokal and Rohlf, 1981). Though multinormality is theoretically needed for factor extraction through eigensystem computation (Legendre and Legendre, 1998), the method seems quite robust to deviations from normality (Glass and Hopkins, 1984). EGVs are then standardised by retrieving means and dividing by standard deviations:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_{x_j}}, \quad (3.3)$$

where x_{ij} is the value of the variable x_j in cell i , \bar{x}_j the mean of this variable over all cells, and s_{x_j} its standard deviation. Let \mathbf{Z} be the $N \times V$ matrix of standardised measurements z_{ij} . The $V \times V$ covariance matrix among standardised variables is then computed as:

$$\mathbf{R}_G = \frac{1}{N} \mathbf{Z}^t \mathbf{Z}, \quad (3.4)$$

where \mathbf{Z}^t is the transposed matrix of \mathbf{Z} . Because of standardisation (equation. 3.3), \mathbf{R}_G is also a correlation matrix.

The N_S lines of \mathbf{Z} corresponding to the N_S cells where the focal species was detected are then stored in a new $N_S \times V$ matrix (say \mathbf{S}), from which the $V \times V$ species covariance matrix is calculated:

$$\mathbf{R}_S = \frac{1}{N_S - 1} \mathbf{S}^t \mathbf{S} \quad (3.5)$$

Note that, in contrast to \mathbf{R}_G , \mathbf{R}_S is not a correlation matrix, since standardisation was performed on the global data set, not on the species subset.

Let \mathbf{u} be a normed vector of the EGV space. The variance of the global distribution on this vector is $\mathbf{u}^t \mathbf{R}_G \mathbf{u}$, while that of the species distribution is $\mathbf{u}^t \mathbf{R}_S \mathbf{u}$. The first specialisation factor should thus maximize

the ratio $\Theta(\mathbf{u}) = \frac{\mathbf{u}^t \mathbf{R}_G \mathbf{u}}{\mathbf{u}^t \mathbf{R}_S \mathbf{u}}$. We also require,

however, that this vector be orthogonal to the marginality factor, \mathbf{m} , given as the vector of means over the V columns of \mathbf{S} :

$$\mathbf{m} = \left\{ \frac{1}{N_S} \sum_{i=1}^{N_S} z_{ij} \right\} \quad (3.6)$$

The problem therefore becomes that of finding the vector \mathbf{u} that maximises $\Theta(\mathbf{u})$ under the constraint $\mathbf{m}^t \mathbf{u} = 0$. This is equivalent to find \mathbf{u} such that

$$\begin{cases} \mathbf{u}^t \mathbf{R}_S \mathbf{u} = 1 \\ \mathbf{u}^t \mathbf{m} = 0 \\ \mathbf{u}^t \mathbf{R}_G \mathbf{u} \quad \max \end{cases} \quad (3.7)$$

A change in variables allows to rewrite the problem:

$$\begin{cases} \mathbf{v}^t \mathbf{v} = 1 \\ \mathbf{v}^t \mathbf{y} = 0 \\ \mathbf{v}^t \mathbf{W} \mathbf{v} \quad \max \end{cases} \quad (3.8)$$

where $\mathbf{v} = \mathbf{R}_S^{1/2} \mathbf{u}$, $\mathbf{y} = \frac{\mathbf{z}}{\sqrt{\mathbf{z}^t \mathbf{z}}}$,

$\mathbf{z} = \mathbf{R}_S^{-1/2} \mathbf{m}$, and $\mathbf{W} = \mathbf{R}_S^{-1/2} \mathbf{R}_G \mathbf{R}_S^{-1/2}$ is a symmetric matrix. It can be shown that the solution is given by the first eigenvector of

$$\mathbf{H} = (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \mathbf{W} (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \quad (3.9)$$

Indeed,

\mathbf{y} is an eigenvector of \mathbf{H} because $\mathbf{H} \mathbf{y} = (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \mathbf{W} (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \mathbf{y} = \mathbf{0}$

\mathbf{H} is symmetrical and thus admits a base of orthonormed eigenvectors so that

$$\mathbf{H} \mathbf{v} = \lambda \mathbf{v} \Rightarrow \mathbf{v}^t \mathbf{y} = 0$$

$\mathbf{v}^t \mathbf{H} \mathbf{v}$ is maximum for the first eigenvector, which also maximises $\mathbf{v}^t \mathbf{W} \mathbf{v}$ since $\mathbf{v}^t \mathbf{y} = 0 \Rightarrow$

$$\mathbf{v}^t \mathbf{H} \mathbf{v} = \mathbf{v}^t (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \mathbf{W} (\mathbf{I}_V - \mathbf{y} \mathbf{y}^t) \mathbf{v} = \mathbf{v}^t \mathbf{W} \mathbf{v}$$

The V eigenvectors of \mathbf{H} are then back transformed, and the new eigenvectors ($\mathbf{u} = \mathbf{R}_S^{-1/2} \mathbf{v}$) are stored in a matrix \mathbf{U} . These vectors are \mathbf{R}_S -orthogonal (all $\mathbf{S} \mathbf{u}$ have variance 1 and are uncorrelated). Furthermore, due to the constraint that \mathbf{u} be orthogonal to \mathbf{m} , this system has one null eigenvalue. The corresponding eigenvector is thus deleted from \mathbf{U} , and \mathbf{m} is substituted instead as the first column. It should be noted that, although all marginality is accounted for by the first factor, this factor is not “pure”, in that the niche of the focal species may also display some restriction on it, in addition to its departure from average. The amount of specialisation on this first axis is provided by the difference be-

tween the traces (sum of all eigenvalues) of \mathbf{W} and \mathbf{H} .

All these procedures are implemented in the software “*Biomapper*” (Hirzel et al., 2001) available at our website (<http://www.unil.ch/biomapper>).

3. Interpretation of the factors

The coefficients m_i of the marginality factor express the marginality of the focal species on each EGV, in units of standards deviations of the global distribution. The higher the absolute value of a coefficient, the further the species departs from the mean available habitat regarding the corresponding variable. Negative coefficients indicate that the focal species prefers values that are lower than average with respect to the study area, while positive coefficients indicate preference for higher-than-average values. An overall marginality M can be computed over all EGV as:

$$M = \frac{\sqrt{\sum_{i=1}^V m_i^2}}{1.96} \quad (3.10)$$

so that the marginalities of different species within a given area can be directly compared.

The coefficients of the next factors receive a different interpretation: the higher the value, the more restricted is the range of the focal species on the corresponding variable. Note that only absolute values matter here, since signs are arbitrary. The eigenvalue I_i associated to any factor expresses the amount of specialisation it accounts for, i.e. the ratio of the variance of the global distribution to that of the species distribution on this axis. Eigenvalues usually rapidly decrease from the second factor to the last one, so that only the few four or five first axes are useful to compute habitat suitability (see below). Different criteria may be used for the selection process, such as direct comparison with the broken-stick distribution, or threshold value for cumulative variance.

A global specialisation index can be computed as

$$S = \sqrt{\frac{\sum_{i=1}^V I_i}{V}} \quad (3.11)$$

and may be used for among-species comparisons, provided the same area is used as reference.

D. Habitat suitability map

A variety of methods can be envisaged to compute the suitability for the focal species of any cell from the study area. Among the several alternatives tested, the following approach turned out to be quite robust and was implemented in *Biomap*-

per. It builds on a count of all cells *from the species distribution* that lay as far or farther apart from the median than the focal cell on a factor axis. This count is normalised in such a way that the suitability index ranges from zero to one.

Practically, this is performed by dividing the species range on each selected factor in a series of classes, in such a way that the median would exactly separate two classes (fig.3.3). For every cell from the global distribution, we count the number of cells from the species distribution that lay either in the same class or in any class farther apart from the median *on the same side* (fig.3.3). Normalisation is achieved by dividing twice this number by the total number of cells in the species distribution. Thus, a cell laying in one of the two

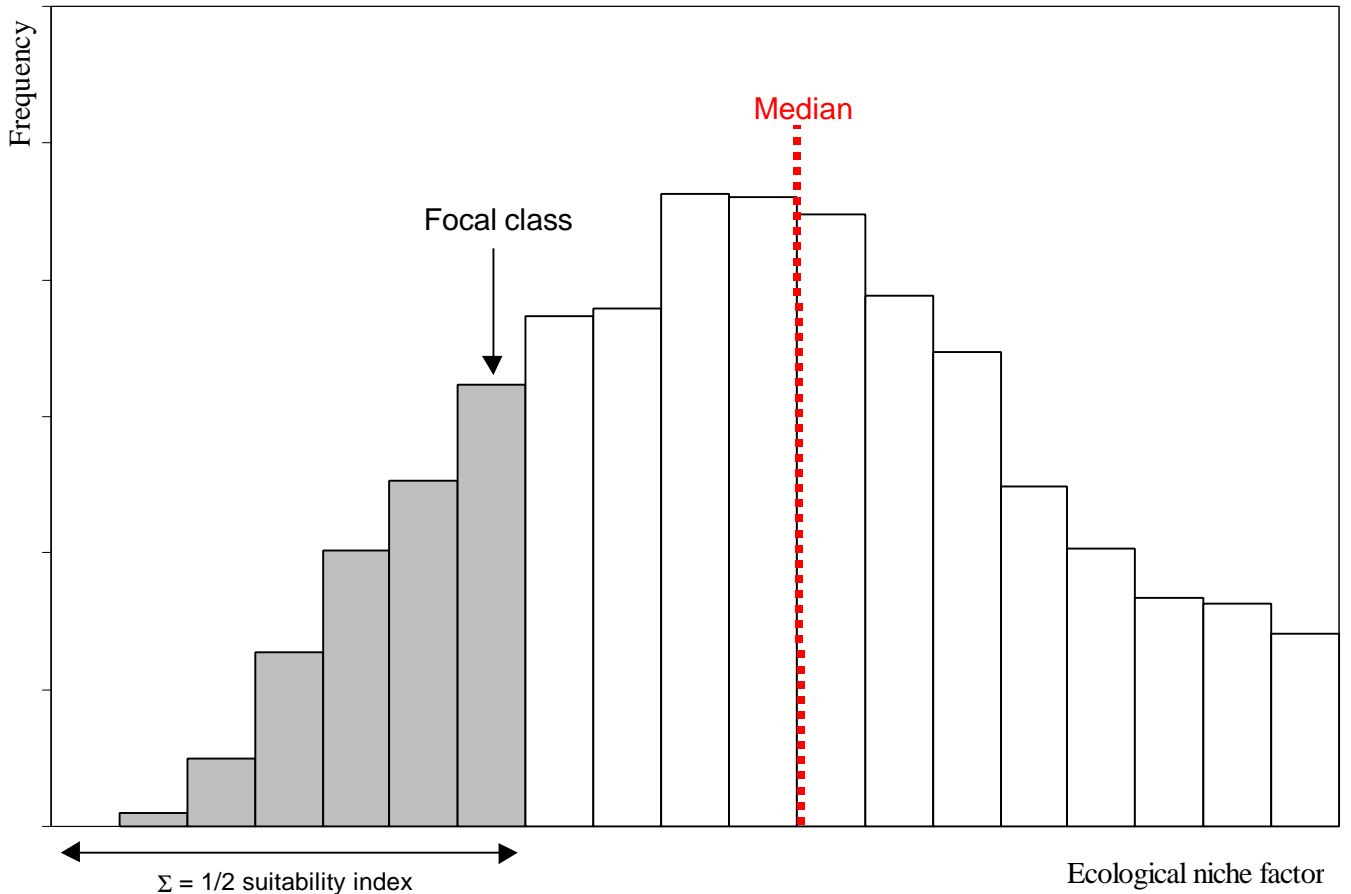


Fig 3.3: The suitability of any cell from the global distribution is calculated from its situation (arrow) relative to the species distribution (histogram) on all selected niche factors. Specifically, it is calculated as twice the dashed area (sum of all cells from the species distribution that lay as far or farther from the median, dashed vertical line) divided by the total number of cells from the species distribution (surface of the histogram).

classes directly adjacent to the median would score one, and a cell laying outside the species distribution would score zero.

An overall suitability index of the focal cell can then be computed from a combination of its scores on each factor. In order to account for the differential ecological importance of the factors, we attribute equal weight to marginality and specialisation, but, while all the marginality component goes to the first factor, the specialisa-

tion component is apportioned among all factors proportionally to their eigenvalue (the marginality factor may thus take more than half of the weight if it also accounts for some specialisation).

Repeating this procedure for each cell allows to produce a habitat-suitability map, where suitability values range from 0 to 1. To convert this quantitative (or semi-quantitative) map into a presence/absence one, a threshold value may be chosen,

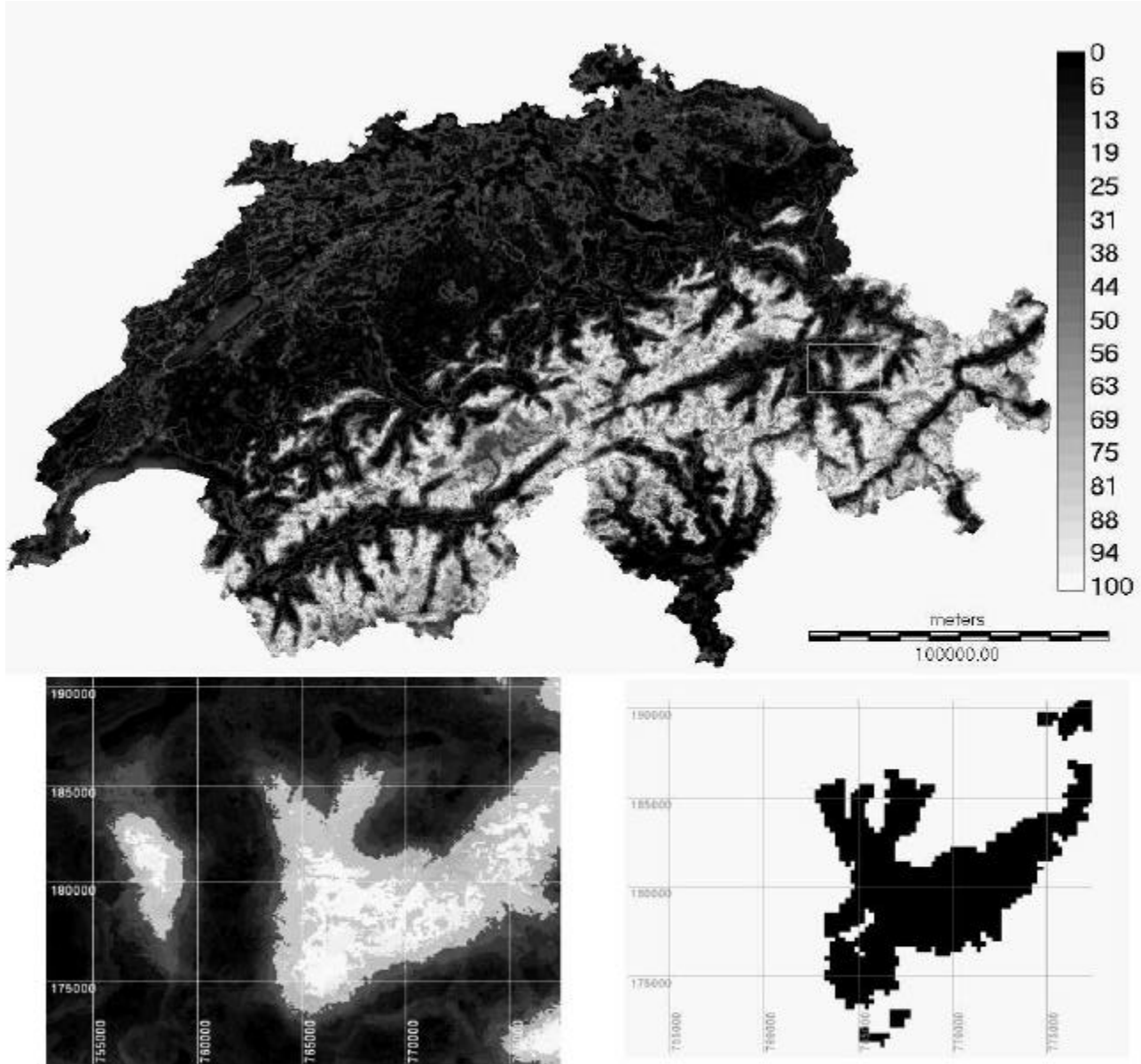


Fig 3.4: Habitat-suitability map for Alpine Ibex in Switzerland, as computed from ENFA. Below is displayed part of Eastern Switzerland at a larger scale, with the HS-map on the left, and presence data on the right. The largest suitable patch is indeed occupied, while the smaller one was not colonised, being either too small or unreachable.

above which the cell will be considered as suitable. A validation data set can be used to find out the best threshold value, e.g. through the ROC plot method (Zweig and Campbell, 1993; Fielding and Bell, 1997), given some a-priori cost values to each type of inferential error. Typically, since our method builds on presence data only, lower costs should be attributed to cells wrongly considered as suitable than to cells wrongly considered as unsuitable.

E. An application to the alpine Ibex

We saw that alpine Ibex (*Capra ibex*) was exterminated from the Swiss Alps in the last century, due to excessive hunting. Reintroduction attempts, starting from 1911, were highly successful, and colonies have since grown rapidly throughout the Swiss Alps. As a legally protected species, Ibex populations were carefully monitored since their reintroduction, so that presence data are highly reliable. However, as their expansion is still hindered by the patchiness of their habitat, they presumably do not occupy yet all suitable regions. Absence data therefore do not necessarily reflect poor-quality habitat, a point which strongly advocates for the use of an analysis relying on presence data only.

The whole Switzerland was chosen as reference area, and modelled as a raster map based on the Swiss Coordinate System (plane projection), comprising 4'145'530 square cells of 1ha (100 x 100 m) each. We used 34 ecogeographical

variables derived from governmental data bases (table 3.1). Topographical data (altitude, slope and aspect) were directly obtained as quantitative variables. Frequency- and distance data were derived from boolean variables describing soil occupancy, as official data bases attribute each cell to one category only (snow, rocks, meadow, forest, building, etc), according to a regular sampling. Distance data express the distance between the focal cell and the closest cell belonging to a given category. Frequency describes the proportion of cells from a given category within a circle of 1200 m radius around the focal cell (circle surface is about 5 km², which corresponds to the mean area explored daily by individual Ibexes (Abderhalden and Buchli, 1997).

The presence database was a digitised map of Ibex home ranges. The polygons drawn by fauna managers were converted in raster format at the same resolution as EGV maps. This raster was then randomly partitioned in two data sets, every cell having a 0.5 chance to belong to one of each set. The first set (101'564 cells) was used to calibrate the model, and the other set (101'550 cells) to validate it. Application of the ENFA method to the calibration set provided an overall marginality of $M = 1.1$ and an overall specialisation value of $S = 2.2$, i.e. a tolerance of 0.45, showing that Ibex' habitat differs drastically from the average conditions in Switzerland, and that Ibex are quite restrictive on the range of conditions they withstand. The five factors retained (out of the 34 computed) accounted for 74% of the total sum of ei-

Table 3.1: Nature and source of the 34 eco-geographical variables used in the ENFA of Ibex distribution. OFT is the swiss federal office of topography, OFS is the swiss federal office of statistics and OFEFP is the swiss federal office of environment, forest and landscape.

Official data base	Source	EGV derived from it
AS85R (Land use)	OFS	Rock, snow, forests, meadows,...: frequency and distance
DHM (Topography)	OFS	Altitude, slope, aspect, SD altitude
GWN (Hydrography)	OFS	Proximity of rivers and lakes
Vector 200 (land use)	OFT	Villages, towns, railways, roads, etc.: distance
SB (Ibex colonies)	OFEFP	Calibrating and validating presence data set

genvalues (that is, 100% of the marginality and 74% of the specialisation). The marginality factor alone accounted for 46% of this total specialisation, a quite important value, meaning that Ibex display a very restricted range on those conditions for which they mostly differ from “background Switzerland”.

Marginality coefficients (table 3.2) showed that Ibexes are essentially linked to high-altitude, steep and rocky slopes, rich in pastures (rock frequency = 0.35, altitude = 0.27, frequency of slopes greater than 30° = 0.23, grass frequency = 0.32, distance to grass = -0.24). By contrast, Ibex tends to avoid forest (frequency = -0.23) and human activities (distance to secondary roads = 0.22, distance to agricultural meadows = 0.23). Aspect (northness, eastness) as well as snow and water (lakes, rivers) had only marginal effects. The very large eigenvalue (76.6) attributed to this first factor means that randomly-chosen cells in Switzerland are about 80 times more dispersed on this axis than are the cells where Ibex was recorded. Or in other words, Ibex are extremely sensitive to shifts from their optimal conditions on

this axis. The next factors account for some more specialisation, mostly regarding agricultural meadow frequency and altitude (2nd factor) as well as forest frequency (3rd factor), showing some sensitivity to shifts away from their optimal values on these variables.

A suitability map was built from these five factors for the whole Switzerland, which is plotted on fig.3.4. Enlargement of a small part of it shows one suitable patch where ibex was indeed recorded (right), as well as one smaller suitable patch (left) where Ibex was not recorded. This patch presumably was not colonised because of its isolation, or too small to sustain a viable colony on the long term. As false positives like those ones give no indication about the quality of our model, standard quality estimators like the kappa index (Monserud and Leemans, 1992), which attribute the same importance to false positives and false negatives, cannot be used to validate it. Instead, we evaluated the distribution of suitability values of cells from the validation set. As shown in fig. 3.5, these cells differ drastically from the global distribution. Predicted suitability exceeds 0.5 in

Table 3.2: Variance explained by the first five (out of 34) ecological factors, and coefficient values for the 13 most important initial variables. EGVs are sorted by decreasing absolute value of coefficients on the marginality factor. Positive values on this factor mean that Ibex prefers locations with higher values on the corresponding EGV than average location in Switzerland. Signs of coefficient have no meaning on the specialisation factors. The amount of specialisation accounted for is given in parentheses.

EGV	Marginality (46%)	Spec. 1 (11%)	Spec. 2 (8%)	Spec. 3 (6%)	Spec. 4 (4%)
Rock frequency	0.350	-0.105	-0.132	-0.186	-0.163
Grass frequency	0.321	-0.021	-0.044	-0.043	-0.039
Altitude	0.269	-0.365	-0.561	0.005	0.111
Distance to grass	-0.245	0.021	-0.096	-0.121	-0.073
Dist.to agric.Meadows	0.231	-0.062	0.017	0.558	-0.784
Slope>30° frequency	0.230	-0.035	0.074	0.011	0.004
Dense forest frequency	-0.228	0.025	-0.727	0.037	-0.001
Distance to 2° roads	0.222	-0.012	-0.049	-0.129	0.115
Agric. meadows freq.	-0.212	-0.907	-0.011	0.236	-0.383
Distance to towns	0.209	0.012	0.049	-0.021	0.002
Std dev. of altitude	0.204	0.000	-0.005	-0.032	-0.044
Distance to forests	0.203	0.015	0.165	-0.068	0.108
Distance to villages	0.200	0.003	0.017	-0.097	0.320

83% of them, which differs highly significantly ($p < 0.0001$, bootstrap test) from the value of 24% expected if cells were randomly chosen from the global distribution.

F. Discussion

Niche factors and distribution maps

The originality of the present approach lies in the fact that it builds on the concept of ecological niche, which is central to the whole field of ecology. A basic tenet of the niche theory is that fitness (or habitat suitability) does not bear monotonic relationships with conditions or resources, but instead decreases from either side of an op-

timum. In this respect, our approach differs fundamentally from other techniques like discriminant functions or first-order regressions, where relationships are assumed linear and monotonic. Accordingly, our analysis directly provides two key measurements regarding the niche of the focal species, namely those of marginality and of specialisation. Outputs thus have intuitive ecological meaning, and allow direct comparisons with the niche of different species. Application of the analysis to evaluate e.g. species packaging or niche-overlap measurement among members of a guild would be straightforward extensions of the present approach (e.g. Dolédec et al., 2000).

Our application to Ibex data, for in-

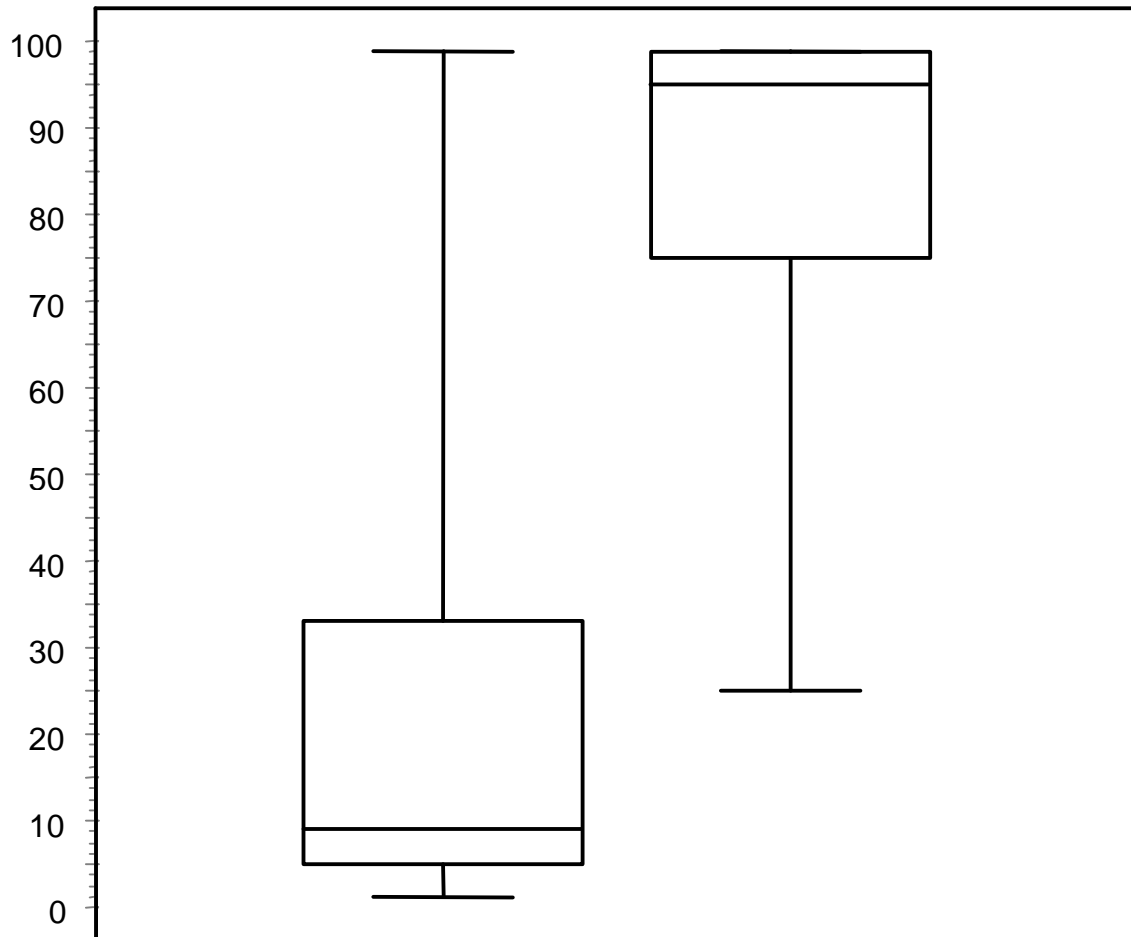


Fig 3.5: Box-plots presenting the distributions of the habitat-suitability values for the whole set of cells (left) and the validation subset (right). The latter is made of the 101'550 cells with Ibex but not included in the analysis. Boxes delimit the interquartile range, the middle line indicating the median; whiskers encompass the 80% confidence interval. The two distributions obviously differ, as the global one is mostly confined to low values (suitability 5-33%), while the validation set concentrates on high-suitability values (75-99%).

stance, provides quantitative estimates of marginality and specialisation for this species which evidence its very peculiar ecological requirements. Furthermore, interpretation of the factors in terms of the EGVs turns out to be very consistent with the experience of field specialists. In particular, the EGVs that correlate with the marginality factor are precisely those most often cited as particularly relevant for *Ibex* ecology (Rauch, 1941; Nievergelt, 1966; Nievergelt and Zingg, 1986; Hausser, 1995; Hindenlang and Nievergelt, 1995).

