N° d'ordre: 120-2008 Année 2008

THÈSE
présentée devant l'Université Claude Bernard — Lyon I

pour l'obtention du
Diplôme de Doctorat
(arrêté du 7 août 2006)

Présentée et soutenue publiquement le 7 juillet 2008
par
Mathieu BASILLE

Habitat selection by lynx (*Lynx lynx*) in a human-dominated landscape — From theory to application —

Jury

M. Dominique ALLAINÉ
M. Reidar ANDERSEN
M. Jean-Michel GAILLARD
M. Éric MARBOUTIN
Mme Evelyn MERRILL
M. Eloy REVILLA
M. Nigel Gilles YOCCOZ

UMR CNRS 5558
Laboratoire de Biométrie et Biologie Évolutive
Université Claude Bernard – Lyon I – Bâtiment G. Mendel
43 boulevard du 11 novembre 1918
69622 Villeurbanne FRANCE

[Mathieu Basille, "Lynx habitat selection in a human dominated landscape. From theory to application.", Université Claude Bernard Lyon I, PhD thesis]
Copyright © Mathieu Basille XX/05/2008
Copyleft : This work is free*, you can copy, spread, and modify it under the terms of the Free Art License.
http://artlibre.org/licence/lal/en/

* At the notable exception of the chapters already published, or going to be published, in scientific journals.

[Mathieu Basille, « Sélection d'habitat du lynx dans un paysage dominé par l'homme. De la théorie à l'application. », Université Claude Bernard Lyon I, Thèse de doctorat]
Copyright © Mathieu Basille XX/05/2008
Copyleft : cette oeuvre est libre*, vous pouvez la redistribuer et/ou la modifier selon les termes de la Licence Art Libre.
http://www.artlibre.org/licence/lal/

* À l'exception notable des chapitres déjà publiés, ou destinés à être publiés, dans des revues scientifiques.

UNIVERSITÉ CLAUDE BERNARD — LYON I

Président de l'Université	**M. le Professeur L. COLLET**
vice-président du conseil scientifique	M. le Professeur J.F. MORNEX
Vice-président du Conseil d'Administration	M. le Professeur J. LIETO
Vice-président du Conseil des Etudes et de la Vie Universitaire	M. le Professeur D. SIMON
Secrétaire Général	M. G. GAY

SECTEUR SANTÉ

Composantes

UFR de Médecine Lyon R.T.H. Laënnec	Directeur : M. le Professeur P. COCHAT
UFR de Médecine Lyon Grange-Blanche	Directeur : M. le Professeur X. MARTIN
UFR de Médecine Lyon-Nord	Directeur : M. le Professeur J. ÉTIENNE
UFR de Médecine Lyon-Sud	Directeur : M. le Professeur F.N. GILLY
UFR d'Odontologie	Directeur : M. O. Robin
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : M. le Professeur F. LOCHER
Institut Techniques de Réadaptation	Directeur : M. le Professeur MATILLON
Département de Formation et Centre de Recherche en Biologie Humaine	Directeur : M. le Professeur P. FARGE

SECTEUR SCIENCES

Composantes

UFR de Physique	Directeur : Mme. le Professeur S. FLECK
UFR de Biologie	Directeur : M. le Professeur H. PINON
UFR de Mécanique	Directeur : M. le Professeur H. BEN HADID
UFR de Génie Électrique et des Procédés	Directeur : M. le Professeur G. CLERC
UFR Sciences de la Terre	Directeur : M. le Professeur P. HANTZPERGUE
UFR de Mathématiques	Directeur : M. le Professeur A. GOLDMAN
UFR d'Informatique	Directeur : M. le Professeur S. AKKOUCHE
UFR de Chimie Biochimie	Directeur : Mme. le Professeur H. PARROT
UFR STAPS	Directeur : M. C. COLLIGNON
Observatoire de Lyon	Directeur : M. le Professeur R. BACON
Institut des Sciences et des Techniques de l'Ingénieur de Lyon	Directeur : M. le Professeur J. LIETO
IUT A	Directeur : M. le Professeur M. C. COULET
IUT B	Directeur : M. le Professeur R. LAMARTINE
Institut de Science Financière et d'Assurances	Directeur : M. le Professeur J.C. AUGROS