These results obviously suffer from the same caveat as any inferential approach: a variable might turn out to correlate with one of the main axes, not because of its intrinsic importance, but because it correlates strongly with another, crucially important variable. ENFA is a purely descriptive method and cannot extract causality relations. Nonetheless, it provides (at worst) important clues about preferential conditions, and remains a powerful tool to draw potential habitat maps.

In this respect, a limitation of our software is that it does not yet include confidence intervals on distribution maps. Increasingly, conservation managers are demanding risk analyses that incorporate uncertainties in model predictions. These could clearly be obtained through the bootstrapping of presence data. Though not yet implemented in *Biomapper*, this procedure will certainly provide an important and useful extension.

A second limitation, less easy to deal with, is that ENFA only handles linear dependencies within the species niche. Multiplicative or non-linear interactions cannot be accommodated in the present state, except through transformations or non-linear combinations of the original eco-geographical variables.

A third limitation is that some EGVs may turn out to be constant in \mathbf{S} , or in linear combination with other EGVs, which makes \mathbf{R}_S singular. This is likely to happen with coarsely measured data or small species data sets. Whenever this happens,

Biomapper identifies the constant or correlated EGVs and removes (one of) them from the analysis. An alternative approach would obviously consist in improving the field sample, either by increasing the presence data set, or by measuring EGVs on a finer scale.

Finally, a last important point to emphasise again is that our approach characterises ecological niches *relative* to a reference area. Marginality and specialisation are thus bound to depend on the geographic limits of the study area. Some species may turn out to occur at the very edge of their distribution, and may thus appear quite specialised in the reference set, however widespread they might be otherwise. Reciprocally, *Ibex* would have appeared much less marginal and specialised, had our sampling area been restricted to the Alps. The same distinction must be applied here as the one made between fundamental and realised niches (Hutchinson, 1957). Our analysis does not investigate fundamental niches, but only their specific realisation within a given geographical and temporal context.

ENFA versus logistic regressions

With respect to more standard techniques, a crucial advantage of ENFA is that it does not require absence data. Presence data are compared instead with “background” environment. This of course implies that presence data should be unbiased samples of actual distributions, which we suspect might not be the case of many available database (sampling efforts are frequently biased with respect to environment). However, though this problem is difficult to circumvent, the point must also be made that database often simply lack any absence data. And when available, these may turn out to be either unreliable (in the case of cryptic or poorly known species) or meaningless (in the case of invading species, or those living in fragmented habitats where some patches have become extinct). As many species enter one of these categories, our approach po-

tentially has a wide application range. In particular, predictions about the expected expansion of invading species seem a promising test bed.

In the case of stable populations from well-known species, one might prefer more classical approaches such as logistic regressions, able to extract relevant information from absence or abundance data. This point deserves proper investigation, in order to localise the threshold where the benefits gained from incorporating absence data are compensated by the costs induced by their possible poor quality. Investigations are presently in progress to compare the power of ENFA to that of classical logistic regression analyses under different biological and sampling scenarios, in order to assess their respective advantages and inconveniences. Preliminary results show ENFA to be more robust than classical logistic regressions with respect to several habitat-occupancy scenarios.

Finally, the point must also be made that the procedures used by standard stepwise analyses to select significant EGVs from the original set turn out to be highly sensitive to the algorithms chosen, as well as to the input order. Consequences are that 1) many trials are needed in order to sort out the “best” model, and 2) variables that bear a causal relationship to the focal species’ presence might well be lost in the process, if other EGVs present spurious correlations. This implies some subjective choices, and requires a good *a priori* knowledge of the focal species’ ecology. In contrast, our factor analysis does not reject any input EGV, but only weights them. The subjective components and *a priori* knowledge required are thereby kept minimal, and correlations among variables and axes are immediately visible and interpretable.

G. Note

A modified version of this chapter has been published in Ecology (Hirzel et al, in press).

H. References

- Abderhalden, W., Buchli, C., 1997. Steinbockprojekt Albris/SNP, Schlussbericht. Zerne, ARINAS & FORNAT.
- Akçakaya, H.R., Atwood, J.L., 1997. A habitat-based metapopulation model of the California Gnatcatcher. *Conserv. Biol.* 11, 422-434.
- Akçakaya, H.R., McCarthy, M.A., Pearce, J.L., 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., Raphael, M.G., 1998. Assessing human impact despite uncertainty viability of the northern spotted owl metapopulation in the northwestern USA. *Biodiversity and conservation* 7, 875-894.
- Begon, M., Harper, J.L., Townsend, C.R., 1996. *Ecology*. Oxford, Blackwell Science.
- Clark, J.D., Dunn, J.E., Smith, K.G., 1993. A multivariate model of female black bear habitat use for a geographic information system. *J. Wildl. Manage.* 57, 519-526.
- Cooley, W.W., Lohnes, P.R., 1971. *Multivariate data analysis*. New-York, John Wiley & sons.
- Corsi, F., Duprè, E., Boitani, L., 1998. A large-scale model of wolf distribution in Italy for conservation planning. in prep.
- Dolédéc, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche separation in community analysis: a new method. *Ecology* 81, 2914-2927.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38-49.
- Glass, G.V., Hopkins, K.D., 1984. *Statistical methods in education and psychology*. London, Prentice Hall int.
- Hausser, J., 1995. *Mammifères de Suisse*. Bâles, Birkhäuser.
- Higgins, S.I., Richardson, D.M., Richard, M.C., Trinder-Smith, T.H., 1999. Predicting the landscape-scale distribution of alien

- plants and their threat to plant diversity. *Conserv. Biol.* 13, 303-313.
- Hindenlang, K., Nievergelt, B. (1995). *Capra ibex* L., 1758. *Mammifères de Suisse*. Birkhäuser. Bâles, 450-456.
- Hirzel, A., Hausser, J., Chessel, D., Perrin, N., in press. Ecological-Niche Factor Analysis: How to compute habitat-suitability maps without absence data? *Ecology*.
- Hirzel, A.H., Hausser, J., Perrin, N., 2001. *Biomapper 1.0*. Lausanne, Lab. for Conservation Biology.
URL: <http://www.unil.ch/biomapper>
- Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harbour symposium on quantitative biology.
- Jongman, R.H.G., ter Braak, C.J.F., Van Tongeren, O.F.R., 1987. *Data analysis in Community and Landscape ecology*. Cambridge, Cambridge university press.
- Kéry, M., submitted. Inferring the absence of a species from sighting data: a case study with three snake species. -.
- Legendre, L., Legendre, P. (1984). *Analyse discriminante. Ecologie numérique*. Québec, Masson. 2, 152-162.
- Legendre, L., Legendre, P., 1998. *Numerical ecology*. Amsterdam, Elsevier Science PV.
- Livingston, S.A., Todd, C.S., Krohn, W.B., Owen, R.B., 1990. Habitat models for nesting bald eagles in Maine. *J. Wildl. Manage.* 54, 1990.
- Manel, S., Dias, J.M., Buckton, S.T., Ormerod, S.J., 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *J. Appl. Ecol.* 36, 734-747.
- McArdle, B.H., 1990. When are rare species not there? *OIKOS* 57, 276-277.
- Mladenoff, D.J., Haight, R.C., Sickley, T.A., Wydeven, A.P., 1997. Causes and implications of species restoration in altered ecosystems. A spatial landscape projection of wolf population recovery. *Bioscience* 47, 21-31.
- Mladenoff, D.J., Sickley, T.A., Haight, R.G., Wydeven, A.P., 1995. A regional landscape analysis and prediction of favorable Gray Wolf habitat in the northern Great Lakes region. *Conserv. Biol.* 9, 279-294.
- Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. *Ecol. Model.* 62, 275-293.
- Nievergelt, B., 1966. *Der Alpensteinbock (Capra ibex L.) in seinem Lebensraum : Ein ökologischer Vergleich verschiedener Kolonien*.
- Nievergelt, B., Zingg, R. (1986). *Capra ibex* Linnaeus, 1758 - Steinbock. *Handbuch der Säugetiere Europas : Paarhufer*. J. Niethammer and F. Krapp. Wiesbaden, AULA-Verlag. 2/II, 384-404.
- Özesmi, S.L., Özesmi, U., 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecol. Model.* 116, 15-31.
- Palma, L., Beja, P., Rodrigues, M., 1999. The use of sighting data to analyse Iberian lynx habitat and distribution. *J. Appl. Ecol.* 36, 812-824.
- Peeters, E.T.H., Gardeniers, J.J.P., 1998. Logistic regression as a tool for defining habitat requirements of two common gammarids. *Freshwater Biology* 39, 605-615.
- Rauch, A., 1941. *Le Bouquetin dans les Alpes*. Paris.
- Roloff, G.J., Haufler, J.B., 1997. Establishing population viability planning objectives based on habitat potentials. *Wildlife Society Journal* 25, 895-904.
- Sanchez-Zapata, J.A., Calvo, J.F., 1999. Raptor distribution in relation to landscape composition in semi-arid Mediterranean habitats. *J. Appl. Ecol.* 36, 254-262.
- Sokal, R.R., Rohlf, F.J., 1981. *Biometry: The principles and practice of statistics in biological research*. New York, W.H. Freeman & co.
- Solow, A.R., 1993. Inferring extinction from sighting data. *Ecology* 74, 962-964.
- Spitz, F., Lek, S., 1999. Environmental impact prediction using neural network modelling. An example in wildlife damage. *J. Appl. Ecol.* 36, 317-326.
- Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39, 561-577.

IV. ASSESSING HABITAT-SUITABILITY MODELS WITH A VIRTUAL SPECIES

A. Introduction

PREDICTION of species distribution is an important element of conservation biology. Management for endangered species (Palma et al., 1999; Sanchez-Zapata and Calvo, 1999), ecosystem restoration (Mladenoff et al., 1997), species re-introductions (Breitenmoser et al., 1999), population viability analyses (Akçakaya et al., 1995; Akçakaya and Atwood, 1997) and human-wildlife conflicts (Le Lay et al., 2001) often rely on habitat suitability modelling. Multivariate models are commonly used to define habitat suitability and, combined to Geographical Information Systems (GIS), allow to create potential distribution maps (Guisan and Zimmermann, 2000).

Numerous multivariate analyses were developed for building habitat suitability or abundance models, but very few studies compare their predictive power (e.g. Lek et al., 1996; Paruelo and Tomasel, 1997; Guisan et al., 1999; Manel et al., 1999; Öziesmi and Öziesmi, 1999). In this paper, we compare a common method, the Generalised

Linear Model (GLM) (e.g. Austin et al., 1984; Augustin et al., 1996; Guisan et al., 1998), to the Ecological Niche Factor Analysis (ENFA, cf. chapter III).

GLM is a generalisation of multiple regression analysis with binomial distribution and logistic link that may fit polynomials of higher degree than linear. The dependent variable (presence/absence of the species) is explained by a sum of weighted ecogeographical predictors. The weights are tuned in order to generate the best fit between the model and the calibration data set (Jongman et al., 1987; Nicholls, 1989).

ENFA compares the ecogeographical predictor distribution for a presence data set consisting in locations where the species has been detected, to predictor distribution of the whole area. Like the Principal Component Analysis, ENFA summarises all predictors into a few uncorrelated factors retaining most of the information. But in this case, the factors have an ecological meaning: the first factor is the *marginality*, and reflects the direction on which the species niche mostly differs from the available conditions in the global area. Subsequent factors represent the *specialisation*. They are extracted successively by computing the direction that maximises the ratio of the

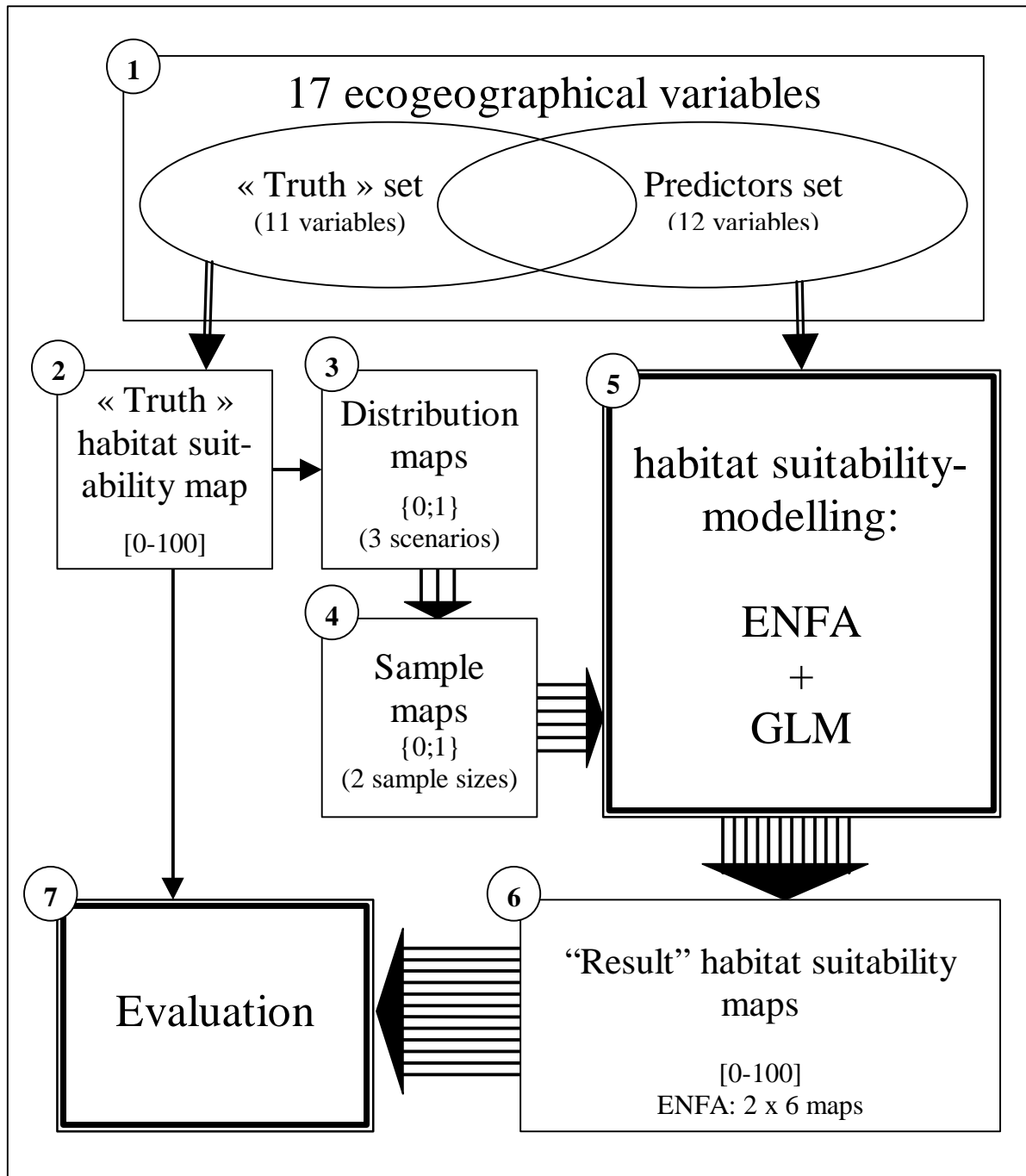


Figure 4.1: This flow chart summarises the steps involved in this study. 1. Seventeen ecogeographical variable (predictor) maps are prepared. 2. A “truth” set of predictors is used to generate a “truth” habitat suitability map. 3. On this basis, three **distribution maps** are generated corresponding to three distribution scenarios (see text). 4. The distribution maps are sampled with two sample sizes (300 and 1200 points), generating **sample maps**. 5. A “predictors” set (partially overlapping the “truth set”, see text) of the predictors is used in conjunction with the sample maps to compute predicting models with both GLM and ENFA methods. 5. These models are used to produce “**result**” **habitat suitability maps** (6). 7. The “result” maps are statistically compared to the “truth” map to assess the predicting power of each analysis in each scenario. Single-framed boxes symbolise set of maps; the type of data is indicated. Double-framed boxes symbolise statistical processes. Arrows symbolise the data flow, each shaft accounting for one map. See text for further explanations.

the species distribution. A large part of the information is accounted for by a few of the first factors. The species distribution on these factors is used to compute a habitat suitability index for any set of descriptor values (cf. chapter III).

Practically, the main difference between these analyses is the quality of input data: GLM needs presence/absence data whereas ENFA only needs presence data. The latter is thus much less demanding than the former and it is interesting to compare their predictive power. Obviously, this power depends on the situation: for example, when absence data are reliable, GLM could get extra power by using this information, but in other situations, it could be misled by false absences ; see also “stochastic zeros” in Welsh et al., 1996).

The goal of this paper is thus to circumscribe the domain of application of both methods from the point of view of absence data quality. It is more complex a task than simply comparing analyses on the same data set. Indeed, measuring their sensitivity to various data qualities entails exploring several distribution patterns of ecologically identical species in a common landscape. But as such species could not live simultaneously in a same place, it is impossible to find such data in the real world; it is therefore necessary to generate simulated species distribution data. Moreover, this method presents the following advantages: 1° the input data set can be fully controlled, qualitatively as well as quantitatively. 2° the “reality” being perfectly known, model accuracy assessment is straightforward and certain. Nevertheless, in order to track reality as closely as possible, the environmental predictors were taken from a real area in the Swiss Alps.

B. Methods

This study implied to build a virtual species completely characterised by its ecological niche, which would be modelled by a “truth” habitat suitability map. Three data sets were then generated, simulating

three different scenarios. These data sets, in conjunction with environmental variables, were fed into the GLM and ENFA analyses, which produced “predicted” habitat suitability maps. Finally, resulting models were evaluated by statistically comparing each “predicted” maps to the “truth” map. These steps (summarised in fig. 4.1) will now be developed in full details.

1. Ecogeographical variables

Although a virtual species is used, environmental data are real and issued from a square region of 25.6 by 25.6 km located in the Swiss Alps (see fig. 4.3), numerically modelled by 17 Geographic Information System (GIS) raster maps of 256 by 256 cells, representing 17 ecogeographical variables. These predictors are topographical, ecological or related to human activities (see table 1). They were derived from land-cover, topography, hydrography and road/rail network GIS databases. Boolean variables were transformed into continuous ones by computing a new map storing either the distance to the nearest cell of this category, or the proportion of those cells within a circular moving-window of 1200-m radius. We used *Idrisi 2.0* (Eastman, 1997) and *Biomapper 1.0* (Hirzel et al., 2001) to achieve these operations and to deal with the predictor maps.

2. Virtual ecological niche: the “truth” habitat suitability map

On this spatial canvas, the virtual species was generated by creating a simulated ecological niche in a n -dimensional space, *sensu* Hutchinson (1957). It was modelled by a niche coefficient H ($H \in [0,1]$), which can be viewed as a probability of each cell to belong to the niche; note that H is *de facto* a habitat suitability index. This value was built as summarised in equation 4.1:

$$H = \frac{1}{\sum w_i} \sum w_i H_i + e \quad (4.1)$$

where H is the habitat suitability of the focal cell, H_i is the value of the i th partial

niche coefficient, w_i is the weight assigned to the i th partial niche coefficient, and e is a random value

Global habitat suitability is composed of a weighted average of *partial niche coefficients* (H_i) and a stochastic coefficient (e). The partial niche coefficients are the habitat suitability engendered by each predictor

value; they were computed from 11 predictors (playing the role of Hutchinsonian environmental-space dimensions), picked out of the 17 available predictors, by 11 *niche-functions* (table 4.1). Three types of functions were used to model three types of environmental optimum: 1° A gaussian function modelled a median optimum, 2° a

Table 4.1: Ecogeographical variables (predictor) used to generate the virtual habitat suitability map. Following equation 1, the global habitat suitability-value is computed by a weighted average of the partial niche coefficients, which are themselves computed from the predictors by niche-functions (See fig.1). This table indicates for each predictor (1st col.) which type of niche function was used to computes its partial niche coefficient (2nd col.) and which was its weight for the global habitat suitability computation (3rd col.).

predictor	Niche-function	Weight (w_i)
Forest frequency	Linear (increasing)	6
Elevation	Gaussian	5
Southern aspect freq.	Linear (increasing)	2
Distance to towns	Truncated linear	2
Distance to forests	Linear (decreasing)	1
Slope>30° frequency	Linear (decreasing)	1
Distance to waters	Linear (decreasing)	1
Distance to villages	Gaussian	1
Dist. to primary roads	Truncated linear	1
Dist. to secondary roads	Truncated linear	1
Distance to railways	Truncated linear	1

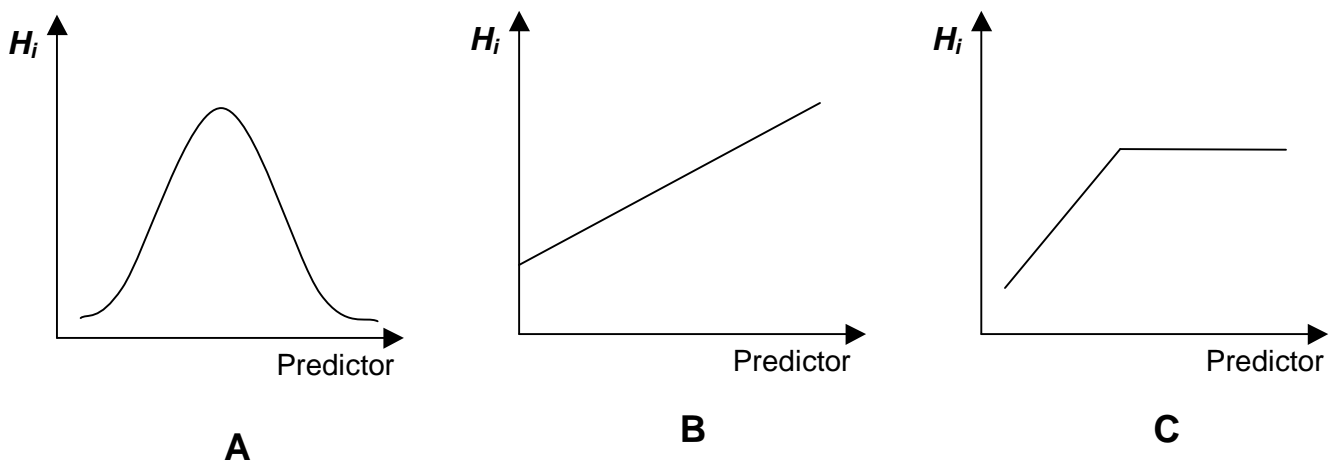


Figure 4.2: The partial niche coefficient H_i ($H_i \hat{I} [0,1]$) is a function of each ecogeographical variable (predictor). Three types of function model three types of niche optimum: in A, the optimal value of the predictor lies somewhere in the middle of the available range and decreases “gaussianly” in either direction. B and C are typically used to model distance related variables, either to disturbance or food sources. In B, the habitat suitability is linearly increasing (or decreasing) as the location goes farther from the source; in C, the truncated linear type shows a buffer zone effect, the influence of the source becoming null above some given range.

linear function modelled an extreme optimum, 3° a truncated linear function modelled a buffer zone effect (see figure 4.2).

Each of this H_i was then weighted by a w_i factor and the global niche coefficient calculated as their weighted average. Finally, a random term e , generated from a uniform distribution in the range $[-0.05, 0.05]$, was added.

The niche-function parameters and the weights were arbitrarily tuned in order to generate about 50% of cells with $H > 0.5$.

This produces the “truth” *habitat suitability map* (fig. 4.3), representing the “real” intrinsic preferences of our virtual species; by “truth” map we are meaning that it represents the kind of information usually unreachable by ecologists, the information they are trying to reveal through field sampling and statistical analysis. The

“truth” map will be constantly used as basis to generate data and as reference to assess the accuracy of habitat suitability analyses. A 3-D view of the landscape is presented in figure 4.4 to provide a better understanding.

3. Distribution maps

Computed on the basis of the “truth” map, the *distribution maps* give the “truth” presence/absence of the virtual species, an information usually unavailable to field ecologists.

Three distribution scenarios were addressed in order to determine the advantages and drawbacks of each habitat suitability analysis. They can be viewed as three historical phases of colonisation; the fundamental niche does not change but the realised one does: 1° A “spreading phase”

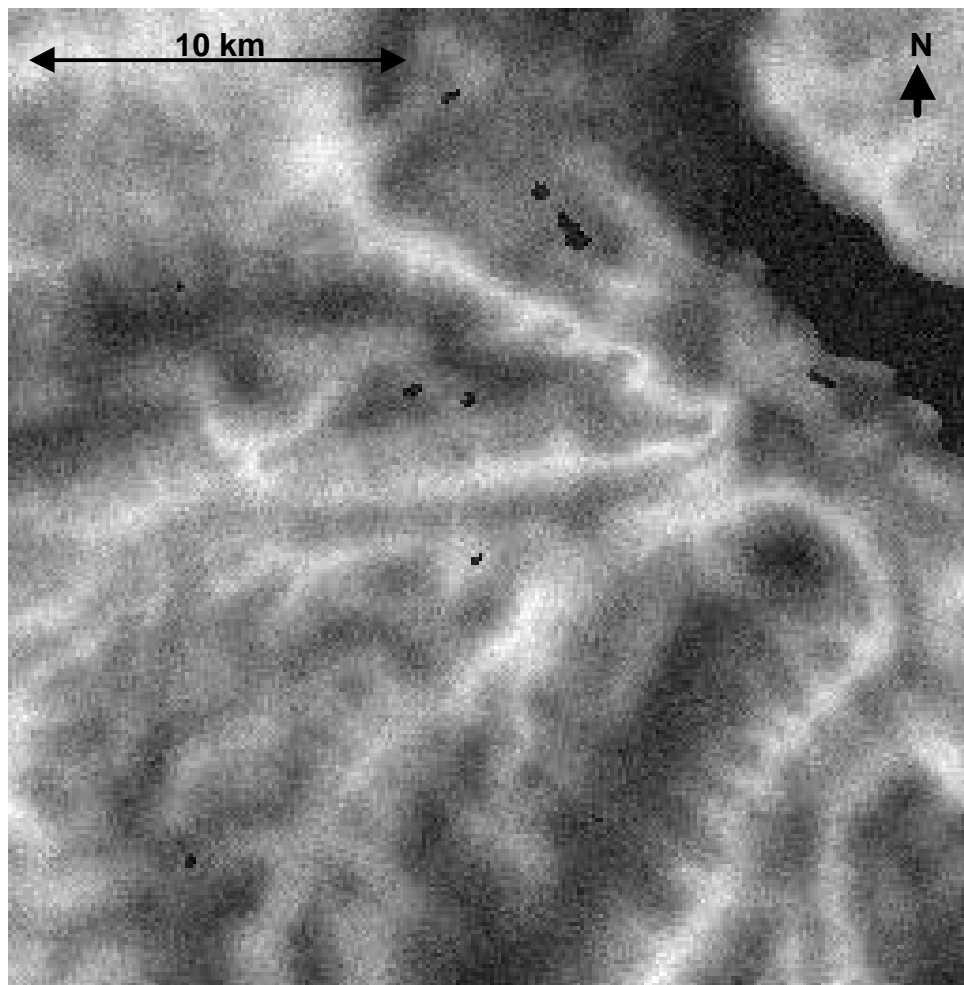


Fig. 4.3: The “truth” *habitat suitability map* generated to model the ecological niche of the virtual species. High suitability areas are indicated by white pixels.