Résumé

Ce travail s'intéresse à la sélection d'habitat à différentes échelles chez le lynx (*Lynx lynx*) au sud de la Norvège. Nous avons d'abord précisé le contexte théorique, pour décrire l'habitat et la niche d'une espèce. Nous avons montré que la sélection d'habitat s'exprimait à l'échelle où les variables d'habitat sont les plus variables. Nous avons dévelopé un cadre d'analyse généralisé appelé GNESFA (englobant trois méthodes complémentaires, l'ENFA, la MADIFA et la FANTER) pour effectuer une description extensive de l'habitat d'une espèce. Dans le cas du lynx, nous avons montré l'existence d'un compromis à large échelle entre la recherche de nourriture et l'évitement des risques de mortalité. Nous avons établi le rôle limitant du lynx sur sa proie principale le long d'un gradient de productivité. Ces travaux ouvrent la voie à une analyse approfondie à échelle plus fine.

Mots-clés : analyses de niche ; *Caproelus capreolus* ; chevreuil ; habitat ; lynx ; *Lynx lynx* ; Norvège ; sélection d'habitat

Abstract

This work aims at studying habitat selection at different scales on lynx (*Lynx lynx*) in southern Norway. We first provided the general theoretical context, in order to describe a species' habitat and niche. We demonstrated that habitat selection occurs at the scale where habitat variable show the highest variability. We developped a general niche-environment framework called GNESFA (which encompasses three complementary methods, the ENFA, the MADIFA and the FANTER) to perform an extensive description of a species' habitat. In the case of lynx, we showed the existence of a trade-off at large scale between abundance of prey and avoidance of human activity. We demonstrated the limiting role of lynx on their favored prey along a gradient of productivity. This research shows the way towards a more thorough fine-scaled analysis.

Keywords: *Capreolus capreolus*; habitat; habitat selection; lynx; *Lynx lynx*; niche analyses; Norway; roe deer

Contents

List of Figures

List of Tables

“Calvin and Hobbes”, by Bill Watterson, *The Days Are Just Packed*, 1992.

Preamble

"An expert has also been defined as someone who knows more and more about less and less until finally he or she knows everything about nothing."
Henry MINTZBERG.

AS ERNST MAYR ONCE SAID IN A LETTER to his friend and colleague William B. Provine in 1979 (Provine, 2005), *"a scientist's achievement may lie in many different areas: As an innovator (new discoveries, new theories, new concepts), as a synthesiser (bringing together scattered information, sharing relationships and interactions, particularly between different disciplines, like genetics and taxonomy), as a disseminator (presenting specialized information and theory in such a way that it becomes accessible to non-specialists [popularizer is a misleading term]), as a compiler or cataloguer, as an analyst (dissecting complex issues, clarifying matters by suggesting new terminologies, etc.), and in other ways"* (his emphasis). I've met many great innovators, disseminators, analysts, etc., in my short but rich scientific life. However, in my opinion, Ernst Mayr forgot one type, maybe less conspicuous, for his list to be complete: the catalyst.

The catalyst is an agent that is able to increase the rate of a chemical reaction or biological process[1], without being consumed in the reaction. Even better, the catalyst may undergo several chemical transformations during the reaction, but at the conclusion of the reaction, the catalyst is regenerated unchanged. From my point of view, the regeneration is however not precisely identical. In the scientific arena, the catalyst takes advantage of his human nature, and may improve his own experience based on the reaction process as much as the products themselves. He may thus regenerate improved in many ways. I would certainly wish to be seen as such an effective catalyst. Here is the product of a highly unstable reaction: a PhD thesis.