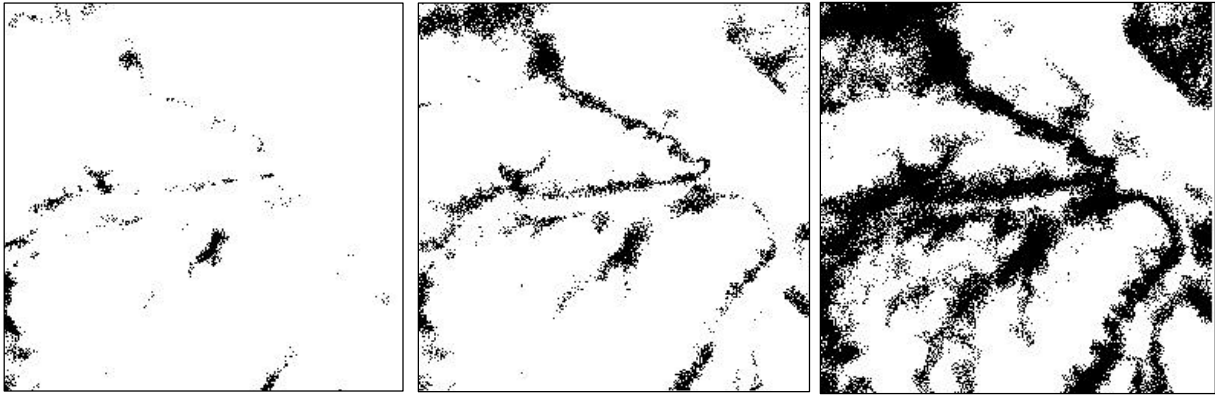


Figure 4.5: Distribution maps of the virtual species for three colonisation scenarios. Black points are the cells where the species is present and the white are those where it is absent. Map A represents the "spreading" scenario: the species entered the area from the south-west and is currently propagating in all directions, settling down in the most suitable areas. Map B shows the "equilibrium" scenario in which the species occupies uniformly all the suitable areas. Map C presents the "overabundance" scenario in which very high densities force the species to occupy less adequate areas.

showing a density gradient from the south-west corner of the map to the north-east corner, 2° an "equilibrium phase" where the species is abundant enough to occupy all the available suitable areas, and 3° an "overabundance phase" where the species is so numerous that it has to spread in less suitable areas. (fig. 4.5).

The "equilibrium" distribution map was computed as follows: To each cell of the "truth" habitat suitability map was added a

random value taken in the range $[-0.2, 0.2]$ (uniform distribution); this was made in order to introduce some stochasticity in the model. If the resulting habitat suitability coefficient was larger than 0.7, the cell was marked as occupied.

The "overabundance" distribution map was computed in a similar way but with a 0.5 habitat suitability threshold to simulate the overflowing density.

The "spreading" distribution needed an additional operation: each cell of the "truth" habitat suitability map was beforehand multiplied by a value decreasing in $1/d^2$, d being the distance to a point arbitrarily placed south-westward of the south-west corner of the map; this gradient-function was tuned to produce values ranging from 0 to 1, 0.5 lying approximately in the middle of the map. This new gradient map was then submitted to the same operations as the "equilibrium" scenario (habitat suitability threshold = 0.7).

This generating method assured to obtain distribution maps with a presence density correlated with area suitability.

4. Sample maps

These distribution maps were then used to simulate "field" sampling data usually resulting from the trapping/detecting/observation activities of field biologists. As the GLM and the ENFA do not need the same kind of data it was necessary to generate two data sets: one presence/absence set for the GLM and

one presence set only for the ENFA. In order to compare results, sampling sizes were identical for all scenarios and analyses. Two sampling sizes were addressed, 300 points and 1200 points.

ENFA data sets were generated by randomly picking points in the distribution maps in order to get the targeted sample size. The probability to pick one cell in a given area was correlated with its density, which was variously correlated with its suitability depending on the scenario. The "spreading" scenario had only 418 occupied cells and it was therefore impossible to get the 1200 points sample size in this case.

The same presence points were used for the GLM but additional absence data were generated as follows: in order to take into account the spatial auto-correlation of the predictors, a 7-cells-radius circular buffer was drawn around each presence point; the absence points were then randomly drawn from the area out of these buffers (A procedure similar to that used by Akçakaya et al., 1997).

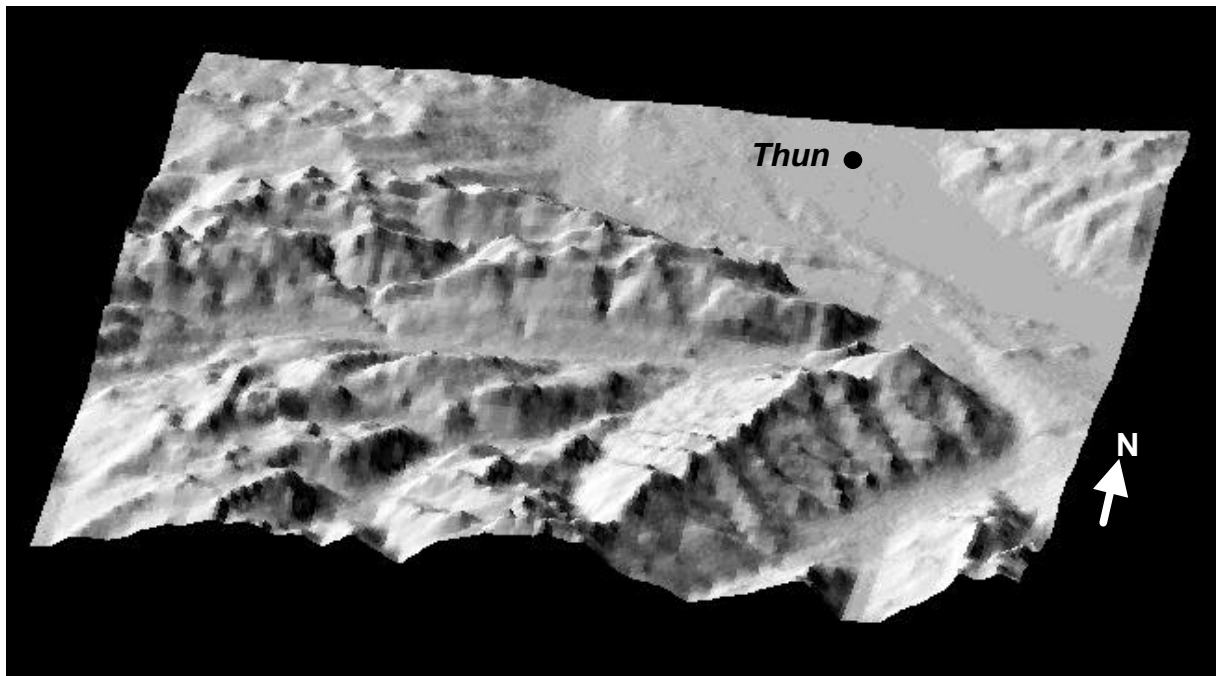


Fig. 4.4: 3D view of the studied region, a 25.6 x 25.6 km area in the Bern Alps (Switzerland). The landscape is viewed from the south. In the north-east corner lie the Aare valley and Thun town; the flat area is the lake of Thun. In the middle of the landscape lies the west-east orientated Simmental valley. Elevations range from 551 to 2637 meter above sea level.

The number of GLM presence/absence points was thus the double of the number of ENFA presence points, but the numbers of "field" data (presence) were identical.

5. Result habitat suitability maps

The simulated data sets were then submitted as dependent variables to the GLM and the ENFA. The independent variables were a set of 12 predictors out of the 17 available; six of them were arbitrarily taken among those used to generate the "truth" habitat suitability map: elevation, southern aspect frequency, distance to towns, distance to forests, slope > 30° frequency and distance to primary roads; six other were new: distance to rivers, distance to lakes, distance to pastures, distance to agricultural meadows, rock frequency and bushes frequency.

ENFA was entirely performed with the *Biomapper* software (Hirzel et al., 2001). The predictors were first normalised by the Box-Cox algorithm (Sokal and Rohlf, 1981). Ecological niche factors were then computed on these normalised predictors and ENFA provided one marginality factor and 11 tolerance factors totally uncorrelated, each factor being a linear combination of the predictors. Among these factors we kept only those explaining a significant amount of total variance by comparison to a broken-stick distribution (always greater than 75%). Factor distributions were computed on six classes and these empirical distributions used to compute the habitat suitability maps (cf. chapter III). A 7x7-gaussian filter was finally applied to them in order to smooth the stepped shape produced by this analysis.

The GLM were calibrated in the *S-PLUS* software (Mathsoft Inc.) by using a binomial distribution and a stepwise variable selection procedure. As the niche coefficient was not a linear function of the predictors we introduced in the input variables not only the 12 predictors already mentioned but also their square power; bell shaped and truncated linear niche-functions could thus be modelled with sat-

isfying accuracy. Because of the high sensitivity of the stepwise process – which eliminates a part of the input predictors to retain only the most relevant – to the input order of the predictors, we tried several orders, retaining the model that explained the highest proportion of the variance. The habitat suitability was then expressed as a linear combination of the predictors and their square terms. The model was then implemented in the GIS and maps were produced, that had finally to be transformed by the inverse logistic function to be scaled between 0 and 1 (For further explanations, see e.g. Guisan et al. (1999)).

6. Evaluation

The accuracy of the predicted habitat suitability maps had finally to be assessed. With a real species we would have used independent evaluation data and calculated various statistics to assess the accuracy of the classification (reviewed in Fielding and Bell, 1997). But here, with a virtual species, the "true" habitat suitability that the models were supposed to reproduce was perfectly known. More adapted statistics based on the Pearson's correlation coefficient between the two maps could thus be used. In order to get round the pseudo-replication engendered by spatial autocorrelation between cells we proceeded as follows: 250 cells were picked randomly and a determination coefficient R^2 (proportion of variance explained by the model) was computed between the values of these cells in the "result" map and the "truth" map (Mesplé et al., 1996). This process was replicated 10 times and the mean and standard deviation of the R^2 were computed. The mean R^2 was used to assess the accuracy of the models. The results obtained by both techniques in each scenario were compared by a bilateral Student t-test. Their sensitivity to distribution scenario and sample size were also assessed with a Student t-test for each method.

Table 4.2: Mean ($\langle R^2 \rangle$) and standard deviation (SD) of the proportion of explained variance obtained by comparing 10 times each “result” maps to the “truth” map. The greater the $\langle R^2 \rangle$, the higher the predictive power of the “result” map. ENFA proved to be better in the “spreading” scenario, whereas GLM was better in the “overabundance” scenario with the small sample size. The probability (P) of the GLM and ENFA to have different predictive power was computed with a bilateral Student *t*-test. When the difference is significant, the best analysis is emphasised. Note that the best R^2 that could be achieved was 0.67. (Ns: non significant; ***: $10^{-4} < P < 10^{-3}$; ****: $P < 10^{-4}$).

Scénario	GLM		ENFA		GLM = ENFA ?	
	$\langle R^2 \rangle$	SD	$\langle R^2 \rangle$	SD	P	
Spreading, 300 pts	0.38	0.03	0.57	0.05	$3.9 \cdot 10^{-9}$	****
Equilibrium, 300 pts	0.53	0.04	0.55	0.03	$1.3 \cdot 10^{-1}$	Ns
Overabundance, 300 pts	0.63	0.02	0.57	0.04	$2.2 \cdot 10^{-4}$	***
Equilibrium, 1200 pts	0.54	0.03	0.56	0.05	$5.0 \cdot 10^{-1}$	Ns
Overabundance, 1200 pts	0.63	0.04	0.60	0.03	$1.2 \cdot 10^{-1}$	Ns

C. Results

Equilibrium and overabundance scenarios were addressed with two sample sizes (300 and 1200 points) for both analyses (ENFA and GLM) and the spreading scenario only with 300 sample points, which makes a total of ten habitat suitability maps.

In order to compare the predictive power of these “result” maps, the proportion of explained variance (R^2) was computed on a sample of 250 pairs of points taken in the “result” map and in the “truth” map. This coefficient was computed ten times for each map and the mean (R^2) and standard deviation (SD) were computed. These average R^2 achieved by each analysis were then compared in pair by mean of a Student *t*-test.

Due to the stochasticity added in the process of building the “truth” habitat suitability map, it was impossible to get $R^2 = 1$. The best model would have been the map computed just before the addition of stochasticity and this one gave $\langle R^2 \rangle = 0.67$ (SD = 0.07). This is to be remembered when looking at the absolute signification of the results presented in table 4.2.

In the “spreading” scenario, the ENFA proved to be significantly more efficient

than the GLM. In the equilibrium scenario, there was no significant difference between the two methods. In the case of “overabundance”, the GLM gave significantly better results with the sample size of 300, but when the sample was larger (1200 points) the difference between the two analyses disappeared.

Another interesting result was the sensitivity analysis of each method to the quality and quantity of the input data. This was achieved by intra-method pair comparisons using the same procedure as above.

Table 4.3 shows that ENFA is generally robust, the predictive power of the maps being significantly different in only 30% of the cases. The “overabundance” scenario was the most sensitive. In contrast, the GLM is quite sensitive to data quality (scenario effect) but not to data quantity. Actually, predictive powers are always highly significantly different, except between maps produced with the same scenario but different sample sizes.

D. Discussion

The three addressed scenarios were modelled with unequal success by the two analyses. The ENFA appeared to be very robust to data quality and quantity, none of the investigated cases presents a signifi-

Table 4.3: Sensitivity analyses of ENFA and GLM. The upper part of the table compares predictive power of maps produced by the ENFA in each scenario. The lower part compares the maps produced by the GLM. ENFA maps have generally a similar predictive power (*t*-test not significant), while the GLM maps are generally different (*t*-test very highly significant). When the predictive powers of two “result” maps do not differ significantly (*Ns*), the *t*-test probability is given between parentheses. Significant results are indicated by * ($0.01 < P < 0.05$), ** ($0.001 < P < 0.01$) and **** ($P < 0.0001$).

		Spreading 300 pts	Equilibrium 300 pts	Overabundance 300 pts	Equilibrium 1200 pts
ENFA	Equilibrium, 300 pts	Ns (0.24)			
	Overabundance, 300 pts	Ns (0.67)	Ns (0.38)		
	Equilibrium, 1200 pts	Ns (0.43)	Ns (0.84)	Ns (0.63)	
	Overabundance, 1200 pts	Ns (0.12)	**	*	*
GLM	Equilibrium, 300 pts	****			
	Overabundance, 300 pts	****	****		
	Equilibrium, 1200 pts	****	Ns (0.30)	****	
	Overabundance, 1200 pts	****	****	Ns (0.95)	****

cantly better or poorer fit; the overall goodness of fit was good with an average explained variance proportion of 0.58 ($SD=0.02$). On the other hand, GLM was moderately sensitive to data quality but not to data quantity (average explained variance: 0.52, $SD=0.11$).

Relying on absence data is both the strength and the weakness of this analysis: when they really reflect low habitat suitability – like in the “overabundance” scenario – the additional information improves the model; but when the absences are due to historical causes – like in the “spreading” scenario – this information is fallacious and decreases the overall predictive power.

ENFA and GLM were not sensitive to sample size, as both analyses produce slightly better results with 1200 points than with 300. An interesting sequel to this study would be to explore more thoroughly the effect of the sample size and particularly the minimal efficient size as it could give useful clues when conceiving a sampling design.

As this experiment was not designed to explore qualitatively the results of these methods, it was not clear which one pro-

duced the best ecological interpretation of the data. GLM stepwise procedure is highly sensitive to predictor input order when these are not fully uncorrelated; adding or removing a predictor often qualitatively modify the resulting model. In contrast, ENFA is not at all sensitive to these “input effects”. Thus, when ecological interpretation is the aim of the study, ENFA could be more useful even for situations in which a GLM should provide a higher correlation to observed data.

Spatial autocorrelation is always problematic with the use of geographical space (Legendre, 1993). In this study, the problem arises when comparing the “result” habitat suitability maps with the “truth” habitat suitability map. Independent data are needed in order to use adequately the *t*-test of significance (Sokal & Rohlf, 1981). To reduce the correlation between sample points, a small sample size (250 points, being separated by an average distance of 913 m) was used to compute the R^2 statistics. Nevertheless, spatial autocorrelation can never be totally removed and it is best to cope with it (Legendre, 1993). Several methods exist either to remove spatial autocorrelation or to take it into account

when computing the significance test (e.g. Clifford et al., 1989; Dutilleul, 1993). Unfortunately none of them were suitable to our case (testing the equality of two R^2) and generalising them to include it was beyond the scope of this paper and would have added little information.

Actually, ENFA-GLM comparison (table 4.2) and GLM sensitivity analysis (lower part of table 4.3) are very clear: the t -tests are either highly significant ($P < 10^{-4}$) or not at all ($P > 0.1$). Taking autocorrelation into account would consist to lower the actual sample size to an effective sample size (Dutilleul, 1993) but it would probably not lower it enough to qualitatively modify these results. Actually, ENFA sensitivity analysis only (upper part of table 4.3) could be qualitatively modified should the degrees of freedom be lower; two hardly significant ($P = 0.03$) pair-comparisons could become non-significant. This would prove ENFA even more robust.

The virtual species approach proved to be most serviceable. When comparing models on a real data basis, it is only possible to make assumptions about what is the true habitat suitability by using various expert and statistical evaluation methods. Many factors are out of reach and may introduce a bias that cannot be accurately assessed (Alldredge and Ratti, 1992; Paruelo et al., 1997; Guisan et al., 1999; Manel et al., 1999): they may be historical, like disturbances, catastrophes, diseases or colonisation events, or spatial, like dispersal barriers or corridors, or they can be ecological, like interspecific competition. Real data are therefore only a snapshot of a dynamical situation and can only give a partial and instantaneous comprehension of the fundamental ecological niche. By generating a virtual species, the "truth" is now completely reachable and resulting models can be accurately compared to it.

Moreover, previous works (Lek et al., 1996; Paruelo et al., 1997; Guisan et al., 1999; Manel et al., 1999; Öziesmi et al., 1999) compare methods on a unique case;

the method application domain may thus hardly be explored. In contrast, our "true" ecological niche approach may be used to generate various kinds of data sets to test different situations. In this paper, we explored the effect of colonisation history on the accuracy of two models, but sensitivity to many other effects could be tested: sampling size, sampling bias, interspecific competition, etc. Field ecologists are proposed a wide panel of statistical analyses but it is often difficult to choose among them. By circumscribing theoretically their application domain, this virtual species approach helps to select the most appropriate model in each situation. The ability to manipulate the virtual species allow to isolate and thus to better understand problems encountered when dealing with real species; this will end up in better-suited analyses.

Although the error-free quality of simulated data has been often used in other domains to qualify method results (Ferré, 1995; Mesplé et al., 1996; Olley and Kochhar, 1996; Delay and Lamotte, 2000) it is less current in ecology (Paruelo et al., 1997; Kendall et al., 1999; Moilanen, 1999) and has never been applied to habitat suitability model assessment.

There is a risk that virtual species do not simulate correctly the reality, introducing errors or biases in the results. In this study, we reduced this risk by several means: 1° We used real ecogeographical data; simulated predictors could be interesting in some cases, for example to explore the model sensitivity to their distribution, but as our study was focused on the quality of presence/absence data, there was no need to do so; the correlation between variables and their spatial auto-correlation, as well as their distribution was therefore representative of what can be found in reality. 2° Niche-function was made of both linear and non-linear components, with some stochasticity; therefore, the resulting niche shape did not favour one particular analysis. 3° Half of the predictors included as predictors in the analyses were not used to generate the virtual species; this simulated

the fact that, in real cases, true predictors are generally unknown or not available and models are basically built on correlated variables

E. Conclusion

1° This paper gives insight on the domains of application of GLM and ENFA. It appears that the robustness of ENFA makes it particularly suitable and efficient when the quality of data is either poor (the absence data are unreliable) or unknown. The GLM offers slightly better results when the available presence/absence data are sufficiently good.

2° Virtual species simulation proved to be useful when assessing analysis predictive power in spatial ecology, allowing to achieve a more accurate evaluation and to better control the experiment parameters.

F. Note

A modified version of this chapter has been published in *Ecological Modelling* (Hirzel et al, in press).

G. References

- Akçakaya, H.R., Atwood, J.L., 1997. A habitat-based metapopulation model of the California Gnatcatcher. *Conserv. Biol.* 11, 422-434.
- Akçakaya, H.R., McCarthy, M.A., Pearce, J.L., 1995. Linking landscape data with population viability analysis : management options for the helmeted honeyeater *Lichenostomus melanops cassidix*. *Biological Conservation* 73, 169-176.
- Allredge, J.R., Ratti, J.T., 1992. Further comparison of some statistical techniques for analysis of resource selection. *J. Wildl. Manag.* 56, 1-9.
- Augustin, N.H., Mugglestone, M.A., Buckland, S.T., 1996. An autologistic model for the spatial distribution of wildlife. *J. Appl. Ecol.* 33, 339-347.
- Austin, M.P., Cunningham, R.B., Fleming, P.M., 1984. New approaches to direct gradient analysing using environmental scalars and statistical curve-fitting procedures. *Vegetatio* 55, 11-27.
- Breitenmoser, U., Zimmermann, F., Olsson, P., Ryser, A., Angst, C., Jobin, A., Breitenmoser-Würsten, C., 1999. Beurteilung des Kantons St.Gallen als Habitat für den Luchs. Bern, KORA.
- Clifford, P., Richardson, S., Hémon, D., 1989. Assessing the significance of the correlation between two spatial processes. *Biometrics* 45, 123-134.
- Delay, F., Lamotte, J.-L., 2000. Numerical simulations of geological reservoirs: improving their conditioning through the use of entropy. *Mathematics and Computers in Simulation* 52, 311-331.
- Dutilleul, P., 1993. Modifying the t-test for assessing the correlation between two spatial processes. *Biometrics* 49, 305-314.
- Eastman, J.R., 1997. Idrisi for Windows 2.0 user's guide. Worcester, Clark University.
- Ferré, L., 1995. Selection of components in principal component analysis: A comparison of methods. *Computational Statistics & Data Analysis* 19, 669-682.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38-49.
- Guisan, A., Theurillat, J.-P., Kienast, F., 1998. Predicting the potential distribution of plant species in an alpine environment. *J. Veget. Sci.* 9, 65-74.
- Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143, 107-122.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147-186.
- Hirzel, A.H., Hausser, J., Perrin, N., 2001. Biomapper 1.0. Lausanne, Lab. for Conservation Biology. URL: <http://www.unil.ch/biomapper>

- Hirzel, A., Helfer, V., Métral, F., in press. Assessing habitat-suitability models with a virtual species. *Ecol. Model.*
- Jongman, R.H.G., ter Braak, C.J.F., Van Tongeren, O.F.R., 1987. *Data analysis in Community and Landscape ecology*. Cambridge, Cambridge university press.
- Kendall, B.E., Briggs, C.J., Murdoch, W.W., Turchin, P., Ellner, S.P., McCauley, E., Nisbet, R.M., Wood, S.N., 1999. Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. *Ecology* 80, 1789-1805.
- Le Lay, G., Clergeau, P., Hubert-Moy, L., 2001. Computerized Map of Risk to Manage Wildlife Species in Urban Areas. *Environmental Management* 27, 451-461.
- Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659-1673.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Model.* 90, 39-52.
- Manel, S., Dias, J.M., Buckton, S.T., Ormerod, S.J., 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *J. Appl. Ecol.* 36, 734-747.
- McArdle, B.H., 1990. When are rare species not there? *OIKOS* 57, 276-277.
- Mesplé, F., Troussellier, M., Casellas, C., Legendre, P., 1996. Evaluation of simple statistical criteria to qualify a simulation. *Ecol. Model.* 88, 9-18.
- Mladenoff, D.J., Haight, R.C., Sickley, T.A., Wydeven, A.P., 1997. Causes and implications of species restoration in altered ecosystems. A spatial landscape projection of wolf population recovery. *Bioscience* 47, 21-31.
- Moilanen, A., 1999. Patch occupancy models of metapopulation dynamics: Efficient parameter estimation using implicit statistical inference. *Ecology* 80, 1031-1043.
- Nicholls, A.O., 1989. How to make biological surveys go further with generalized linear model. *Biol. Conserv.* 50, 51-75.
- Olley, P., Kochhar, A.K., 1996. Case Simulation to Assess Learning Systems. *Engineering Applications of Artificial Intelligence* 9, 285-300.
- Özesmi, S.L., Özesmi, U., 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecol. Model.* 116, 15-31.
- Palma, L., Beja, P., Rodgrigues, M., 1999. The use of sighting data to analyse Iberian lynx habitat and distribution. *J. Appl. Ecol.* 36, 812-824.
- Paruelo, J.M., Tomasel, F., 1997. Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. *Ecol. Model.* 98, 173-186.
- Sanchez-Zapata, J.A., Calvo, J.F., 1999. Raptor distribution in relation to landscape composition in semi-arid Mediterranean habitats. *J. Appl. Ecol.* 36, 254-262.
- Sokal, R.R., Rohlf, F.J., 1981. *Biometry: The principles and practice of statistics in biological research*. New York, W.H. Freeman & co.
- Solow, A.R., 1993. Inferring extinction from sighting data. *Ecology* 74, 962-964.
- Welsh, A.H., Cunningham, R.B., Donnelly, C.F., Lindenmayer, D.B., 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecol. Model.* 88, 297-308.

V. LANDSCAPE EXPLICIT POPULATION DYNAMICS: WHEN GIS COME TO LIFE!

A. Introduction

Starting with very theoretical models (Levins, 1969; Murdoch, 1994), meta-population dynamics has progressively incorporated more and more realism (Hanski and Gilpin, 1991; Hanski and Gilpin, 1997; Akçakaya and Raphael, 1998; Hanski, 1998b) in order to address questions from conservation or fundamental biology. With the emergence of spatially explicit models, the parallel with landscape ecology became easier to draw and many attempts were made either to explore the impact of landscape structure on classical metapopulation parameters (e.g. With and Crist, 1995; Gustafson and Gardner, 1996; Clarke et al., 1997; Hill and Caswell, 1999; Bunn et al., 2000) or to directly include landscape information into the model (reviewed in Wiens (1997).

Analytical models tend to neglect effects of the inter-patches landscape; these effects are at best integrated into symmetric migration rates and dispersal mortality, both in occupancy models (Hanski, 1994; Day and Possingham, 1995; Hanski, 1998a; Moilanen and Hanski, 1998; Han-

ski and Ovaskainen, 2000) and interconnected local population models (Akçakaya et al., 1995; Akçakaya and Atwood, 1997; Akçakaya et al., 1998). Inter-population landscape structure is actually better modelled by individual-based models (see for instance, Judson, 1994; Tischendorf, 1997; Beecham and Farnsworth, 1998; Letcher et al., 1998; Wiegand et al., 1999) but with an inherent loss of population-wide view and a difficulty to address large systems.