This thesis is actually the result of many questions, some are mine,

---

[1] And I believe that research implies both.

others are from great minds and were only adopted. I guess that the way one answers questions highly depends on the people you meet, the culture you share, the experience you've had. There are thus several possible readings for this text. One possibility would be, just as Calvin, to question everything about it: "Whence does it come? What is it? Where is it going?" This directly leads us to question the very causality of this work, following Tinbergen's four questions (Tinbergen, 1963), all asking the same one at different levels: "Why this thesis?" According to Tinbergen, we can answer to this question with different points of view, from evolutionary explanations (ultimate function and phylogeny), to proximate explanations (ontogeny and causation).

Ultimate function

This thesis originated in the beginning of the year 2005, during a meeting with Reidar ANDERSEN, John LINNELL and John ODDEN, three *disseminators* from Norway. These people were working on lynx (*Lynx lynx*, see Fig. 1), and were disseminating huge knowledge about their object of study. But they were still wondering —this is the first quality of a scientist, ask oneself many questions; the ultimate goal is not to answer them, it is the process that is valuable. After all, we're labeled "researchers" and not "finders" with reason— they were wondering, as I was saying, about the exact place of lynx within their ecosystem. They were wondering about lynx predation in a human dominated landscape, about their role in the threefold system lynx–roe deer–people. That's where it all started, and from that very moment, I knew it'll be my primary activity, and my hobby, for the coming years.

Phylogeny

As a matter of fact, this is not exactly true: Things started long before the thesis itself. Maybe while I was a little kid, knowing nothing about science and ecology, but yet fascinated by big cats. It lasted long before I could specialized on ecology and direct my efforts towards these fascinating animals. Many years later, I went at the University of Lyon, to learn about ecology. I was helped in that mission by Sébastien DEVILLARD, who took me in charge during my Master, and initiated me into the study of real cats.

Figure 1: A lynx (*Lynx lynx*) in Norway. Photo credit: Scandlynx.

They were real, feral, cats! He latter introduced me to Dominique Pontier and Jean-Michel Gaillard, both *synthesisers* in the team "Écologie Évolutive des Populations" (*Population Evolutive Ecology*), the latter being at the same time an incredible bibliographic *compiler*. They accepted to supervise me for my Master work, together with Éric Marboutin from the French wildlife services (Office national de la chasse et de la faune sauvage), on the exploration of habitat selection of French lynx. Big cats! Finally, the realization of this kid's early dream. To be (almost) complete about this thesis phylogeny, I should also mention Daniel Chessel, an old-fashioned *analyst*, but still always so accurate, whose teaching and discussions on statistics greatly influenced me and trained me in the use of multivariate analyses.

Ontogenesis

The development of this thesis was not exactly linear. The ultimate function of this work urged me to focus first on two complementary issues : "What are we modelling?", and "How do we model it?". These two prerequisites problems were tackled with the cooperation of two persons, two

Name	Description
Ubuntu	A free Operating System, based on Debian
R	Statistical analyses; enhanced in many ways by numerous packages
Emacs	Everything but a text editor…
ESS	Emacs Speaks Statistics!
GRASS	GIS processing and analyses
QGIS	GIS displaying and interface to GRASS
LaTeX	Document processing
Inkscape	Vector graphics editing
Gimp	Image editing
Firefox	Web-browser, vital for any research!
Thunderbird	E-mail client
Sunbird	A calendar: very useful to remind dates…

Table 1: Some free software in the scientist's toolbox.

innovators of the century to come. Bram VAN MOORTER, fond of theory and simulations, helped me to build both ~~my character~~ my mind and the first part of this work on habitat selection theory. He owns at least half of this work. On the other hand, Clément CALENGE was the leader on niche analyses in the second part of this manuscript. Our exchanges were very rich and constructive, and even if I'm still unable to invert a matrix in my head, I've learned a lot through contact with him, and I hope the reverse is true.

Every person previously cited was involved in this thesis' development, in many different ways. This work, as a result of all these interactions, is partly their as well. They were indeed all reactants of the chemical reaction (see for example Appendix B).