A third way to link population and landscape ecology is the cellular automaton model (CA). These are discrete dynamical systems consisting of an array of cells, each of which can be in one of a finite number of possible states, updated according to a local, identical neighbourhood-controlled transition rule (Hogeweg, 1988; Sipper and Tomassini, 1998). Conceived as mimics of biological cellular tissues (Von Neumann, 1966), CA have been applied to a wide range of domains where highly complex non-linear dynamics arise from local interactions of many simple components. Natural phenomena correspond commonly to this kind of processes (Wolfram, 1984). Since Hogeweg (1988) discussed how CA could weld together the ecological theories of processes and spatial

patterns, they have been more widely used. Whilst theoretical studies addressed some special issues of biological models (e.g. Ruxton, 1996; Ruxton and Saravia, 1998; Schönfisch and de Roos, 1999) CA were mainly applied to vegetation dynamics (e.g. Hogeweg, 1988; Balzter et al., 1998; Cannas et al., 1999; Alonso and Solé, 2000), epidemiology (e.g. Grimm et al., 1996; Keeling and Gilligan, 2000; Sirakoulis et al., 2000) and animal population dynamics (e.g. Molofsky, 1994; Kizaki and Katori, 1999; Hiebeler, 2000). Conformably with earlier work, these CA modelled qualitative phenomena: cell state represented for instance presence/absence, vegetation stage or behavioural phase. Recently, quantitative issues have been addressed by assigning to cell state a numerical value representing abundances (Darwen and Green, 1996; Rushton et al., 1997; South, 1999) or cell saturation (Karafyllidis, 1998). These were however theoretical studies deterministically exploring spatially- but not landscape-explicit dynamical questions; indeed, whilst Darwen and Green (1996) as well as South (1999) modelled blank landscapes, Karafyllidis's (1998) spatial pattern was the latitudinal temperature gradient of a rectangular zone in the northern hemisphere.

In this paper, we present a quantitative stochastic landscape-explicit CA. Cell state represents population density and transition rules reflect local dynamics as well as dispersal. Both habitat suitability and dispersal barriers are included into the numerical landscape. Moreover we demonstrate how various Geographic Information Systems (GIS) tools may be used to calibrate the CA to model real populations with an application to Ibex (*Capra ibex ibex*, ungulates) in the Swiss Alps. The program *HexaSpace* that implements model calibration and simulation may be applied to any species.

B. Model description

Model structure and functioning will first be exposed, followed by parameter calibration.

1. Cell lattice

In our CA, the landscape is modelled by a lattice of C identical hexagonal cells (hexs). The dynamics of the whole metapopulation follow from hex local behaviour and interactions. Each hex corresponds therefore to a small part of the landscape and is individually characterised by a set of fixed structural parameters and a population density, which varies following both reproduction and dispersal events.

The parameters represent environmental features of the landscape, immutable at the considered temporal scale, and are therefore initialised once at the beginning of the simulations. The carrying capacity parameter K_i (The subscript ($i=1, 2, \dots, C$) refers to the i th hex.) represents the equilibrium population density of the hex; it is a function of the habitat suitability and availability for the focal species. Six im-

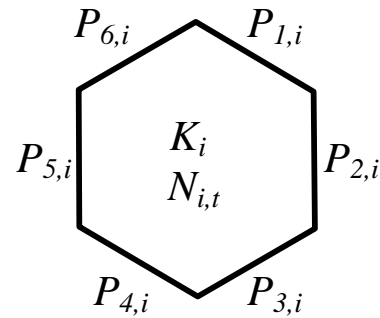


Fig. 5.1: Hexagonal cells (or hexs) are the basic component of the cellular automaton model. Representing a small portion of the landscape, they are organised adjacently into a hexagonal tessellation. Each hex is characterised by a carrying capacity (K_i) and wall impermeabilities ($P_{j,i}$), which represent the energy cost linked to interconnection with adjacent hexs. Population density ($N_{i,t}$) is the only value that varies with time.

permeability parameters $P_{j,i}$ describe how easily individuals can disperse from the focal hex i to each adjacent hex in the j th direction (cf. fig.5.1). Impermeability quantifies the energy cost needed to go from the centre of one hex to the other by the easiest way; it does not depend on the sense of travel.

2. Transition rules

The hex state is its population density $N_{i,t}$, i.e. the numbers of individuals living in the hex; it is the sole hex value varying with time. Density changes (or hex transition) happen synchronously within all cells, following a precise schedule. A simulation is composed of a sequence of T time steps representing seasons, years, generations, etc.. Each time step is itself composed of two kinds of phases, one reproduction phase and T_d dispersal phases. The whole procedure is better explained in pseudo-code:

1. Set hex densities to their initial values.
2. For each of the T time-steps do:
 3. Compute hex density-dependent rates of increase.
 4. Update hex densities.
 5. For each of the T_d dispersal phases do:
 6. Compute the amount of successful dispersers of each hex.
 7. Update hex densities.
 8. End of the current dispersal phase.
9. End of the current time-step.

3. Reproduction phase

During the reproduction phase, the density varies only as a response to local hex saturation at time t $S_{i,t} = \frac{N_{i,t}}{K_i}$. The relation between saturation and rate of increase may be one of the following (fig.5.2):

Truncated exponential (or ceiling):

$$r_{i,t} = \begin{cases} r_0 & \text{if } S_{i,t} < 1 \\ 0 & \text{if } S_{i,t} \geq 1 \end{cases} \quad (5.1a)$$

Logistic: $r_{i,t} = r_0(1 - S_{i,t})$ (5.1b)

Logistic with Allee effect:

$$r_{i,t} = r_{\max} (1 - S_{i,t}) - e^{-\frac{S_{i,t}}{f}} (r_{\max} - r_0) \quad (5.1c)$$

where $r_{i,t}$ is the per capita rate of increase in the i th hex at time t , r_0 is the intrinsic rate of increase (limit of r when $N_{i,t}$ tends to 0). To add Allee effect – i.e. lower-than-expected rate of increase at low density, e.g. due to consanguinity or lack of mates – two more parameters are needed, namely f , which controls the “flatness” of the curve, and r_{\max} , the maximal rate of increase that would have

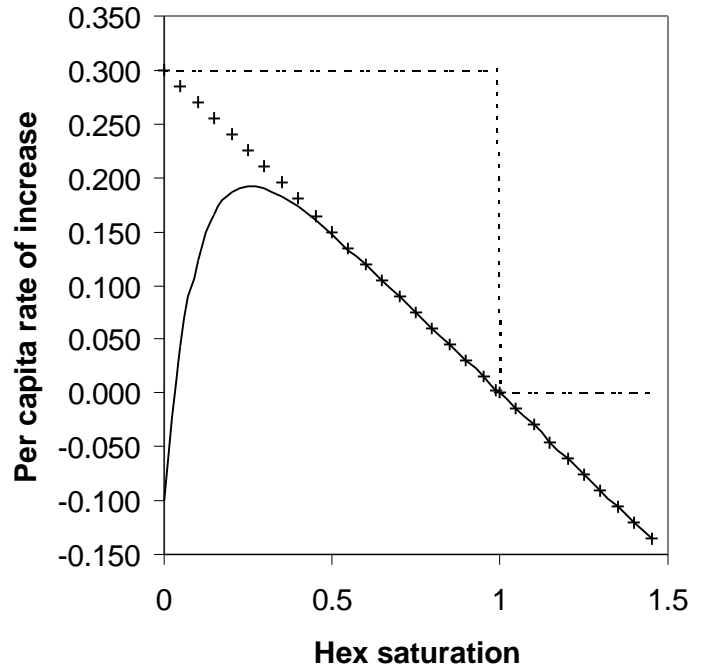


Fig. 5.2: Reproduction transition functions: Hex per capita rate of increase (r) varies as a function of its saturation (proportion of carrying capacity, N/K) following a truncated exponential (dashed line), a logistic (crossed line) or a logistic with Allee effect (plain line).

been achieved without Allee effect (pure logistic function).

The new density is then computed by:

$$N_{i,t+1} = r_{i,t} N_{i,t} \quad (5.2)$$

As hex density is discrete and transition functions are continuous, the results are randomly rounded up or down by mean of the following function:

$$\text{RandomRound}(x) = \text{Trunc}(x + \mathbf{e}) \quad (5.3)$$

where x is the value to be rounded, \mathbf{e} is a random number drawn from a uniform distribution in the interval $[0,1[$, and Trunc is a function that remove all decimal digits. For example, the number 2.8 will be rounded up to 3 with a 80% probability and rounded down to 2 with 20% probability.

4. Dispersal phase

During a dispersal phase, individuals can migrate from their hex to an adjacent one. The dispersal radius is therefore only one hex wide, but by repeating sequentially several (T_d) dispersal phases, further hexs may be colonised during a single time step. Three assumptions are made: 1° Dispersal is function of the focal hex saturation, 2° Animals cannot see neighbouring hex saturation rates and choose therefore their dispersal direction according only to obstacles surrounding their current hex, 3° some dispersers will die in the process as a result of obstacles they have to cross. This results in the following dispersal transition procedure:

The amount of disperser D is determined by

$$D = m_{i,t} N_{i,t} \quad (5.4)$$

where $N_{i,t}$ is the density of the hex and $m_{i,t}$ the dispersal rate, which is related to hex saturation $S_{i,t}$ by a sigmoid function; the more crowded an hex, the more individuals will tend to emigrate. The sigmoid is given by the following equation:

$$m_{i,t} = \frac{\mathbf{a}}{1 + e^{-b(S_{i,t} - \mathbf{g})}} + \mathbf{d} \quad (5.5)$$

where \mathbf{a} is the amplitude of the sigmoid (range between the lower and upper asymptotes), \mathbf{b} is the slope of the curve at the inflexion point, \mathbf{g} is the position of the inflexion point and \mathbf{d} is the lower asymptote. This function was chosen for it offered a high versatility and was therefore likely to fit to any situation.

The result of the dispersal function is randomly rounded by the Randomround function (5.3). This amount is removed from the hex density.

The D dispersers are dispatched among the six directions according to inverse proportion of their impermeabilities. An impermeability threshold may be set, above which no individual can cross the frontier. Thus, $D_{j,i}$ individuals leave the hex i and take the direction j . “Fractional individuals” are randomly assigned to a direction by a process analogue to Randomround function in a way guaranteeing that

$$D = \sum_{j=1}^6 D_{j,i} .$$

A part of the dispersers die on the way, affected by dispersal mortality $d_{j,i}$, a linear function of the impermeability.

$$D_{j,i}^* = (1 - d_{j,i}) D_{j,i} \quad (5.6)$$

Finally, these $D_{j,i}^*$ successful dispersers are added to their target hex density.

5. Stochasticity

Dispersal and reproduction may both be affected by stochasticity. Dispersal stochasticity affects the number of dispersers leaving a hex.. It is achieved by adding a gauss-distributed coefficient with mean zero and standard deviation (SD) \mathbf{e}_m to the $m_{i,t}$ parameter in equation (5.4). The resulting value is truncated to stay in the range from zero and one.

For reproduction, three levels of stochasticity are considered: 1° Local stochasticity is the part of environmental stochasticity acting at the hex level; it is com-

pletely independent of the adjacent hexs.
 2° Regional stochasticity (spatially correlated environmental stochasticity) affects reproduction of individuals at the landscape scale; it is mainly due to climatic phenomena affecting widespread area and is therefore spatially autocorrelated.
 3° Demographic stochasticity acts at the individual level and its importance decreases as the inverse of the density.

Stochasticity is therefore driven by three user-defined coefficients, namely \mathbf{e}_e for environmental stochasticity (local and regional), \mathbf{e}_d for demographic stochasticity and \mathbf{e}_m for dispersal rate stochasticity. The gaussian-distributed pseudo-random numbers are generated by mean of the procedure *gasdev* (Press et al., 1989).

Stochasticity in the rate of increase without spatial autocorrelation is achieved by adding to the $r_{i,t}$ parameter of equation (5.1) a gauss-distributed coefficient with mean zero and $SD = \mathbf{e}_e + \mathbf{e}_d / N_{i,t}$.

When spatial autocorrelation is involved, the procedure is more complex. Hexs cannot be processed individually anymore; an autocorrelated random-map must therefore be built at the beginning of each reproduction phase. For each hex i , a random number is drawn from a gaussian distribution with mean $r_{i,t}$ and $SD = \mathbf{e}_e$. Spatial autocorrelation is achieved by a sequential gaussian simulation algorithm (Gomez-Hernandez and Srivastava, 1990; Goovaerts, 1998); this algorithm is based on a gaussian correlogram with two parameters: maximal spatial correlation (geostatisticians “nugget effect”) and range, i.e. distance at which spatial correlation becomes lower than 0.05. The sequential gaussian simulation algorithm is detailed at the end of this chapter. To this regional stochasticity is further added a gauss-distributed demographic stochasticity with mean zero and $SD = \mathbf{e}_d / N_{i,t}$.

C. Model parameterisation

Model parameters are mainly calibrated on the basis of Geographical Information System (GIS) analyses. All input maps must be stored as raster data structures, at an identical resolution. Four kinds of input data are needed: 1° A *habitat suitability (HS) map*, quantifying the habitat quality of each pixel of the map, 2° A *friction map*, representing the energy cost expended to cross each pixel. Friction is a land-cover-based quantitative value representing the resistance opposed to movement of an individual or energy cost needed to cross a pixel (Eastman, 1999a); it is correlated with the amount and the nature of obstacles lying in a pixel. 3° A *region map* circumscribing areas of interest (e.g. existing populations or potential colonisation areas). 4°. Demographic informations of a sample of the populations represented in the region map, including at least *carrying capacity*, *intrinsic rate of increase* and *dispersal rate*.

As hex size will affect all further operations, it must be carefully chosen as a trade-off between landscape patchiness, species average home range and dispersal behaviour, desired spatial and temporal scale as well as resolution of available data, and computation time. Smaller hex sizes induce finer spatial resolution, more dispersal phases and longer computations. Hex size is geometrically defined by the radius of the circumscribed circle and numerically as the half length of the bounding-box side. Each pixel of the raster map can thus be associated to one and only one hex, and each hex is composed of the same number of pixels (fig.5.3). The hexs are contiguous and only a few pixels lying close to the borders of the map are left aside. Once the CA has been calibrated, GIS maps are no longer used.

1. Regions

The hexs are a rather artificial structure. Field biologists and wildlife managers use to work with individuals, populations or

administrative regions; an interface is therefore needed between the CA model and the “real world”. This interface is provided by the concept of *region*, which is simply a set of hexs, adjacent or not, associated for statistical and parameterisation purposes. A region might correspond to a biological population home range, a national park area, a game reserve or any other relevant landscape partitioning sub-

region carrying capacity K_R as a function of global quality Q_R is then fitted; this allows to derive for each hex a carrying capacity K_i from its quality Q_i .

Moreover, knowledge provided by population dynamics models is used to choose the reproduction function (among equations 5.1a, b and c), set the intrinsic rate of increase and possibly Allee parameters, as well as stochasticity coeffi-

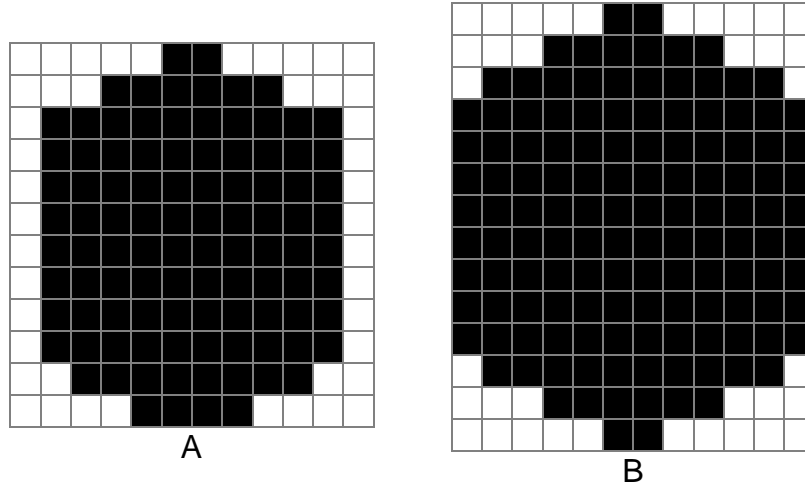


Fig. 5.3: Examples of a 6-pixel-radius hex (A, 100-pixel area) (A) and a 7-pixel-radius hex (B, 132-pixel area) pixel layout. The radius refers to the hexagon’s circumscribed circle. In order to assure that all pixels are attributed to one and only one hex, the lower part of the hex must have a shape complementary to that of the upper part; this may generate a slightly asymmetrical layout, as in A.

jected to management. Hexs are attributed to each region on the basis of the region raster map. Each hex is associated to the region that covers most of its surface.

2. Local population dynamics

Hex carrying capacities K_i are inferred from the HS map and the patch carrying capacities. The HS map yields a habitat quality information for each cell of the map; it may be computed by various statistical methods (extensively reviewed by Guisan and Zimmermann, 2000). The population dynamics models give an estimation of each population carrying capacity K_R .

For each hex, a quality Q_i is computed as the average HS of its pixels. Then, each region global quality Q_R is computed as the sum of the qualities of its hexs. A model of

cients.

3. Impermeabilities

Hex impermeabilities are computed from the friction map. There is no easy way to compute such a friction map. We are currently developing a statistical method allowing building such a map from either genetic or field observations data (cf, Appendix B; Patthey & Hirzel, in prep.).

Impermeability covers the same concept linked to hex interconnections. The impermeability between two hexs is computed as follows: on the friction map, 500 random paths are generated linking their centres; path costs are computed by summing the frictions of the pixels they cross; the lowest of these costs is the looked-for impermeability. This process allows the species to walk around any local obstacle

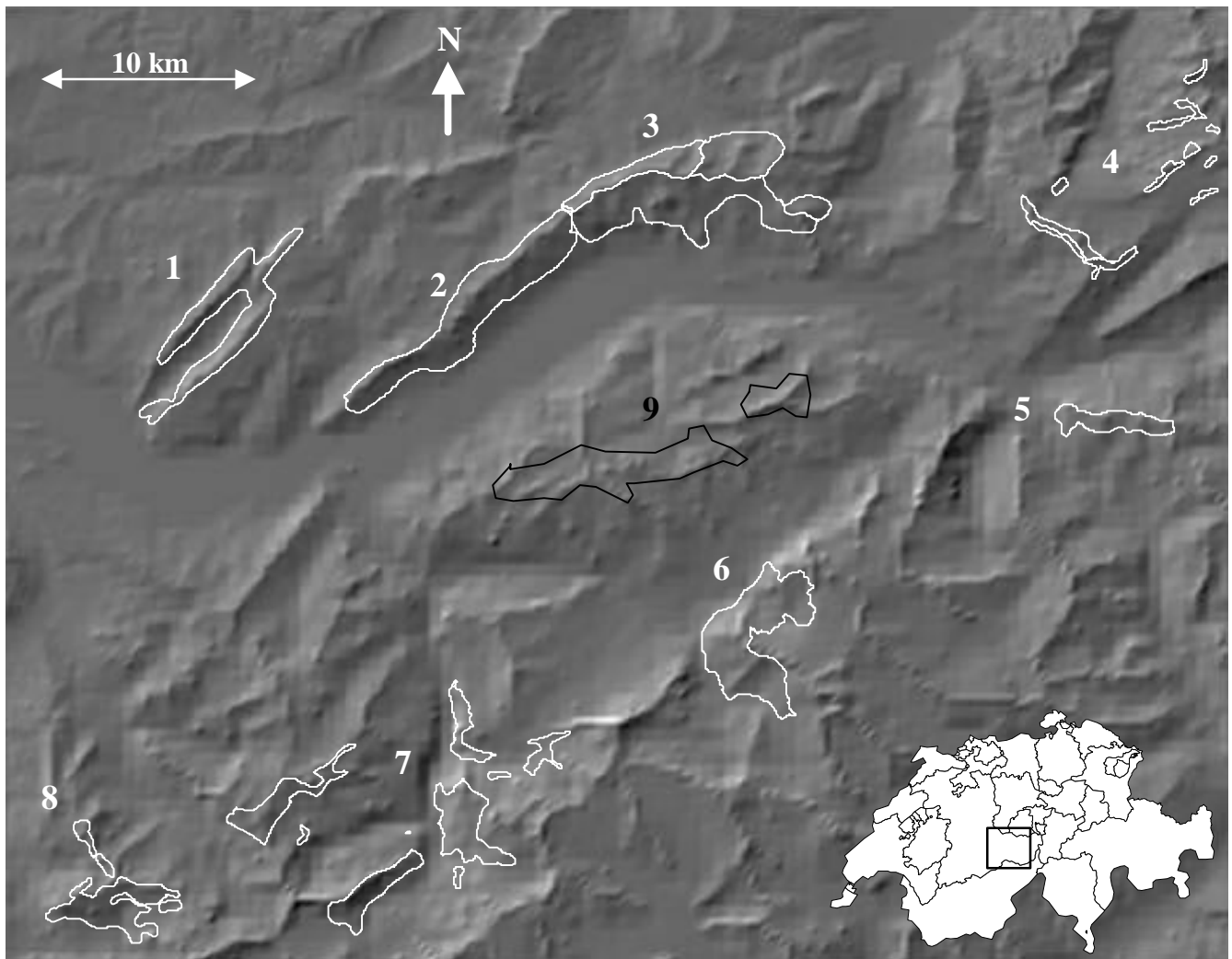


Fig. 5.4: Relief-shaded map of the studied area. It is located in the Bern Alps (framed in the Switzerland map in the lower-right part of the figure). The northern massif is separated from the southern one by the Aare valley, where are located most human structures as well as the Aare river and the Thoune and Brienz lakes. The area lying between populations 5, 6, 7 and 9 is occupied by several glaciers. In white are delineated the current Ibex populations: (1)Justistal, (2)Augstmatthorn, (3)Rothorn, (4)Hutstock, (5)Blattenstock, (6)Wetterhorn, (7)Schwarzmönch and (8)Bire-Öschinen. A smaller massif is yet unoccupied by Ibex although it contains highly suitable habitat; a potential population – Schwarzhorn (9) – has been circumscribed in black.

(e.g. a lake or a village) lying in between two hexs, but not to bypass a long linear barrier (e.g. a river or a highway).

4. Dispersal mortality

Dispersal mortality is modelled as a linear function of impermeability. The slope and intercept of this function must be chosen to fit to species behaviour. For example, amphibians are not very choosy about their dispersal path and may suffer from high dispersal costs; by contrast, ungulates will try to walk around obstacles and thus

lower the travel risks. It is difficult to calibrate these parameters analytically, but expert knowledge allows to fix lower and upper bounds to dispersal mortality. This information is sufficient to tune the mortality-impermeability simple linear function.

5. Initial values

Finally, the density of each hex must be initialised. For theoretical studies this can be done by setting a fixed or random value, or a fixed hex saturation. When real populations are addressed, hex densities must

reproduce monitored data. Once more, the region map is useful to parameterise the CA model: region monitored numbers are dispatched among their hexs in proportion with their relative carrying capacity.

6. Calibration and execution of the model

The CA model is implemented into a Delphi-written (Borland, 1998) software named *HexaSpace*. This program allows 1° to calibrate model parameters – it is linked to the GIS softwares *Idrisi* (Eastman, 1999b) and *Biomapper* (Hirzel et al., 2001) – and 2° to run simulations. It may be freely downloaded at <http://www.unil.ch/biomapper/hexaspace>.

D. Application to Ibex

1. Input data

The study area was a 40.6 x 50.8 km region in the Bern Alps centred on Brienz lake (by 8°00' lg.E and 46°40' lat. N). This region was chosen because several of its Ibex populations were the result of spontaneous colonisation events (Nievergelt, 1966) (in contrast to artificial reintroduction, the general case in other parts of Switzerland). Previous works have produced local dynamics models for all Swiss Ibex populations (Chapter II; see also Hirzel and Perrin, 1998), a HS map (cf. chapter III; Hirzel et al., in press) and a friction map (cf. appendix B). These works will now be briefly summarised. Population dynamics models were age- and sex-structured logistic models; such a detailed information was nevertheless irrelevant here and we kept only the intrinsic rate of increase, the carrying capacity of each population and the stochasticity coefficients. The eight populations involved into this study were the following: Justistal, Augstmatthorn, Brienzer Rothorn, Hutstock, Blattenstock, Wetterhorn, Schwarzmönch and Bire-Öschinen (fig.5.4). Moreover, a “potential population” was circumscribed and added to the model on the massif of the Schwarzhorn;

indeed, although this region contains highly suitable areas, it has not yet been colonised by Ibex (but one individual issued from Wetterhorn was observed wandering in this region from 1989 to 1992 (C. Siegenthaler, pers. comm.)). It was then interesting to see if this would be reproduced by the model.

The region map was rasterised from the OFEFP “summer Ibex colonies” polygon vector map published by Federal Office of Statistics (Geostat) and Schwarzhorn “potential population” was manually added. It was linked with the modelled populations.

The Ibex HS map was the result of an Ecological Niche Factor Analysis (ENFA, chapter III; Hirzel et al., in press) performed over 34 ecogeographical variables at a 100 x 100 m resolution, on the whole Switzerland; the ENFA was chosen because it can build a HS model without species-absence data, which are unreliable in the case of Ibex (chapter IV). Habitat quality was quantified by a suitability index varying from 0 (poor quality) to 100 (high quality) (fig.5.5).

Connectivity among existing populations was derived from historical data about spontaneous colonisations. Nievergelt (1966) relates two such events: 1° from Blattenstock to Wetterhorn (in 1929), 2° from Augstmatthorn to Brienzer Rothorn (in 1947). He notes also that population exchanges are highly probable between Schwarzmönch and Wetterhorn populations. On the other hand, no spontaneous dispersal was ever observed from Augstmatthorn to Justistal, but individuals captured in the former and dropped into the latter rapidly returned home. Moreover, though Hutstock and Rothorn colonies are quite close (7.5 km apart), no population exchange has ever occurred between them.

The friction map was built by confronting this connectivity information to landscape features lying between the populations. We are developing a method based on multiple regressions, by which a friction coefficient may be assessed for

each landscape feature; positive coefficients indicate obstacles to movement, whilst negative values indicate condition favouring dispersal; the friction map is computed on the base of these coefficients (Appendix B; Patthey & Hirzel, in prep.). Landscape features having a significant impact (either positive or negative) on Ibex dispersal included topographic (steep slopes, lowest and highest altitudes), hydrographic (large rivers) and ecological (low HS) information.

2. Model parameters

Hex radius was set to 7 pixels (or 700 meters) as the best trade-off; a 1000-m radius proved to be too coarse a resolution to comply with some landscape peculiarities; in contrast, a 500-m radius generated hexs too small to hold more than three individuals. Each 7-pixel-radius hex covered therefore a 1.32 km² area, which is about half of the female Ibex mean summer home range (Abderhalden and Buchli, 1997). This generated a lattice of 36 by 42 cells. Hex HS ranged from 0.3 to 95.8 and impermeabilities from 12 to 47. Impermeability threshold over which no dispersal could occur was set to 25 to reflect known interconnections between local populations (Nievergelt, 1966). Each time step, one phase was first devoted to reproduction

and 8 phases were then devoted to dispersal in order to allow a theoretical maximal yearly dispersal range of 11.2 km, a value consistent with field knowledge.

Carrying capacity of all Swiss Ibex populations were assessed on the base of monitored data (chapter II). These results were used to fit the relation between HS and carrying capacity (fig.5.6 and equation 5.7) (for an application of this method to Roe Deer (*Capreolus capreolus*), see Saudan & Hirzel, 2000). As we were only interested in the lower value prediction of the model (hex HS and carrying capacity) and as it was observed that the linear model tended to overestimate the carrying capacity of the largest populations, only populations with K less than 500 individuals were fed into the linear regression. The integer carrying capacity K_i of each hex was therefore computed as a linear function of its suitability Q_i by:

$$K_i = \text{Round}(0.077 Q_i - 0.77) \quad (5.7)$$

This implies that all hexs with a HS lower than 10 were assigned a null carrying capacity. Following this model, hex carrying capacity varied between 0 and 7.

Dispersal mortality was modelled as linear function of impermeability. According to Ibex experts (Rauch, 1941; Nievergelt, 1966) and J.-C. Roch, P. Ratti,

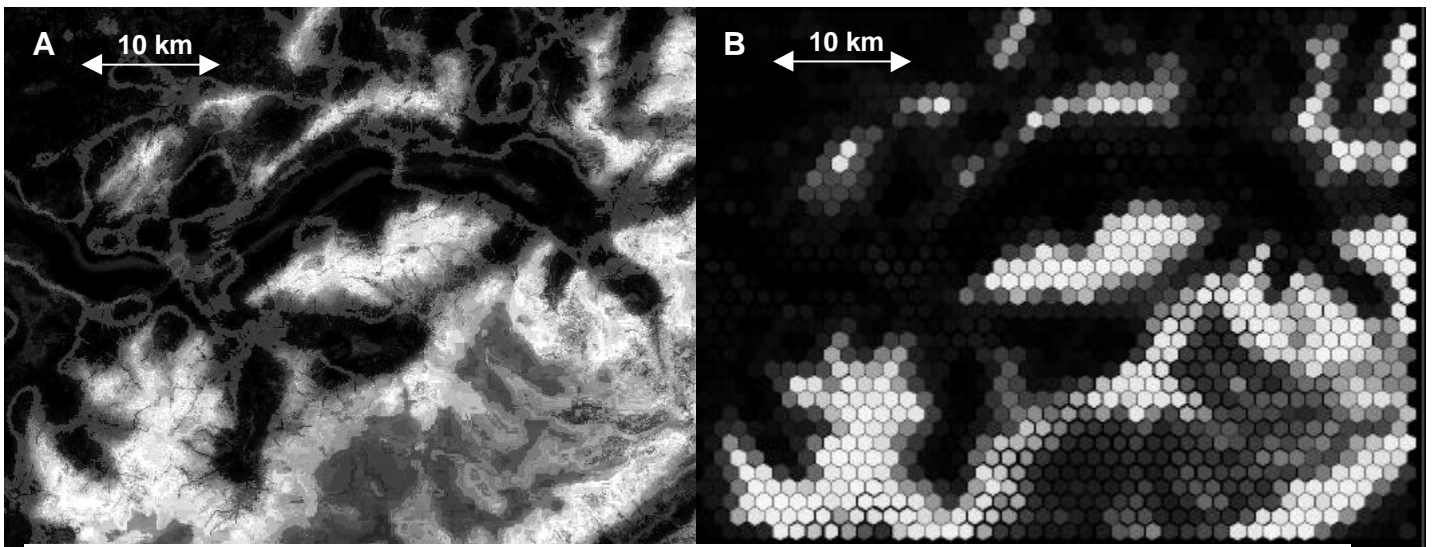


Fig. 5.5: Pictures A and B represent the same area as in fig.5.4. Picture A shows the Ibex habitat suitability; the brighter the pixel, the higher the suitability. Picture B is the derived cellular automaton; the brighter the hexagonal cell, the higher its carrying capacity, and the thicker the cell wall, the higher its impermeability.

pers. comm.), mortality is very low during dispersal. The model was therefore set up in such a way that dispersal mortality d varied as a function of impermeability P from 0 to 0.1:

$$d = 0.003 P - 0.04 \quad (d_{max} = 0.1) \quad (5.8)$$

Several values of these parameters were tried to assess simulation sensitivity. Whilst much higher dispersal mortality rates ($d_{max} > 0.5$) tended to slow down colonisation processes or even prevent it, simulation results did not qualitatively varied with d_{max} lower than 0.2.

The population dynamics was modelled by the logistic function (5.1b), where the intrinsic rate of increase was set to $r_0 = 0.3$ (chapter II). Dispersal rate was modelled

by the sigmoid function (5.5). Environmental and demographic stochasticity were set respectively to $e_e = 0.05$ and $e_d = 0.05$ (Hirzel et al., 1998), chapter II); no regional stochasticity was included. Ibexes are not great dispersers and they tend to leave a patch only when it becomes overcrowded (Rauch, 1941); the parameters were therefore tuned in such a way to constrain dispersal rate between 0 and 0.1 with an amplitude $a = 0.1$, a slope $b = 1$ at the inflexion point $g = 1.3$ and a minimal value $d = 0$. The high value of g assured that there would be no noticeable dispersal before an over-saturation of 120% would be achieved. Dispersal stochasticity was set to $e_m = 0.05$.

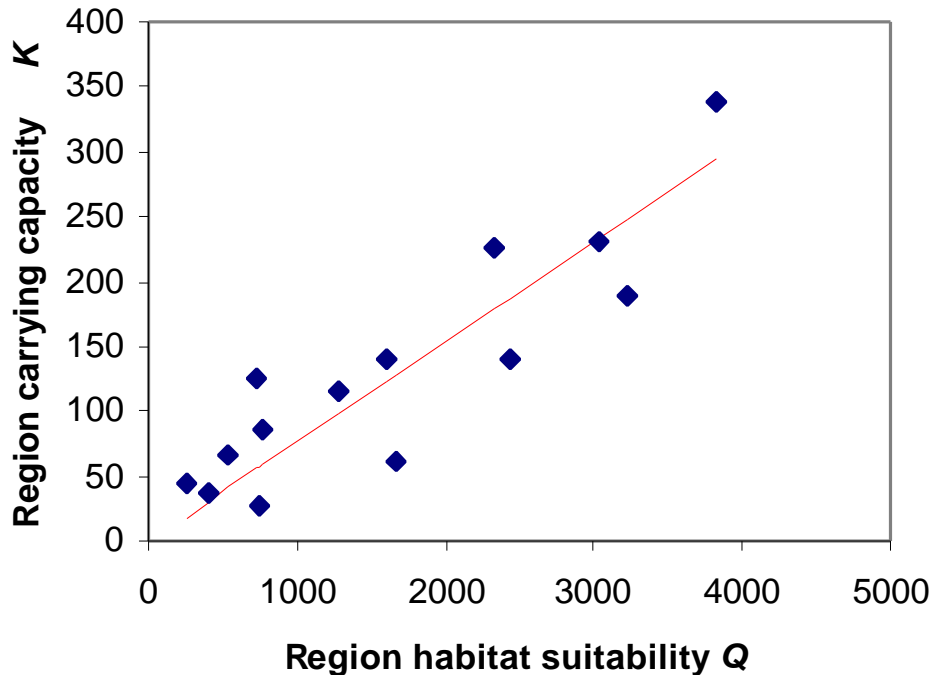


Fig.5.6: Carrying capacity as a function of habitat suitability. The linear model was fitted on 14 Swiss Ibex populations smaller than $K=500$. The determination coefficient is $R^2 = 0.93$. This model was used to compute hex carrying capacity from the quality of underlying landscape.

3. Results

Once the parameters have been fixed, 1000 replicates were run over a 200-time-steps period. A “snapshot” of the situation

was recorded every 5 time steps. On the screen, hex and region dynamics could be played back, showing, for each recorded time unit, density quantile statistics of the 1000 replicates. This allowed to study the

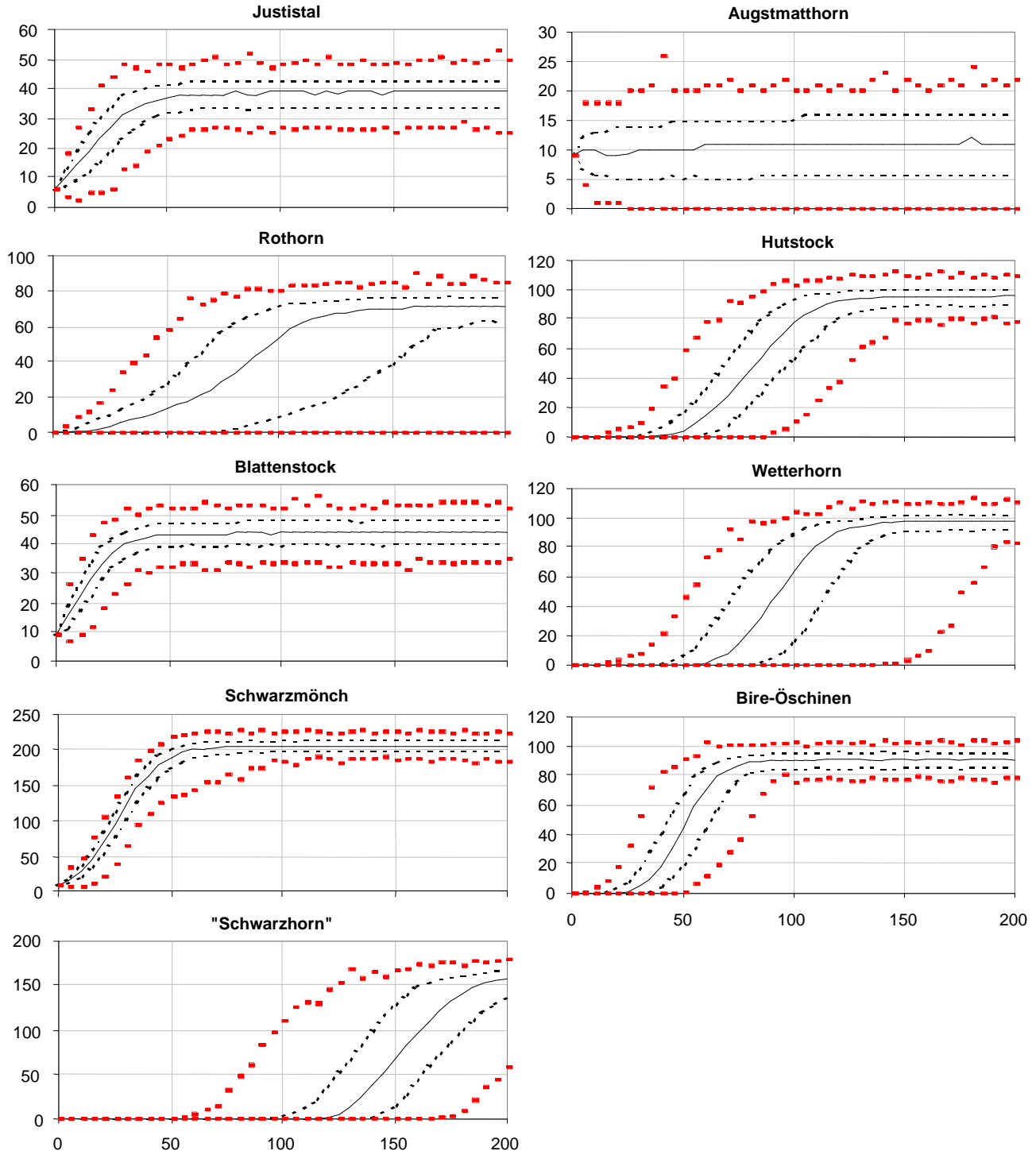


Fig.5.7: Population dynamics in the 9 recorded regions over a 1000-replicates simulation. Plain line traces median density evolution, dashed lines encompass then 80%-confidence interval and the small sticks are the minimum and maximum densities attained at each time unit. Initially, 10 individuals were introduced in each of the following regions: Justistal, Augstmatthorn, Blattenstock and Schwarzmönch (fig.5.4). The “Schwarzhorn” region is a potential patch presently unoccupied.

spreading across the landscape and the dispersal paths. Only the region dynamics are presented here (fig.5.7) and major results are summarised in table 5.1.

It was impossible to remake the exact history of the whole area as local populations have been perturbed by several reintroduction events as well as by management culling. Our aim was to generate plausible patterns of spontaneous colonisation and equilibrium densities similar to observations.

Apart from a few wandering individuals, Justistal region was, in the simulations, quite isolated from the rest of the regions. After about 70 years, when approaching saturation, it colonised a near small patch north-eastward; the constant exchange between these two patches made density of the source population to decrease slightly to attain an equilibrium between 34 and 43 individuals, slightly less than its nominal carrying capacity ($K=48$). In the field, Justistal population was reintroduced from 1949 to 1957. It is now fluctuating about its carrying capacity. Colonisation of the northern patch has not yet been observed but animals were known to wander eastward to Augstmatthorn region (Nievergelt, 1966).

In the simulations, the relatively poor quality of the Augstmatthorn region induced a fast colonisation (between 5 and 70 years in 80% of the replicates) of the adjacent Rothorn region. Emigration toward the latter caused its equilibrium density to stay lower (6 to 16 individuals) than its nominal carrying capacity ($K=23$). Rothorn density stabilised itself between 62 and 77, about its nominal carrying capacity ($K=75$). These two populations were isolated from all other except a few exchanges with Justistal. In the field, Augstmatthorn population was reintroduced from 1921 to 1924. Emigration causes Augstmatthorn equilibrium density to stay lower than its carrying capacity. First Rothorn spontaneous colonisation happened about 25 years later, in 1947 (In the simulations, colonisation happened such early with a prob-

ability of $P = 0.75$). It the stabilised about its carrying capacity.

In the simulations, Blattenstock progressively spread to colonise the Hutstock region after 30 to 60 years (80% confidence interval), stabilising themselves respectively between 40 and 48 ($K=43$), and between 90 and 101 ($K=97$). The westbound spreading attained the potential Schwarzhorn region after 90 to 140 years and mixed with the eastbound wave coming from Wetterhorn region. Schwarzhorn attained an equilibrium between 136 and 166 individuals ($K=158$). In the field, we found no historical information about Blattenstock and Hutstock colonies before 1990 when the yearly monitoring began. They are now both fluctuating about their carrying capacity. Schwarzhorn has not been colonised so far, with the exception of one isolated individual from 1989 to 1992 (C. Siegenthaler, pers. comm.).

Exchange with adjacent populations caused, in the simulations, Schwarzmönch population equilibrium density to stay higher than its carrying capacity ($K = 186$), between 198 and 214 individuals. In the field, Schwarzmönch population was reintroduced from 1924 to 1930 and in 1949 and 1950. After an initial slow increase, population jumped up in the early '90 to fluctuate above its carrying capacity.

In the simulations, the Schwarzmönch population rapidly spread in all directions to colonise first the Bire-Öschinen region after 15 to 35 years and then the Wetterhorn after 40 to 85 years. They stabilised themselves respectively between 85 and 95 ($K = 84$) and between 92 and 103 individuals ($K = 90$); both attained therefore a slightly higher density than their carrying capacity. Spreading continued then toward Schwarzhorn massif. In the field, Ibexes were reintroduced in Bire-Öschinen in 1961 and 1962 (30 to 36 years after Schwarzmönch); data are not available before that date. After a stepwise increase probably caused by immigrations, it attained its carrying capacity.

Table 5.1: Population nominal carrying capacity compared to simulated density, given by the 80% confidence interval limits. The 80% confidence limits of colonisation time are also given. Initial populations are typed in boldface.

Region	Carrying capacity [individuals]	Equilibrium den- sity, 10% limit [individuals]	Equilibrium den- sity, 90% limit [individuals]	Colonisation time, 10% limit [years]	Colonisation 90% limit [y]
Justistal	48	34	43	0	0
Augstmatthorn	23	6	16	0	0
Rothorn	75	62	77	5	70
Hutstock	97	90	101	30	60
Blattenstock	43	40	48	0	0
Wetterhorn	90	92	103	40	85
Schwarzmönch	186	198	214	0	0
Bire-Öschinen	84	85	95	15	35
“Schwarzhorn”	158	136	166	90	140

E. Discussion

1. Ibex population dynamics

Whilst many features of the real populations were reproduced by the simulated dynamics, it is difficult to quantify the agreement between model and reality. Due to several reasons that are listed below, we had to evaluate the model on the base of qualitative arguments such as population connectivity, spontaneous colonisation time and source/sink dynamic effects. Among the good agreements, some result from the calibration process, as, for example, the Justistal population isolation or the dispersal barrier between Rothorn and Hutstock. But, and this is more interesting, in other cases, uncalibrated dynamic features are reproduced by the model: the Augstmatthorn population decrease caused by the attraction of the Rothorn may be observed in the compared dynamics of these two populations. Moreover, the absence of Ibex in the Schwarzhorn potential population could be explained by a time delay: without human influence, our model predicts that Schwarzhorn would not be colonised before 90 years after Blattenstock reintroduction, that is at earliest (0.1 probability) in 2010, more probably (0.5) in 2040. Equilibrium density of Schwarzmönch, Wetterhorn and Blatten-

stock lay slightly over their carrying capacity, a tendency also observed in the real data; this could be due to a source-sink dynamics effect.

Nevertheless, several simplifications made it impossible to make detailed quantitative predictions about the history of these populations: Ibex behaviour has been reduced to a reproduction season and a dispersal season; reality is more complex as Ibexes are known to migrate from summer to winter settlements, dispersal happening mainly during the mating season (Nievergelt, 1966; Abderhalden et al., 1997). Males and females are also living separately most of the time, and their dispersal behaviours differ notably. In this paper all parameters were related to females, as they are the active principle of population dynamics and as generally only a few males mate with all females of the population. Allee effect, although known to be present in this species, was not introduced in the model, as there are no sufficient low-density data to tune the Allee parameter. Preliminary results show that Allee effect main influence is to delay colonisation of new areas. But all these factors may be neglected at first approximation. More critical is the impact of human management: most populations were either founded or at least maintained by several artificial reintroductions (Nievergelt, 1966). Individuals were captured in some populations to be released in another. Fi-

nally, since 1990, all colonies are submitted to management culling. The impact of these actions is difficult to take into account. What may be modelled here anyway are the interconnections between populations and their impact on local equilibrium, dispersal paths and colonisation probabilities. It would be interesting to apply this CA model to a more “natural” case.

We tested the sensitivity of Ibex simulations to regional stochasticity. Maximal autocorrelation was fixed to 1 and range of 5, 10 and 15 hexs (respectively 7, 14 and 21 km) were explored. The same set of simulations as described above was run with these values but no effect was detected. This is not surprising as the Ibex has a relatively large rate of increase and its populations grow quickly largely above their extinction threshold. Regional stochasticity would probably have a much more important impact on the dynamics of a less hardy species.

2. Cellular automaton model

2D-cellular automata are generally based on a square lattice. This is clearly for the sake of implementation simplicity and speed of execution. Moreover, in most CA applications, what is modelled is the cell dynamics as a result of interaction with neighbouring cells; these interactions can easily be weighted to correct the diagonal-length bias. In our case, the cells are used as pathway nodes and bias correction is impossible. This is why we chose to base our CA on a hexagonal lattice. The six neighbours are therefore equidistant (by contrast with an eight-cell square neighbourhood) and more regularly distributed (by contrast with a four-cell square neighbourhood).

Hex size is the basic parameter of the whole model. It conditions the whole system. First, the spatial resolution is evidently affected. But consequently, temporal resolution is also conditioned: larger hexs means wider neighbourhood effect and thus, less dispersal phases are needed to cover the same distance. This in turn

will affect the dispersal rate function. From a more technical point of view, smaller hexs imply more of them in the lattice, more dispersal phases, and thus, longer computations (The 1000-replicates simulation described above took about one hour to complete on a Pentium III, 933 MHz processor; when including regional stochasticity, it jumped to 36 hours). However, hex HS and impermeabilities are more sensible to spatial resolution of input raster map than to hex size. In particular, as impermeabilities are computed on the base of the least-cost path between the hex centres implies that even a thin linear barrier (e.g. a river, a highway) might have a high weight.

This model could apply to any species but its current implementation has been optimised for ground animals. Indeed, for flying dispersers (e.g. insects, birds or plants), dispersal ranges may be far higher than hex size and obstacles are less relevant. With the current model, it would be achieved by setting a high number of dispersal phases, which would generate high computation time. A better implementation would be to extend the neighbourhood radius, namely to allow dispersers to cross more than one hex during a single phase. This would allow larger dispersal range with the same amount of dispersal phases and thus keeping computation time low.

Population dynamics models are generally either population-based or individual-based. Both types of models have positive and negative properties that make them useful in different cases. As population-based models often aggregate large and complex areas into a single patch their spatial resolution is voluntarily coarse; interconnections between patches are also generally very simplified (Wiens, 1997). This lack of resolution makes the impact of landscape-level modifications difficult to evaluate. On the other hand, as management and monitoring are generally done at population level, application of this approach is straightforward. By contrast, individual-based models have a high spatial

resolution but it is difficult then to generalise information at population level; by the way, this kind of models can be practically intractable when large (not endangered) populations are addressed. The CA model exposed in this paper gives an alternative and intermediate approach able to model finely spatial phenomena while still keeping control at population level. Indeed, our CA is a kind of population-based model as it may be viewed as composed of a large number of tiny local populations strongly connected to their neighbourhood.

Including inter-population landscape structure into a model has proven to have both quantitative and qualitative effects. Quantitatively, (Moilanen et al., 1998) have shown that using environmental factors both for patch quality and dispersal rates assessment improved slightly but significantly the predictive power of their metapopulation occupancy model; qualitatively, (Gustafson et al., 1996)'s theoretical study have shown how landscape patterns could introduce asymmetry into the dispersal rates (funnel effect) and add conservation importance to small or poor-quality patches lying between more important ones.

However, this fine spatial resolution and its underlying advantages have a cost: much more parameters have to be tuned than in a simpler mathematical model. Still, by its lattice structure, the CA is very similar to GIS raster files and may benefit from its large toolkit of powerful spatially-explicit analyses. This paper has shown how such a model may be tuned both on GIS statistical data and empirical knowledge. Whilst GIS are powerful descriptive and analytical tools, they dynamical capabilities are limited; this paper has demonstrated that CA might be viewed as their direct extension toward process modelling. According to GIS wide and fast-growing success in applied and fundamental ecological sciences, there is no doubt that CA models are promised a brilliant future.

To conclude, the CA approach appears to fill the gap between mathematical

population-based and numerical individual-based models. We intend now to explore these new theoretical fields and refine the parameter-tuning analyses. By its fine spatial resolution linked to population-wide understanding make of our CA model an ideal tool to address various kinds of population movements, as migration, invasion and recolonisation, all topics of crucial importance in many conservation biology problems, either for pests management or endangered species protection or reintroduction.

F. Sequential gaussian algorithm

The sequential gaussian simulation algorithm used for spatially correlated stochasticity was derived from (Gomez-Hernandez et al., 1990; Goovaerts, 1998) with a few modifications designed to speed it up. The principle is to generate for each hex i a spatially correlated and normally distributed (mean=0, SD= e_e) random value X_i . The sequential gaussian simulation uses a kriging interpolation based on the Gaussian correlogram model, i.e.:

$$\mathbf{r}(h) = \begin{cases} 1 & \text{if } h = 0 \\ \mathbf{r}_{\max} e^{-\frac{3h^2}{a^2}} & \text{if } h > 0 \end{cases} \quad (6.9)$$

where $\mathbf{r}(h)$ is the spatial correlation between two hexs separated by a distance h , \mathbf{r}_{\max} is the maximal correlation for $h>0$ and a is the range, i.e. the distance at which $\mathbf{r}(a) = 0.05$.

The algorithm proceeds as follows:

1. Define a random path visiting only once each hex.
2. For each hex i taken sequentially along the path do:
 - 2a. Interpolate the value \mathbf{m} and its SD \mathbf{s}_i using simple kriging. Interpolation is done over a maximum of 20 nearest hexs inside a circular search area of radius a . If no hex has yet been generated inside this area, assume $\mathbf{m} = 0$ and $\mathbf{s}_i = 1$.

- 2b. Generate the value X_i of the hex, drawn from a Gaussian distribution with mean μ_i and SD σ_i .
- 2c. Proceed to the next hex along the path and repeat steps 2a to 2c until all hexs have been addressed.
3. Multiply each X_i by the environmental stochasticity coefficient e_e .

G. References

- Abderhalden, W., Buchli, C., 1997. Steinbockprojekt Albris/SNP, Schlussbericht. Zerne, ARINAS & FORNAT.
- Akçakaya, H.R., Atwood, J.L., 1997. A habitat-based metapopulation model of the California Gnatcatcher. *Conserv. Biol.* 11, 422-434.
- Akçakaya, H.R., McCarthy, M.A., Pearce, J.L., 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., Raphael, M.G., 1998. Assessing human impact despite uncertainty viability of the northern spotted owl metapopulation in the northwestern USA. *Biodiversity and conservation* 7, 875-894.
- Alonso, D., Solé, R.V., 2000. The divgame simulator: A stochastic cellular automata model of rainforest dynamics. *Ecol. Model.* 133, 131-141.
- Balster, H., Braun, P.W., Köhler, W., 1998. Cellular automata models for vegetation dynamics. *Ecol. Model.* 107, 113-125.
- Beecham, J.A., Farnsworth, K.D., 1998. Animal foraging from an individual perspective: an object oriented model. *Ecol. Model.* 113, 141-156.
- Borland, 1998. Borland Delphi Professional, Inprise Corporation. URL: <http://www.inprise.com/delphi/>
- Bunn, A.G., Urban, D.L., Keitt, T.H., 2000. Landscape connectivity: a conservation application of graph theory. *Journal of Environmental Management* 59, 265-278.
- Cannas, S.A., Páez, S.A., Marco, D.E., 1999. Modeling plant spread in forest ecology using cellular automata. *Computer Physics Communications* 121-122, 131-135.
- Clarke, R.T., Thomas, J.A., Elmes, G.W., Hochberg, M.E., 1997. The Effects of Spatial Patterns in Habitat Quality On Community Dynamics Within a Site. *Proceedings of the Royal Society of London - Series B: Biological Sciences* 264, 347-354.
- Darwen, P.J., Green, D.G., 1996. Viability of populations in a landscape. *Ecol. Model.* 85, 165-171.
- Day, J.R., Possingham, H.P., 1995. A Stochastic Metapopulation Model With Variability in Patch Size and Position. *TPB* 48, 333-360.
- Eastman, J.R., 1999a. Idrisi32, Guide to GIS and image processing. Worcester, Clark University.
- Eastman, J.R., 1999b. Idrisi 32. Worcester, Clark University. URL: <http://www.clarklabs.org>
- Gomez-Hernandez, J.J., Srivastava, R.M., 1990. ISIM3D: An ANSI-C three-dimensional multiple indicator conditional simulation program. *Computers & Geosciences* 16, 395-440.
- Goovaerts, P., 1998. Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the spaces of uncertainty of flow predictions. Stanford, Stanford University. 2001.
- Grimm, V., Frank, K., Jeltsch, F., Brandl, R., Uchmaski, J., Wissel, C., 1996. Pattern-oriented modelling in population ecology. *The Science of The Total Environment* 183, 151-166.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147-186.
- Gustafson, E.J., Gardner, R.H., 1996. The effect of landscape heterogeneity on the probability of patch colonization. *Ecology* 77, 94-107.
- Hainard, R., 1962. Mammifères sauvages d'Europe. Paris, Delachaux et Niestlé.
- Hanski, I., 1994. A practical model of metapopulation dynamics. *J. Anim. Ecol.* 63, 151-162.