Proximate causation

In the end, I could not have accomplished my work without any tools. These were only free software[2], and I'm proud I was able to never use any proprietary, closed-source software (at least 99.9% of the time). They were useful (and actually used) at every step of my work, from data processing to report writing, through communication or diagram editing (see Table 1). A handful of incredibly powerful software is now available to the scientist (Tufto and Cavallini, 2005), and I'm convinced they will play a major role

---

[2] *Free* as in free speech, not as in free beer…

in the future, as they are more and more integrated in their development (e.g. the use of R within Emacs through ESS, or the concurrent use of GRASS and R within QGIS, etc.). Maybe even more important than the tool is the format itself. While it is still difficult to avoid the classical (yet proprietary) Word document (.doc), I tried as much as possible to rely on open and standardized document formats. This thesis itself is made with a customed LaTeX class, and delivered (purely symbolically since it does not cover the chapters already published, or going to be published, in scientific journals) as a free document: anyone can copy, spread, and modify it under the terms of the Free Art License[3].

Altogether, I met many amazing people during this thesis, *innovators*, *synthesisers*, *disseminators*, *compilers*, *cataloguers*, *analysts*, etc. They were not all included in the thesis reaction (thus not cited in these pages), but all of them did have an impact on the catalyse anyway. I believe it was a great opportunity to work with everyone of them.

Looking back to the history of this thesis, it is quite funny to notice that I was long fascinated by big cats in Africa, especially panthers, but in some sort of great irony, I ended up (well, it's not an *end*, rather a beginning) high North, studying Norwegian lynx instead. After almost four years of work in my ivory tower (Bunnefeld et al., 2007), many questions remained unanswered, probably even more emerged, but I hope that I contributed a little to science in this way. *Well, I don't know about anyone else, but I came from my room, I'm a kid with big plans, and I'm going outside! See you later!*

---

[3] http://artlibre.org/licence/lal/en/

Part I

Habitat selection theory

Chapter 1

Habitat selection theory: from species to individuals

> *"Whether you can observe a thing or not depends on the theory which you use. It is the theory which decides what can be observed."*
>
> Albert EINSTEIN.

THE STUDY OF A SPECIES' HABITAT is fundamental to understand and to explain the ecology of a species, as it expresses the connection between the species, its environment, and its performance in this context (Whittaker et al., 1973). This in turn has tremendous effects on large parts of biology, like evolution (speciation: Rosenzweig, 1978, adaptation: Davies et al., 2007), demography (population ecology: Morris, 1987), synecology (species diversity: Hutchinson, 1959), etc. Despite this central place of what might be called *habitat ecology*, it is surprisingly deficient in theory (and theories) and lacks standard definitions.

This may be explained by the relative youth of this field of ecology. The subject became really topical at the beginning of the 80s (Johnson, 1980; Rosenzweig, 1981), and since then is experiencing an exponential growth before reaching its current importance (see Fig. 1.1). It is interesting to note that in one of the first occurrences of the term "*habitat selection*" (we can trace it up to 1933), Thorpe (1945) already recognized the evolutionary dimension of the subject, as an isolational mechanism of speciation. However, the second birth of habitat selection studies, that seemed to happen in 1978, coincided with an increasing interest in species distribution (Guisan and Thuiller, 2005), and since then the evolutionary significance of habitat

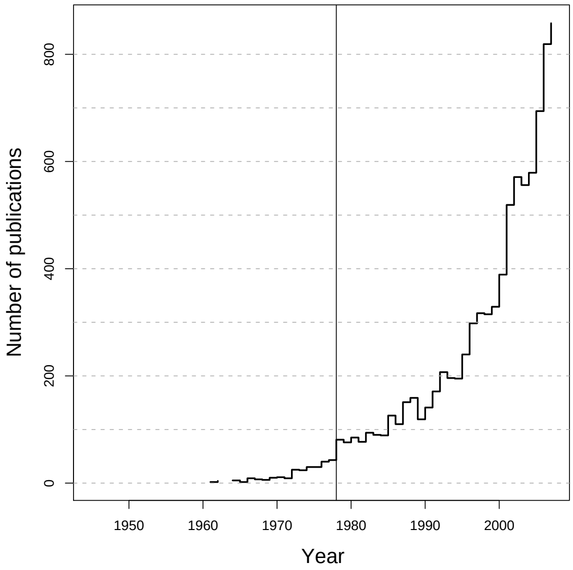


Figure 1.1: Trends in the use of the terms "habitat selection" in scientific publications. The results comes from a search in the ISI Web of Knowledge (`http://isiknowledge.com/`), with "habitat selection" as a topic keyword. The vertical line is set at the year 1978.