- Hanski, I., 1998a. Connecting the parameters of local extinction and metapopulation dynamics. *Oikos* 83, 390-396.
- Hanski, I., 1998b. Metapopulation dynamics. *Nature* 396, 41-49.
- Hanski, I., Gilpin, M.E. (1991). Metapopulation dynamics: brief history and conceptual domain. *Metapopulation dynamics: empirical and theoretical investigations*. M. Gilpin and I. Hanski. London, Academic Press, 3-16.
- Hanski, I., Gilpin, M.E. (1997). Metapopulation dynamics: From concepts and observations to predictive models. *Metapopulation biology : ecology, genetics, and evolution*. I. Hanski and M.E. Gilpin. San Diego, Academic Press, 69-91.
- Hanski, I., Ovaskainen, O., 2000. The metapopulation capacity of a fragmented landscape. *Nature* 404, 755-758.
- Hiebeler, D., 2000. Populations on fragmented landscapes with spatially structured heterogeneities: landscape generation and local dispersal. *Ecology* 81, 1629-1641.
- Hill, M.F., Caswell, H., 1999. Habitat fragmentation and extinction thresholds on fractal landscapes. *Ecology Letters* 2, 121-127.
- Hirzel, A., Hausser, J., Chessel, D., Perrin, N., in press. Ecological-Niche Factor Analysis: How to compute habitat-suitability maps without absence data? *Ecology*.
- Hirzel, A.H., Hausser, J., Perrin, N., 2001. Biomapper 1.0. Lausanne, Lab. for Conservation Biology. URL: <http://www.unil.ch/biomapper>
- Hirzel, A.H., Perrin, N., 1998. Modeling the dynamics and management of Ibex populations in Switzerland. *Proc. 2nd World Conf. Mt. Ungulates*, 139-144.
- Hogeweg, P., 1988. Cellular automata as a paradigm for ecological modeling. *Applied Mathematics and Computation* 27, 81-100.
- Judson, O.P., 1994. The rise of the individual-based model in ecology. *Trends in Ecology & Evolution* 9, 9-14.
- Karafyllidis, I., 1998. A model for the influence of the greenhouse effect on insect and microorganism geographical distribution and population dynamics. *Biosystems* 45, 1-10.
- Keeling, M.J., Gilligan, C.A., 2000. Bubonic plague: A metapopulation model of a zoonosis. *Proceedings of the Royal Society of London B* 267, 2219-2230.
- Kizaki, S., Katori, M., 1999. A stochastic lattice model for locust outbreak. *Physica A* 266, 339-342.
- Letcher, B.H., Priddy, J.A., Walters, J.R., Crowder, L.B., 1998. An Individual-Based, Spatially-Explicit Simulation Model of the Population Dynamics of the Endangered Red-Cockaded Woodpecker, *Picoides borealis*. *Biological Conservation* 86, 1-14.
- Levins, R., 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society America* 15, 237-240.
- Moilanen, A., Hanski, I., 1998. Metapopulation dynamics: Effects of habitat quality and landscape structure. *Ecology* 79, 2503-2515.
- Molofsky, J., 1994. Population Dynamics and Pattern Formation in Theoretical Populations. *Ecology* 75, 30-39.
- Murdoch, W.W., 1994. Population regulation in theory and practice. *Ecology* 75, 271-287.
- Nievergelt, B., 1966. Der Alpensteinbock (*Capra ibex* L.) in seinem Lebensraum : Ein ökologischer Vergleich verschiedener Kolonien.
- OFEFP, 1991. Commentaires relatifs à l'ordonnance sur la régulation des Bouquetins en Suisse (ORB), OFEFP.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1989. Numerical recipes in Pascal : The art of scientific computing. Cambridge, Cambridge University Press.
- Ratti, P., 1994. Stand von Hege und Erforschung des Steinwildes im Kanton Graubünden (Schweiz). *Zeitschrift für Jagdwissenschaft* 40, 223-331.
- Rauch, A., 1941. Le Bouquetin dans les Alpes. Paris.
- Rushton, S.P., Lurz, P.W.W., Fuller, R., Garson, P.J., 1997. Modelling the distribution of the red and grey squirrel at the land-

- scape scale : a combined GIS and population dynamics approach. *J. Appl. Ecol.* 34, 1137-1154.
- Ruxton, G.D., 1996. Effects of the spatial and temporal ordering of events on the behaviour of a simple cellular automaton. *Ecol. Model.* 84, 311-314.
- Ruxton, G.D., Saravia, L.A., 1998. The need for biological realism in the updating of cellular automata models. *Ecol. Model.* 107, 105-112.
- Saudan, M., Hirzel, A., 1999. Corrélation entre facteurs écogéographiques et capacité de soutien chez le chevreuil (*Capreolus capreolus*) au Valais. Lausanne, Laboratory for Conservation Biology.
- Schönfisch, B., de Roos, A., 1999. Synchronous and asynchronous updating in cellular automata. *Biosystems* 51, 123-143.
- Sipper, M., Tomassini, M. (1998). An introduction to cellular automata. *Bio-inspired computing machines: Toward novel computational architectures*. D.M.a.M. Tomassini, PPUR, 49-58.
- Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A., 2000. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecol. Model.* 133, 209-223.
- South, A., 1999. Dispersal in spatially explicit population models. *Conserv. Biol.* 13, 1039-1046.
- Tischendorf, L., 1997. *Modelling Individual Movements In Heterogeneous Landscapes - Potentials Of a New Approach*. *Ecol. Model.* 103, 33-42.
- Von Neumann, J., 1966. *Theory of selfreproducing automata*, Univ. of Illinois Press.
- Wiegand, T., Moloney, K.A., Naves, J., Knauer, F., 1999. Finding the missing links between landscape structure and population dynamics: a spatially explicit perspective. *Am. Nat.* 154, 605-627.
- Wiens, J.A. (1997). *Metapopulation dynamics and landscape ecology*. *Metapopulation biology : ecology, genetics, and evolution*. I. Hanski and M.E. Gilpin. San Diego, 43-62.
- With, K.A., Crist, T.O., 1995. Critical thresholds in species' responses to landscape structure. *Ecology* 76, 2446-2459.
- Wolfram, S., 1984. Cellular automata as models of complexity. *Nature* 311, 419-424.

VI. GENERAL CONCLUSION

Through the problematic of Ibex population in Switzerland, tackled from the demographic, landscape and management angles, this work produced several tools that might be used for various other large populations. A chain of models and softwares was forged, whose each link allows putting together theoretical knowledge and field data to achieve a global comprehension of the species in its environment. Three fields of study were covered – population dynamics, environmental management and landscape ecology – to be finally unified into a global approach (fig. 6.1).

Population ecology

The local population model (*SIM-Ibex*, cf. chapter II) was designed to support Ibex culling management. It was therefore very detailed (e.g. it was structured by age and sex classes). But for the general purpose of my thesis, the *SIM-Ibex* software most useful

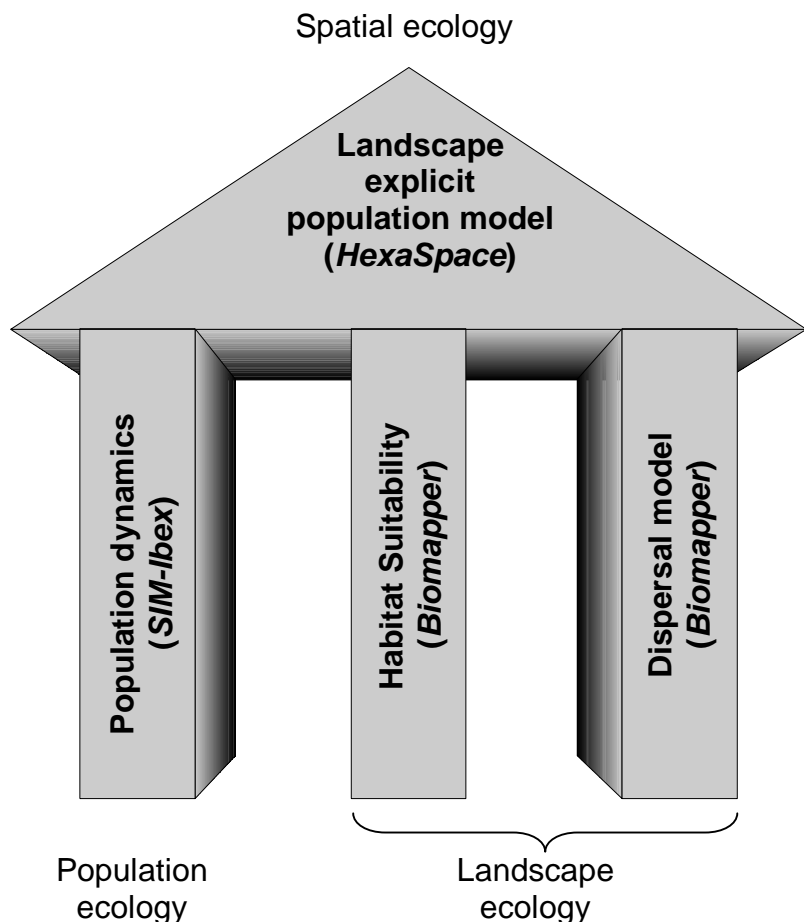


Figure 6.1: This work linked a population ecology model (SIM-Ibex) to two landscape ecology models (built with Biomapper) into a global landscape-explicit population model (HexaSpace).

property was automated and robust assessment of population carrying capacity on the base of survey and culling data. As all Swiss colonies have been monitored at least since 1990, and as these data are collected by the Federal Office of Environment, Forest and Landscape, it was possible to compute their carrying capacities.

Although the software is Ibex-specific, its underlying population dynamics model might be extended to other species. It is difficult to build population models generalisable to any species, as they must fit tightly to life-history peculiarities and available data. Nevertheless, our work has shown how it was possible to realise a robust management-support software.

Landscape ecology

By the development of *Biomapper* (cf. chapters III and IV), we contributed several GIS-methods to landscape ecology. Indeed, GIS are mainly conceived for human-linked applications (e.g. urbanism, agriculture, cadastre or satellite imagery) and natural sciences must therefore use existing general-purpose tools or develop their own modules. Our project was incidentally to produce ecology-oriented analyses and methods.

Whilst *Biomapper* modules cover GIS operations along the whole process of habitat suitability mapping, from map conversion and preparation to results validation (cf. appendix B), its main contribution is the Ecological Niche Factor Analysis (ENFA). Being based on the fundamental concept of ecological niche, it provides two key informations about the niche of the focal species, namely those of marginality and of specialisation. A crucial advantage of ENFA is that it does not require absence data. Presence data are compared instead with “background” environment, making this method fairly robust to data quality. This property is important because database often simply lack any absence data or, when available, these may turn out to be either unreliable (in the case of cryptic or poorly known species) or meaning-

less (in the case of invading species, or those living in fragmented habitats where some patches have become extinct). As many species enter one of these categories, our approach potentially has a wide application range.

Biomapper outputs may have many applications. On the one hand, they may be used as sheer results: as in the following examples: Habitat suitability maps are powerful tools to predict 1° areas that could be invaded by an expanding population or 2° best locations where to reintroduce individuals of an endangered species. 3° They may also help field biologist to circumscribe home range of cryptic species by focalising their exploring efforts on areas with high presence probability. 4° Potential distribution maps that may be derived from the former are frequently used in atlases. 5° Ecological niche marginality and specialisation scores give precious information about a species realised niche. 6° Friction maps may be used to infer dispersal paths. On the other hand, *Biomapper* outputs may also be used as temporary results and fed into other models for further analysis, as in the following examples: 1° A species habitat suitability map may be used as predictor into the habitat model of one of its predators. 2° Although *Biomapper* results are a static snapshot of a process, both habitat suitability and friction maps may be used to parameterise dynamical models. 3° Friction maps may be used to compute “ecological distance” between populations, which seems to better explain genetic distances than classical geographic distance (Fontanillas & Hirzel, unpublished results).

Spatial ecology

Simple population dynamics models are often said to be lacking applicability because they neglect spatial processes. In contrast, ecological GIS-based models, despite their high space-awareness, are often criticised for they are only static representation of reality and do not give any clue to driving processes. These cases are ex-

tremities of a continuum ranging from purely local dynamical to purely spatial static models; our Ibex management models and habitat suitability mapping methods obviously may be pinned at both ends of this spectrum. Attempts have been made to introduce spatiality into dynamical models and dynamics into spatial models (cf. chapter V) but they generally keep cautiously the standardised methods used in their speciality and therefore do not adventure themselves very far from their end of the continuum.

The cellular automaton base of our model *HexaSpace* achieves a better merging: its lattice spatial structure is borrowed from GIS and the cellular dynamics and interactions derive directly from mathematical population dynamics. This double heritage is used during both calibration and prediction phases.

Parameter calibration benefits from the well-known assessment methods of population dynamics and from the latest developments in GIS technology and landscape ecology. We have shown in chapter V how techniques developed for demographic modelling (cf. chapter II) and habitat suitability mapping (cf. chapter III) might be used to tune our landscape-explicit model.

From the predictive power point of view, both approaches conjugate their qualities and compensate their weaknesses. Actually, combination of dynamics and landscape is generally made through individual-based models, with an inherent loss of population-wide understanding. By taking both of these aspects into account, *HexaSpace* allows addressing several types of problems, as for example: 1° Where an invading species will be spreading? 2° when and how long? 3° What are the crucial dispersal corridors between populations and 4° how to improve or suppress them? 5° What is the source-sink dynamics of the metapopulation, due to both demographic and geometric causes? 6° What is the importance of dispersal dynamics on habitat colonisation?

Visions

We saw how linking landscape- to population ecology could improve our understanding and predictive power of large and complex ecological processes. However, our work is currently restricted to small temporal scale, at which the landscape may be considered invariant. At the evolution scale, or even at smaller scale if the most pessimistic predictions of global changes prove to be true, this is obviously incorrect. Including landscaped dynamics, i.e. implementing temporal variation of carrying capacity and, possibly, of friction, in our cellular automaton would therefore entail powerful prediction potentialities.

Moreover, merging landscape and population ecology opens the doors toward unification of population ecology and genetics. Both domains are till now rather disconnected, although they are fundamentally addressing the same questions: how and why are populations structured?

Collaborations between ecology and genetics might produce improvements in both fields. Population genetics, by comparing allelic frequencies, allows quantifying connectivity between populations; this information is obviously interesting for dispersal rates assessment or friction mapping. Reciprocally, landscape-explicit dynamics are important to explain observed population structures, to reconstruct species expansion histories or, more prosaically, to help focusing sampling campaigns on crucial locations in order to prevent time and money waste.

Still more interesting could be to fuse more completely these approaches together, for example by extending the *HexaSpace* cellular automaton principle to model directly allele propagation. This “landscape genetics” should produce interesting results for both conservation and theoretical purposes.

Finally, linking population dynamics, population genetics and landscape dynamics would be a crucial mean to address problems of speciation by geographic isolation. Effects of large migration barriers

on population movements induced by climatic changes could be simulated and their results confronted to observed genetic patterns.

Large population modelling often implies working on wide areas and therefore complex landscapes. This entails several technical difficulties that have delayed such modelling efforts. Now that we can take advantage of powerful enough computers to deal with them, a wide field of study is opening before us. Scientists are

no longer constrained to deal with individuals or small numbers but can tackle any fundamental or applied question. Moreover, when addressing large populations, they benefit from the statistical facilities of large numbers and have therefore access to the capabilities of classical statistics. Links with population genetics should allow boarding complex evolutionary processes involving population movements and density fluctuations.

ACKNOWLEDGEMENTS

This thesis has benefited from the help or inspiration of many people during all the stages of its maturation. Coming at the end of this work, I wish to list here all the contributions that made this work possible, by order of apparition:

First came Nicolas Perrin and Jacques Hausser whose teaching in theoretical and quantitative ecology generated my enthusiasm for these domains. In the same time, Laurent Keller taught me rudiments of scientific presentations and reports, which were most useful in all my further works.

During my Masters, Roger Ardit and particularly Jurek Michalski taught me the particularities of theoretical research and scientific publication. I am also grateful to Antony Starfield whose modelling advises and personality were greatly inspiring.

How to express all my gratitude to Nicolas Perrin? This work would not have been possible without his kind help and his receptiveness. All along my thesis, Nicolas generously provided me his keen advices upon all facets of scientific work, from modelling to English writing. His example was a constant inspiration for me. Moreover, he proved always to be a joyful fellow.

My consideration goes also to Cornelis Neet who was among the initiators of the Ibex project and was also my principal link with applied wildlife management, from

the conceptualisation phase of my thesis to its conclusion as jury member.

I wish also to thank all the members of the Ibex workgroup, Hans-Jörg Blankenhorn, Cornelis Neet, Reini Schnidrig-Petrig, Edy Hollenweg, Cornelia Schütz, Peider Ratti, Patrick Durand, Marco Giacometti and Cornelia Gallmann, who contributed actively to *SIM-Ibex* improvement.

I am also very grateful to Jérôme Goudet with whom I had a lot of productive connections in the domain of statistics and algorithmic.

As my work was covering the whole Switzerland, it was impossible for me to collect the data. I had therefore to rely upon other persons for this task. My gratitude goes to all people who contributed to Ibex survey in whole Switzerland. In particular, I want to thank Jean-Claude Roch for the field knowledge he accepted to share with me directly in the heart of Vaud Ibex colonies.

I greatly appreciated Harald Brunner's wide biological knowledge and he was also very helpful to palliate my German deficiencies both for reading papers and translating *SIM-Ibex*.

I want also to thanks Bertrand Gentizon, Régis Caloz, Antoine Guisan, Anthony Lehman, Gwenaëlle Le Lay and Séverine Vuilleumier for all the interesting discus-

sions we had about spatial ecology and GIS use.

Thanks to several thorough testers and reviewers, bug density in my softwares was greatly lowered and clarity of my papers was highly improved. For that, I am grateful to Nicolas Perrin, Annie Marineau, Etienne Bertouille, Sébastien Sachot, Patrick Patthey, Isabelle Dunand and Gwénaëlle Le Lay. My thanks fly also to all those around the World who tested *Biomapper* and sent me their comments and bug reports.

During my thesis, I benefited from the enthusiastic help of several students: Zoé Auguste-Bernex, Karine Fattebert, Myriam Saudan, Véronique Helfer, François Métral, Patrick Presi, Raoul Vega, Marion Tanner and Séverine Grosjean. May the Force be with them.

I do not forget all the practical and administrative help I received from France Pham, Patrick Moratal, Christian Koenig and Michel Schupbach.

My sister Laurence Mbayo does not only merit thanks for the *Biomapper*'s logo realisation but she worked hard too, along with Martine Théraulaz and Janine Hirzel, to keep high my glycemia and vitamin rate all along my thesis.

May all my kind co-workers of the Institute of Ecology be thanked here for the good ambience they distillated around me at the university or during the unforgettable Roscoff training courses.

Finally, I want to thank Jérôme Goudet, Cornelis Neet and Antoine Guisan who kindly accepted to expertise my work.

Working and discussing with all those persons made my PhD time both an exciting and very instructive experience. For that, I thank again all of them. Nevertheless, should the reader find any error or inaccuracy in this work, it would be due to my very own responsibility.

This research was supported by the Swiss Federal Office for Environment, Forest and Landscape (OFEFP), Wildlife section, grant 0310.3600.305.

APPENDICES

Appendix A: Ibex official forms

1. Monitoring form

BOUQUETIN Formulaire I

Annexe 1
(art. 2, 2e al.)

Recensement de la population

Boucs		Chèvres	
32	de 11 ans et plus		
25	de 6 à 10 ans		45
29	de 3 à 5 ans		
Total jeunes animaux		33	de 1 et 2 ans
Total faons		16	1) <input checked="" type="checkbox"/> calculé <input type="checkbox"/> relevé
Total boucs		86	
Population (Boucs, chèvres, faons)		180	45
Total chèvres			

SR 3) 1 : 0,52

Progéniture 16

Accroissement 2) 16.5

1)

Voir commentaires au verso.

Colonie: Cape-mr-Moine
Canton:VD.....
Année:1951.....
Date du recensement:6. juillet. 1951.....
Effectué par: M. M. Cathélas, S.P.F.

2. Culling plan form

Plan de tir 1991 = 8.12 % du cheptel en 183 lots (moyenne 1980)
Plan de tir 1992 = 4.1 % du cheptel en 180 lots

BOUQUETIN Formulaire II

Annexe 2
(art. 6, 2e al.)

Planification des tirs

Boucs	Chèvres
<div>3 de 11 ans et plus</div> <div>2 de 6 à 10 ans</div> <div>3 de 3 à 5 ans</div>	<div>4 de 3 ans et plus</div>
Jeunes animaux	de 1 et 2 ans
Faons	

Total boucs: 8 (12) + 1/2 = 8.5

Total chèvres: 4

Ensemble des tirs: 12.5

SR: 1:4.12 lors des tirs

Quota de tirs en % de VM des effectifs des 3 dern. années: 6.85 %

VM de la progéniture des 3 dernières années

VM du gibier péri des 3 dernières années

Quota de tirs en % de VM des effectifs des 3 dern. années

Abréviations
 SR: sex ratio
 VM: (valeur) moyenne

Voir commentaires au verso.

Colonie: *Lape. au Maine*
 Canton: *Vaud*
 Année: *1992*

3. Culling results form

BOUQUETIN Formulaire III Contrôle de la planification des tirs
(données de l'année précédente)

Annexe 3
(art. 9, 4e al.)

Colonie: *Capr-95: Hous...*

Canton: *VD.....*

Année: *1951.....*

boucs			chèvres		
tirs prévus	tirs	gibier péri	âge	gibier péri	tirs prévus
2			13 +	1	
			12		
			11		1
2	2		10		
	1		9		
	1		8		2
2		1	7		
			6		
			5		
2			4		
	2		3		1
			2		
2	3		1		
	1		faons		
			indéter. 1)	2)	
0			Total	1	4
Total tirs et gibier péri ♂		11	Total tirs et gibier péri ♀		5
Total pertes		16	Total pertes		

Total, pertes (tirs, gib. péri)

♂	11	16	5	♀
Total				
SR	1 : 0.45			

Total, tirs

♂	10	14	4	♀
Total				
SR	1 : 0.44			

Total, gib. péri

♂	1	2	1	♀
Total				
GV	1 : /			

Voir commentaires au verso.

Appendix B: Biomapper modules

Biomapper is a kit of GIS- and statistical tools designed to build habitat suitability (HS) models and maps. It is centred on the Ecological Niche Factor Analysis (ENFA) but comprises several other modules covering operations all along the process:

1° As biological data are generally composed of observation lists and vegetation relevés, they must be converted into a GIS map. 2° Available GIS-databases have not been conceived for ecological applications; thus, more relevant variables according to focal species ecology must be derived from them (e.g. local rock frequency, local forest border length, grouping of scattered observations into a single patch). 3° Variables have to be explored statistically and normalised through appropriate transformations. 4° Observation data sets may be parted into calibration and validation sets. 5° Ecological Niche Factor Analysis was specifically conceived for environmental studies. 6° Habitat suitability mapping and validation. 7° Friction map computation on the basis of genetic or field data.

The modules are the following, sorted by alphabetical order.

1. Agglomerator

This module is used to build maps from vegetation relevés. These relevés are a list of record defining x and y coordinates, a plant species and a set of attributes of this species at this location (e.g. herb, bush and tree stage coverage); therefore, a same location occurs as many time as they contain species. *Agglomerator* builds a table where rows represent locations and columns represent species attributes. This table may then be used by the module *Convertor* to build GIS maps.

2. BigGroup

BigGroup determines contiguous groups of identically valued integer cells in a map. Cells belonging to the same contiguous

group are given a unique integer identifier (in fact, of real type), numbered from 1 and consecutively in the order found.

This module is similar to the Idrisi *GROUP* module but is more efficient (more groups may be processed and faster, a background group may be defined, discontinuous cells may be grouped together).

3. Booleanisator

This module is designed to convert a quantitative or qualitative map into several boolean maps covering a given value interval. For instance, converting a digital elevation model into three boolean maps, elevation less than 500 meters, from 500 to 1000 m, 1000 m and above.

4. CircAn

This contextual operator produces an image where each cell contains a value integrating variously data lying in a surrounding circular area. This can be viewed as a circular-moving window. The integration consist of one of the following statistics:

- Frequency of occurrence
- Mean (smoothing)
- Gaussian smoothing (Weighted mean, weights being binormally distributed)
- Maximum or minimum
- Border length
- Compactness ratio
- Border length divided by area
- Sum of differences with central cell value
- Convexity-concavity

The user may change the radius of the circular window.

5. Convertor

Convertor is useful to create a map from a list of locations given by their coordinates and thematic attributes (e.g. species observation lists).

6. DistAn

On the basis of a boolean map of observation points, Distance Analysis computes some distance-related statistics used for instance for home-range analyses. The following analyses are implemented:

- Minimum distance
- Distance harmonic mean
- Distance geometric mean
- Minimum convex polygon
- Voronoi tessellation

7. Frictionator

This module builds friction maps from population connectivity information and a set of landscape feature occurrence maps.

In a friction map, cell values represent the energy cost needed to cross it. A friction value equal to one is the standard cost; higher friction values represent dispersal barriers whilst lower values indicate dispersal corridors. Obviously, friction depends on landscape features and the aim of this module is to assign a friction weight to each relevant predictor.

Population connectivity may be assessed by two types of methods: 1° field observations (e.g. signs, radio-tracking, GPS-collar) and 2° population genetics (e.g. microsatellites, allozymes). It is given by a distance matrix assigning a distance to each pair of populations. Populations are spatially referenced by their centre.

First, *Frictionator* computes a frequency histogram of the landscape features lying between each population pair; this may be done by two different algorithms: 1° the *straight path* algorithm scans a rectangular area between the two populations. This rectangle is oriented along the connecting line and may have any odd width (given in number of cells). 2° the *least cost path* algorithm needs an input friction map (which may be made by the straight path algorithm) that will be used to draw N least cost paths. These paths are scanned to compute the histogram. Then, a multiple linear regression is computed with connectivity distance as dependent variable and landscape feature frequencies as predictors.

The multiple regression computes a weight for each landscape predictor; negative weights mean that the landscape feature presence enhances dispersal whilst positive weights indicate that it acts as a barrier to dispersal. A backward stepwise variable selection procedure may be used.

Finally, the friction map is computed by assigning to each cell (x,y) a friction value $F_{x,y}$ computed as follows:

$$F_{x,y} = \sum_{i=1}^P o_{x,y,i} e^{w_i}$$

where P is the number of predictors, $o_{x,y,i}$ is 1 when the i th landscape is present in the cell and is 0 otherway, and w_i is weight assigned to the i th predictor.

Friction cannot therefore be negative.

8. Grid Converter

Grid Converter allows converting ESRI grids into Idrisi rasters, and conversely. Several files may be converted in one operation. Remember that Biomapper use the same file format as Idrisi.

WARNING! Due to ESRI proprietary politics, GridConverter does only work if you have ArcView with Spatial Analyst installed on your computer. (Tested with ArcView 3.0 and 3.1)

9. GroupStat

Given a map with cell value representing groups (e.g. obtained by *GROUP* or *BigGroup*), an optional map containing some quantitative values, *GroupStat* associates to each group a value obtained by some statistical summarising operation on all the cells of that group.

The Output may be in the form of a summary table or a new map where each cell takes the value of the whole group to which it belongs.

The available operations are:

- Area of the group.
- Sum of the values in the group
- Mean of the values in the group

Area does the same operation as Idrisi AREA operator but it can handle more categories (2'147'483'648 in place of

16'000) and can read all three types of input map.

10. Map-3D

The purpose of *Map-3D* is to provide a pleasant 3-D display of a landscape. This allows the user to better understand spatial structures. It uses a digital elevation model (DEM) to display the landscape from a customisable point of view, with shade induced by a parallel light. Another Idrisi raster may be draped over the landscape, using either a pre-existing colour palette or palette file built with Idrisi. *Map-3D* has thus the same purpose as Idrisi module *ORTHO* but is easier and faster to use.

The final image may be saved in BMP format.

11. Sampler

Sampler generates two boolean maps where cells are set to “true” or “false” following a random uniform distribution. The two maps are complementary. The user may choose the number of cells to sample and the proportion of cells to assign to each map. The purpose is to part a species observation map into a calibration and a validations set.

12. Value Extractor

This module extracts values from a map. The output file is an ASCII formatted file giving X and Y coordinates and cell value. It does the reverse operation of the module *Convertor*.

Appendix C: SIM-Ibex key procedures

This appendix presents some of SIM-Ibex key procedures. They are coded in Borland Delphi 4. Procedures were directly pasted from the original code to prevent typing errors. Comments are in French.

1. Parameter assessment

```
//===== Estimation =====
{Estimation des paramètres logistiques (r,K et C) et
des paramètres de Leslie moyens (fm,sjm,sam,svm)
et densité-dépendants (f1,f2,j1,j2,a1,a2,v1,v2)}
Function Estimation:boolean;
Const
  NRm =AnMax-1950;// Nb maximal d'années recensées cplmt
var
  intercept,pente :double;
  msg :string;
  ndata,avant :integer;
  lambda :double;
  nt,n0,n12,n3,n6,n11 :array[0..NRm] of double;
  tj,ta,tv,tt :array[0..NRm] of double;
  sj,sa,sv,f :array[0..NRm] of Double;
  nbs,nbf :integer;
  rec,cal :array[0..NRm] of Boolean;
  Nr, Nr1 :integer;
  i, j, dum :integer;
  prod,n2,n10,lp :double;
  SDf,SDsj,Sdsa,SDsv :double;
  recr, mort :double;

  Ssj,Ssj2,Ssjn :double;
  Ssa,Ssa2,Ssan :double;
  Ssv,Ssv2,Ssvn :double;
  Sn,Sn2 :double;
  Sx,Sx2,Sy,Sy2,Sxy:double;
  xx,yy :double;

  a,X,Y,sig :TvecteurR;
  covar :TmatriceR;
  lista :TvecteurI;
  khi2,ancKhi2,alamda,X1 :double;
  ma,mfit,err :integer;
  SelVide :boolean;
  HorReg :Single;
begin
  Result:=false;
  if ParamCalcules then begin
    Kcalcule:=true;
    Result:=true;
    exit;// nul besoin de recalculer
  end;
  ParamDouteux:=0;
  Kcalcule:=false;
  SelVide:=false;
  for i:=AnMin to AnMax do if Selct[i] then begin
    SelVide:=true;
    Break;
  end;//if
```

When GIS come to life

```
if not SelVide then Selct:=Cens;

//Initialisation
ndata:= 0;
Sx:=0;Sy:=0;Sx2:=0;Sy2:=0;Sxy:=0;
r:=0; K:=0; dum:=0;
Nr:=0;Nr1:=0;avant:=0;
for i:=0 to Nrm do rec[i]:=false;

//Recherche des années recensées complètement
For i:= AnMin To AnMax do begin
  If (Selct[i]) Or (Nr > 0) Then begin
    If (Selct[i]) And (avant<>0) Then begin
      Inc(Ndata);
      yy:=N[avant,NtotG]{-N[avant,5]}; //Calcul sans les cabris
      xx:=ln((N[i,NtotG]+T[avant,TtotG]-T[avant,4]-
      T[avant,8])/N[avant,NtotG])/(i-avant);
      Sx:=Sx+xx;Sy:=Sy+yy;
      Sx2:=Sx2+sqr(xx);Sy2:=Sy2+sqr(yy);
      Sxy:=Sxy+xx*yy;
    End; //If
    If Selct[i] Then avant:= i;
    If (NbDonnees(i)≥5)and Selct[i] Then begin
      Nr:= Nr1;
      If Nr>Nrm Then begin
        MessageDlg(ctIBXtropDonnees1+Txt(Nrm)+ctIBXtropDonnees2, mtInfo r-
        mation,[mbOK],0);
        Break; // For i
      end;
      n0[Nr]:=(0.5*N[i,3]);
      n12[Nr]:=N[i, 4];
      n3[Nr]:=N[i, 3] + N[i, 2] + N[i, 1] + N[i, 0];
      n6[Nr]:=N[i, 1];
      n11[Nr]:=N[i, 0];
      nt[Nr]:=n12[Nr]+n3[Nr];
      tj[Nr]:=T[i,3]+T[i,4]+T[i,7]+T[i,8];
      ta[Nr]:=T[i,0]+T[i,1]+T[i,2]+T[i,6];
      tv[Nr]:=T[i,0];
      tt[Nr]:=T[i,TtotG];
      if {n0[Nr]*}nt[Nr]<>0 then rec[Nr]:=True;
      Inc(Nr);
      Inc(Nr1);
    end
    Else begin
      rec[Nr]:=false;
      If Nr > 0 Then Inc(Nr1);
    end; //else
  End; // If
end; //for i
Dec(Nr);
If ndata=0 Then begin
  MessageDlg(ctIBXerrDonneesInsuffisantes,mtError,[mbOK,mbHelp],103);
  Exit;
end;

// Calcul de r, K
Reglin(Sx,Sy,Sx2,Sy2,Sxy,Ndata,pente,intercept,xx);
r:=RMAX;
K:=-intercept/pente;
Kcalcule:=True;
FOPTsim.Edexp.text:=Txt(Round(K*2/100)*100);
```

```

If K=0 Then FOPTsim.Edexp.text:='500';

// Calcul des taux de survie annuels
For i:= 1 To Nr do begin
  sj[i]:=0; sa[i]:=0; sv[i]:=0;
  If rec[i] and rec[i - 1] And rec[i + 1] Then begin // Nécessité d'une
    année avant et après
    sj[i]:=(-n0[i]+sqrt(n0[i]*n0[i]+4*(n12[i+1]+tj[i])*n0[i-1]))/(2*n0[i-
    1]);
    lambda:= nt[i+1]/nt[i];
    prod:= (nt[i+1] + tt[i])/nt[i];
    n2:= sj[i] * n12[i] / (sj[i] + prod);
    sa[i]:= prod*(n3[i]+ta[i])/(n2+(n3[i]+ta[i]));
    lp:= lambda / sa[i];
    n10:= n6[i] / (power(lp,4) + power(lp,3) + power(lp,2) + lp + 1);
    sv[i]:= prod * (n11[i]+tv[i]) / (n11[i]+tv[i]+n10);
  End
  else begin
    rec[i]:=false;
  end;// If rec[i]
end;// for i

// Régression linéaire pour la densité-dépendance
Ssj:=0;Ssj2:=0;Ssjn:=0; // Somme, somme des carrés, somme des produits
Ssa:=0;Ssa2:=0;Ssan:=0;
Ssv:=0;Ssv2:=0;Ssvn:=0;
Sn:=0;Sn2:=0; // Effectif total
nbs:=0;
for i:=1 to Nr-1 do begin //Calcul des sommes
  if rec[i] then begin
    Ssj:=Ssj+sj[i];
    Ssj2:=Ssj2+Sqr(sj[i]);
    Ssjn:=Ssjn+sj[i]*nt[i];
    Ssa:=Ssa+sa[i];
    Ssa2:=Ssa2+Sqr(sa[i]);
    Ssan:=Ssan+sa[i]*nt[i];
    Ssv:=Ssv+sv[i];
    Ssv2:=Ssv2+Sqr(sv[i]);
    Ssvn:=Ssvn+sv[i]*nt[i];
    Sn:=Sn+nt[i];
    Sn2:=Sn2+Sqr(nt[i]);
    Inc(nbs);
  end;//if
end;//for i
if nbs<=1 then begin
  MessageDlg(ctIBXDonneesInsuffl,mtError,[mbOK,mbHelp],103);
  Exit;
end;//if

f1:=0;
f2:=0.5;
RegLin(Ssj,Sn,Ssj2,Sn2,Ssjn,nbs,j1,j2,jr);
RegLin(Ssa,Sn,Ssa2,Sn2,Ssan,nbs,a1,a2,ar);
RegLin(Ssv,Sn,Ssv2,Sn2,Ssvn,nbs,v1,v2,vr);

// Calcul des valeurs moyennes
fm:=0.5;
sjm:=j1*Sn/nbs+j2; // Valeur pour une densité moyenne
sam:=a1*Sn/nbs+a2;
svm:=v1*Sn/nbs+v2;

```

When GIS come to life

```
{Plausibilité biologique des paramètres estimés}
ParamDouteux:= Ord(abs(j1)>0.007)+Ord(abs(a1)>0.007)+Ord(abs(v1)>0.007)+
  {Pente trop raide}
  Ord(abs(j1)>0.015)+Ord(abs(a1)>0.015)+Ord(abs(v1)>0.015)+ {Pente e n-
    core plus raide}
  Ord(abs(j1)>0.03)+Ord(abs(a1)>0.03)+Ord(abs(v1)>0.03)+ {Pente vraiment
    trop raide}
  Ord(j1>0.0001)+Ord(a1>0.0001)+Ord(v1>0.0001)+ {Pente positive}
  Ord(j2<-1)+Ord(a2<-1)+Ord(v2<-1); {Intercept négatif}
ParamDouteux:=Round(ParamDouteux/15*10)*10;
if ParamDouteux>0
then begin
  ShowMessage(Format(ctIBXParamDouteux,[ParamDouteux]));
  if fTir.Rgstrategie.ItemIndex=0 then fTir.Rgstrategie.ItemIndex:=2;
  {Stratégie OFEFP}
end;//if

{Correction pour les pentes positives}
if fStat.ckCorrPentes.Checked then begin
  if j1>0 then begin j1:=0; j2:=sjm; end;
  if a1>0 then begin a1:=0; a2:=sam; end;
  if v1>0 then begin v1:=0; v2:=svm; end;
end;//if

Result:=true;
ParamCalcules:=true;
End;
```

2. Culling plan

```
//===== Plan de Tir =====
procedure PlanDeTir(var t:TclassesdAge; ideal,accr:double;
  const t0:integer; Xtot:double; const Auto:boolean;const Moyenne:boolean);
{Renvoie dans t le plan de tir. T est vide au départ}
var
  i,tot,bas,haut      :integer;
  Atirer,total        :double;
  X                   :Tlpage;
  ideal2,fin          :integer;
  act                 :TclassesdAge;

begin
  If Auto then begin
    if Moyenne
    then PyramideMoyenne(t0,act,Xtot,i) // Etablir la situation
    else PyramideCourante(t0,act,Xtot);
    end //if
  else for i:=0 to 5 do act[i]:=t[i];      // actuelle

//Calcul du nombre d'animaux à tirer
  If accr ≤ 0 Then accr:= 1;
  ideal2:= Round(ideal * 0.9);
  bas:= Round(ideal2 * 0.8);
  haut:= Round(ideal2 * 1);
  ideal2:=Round(ideal);
  if Xtot<ideal2 then Atirer:= 0
  else if (Xtot≥bas)and(Xtot≤haut) then Atirer:=(accr-1)*Xtot
  else if Xtot>ideal2 then Atirer:=Xtot-ideal2+Round(accr-1)*Xtot;
  If Atirer<0 Then Atirer:= 0;
  If Atirer>TirMax*Xtot Then Atirer:= TirMax*Xtot;
  If Atirer=0 Then begin
    For i:=0 To 5 do t[i]:= 0;
```

```

Exit;// pas besoin de calculer plus loin
end;//if

Case Ftir.Rgstrategie.ItemIndex of
0:begin //***** Stratégie SIM-Ibex *****
// Calcul effectifs théoriques souhaités
ConstrPyramideTheorique (t,Xtot);
fin:=5;
if not fAuto.ckChasseCabris.Checked then begin
fin:=4;
total:=0;
for i:=0 to 4 do total:=total+t[i];
for i:=0 to 4 do t[i]:=t[i]/total;
t[5]:=0;
end;//if
For i:= 0 To 5 do t[i]:= t[i]*ideal2;

// Répartition des tirs
tot:= 0;
For i:= 0 To fin do begin
t[i]:=Trunc(act[i]-t[i]);
If t[i]<0 Then t[i]:= 0;
tot:= tot + Round(t[i]);
end;//for i
if tot>0 then For i:= 0 To fin do t[i]:=Round(t[i]/tot*Atirer);
end;//case 0

1:begin //***** Stratégie par taux *****
for i:=0 to 5 do t[i]:=Taux[i]*Atirer;
end;//case 1

2:begin //***** Stratégie OFEFP *****
bas:= Round(ideal2 * 0.8);
haut:= Round(ideal2 * 1.2);
if Xtot<bas then exit;
if (Xtot≥bas)and(Xtot≤haut) then Atirer:=0.13*Xtot;
if Xtot>haut then Atirer:=0.17*Xtot;
if fAuto.ckChasseCabris.Checked then begin
t[0]:=0.075*Atirer;
t[1]:=0.075*Atirer;
t[2]:=0.15*Atirer;
t[3]:=0.3*Atirer;
t[4]:=0.28*Atirer;
t[5]:=0.12*Atirer;
end
else begin
t[0]:=0.075*Atirer;
t[1]:=0.075*Atirer;
t[2]:=0.15*Atirer;
t[3]:=0.3*Atirer;
t[4]:=0.4*Atirer;
t[5]:=0;
end;
end;//case 2

end;//case of
End;

```

3. Simulations

```

procedure Tfsim.Simulations;
Label

```

```
Fin;
var
  t,i,j,R           :integer;
  dum,accr2,epsilon  :double;
  Nbexp,Nbext        :integer;
  TailleMoy,tirms,tirms Moyens,tirtot,Xtot1 :double;
  TirsReels          :double;
  Off                :byte;//classe d'âge officielle
  coul               :longint;
  comp               :array [0..2]of byte;// couleur du tracé (aléa-
    toire)
  Tc,logFNR0         :double;
  Graph,Gph          :Tbitmap;
  Nmoy,Chc,Chc2,Ntmp  :TclassesdAge;
  Nban2              :integer; //Années non prises en compte pour les
    stat. De chasse
  denom              :integer;
  Explosion           :boolean;
  //pSerie            :^TchartSeries;
begin
  /***** Initialiser les variables *****/
  TailleMoy:= 0;
  tirms Moyens:= 0;
  NbExt:=0;
  NbExp:=0;
  Chmoy:=VecteurNul;
  Nmoy:=VecteurNul;
  Chc:=VecteurNul;
  Chmoy2:=VecteurNul;
  Chc2:=VecteurNul;

  /***** Lire les options *****/
  with FOPTsim do begin
    e:=ValF(Edepsilon.text);
    Nbr:=Val(Ednbr.text);
    Nban2:=val(Edexclu.text);
    if Nbr<1 then Nbr:=1;
    if not Ckstochastique.checked then begin
      e:=0.0;
      Nbr:=1;
    end;
    tf:=Val(Edhorizon.text);
    if tf<=0 then tf:=20;
    SeuilExt:=Val(Edext.text);
    if SeuilExt<0 then SeuilExt:=0;
    SeuilExp:=Val(Edexp.text);
    if SeuilExp<=10 then SeuilExp:=10;
    ideal:=fAuto.seIdeal.Value;
    if ideal<0 then ideal:=0;
  end;//with
  {$ifdef biologiste}
  if fOptSim.ckNumerique.Checked then begin
    fResultats.mmRes.Lines.Add('Simulations');
    fResultats.mmRes.Lines.Add('*****');
    fResultats.mmRes.Lines.Add('Colonie :'+#9+Colonie);
    fResultats.mmRes.Lines.Add('Année initiale :'+#9+IntToStr(an));
    fResultats.mmRes.Lines.Add('Seuil Extinction
      :'+#9+IntToStr(Round(SeuilExt)));
    fResultats.mmRes.Lines.Add('Seuil Explosion
      :'+#9+IntToStr(Round(SeuilExp)));
  end;
```

```

fResultats.mmRes.Lines.Add('Effectif cible
    :'+#9+IntToStr(Round(ideal)));
fResultats.mmRes.Lines.Add('Nombre de répliqués :'+#9+IntToStr(Nbr));
fResultats.mmRes.Lines.Add('Stratégie de chasse
    :'+#9+Ftir.Rgstrategie.Items[Ftir.Rgstrategie.ItemIndex]);
fResultats.mmRes.Lines.Add('Seuil de chasse :'+#9+TxtF(TirMax,2));
fResultats.mmRes.Lines.Add('Stochasticité
    :'+#9+IntToStr(Ord(ckStochastique.Checked)));
fResultats.mmRes.Lines.Add('Coeff.stochasticité :'+#9+TxtF(e,2));
fResultats.mmRes.Lines.Add('Chasse
    :'+#9+IntToStr(Ord(ckChasse.Checked)));
fResultats.mmRes.Lines.Add('');
end; //if
{$endif}

//***** Préparer le graphique *****
With G do begin
    Graph:=Tbitmap.create;
    If Fsim.Cktrace.checked Then begin
        larg:= Imgraph.Width;
        haut:= Imgraph.Height;
        Graph.width:=larg;
        Graph.height:=haut;
        with Graph.Canvas do begin
            Brush.Color:=clWindow;
            FillRect(Rect(0,0,Graph.width,Graph.Height));
        end; //with
    // Fixer les bornes
        minX:= an;
        maxX:= an + tf;
        minY:= 0;
        if UpperCase(FOPTsim.EdmaxY.text)='AUTO' then
            maxY:=Round(SeuilExp/100+0.5)*100
        else maxY:= Val(FOPTsim.EdmaxY.text);
        ll:= maxX - minX; // largeur graphique en unités
        hh:= maxY - minY; // hauteur graphique en unités
        pasX:= Arr(ll/5); // Ecart entre les graduations
        pasY:= Arr(hh/5); //
        minY:= Round(Trunc(minY/pasY)*pasY);
        maxY:= Round(Trunc(hh/pasY+0.9)*pasY)+minY;

    // Structurer la fenêtre
        with Graph.canvas do begin
            mg:= TextWidth(Txt(maxY))+20; // Marge gauche en pixels
            mb:= TextHeight(Txt(maxY))+20; // Marge basse en pixels
            mh:= 30;
            md:= 40;
            eX:= (larg - mg - md)/ll; // échelles en pixels/unité
            eY:= -(haut - mb - mh)/hh; //
            DessineAxe (Graph,ctSimSimulations,ctSimAnnees,'N');
        end; //with
        Imgraph.picture.Graphic:=Graph;
        //Imgraph.refresh;
    end
    Else Fsim.Cktrace.Enabled:= False;
end; //with G

//-----
//----- SIMULATIONS -----
//-----
Tc:= TempsGeneration(ideal);

```

When GIS come to life

```
if ckChasse.Checked then begin
  logFNR0:= FNR0(f1*ideal+f2,j1*ideal+j2,a1*ideal+a2,v1*ideal+v2);
  If logFNR0 > 0 Then accr:= Log10(logFNR0)/Tc+1 Else accr:= 1;
end
else For i:= 0 To 5 do Ch[i]:= 0;
progression.minValue:=0;
Progression.MaxValue:=Nbr;
Progression.Progress:=0;
For R:= 1 To Nbr do begin// -- Boucle des réplicats ---
  q:= 0;// On commence avec la page 0, puis 1, puis 0, puis...
  p:= 1-q;// Page d'avant
  Explosion:=False; //Détecter les explosions en cours de simulation
  For i:= 0 To 20 do begin // Initialisation
    For j:= 1 To 2 do begin
      Xt[i, j, q]:= Round(X[i,j]);
      Xt[i, j, p]:= Round(X[i,j]/accr);
    end;//for j
  end;//for i
  Xtot:= Somme3(Xt,q);
  Xtot1:= Somme3(Xt,p);
  tirTot:= 0.0;
  tirsM:= 0.0;
  Progression.progress:=R;

  {Couleur du tracé}
  if Cktrace.checked then with Graph.canvas do begin
    comp[0]:= 0;
    comp[1]:= 0;
    comp[2]:= 0;
    i:= random(2);
    comp[i]:= Random(200)+55;
    coul:= RGB(comp[0],comp[1],comp[2]);
    Pen.color:=coul;
    Pen.color:=coul;
    MoveTo (CoorX(an),CoorY(Xtot));
  end;//with et if

  {-- Boucle du temps --}
  For t:= 1 To tf do begin
    {Chasse}
    If Ckchasse.checked Then begin
      logFNR0:= FNR0(f1*ideal+f2,j1*ideal+j2,a1*ideal+a2,v1*ideal+v2);
      If logFNR0 > 0
      Then accr2:= Log10(logFNR0)/Tc+1
      Else accr2:= (Xtot+tirtot)/Xtot1;
      {accr2 = (Xtot + tirtot) / Xtot1}
      Cont2Off(Xt,q,Ch);
      PlanDeTir (Ch,ideal,accr2,t0,Xtot,false,false);//établir le plan de
      tir théorique
      Xtot1:= Xtot;
    End;//if Ckchasse

    {Dynamique de la population}
    q:= 1-q;// On modifie l'autre page
    p:= 1-q;
    epsilon:= Random(1000)/1000*e*2-e;// Tiré uniformément entre -e et +e
    {Survie}
    For j:= M To F do begin
      For i:= 1 To 20 do begin
        Xt[i,j,q]:= Xt[i-1,j,p]*(s(i-1,j)+epsilon+Random(1000)/1000*e-
        e/2);
```



```

        If Xt[i,j,q]<0 Then Xt[i,j,q]:= 0;
    end; //for i
end; //for j
Xt[20,F,q]:=0;
Xt[20,M,q]:=0;
{Recrutement}
Xt[0,F,q]:=0;
Xt[0,M,q]:=0;
For i:= 1 To 20 do begin //Contr. Des F de l'an q aux naissances de
    cette même année
        Xt[0,M,q]:=Xt[0,M,q]+SRm*Xt[i,F,q]*(fe(i)+epsilon+ Ra n-
            dom(1000)/1000*e-e/2); //(gasdev*epsilon+fe(i));
        Xt[0,F,q]:=Xt[0,F,q]+SRf*Xt[i,F,q]*(fe(i)+epsilon+ Ra n-
            dom(1000)/1000*e-e/2); //(gasdev*epsilon+fe(i))
    end; //for i
    {Chasse}
    tirtot:=0;
    if Ckchasse.Checked then begin
        For j:= M To F do begin // Chasse
            For i:= 0 To 20 do begin
                TirsReels:=Fchasse(i,j);
                Xt[i,j,q]:= Xt[i,j,q]-TirsReels;
                Xt[i,j,q]:= Round(Xt[i,j,q]);
                Off:=Officielle(j,i);
                Tirtot:=Tirtot+TirsReels;
                Chc[Off]:=Chc[Off]+TirsReels; //pour la pyramide de chasse
                if t>Nban2 then Chc2[Off]:=Chc2[Off]+TirsReels; //ne pas tenir
                compte des lères années
                If Xt[i,j,q]<0 Then Xt[i,j,q]:= 0;
            end; //for i
        end; //for j
    end; //if
    TirsM:=TirsM+TirTot;

    Xtot:= Somme3(Xt,q);
    If Xtot<0 Then Xtot:= 0
    else if Xtot>seuilEXP then Explosion:=True;

{Jauges et tracés}
{$ifdef biologiste}
if fOptSim.ckNumerique.Checked then begin
    fResultats.mmRes.text:=fResultats.mmRes.text+#9+
    IntToStr(Round(Xtot));
end; //if
{$endif}
If Cktrace.checked Then begin
    with Graph.canvas do begin
        If (Xtot>=G.minY)And(Xtot<3.4E+38) Then
            LineTo (CoorX(an+t),CoorY(Xtot))
        Else If (Xtot<G.minY) Then
            LineTo (CoorX(an+t),CoorY(G.minY))
        Else
            MoveTo(CoorX(an+t+0.5),CoorY(G.minY));
    end; //with
    Imgraph.picture.graphic:=Graph;
End; //if Cktrace
Application.ProcessMessages;
If not Encours Then goto Fin;
If (Xtot<seuilExt)Or(Xtot>SeuilEXP) Then break; //for t
males:= 0;
For i:= 0 To 20 do males:= males+Xt[i,M,q];

```

When GIS come to life

```
end; //for t

If Xtot<seuilExt Then Inc(Nbext)
Else begin
  TailleMoy:= TailleMoy + Round(Xtot);
  tirms Moyens:=tirms Moyens+tirms/t; //moyenne des tirs moyens par année
end; //if
If Explosion Then Inc(Nbexp);
Lbext.caption:= TxtF(Nbext/R*100,2)+'%';
Lbexp.caption:= TxtF(Nbexp/R*100,2)+'%';
try
  Lbtaille.caption:= TxtF(TailleMoy/(R-Nbext),2);
Except // au cas où il n'y a que des extinctions
  Lbtaille.caption:='0';
end; //except
try
  Lbtirs.caption:= TxtF(tirms Moyens/(R-Nbext),2);
except //au cas où il n'y a que des extinctions
  Lbtirs.caption:='0';
end; //except
Cont2Off(Xt,q,Ntmp);
denom:=t-Nban2;
if denom<=0 then denom:=1;
for i:=0 to 5 do begin
  Chmoy[i]:=Chmoy[i]+Chc[i]/t;
  Chmoy2[i]:=Chmoy2[i]+Chc2[i]/denom;
  Nmoy[i]:=Nmoy[i]+Ntmp[i];
end; //for i
Chc:=VecteurNul;
Chc2:=VecteurNul;
{$ifdef biologiste}
if fOptSim.ckNumerique.Checked then begin
  fResultats.mmRes.Lines.Add('');
end; //if
{$endif}

end; //for R
Dec®;

Fin:
if R<=0 then R:=1;
for i:=0 to 5 do begin
  Chmoy[i]:=Chmoy[i]/R;
  Chmoy2[i]:=Chmoy2[i]/R;
  Nmoy[i]:=Nmoy[i]/R;
end; //for i
with G do begin
  larg:= Imchasse.Width;
  haut:= Imchasse.Height;
  CLS(Graph);
  Graph.width:=larg;
  Graph.height:=haut;
end;
if Ckchasse.Checked then begin
  if Rbtoutes.checked
  then Pyramide(Chmoy,ctSimPyramideDesTirs,' ',True,Graph)
  else Pyramide(Chmoy2,ctSimPyramideDesAges,' ',True,Graph);
end; //if
else begin
  CLS(Graph);
end; //if
```

```

Imchasse.picture.graphic:=Graph;
Gph:=Tbitmap.Create;
with G do begin
    larg:= Images.Width;
    haut:= Images.Height;
    CLS(Gph);
    Gph.width:=larg;
    Gph.height:=haut;
end;//with
Pyramide(Nmoy,ctSimTitrePyramide2,' ',True,Gph);
Images.picture.graphic:=Gph;
{$ifdef biologiste}
if fOptSim.ckNumerique.Checked then begin
    fResultats.mmRes.Lines.Add('-----');
    fResultats.mmRes.Lines.Add(' ');
end;//if
{$endif}

End;

```

Appendix D: Biomapper key procedures

This appendix presents some of *Biomapper* key procedures. They are coded in Borland Delphi 4. Procedures were directly pasted from the original code to prevent typing errors. Comments are in French.

1. Ecological niche factor extraction

```
{===== Matrice de Transformation Niche Ecologique =====}
function FacteursNiche(const Rg:TdblMatrix;const Moy:TdblVector;
                      const Rs:TdblMatrix;const MoyP:TdblVector;
                      const V:integer;
                      var Fact:TdblMatrix;var ValP:TdblVector):boolean;
{Calcule le facteur de marginalité et les facteurs de restriction de niche
 orthonormés
 selon la méthode Perrin-Hausser-Chessel (cf. article ENFA). La matrice
 Fact
 est la matrice de transformation permettant d'exprimer les variables selon
 les actes
 factoriels. Les facteurs sont classés par ordre de significativité d'é-
 croissante.
 Rg = matrice de corrélation globale
 Moy = vecteur moyennes global
 Rs = matrice de covariance partielle (pour l'espèce)
 MoyP = vecteur moyennes partiel
 Fact = matrice des axes factoriels
 ValP = valeurs propres correspondants aux axes factoriels }
var
  Rs2,S,W,H :TdblMatrix; // matrices de travail (v. article ENFA)
  nM          :TdblVector; // vecteur de travail
  y           :TdblVector; // vecteur de travail
  ztz         :extended; // sqrt(z'z)
  i,j         :integer;
  Trace,TraceW:extended;
begin
  {Initialisation}
  Result:=False;

  {Préparation des matrices de travail}
  Rs2:=PowerMat(Rs,V,-0.5); // S = Rs^(-1/2)
  W:=ProduitMat(ProduitMat(Rs2,Rg,V,V,V),Rs2,V,V,V); //W = S' Rg S
  TraceW:=TraceMat(W,V);

  {Extraction du facteur de marginalité de la matrice W afin que les fa c-
   teurs
   de restriction de niche lui soient orthogonaux}
  nM:=NormaliserVecteur(MoyP,V); // Normaliser M
  y:=MatxVec(Rs2,MoyP,V); // z = Rs^(-1/2) * nM
  ztz:=sqrt(ProduitVec(y,y,V));
  for i:=1 to V do y[i]:=y[i]/ztz; // y = z/sqrt(z'z)
  for i:=1 to V do begin // S = I-yy'
    for j:=1 to V do begin
      S[i,j]:=Ord(i=j)-y[i]*y[j];
    end; //for j
  end; //for i
  H:=ProduitMat(S,ProduitMat(W,S,V,V,V),V,V,V); // H = (I-yy')W(I-yy')

  {Calcul des val. et vec. propres : facteurs de restriction de niche}
  ValVecPropres(H,V,ValP,W); // Calculer Vecteurs propres de H
```

```
Fact:=ProduitMat(Rs2,W,V,V,V);

{Insertion du facteur de marginalité comme première colonne de la matrice
  Fact}
Trace:=0;
for i:=V downto 2 do begin
  ValP[i]:=ValP[i-1];
  Trace:=Trace+ValP[i];
  for j:=1 to V do Fact[j,i]:=Fact[j,i-1];
end; //for i
for i:=1 to V do Fact[i,1]:=nM[i];
ValP[1]:=TraceW-Trace;

Result:=True;
end;
```

2. Habitat suitability map computation

```
{===== Carte Probabiliste 2 =====}
function CarteProbabiliste2(const ProbName:string;const
  img:tRasterList;const V:word;
  const Espece:tRaster;
  const Covar:TdblMatrix;const Moy:TdblVector;
  const ValP:TdblVector;const Fact:TdblMatrix;const nF,nbCat:word;
  const MargWeight:single;
  const ShowDistrib:boolean):boolean;
{Calcule une carte probabiliste en se basant sur une série d'hist ogrammes
  de fréquences unidimensionnels}
var
  i,F          :integer;
  Fvar         :array[1..NbVarMax]of tRasterFile;
  Fprob,Fesp,Ftemp :tRasterFile;
  CarteProb    :tRaster;
  CarteTemp    :tRaster;
  Fic          :TextFile;
  Tmp          :array[1..NbVarMax]of single;//tampon de lecture pour
    les variables
  Obs          :single;//tampon de lecture pour le fichier "espèce"
  Sig          :array[1..NbVarMax]of double;//ecart-type
  MaxF,MinF    :array[1..NbVarMax]of double;//Extrema des facteurs, Pas
    de l'histogramme
  MedF         :array[1..NbVarMax]of double;//médianes des distribu-
    tions factorielles
  FreqS,FreqG   :array[1..NbVarMax]of tHistogram;//Distrib. de
    l'espèce selon chaque facteur
  Cat          :array[1..NbVarMax]of cardinal;//
  Poids        :TdblVector;
  PoidsTotal    :double;
  nbGlobal     :cardinal;// nombre de quadrats occupés par l'espèce
  Pt,Pm        :longint;
  Lu,LuS       :boolean;
  MaxV,MinV    :single;
  Sc,Pr,PrMax  :single;
  Back         :single;
  //.....
function Score(const F:word):single;
{Caclule la coordonnée sur l'axe factoriel traité}
var
  i :integer;
begin
  Result:=0;
```

When GIS come to life

```
    for i:=1 to V do Result:=Result+(Tmp[i]-Moy[i])/sig[i]*Fact[i,F] //Les
        données brutes doivent être centrées-réduites ?
end; //function
//.....
function Probabilite:double;
{Calcule la probabilité que le quadrat appartienne au domaine de l'espèce
 en
 utilisant la méthode des médianes (Hausser, 1995)}
var
    Prob          :int64;
    i,j,F         :integer;
    deb,fin       :integer;
begin
    Result:=0;
    for F:=1 to nF do begin
        if Cat[F]=MedF[F]
        then Prob:=nbGlobal div 2
        else begin
            Prob:=0;
            if Cat[F]>medF[F] then begin // Queue dextre
                deb:=Cat[F];
                fin:=nbCat-1;
            end
            else begin // Queue sénestre
                deb:=0;
                fin:=Cat[F];
            end; //else
            for j:=deb to fin do Prob:=Prob+FreqS[F].BarByIndex[j]; // Prob-
                abilité de la queue
            end; //else
            Result:=Result+(Prob/nbGlobal)*Poids[F]*2;
        end; //for F
    end; //function
//.....
function Probabilite2:double;
{Calcule la probabilité que le quadrat appartienne au domaine de l'espèce
 en
 comparant la distribution spécifique à la distribution globale}
var
    Prob          :extended;
    F             :integer;
begin
    Result:=0;
    for F:=1 to nF do begin
        try Prob:=FreqS[F].BarByIndex[Cat[F]]/FreqG[F].BarByIndex[Cat[F]];
        except Prob:=0; end; //except
        Result:=Result+Prob*Poids[F];
    end; //for F
end; //function
//::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
begin
    {Initialisation}
    Idrisi.TailleTampon:=OptimalBuffer(V);
    Result:=false;
    CarteProb:=tRaster.CreateCopy(img[1]);
    try
        Back:=255;
        with CarteProb do begin
            FileName:=ProbName;
            FileTitle:='Habitat Suitability map for '+Espece.Name;
            Comment.Add('Computed with Biomapper v.'+dlgBioMapperU.Version);
```

```

Comment.Text:=Comment.Text+'Ecogeographic variables: ';
for i:=1 to V do Comment.Text:=Comment.Text+(img[i].Name)+' , ';
Comment.Add(' ');
DataType:=idByte;
ValueUnits:='percent';
BackGround:=cbChecked;
BackGroundValue:=Back;
aFaire:=CarteProb.nbPixels;
Fait:=0;
end;//with
CarteTemp:=tRaster.CreateCopy(CarteProb);
CarteTemp.FileName:='~ProbMap';
CarteTemp.FileTitle:='Temporary map';
CarteTemp.DataType:=idReal;
for F:=1 to nF do begin
    MaxF[F]:=MinSingle;
    MinF[F]:=MaxSingle;
end;//for F
nbGlobal:=0;
MaxV:=-MaxSingle;
MinV:=MaxSingle;
for i:=1 to V do sig[i]:=sqrt(Covar[i,i]); // calcul des écarts-types
    (Covar est la matrice des covariances)

{Recherche des extrema selon chaque facteur}
try
    Tache:='1/4:Looking for factors extrema';
    for i:=1 to V do img[i].OpenR(Fvar[i]);
    while not Fvar[1].eof do begin
        try
            Inc(Fait);
            Lu:=True;
            for i:=1 to V do Lu:=ReadRaster(Fvar[i],Tmp[i]) and Lu ;//lire
chaque variable
            if Lu
            then begin
                for F:=1 to nF do begin
                    Sc:=Score(F);
                    if Sc>MaxF[F] then MaxF[F]:=Sc;
                    if Sc<MinF[F] then MinF[F]:=Sc;
                end;//for F
            end;//if
        except
            On EInvalidOp do begin
                MessageDlg('Bad value encountered while looking for e x-
trema.',mtError,[mbOK],0);
                raise;
                exit;
            end;//on
        else begin
            MessageDlg('An error happened while looking for e x-
trema.',mtError,[mbOK],0);
            raise;
            exit;
        end;//else
    end;//except
    Application.ProcessMessages;
    if not CalculEnCours then exit;
end;//while
finally
    for i:=1 to V do CloseRaster(Fvar[i]);

```

When GIS come to life

```
end;//finally
for F:=1 to nF do begin
  FreqG[F]:=tHistogram.Create(nbCat,MinF[F],MaxF[F]);
  FreqG[F].Title:='Global distr. of factor '+IntToStr(F);
  FreqS[F]:=tHistogram.Create(nbCat,MinF[F],MaxF[F]);
  FreqS[F].Title:='Species distr. of factor '+IntToStr(F);
end;//for

{Calcul de la distribution de l'espèce selon chaque facteur}
try
  Tache:='2/4:Computing species distribution';
  Fait:=0;
  Application.ProcessMessages;
  for i:=1 to V do img[i].OpenR(Fvar[i]);
  Espece.OpenR(Fesp);
  while not Fvar[1].eof do begin
    try
      Inc(Fait);
      Lu:=True;
      for i:=1 to V do Lu:=ReadRaster(Fvar[i],Tmp[i]) and Lu ;//lire
chaque variable
      LuS:=ReadRaster(Fesp,Obs)and(Obs<>0);
      if Lu
      then begin
        for F:=1 to nF do begin
          Sc:=Score(F);
          FreqG[F].Add(Sc);//Calcul de la fréquence globale
          if LuS then FreqS[F].Add(Sc);//Calcul de la fréquence spéc i-
frique
        end;//for F
      end;//if
    except
      On EInvalidOp do begin
        MessageDlg('Bad value encountered while computing distrib u-
tion',
          mtError,[mbOK],0);
        raise;
        exit;
      end;//on
    else begin
      MessageDlg('An error happened while computing distribution',
        mtError,[mbOK],0);
        raise;
        exit;
      end;//else
    end;//except
    Application.ProcessMessages;
    if not CalculEnCours then exit;
  end;//while
finally
  for i:=1 to V do CloseRaster(Fvar[i]);
  CloseRaster(Fesp);
end;//finally
nbGlobal:=FreqS[1].N;

{Calcul des médianes}
for F:=1 to nF do medF[F]:=FreqS[F].ValueToBar(FreqS[F].Median);

{Calcul de la pondération des facteurs}
PoidsTotal:=0;
for F:=1 to nF do PoidsTotal:=PoidsTotal+ValP[F];
```



```

for F:=1 to nF do Poids[F]:=ValP[F]/PoidsTotal;//1/F;
//:=ValP[F]/PoidsTotal;
Poids[1]:=Poids[1]+MargWeight;

{Calcul de la carte probabiliste}
try
  PrMax:=-MaxSingle;
  Tache:='3/4:Computing probability map';
  Fait:=0;
  Application.ProcessMessages;
  for i:=1 to V do img[i].OpenR(Fvar[i]);
  CarteTemp.OpenW(Ftemp);//dans un fichier temporaire pour pouvoir la
  standardiser ultérieurement
  while not Fvar[1].eof do begin
    try
      Inc(Fait);
      Lu:=True;
      for i:=1 to V do Lu:=ReadRaster(Fvar[i],Tmp[i]) and Lu ;//lire
      chaque variable
      if Lu then begin
        for F:=1 to nF do begin
          Sc:=Score(F);
          Cat[F]:=FreqS[F].ValueToBar(Sc);//Catégorie à laquelle appa r-
          tient ce quadrat
        end;//for F
        Pr:=Probabilite;
        if Pr>PrMax then PrMax:=Pr;
        WriteIdrisi(Ftemp,Pr);
      end//if
      else WriteIdrisi(Ftemp,Back);
    except
      On EInvalidOp do begin
        MessageDlg('Bad value encountered while computing probability
        map',
          mtError,[mbOK],0);
        raise;
        exit;
      end;//on
    else begin
      MessageDlg('An error happened while computing probabi lity map',
        mtError,[mbOK],0);
        raise;
        exit;
      end;//else
    end;//except
    Application.ProcessMessages;
    if not CalculEnCours then exit;
  end;//while
finally
  for i:=1 to V do CloseRaster(Fvar[i]);
  CloseRaster(Ftemp);
end;//finally

{Standardisation de la carte probabiliste, pour que les prob. soient
entre 0 et 1}
try
  Tache:='4/4:Standardizing probability map';
  Fait:=0;
  Application.ProcessMessages;
  CarteTemp.OpenR(Ftemp);
  CarteProb.OpenW(Fprob);

```

When GIS come to life

```
while not Ftemp.eof do begin
  try
    Inc(Fait);
    if ReadRaster(Ftemp,Pr)then begin
      WriteIdrisi(Fprob,Round(Pr/PrMax*100));
      if Pr/PrMax<MinV then MinV:=Pr/PrMax;
      end
    else WriteIdrisi(Fprob,Back);
  except
    On EInvalidOp do begin
      MessageDlg('Bad value encountered while standardizing probability map',
        mtError,[mbOK],0);
      exit;
      end;//on
    else begin
      MessageDlg('An error happened while standardizing probability map',
        mtError,[mbOK],0);
      raise;
      exit;
      end;//else
    end;//except
  Application.ProcessMessages;
  if not CalculEnCours then exit;
end;//while
finally
  CloseRaster(Fprob);
  CloseRaster(Ftemp);
  DeleteFile(CarteTemp.FileName+ImgExt);// détruire le fichier temporaire
end;//finally

{Création du fichier doc}
CarteProb.MaxValue:=100;
CarteProb.MinValue:=Round(MinV*100);
CarteProb.Lineage.Clear;
CarteProb.Lineage.Add('Suitability map');
try
  if not CopyFile(img[1].FileName+DocExt,CarteProb.FileName+DocExt)
  then begin
    DeleteFile(CarteProb.FileName+DocExt);
    CopyFile(img[1].FileName+DocExt,CarteProb.FileName+DocExt);
  end;//if
  CarteProb.WriteDoc;
except
  MessageDlg('An error happened while creating '+CarteProb.FileName+DocExt,
    mtError,[mbOK],0);
  raise;
  exit;
end;//except
DeleteFile(CarteTemp.FileName+ImgExt);// détruire le fichier temporaire
Result:=true;

{Montrer, le cas échéant, de la distribution de chaque facteur}
if ShowDistrib then begin
  for F:=nF downto 1 do begin
    FreqG[F].GraphCreate('Factor '+IntToStr(F),'Distribution of factor '+IntToStr(F),
      'Factor '+IntToStr(F),'Frequency');
```

```
        FreqG[F].GraphAdd(FreqS[F]);
        FreqG[F].GraphShow;
        FreqG[F].Free;
        FreqS[F].Free;
    end; //for
end; //if
Finally
    CarteProb.Free;
    CarteTemp.Free;
end; //finally
end;
```

Appendix E: HexaSpace key procedures

This appendix presents some of *HexaSpace* key procedures. They are coded in Borland Delphi 4. Procedures were directly pasted from the original code to prevent typing errors. Comments are in English.

1. Simulation main procedure

```
procedure tHexAutomaton.Run(const AnbCycles: integer);
{----- Run -----}
var
  Cycle, Step, R           :integer;
  nbSteps1, nbSteps2       :integer;
  Transition1               :tTransitionFunction;
  Transition2               :tTransitionFunction;
begin
  {Initialization}
  if not RunIsInitialized then InitializeRun;
  Application.ProcessMessages;

  if Timing.DispersalFirst then begin
    nbSteps1:=Timing.Dispersal;
    nbSteps2:=Timing.Reproduction;
    Transition1:=DispersalTransition;
    Transition2:=ReproductionTransition;
  end//if
  else begin
    nbSteps2:=Timing.Dispersal;
    nbSteps1:=Timing.Reproduction;
    Transition2:=DispersalTransition;
    Transition1:=ReproductionTransition;
  end//else
  with Dynamics do if Stochastic and (RangeSpatial>0)
  then InitializeSpatialStochasticity;

  {Runs}

  try
    for Cycle:=1 to AnbCycles do begin
      for Step:=1 to nbSteps1 do begin
        for R:=0 to (Runs.RunCount-1)*Ord(Dynamics.Stochastic) do begin
          CurrentRun:=R;
          Transition1;
        end//for
        t0:=1-t0;
        t1:=1-t0;
        Application.ProcessMessages;
      end//for
      for Step:=1 to nbSteps2 do begin
        for R:=0 to (Runs.RunCount-1)*Ord(Dynamics.Stochastic) do begin
          CurrentRun:=R;
          Transition2;
        end//for
        t0:=1-t0;
        t1:=1-t0;
        Application.ProcessMessages;
      end//for

      if not Working then exit;
    end;
  end;
end;
```

```

        Inc(WorkDone);
    end; //for
finally
    fRan:=nil;
end; //finally
end;

```

2. Reproduction phase

```

procedure tHexAutomaton.ReproductionTransition;
{----- Reproduction Function -----}
{ Transition function performed during the reproduction steps}
var
    Hex :tHex;
    N0 :integer;
    r,e :double;
begin
    if not Dynamics.Stochastic then begin {== Without stochasticity ==}
        Hex:=FirstHex;
        while Hex<>nil do begin
            with Hex do begin
                try
                    N0:=Nsim[t0];
                    r:=dynamicsFunction(N0/K);
                    Nsim[t1]:=N0+RNDRound(N0*r);
                except
                    Nsim[t1]:=0; // Lorsque K=0
                end; //except
                if Nsim[t1]<0 then Nsim[t1]:=0;
                Hex:=NextHex;
            end; //with
        end; //while
    end //if

    else if Dynamics.RangeSpatial=0 then begin {== Without spatial autocorrelation ==}
        Hex:=FirstHex;
        while Hex<>nil do begin
            with Hex do begin
                try
                    N0:=Nsim[t0];
                    r:=dynamicsFunction(N0/K)+RNDgauss(0,Dynamics.ee+Dynamics.ed/N0);
                    Nsim[t1]:=N0+RNDRound(N0*r);
                except
                    Nsim[t1]:=0; // Lorsque K=0
                end; //except
                if Nsim[t1]<0 then Nsim[t1]:=0;
                Hex:=NextHex;
            end; //with
        end; //while
    end //if

    else begin {== With spatial autocorrelation ==}
        RNDmap(Rows,Cols,0,Dynamics.ee,Dynamics.MaxSpatial, Dynamics.RangeSpatial,fRan);
        Hex:=FirstHex;
        while Hex<>nil do begin
            with Hex do begin
                try
                    N0:=Nsim[t0];
                    r:=dynamicsFunction(N0/K);
                    r:=r+fRan[Hex.Row,Hex.Col]+RNDgauss(0,Dynamics.ed/N0);

```

```
        Nsim[t1]:=N0+RNDRound(N0*r);
    except
        Nsim[t1]:=0; // Lorsque K=0
    end;//except
    if Nsim[t1]<0 then Nsim[t1]:=0;
    Hex:=NextHex;
end;//with
end;//while
end;//else

end;
```

3. Dispersal phase

```
procedure tHexAutomaton.DispersalTransition;
{----- Dispersal Function -----}
{ Transition function preformed for the dispersal steps}
var
    Hex          :tHex;
    i,x          :integer;
    xx           :double;
    nbDisp       :integer;
    Disp         :integer;
    dDisp        :double;
    Neigh:tHex;
    Frac :array[0..5]of double;
    nFrac:double;
    Sortant,Entrant :integer;
    m            :double; //Dispersal rate
begin
    Hex:=FirstHex;
    while Hex<>nil do begin
        with Hex do begin
            Nsim[t1]:=Nsim[t0];
            Hex:=NextHex;
        end;//with
    end;//while

try
    Hex:=FirstHex;
    while Hex<>nil do begin
        with Hex do begin
            try
                m:=RNDgauss(DispersalFunction(Hex.Nsim[t0]/Hex.K), Dispersal.Stochasticity);
            except
                m:=RNDgauss(DispersalFunction(5),Dispersal.Stochasticity);
            end;//except
            if (Hex.Nsim[t0]<0)or(m<0)
            then begin
                nbDisp:=0;
                Hex:=NextHex;
                continue;
            end//then
            else nbDisp:=RNDRound(m*Hex.Nsim[t0]);

            if nbDisp>0 then begin
                try
                    nFrac:=0;
                    for i:=0 to 5 do begin {"Integer" dispersers}
                        Frac[i]:=-1;
                        Neigh:=Hex.Neighbour(i);
```

```

        if (Neigh<>nil) and (Hex.Frictions[i]<Dispersal.Threshold) then
begin
    dDisp:=(1-Hex.Frictions[i]/Hex.SumFrictions)/(Hex.NeighCount-
1)*nbDisp; // Number of dispersers in this direction
    Disp:=Trunc(dDisp);
    nFrac:=nFrac+dDisp-Disp;
    Frac[i]:=nFrac;
    Hex.Nsim[t1]:=Hex.Nsim[t1]-Disp; // Departure of di spersers
    Disp:=RNDRound(Disp*(1-Hex.d[i])); // Dispersers mo rtality
    Neigh.Nsim[t1]:=Neigh.Nsim[t1]+Disp; // Arrival of di spersers
end//if
end;//for i
for i:=1 to Round(nFrac) do begin {"Fractional" dispersers}
    xx:=RND*nFrac;
    for x:=0 to 4 do if (xx<Frac[x])then break;
    Neigh:=Hex.Neighbour(x);
    Hex.Nsim[t1]:=Hex.Nsim[t1]-1; // Departure of the dispe rser
    Disp:=RNDRound(1*(1-Hex.d[x])); // Disperser risk of death
    Neigh.Nsim[t1]:=Neigh.Nsim[t1]+Disp; // Arrival of the dispe r-
ser
end;//for i
except
    // Cas où Hex.K=0
end;
end;//if
Hex:=NextHex;
end;//with
end;//while
finally
end;//finally
end;

```

4. Hex suitability computing

```

procedure tHexAutomaton.BuildSuitability(const HSMaP: tRaster);
{----- Build Suitability -----}
    Assigns to each hex its suitability value, computed as the average of
    all the
    included cells of the HSMaP. The values of the HSMaP must be comprised
    between 0 and 100}
var
    x,y :integer;
    r,c :integer;
    i :integer;
    F :tRasterFile;
    Map :array of array of single;
    Mean :Extended;
    NoData :boolean;
    Prec :tHex;
    Temp :tHex;
begin
    if (HSMaP.Columns<>Cols)or(HSMaP.Rows<>Rows)
    then CreateLattice(Geometry.OutRadius,HSMaP.Rows,HSMaP.Columns);
    Temp:=nil;

    SetLength(Map,HSMaP.Rows,HSMaP.Columns);
    try
        HSMaP.OpenR(F);
        try
            for y:=0 to HSMaP.Rows-1 do begin
                for x:=0 to HSMaP.Columns-1 do begin
                    ReadRaster(F,Map[y,x]);

```

When GIS come to life

```
        end; //for
    end; //for
finally
    CloseRaster(F);
end; //finally
with Metadata do begin
    MinX:=HSmap.minX;
    MaxX:=HSmap.maxX;
    MinY:=HSmap.minY;
    MaxY:=HSmap.maxY;
    ModifiedDate:=DateTimeToStr(Now);
end; //with

WorkToDo:=Rows*Cols;
Temp:=tHex.Create(Self); // Sert uniquement à éviter l'AV produite lors
    de la première itération
fHexCount:=0;
Prec:=Temp;
MinHS:=MaxSingle;
MaxHS:=MinSingle;
for r:=0 to Rows-1 do begin
    for c:=0 to Cols-1 do begin
        x:=CentreX(r,c);
        y:=CentreY(r,c);
        Mean:=0;
        NoData:=False;
        for i:=0 to Geometry.Area-1 do begin
            if (HSmap.BackGround=cbChecked)and
                (Map[y+HexWindow[i].y,x+HexWindow[i].x]=HSmap.BackGroundValue)
            then begin // Incomplete hex
                NoData:=True;
                Break;
            end; //if
            Mean:=Mean+Map[y+HexWindow[i].y,x+HexWindow[i].x];
        end; //for i
        Mean:=Mean/Geometry.Area;
        if not NoData then begin
            Inc(fHexCount);
            Hexs[r,c]:=tHex.Create(Self);
            try Prec.NextHex:=Hexs[r,c] except end;
            Prec:=Hexs[r,c];
            with Hexs[r,c] do begin
                HS:=Mean; // Geometry.Area;
                if HS>MaxHS then MaxHS:=HS;
                if HS<MinHS then MinHS:=HS;
                Row:=r;
                Col:=c;
                NextHex:=nil;
            end; //with
        end; //if
        else Hexs[r,c]:=nil;
        Inc(WorkDone);
        Application.ProcessMessages;
        if not Working then exit;
    end; //for
end; //for

FirstHex:=nil;
for r:=0 to Rows-1 do begin
    for c:=0 to Cols-1 do begin
        if Hexs[r,c]<>nil then begin
```



```

        FirstHex:=Hexs[r,c];
        break;
    end;//if
end;//for
Inc(WorkDone);
Application.ProcessMessages;
if FirstHex<>nil then break;
end;//for
Self.HSMap:=HSmap.Filename;
SuitabilityIsComputed:=True;
FrictionsAreComputed:=False;
DynamicsIsComputed:=False;
InitialIsComputed:=False;
RegionsAreComputed:=False;
finally
    Map:=nil;
    if Temp<>nil then Temp.Destroy;
end;//finally
end;

```

5. Hex impermeabilities computing

(N.B.: In the early phases of HexaSpace development, impermeabilities were named “frictions”. This is still the case in the code)

```

procedure tHexAutomaton.BuildFrictions(const FrictionMap:
    tRaster;nbReplicates:integer);
{----- Build Frictions -----}
{ Assigns to each wall of a hex a crossing difficulty. This is computed by
  computing the least-cost path between the centres of the two involved
  hexs.}
var
    x,y          :integer;
    x0,y0,x1,y1  :integer;
    i            :integer;
    F            :tRasterFile;
    Map          :tFrictionMap;
    Dummy        :tPath;
    Hex,Hex1     :tHex;
    Length       :longint;
    Cost         :extended;
begin
    SetLength(Map,FrictionMap.Columns,FrictionMap.Rows);
    try
        FrictionMap.OpenR(F);
        try
            for y:=0 to FrictionMap.Rows-1 do begin
                for x:=0 to FrictionMap.Columns-1 do begin
                    ReadRaster(F,Map[x,y]);
                end;//for
            end;//for
        finally
            CloseRaster(F);
        end;//finally
        Maxfriction:=MinExtended;
        MinFriction:=MaxExtended;
        Self.FrictionMap:=FrictionMap.FileName;

        WorkToDo:=HexCount;
        Hex:=FirstHex;
        while Hex<>nil do begin

```

When GIS come to life

```
x0:=CentreX(Hex.Row,Hex.Col);
y0:=CentreY(Hex.Row,Hex.Col);
for i:=0 to 3 do begin
  Hex1:=Hex.Neighbour(i);
  if Hex1<>nil then begin
    x1:=CentreX(Hex1.Row,Hex1.Col);
    y1:=CentreY(Hex1.Row,Hex1.Col);
    LeastCostPath(x0,y0,x1,y1,nbReplicates,Map,Dummy,
    Length,Cost,False);
    Hex.Frictions[i]:=Cost;
    if Cost>Maxfriction then Maxfriction:=Cost;
    if Cost<MinFriction then MinFriction:=Cost;
  end; //if
end; //for i
Hex:=Hex.NextHex;
Inc(WorkDone);
Application.ProcessMessages;
if not Working then exit;
end; //while

finally
  Map:=nil;
end; //finally
FrictionsAreComputed:=True;
DynamicsIsComputed:=False;
InitialIsComputed:=False;
end;
```