selection was consistently elusive, if not avoided.

In this section, I will introduce the habitat concept, and some basic elements related to this concept. I will focus mainly on habitat selection, and try to provide commonly accepted definitions of the relevant terms.

1.2 Habitat selection theory: some hints

1.2.1 Habitat vs. niche

The habitat and niche concepts have long been debated and confused, from Whittaker et al. (1973) to Kearney (2006). Several authors attempted to provide definitions of the habitat concept, from the simplest to more complex ones. The habitat of a species is often related to the characteristics of the environment ("*By the habitat of a plant or plant association is meant the kind of situation in which the organism lives.*", McDougall, 1927) and to the performance of animals within the environment (Habitat as the "*resources and conditions present in an area that produce occupancy –including survival and reproduction– by a given organism*", Hall et al., 1997). It is also often used as the specific place in the geographic space occupied by the species, as defined by Morrison et al. (2006) who proposed the simplest definition, the habitat being *"a place where an animal lives."* Following the same approach, habitats are *"spatially bounded subsets of physical and biotic conditions among which population density of a focal species varies from other adjacent subsets"* (Morris, 2003b) or *"wherever an organism is provided with resources that allow it to survive."* (Hall et al., 1997).

Unfortunately, and despite this effort towards standardization, the habitat concept never reached a state of perfect unambiguity, so that every one may use it adapted to his very purpose. This led eventually Mitchell (2005) to question the relevance of the term itself. Indeed, there is no simple and straightforward definition of the habitat, as we will see in Chapter II. The habitat is a complex concept, and we will try to disentangle its use together with the concept of ecological niche. For now, we may thus define habitat in a complex way, quoting Whittaker et al. (1973): *"The* m *variables of physical and chemical environment that form spatial gradients in a landscape or area define as axes a habitat hyperspace. The part of this hyperspace a given species occupies is its habitat hypervolume. The species' population response to habitat variables within this hypervolume, as expressed in a population measure, describes its habitat."*

1.2.2 Basic definitions

The definitions detailed here follow the work of Johnson (1980). Although they were defined in the context of resource preference, they can be applied indifferently to any kind of component, (i.e. resources in a broad sense, not necessarily depletable) to encompass every environmental conditions, and thus habitat variables. As a matter of fact, Johnson (1980) specifically defined resources (in the context of habitat selection) as *"the kind of foods* [an animal] *consumes, and the varieties of habitats it occupy."*

Habitat availability is given by the quantity of any habitat variable ("component") that is accessible to an animal. It is a result of the physical accessibility of that component, weighted by all additional constraints appended to that animal (either biotic or abiotic).

Habitat use is measured by the quantity of a habitat variable an animal actually uses, in a given amount of time. The animal may deplete the resource, or merely consume it, in a generic sense (Hall et al., 1997).

The selection of a component is "a process in which an animal actually chooses that component" (Johnson, 1980). *Habitat selection* is thus this process of selection applied to every habitat variable which result in the choice of specific habitat conditions. Formally, habitat selection is given by the comparison of use and availability. In concrete words, the use of a component is said to be selective when it is disproportionate to its availability, either in a positive way (selection for this component) or in a negative way (selection against this component, i.e. avoidance of this component). Thus, habitat selection does occur when at least one habitat variable is used disproportionately to its availability to the animal.

1.2.3 Selection vs. preference

Habitat preference has been defined first by Johnson (1980) as "*the likelihood of* [a] *component being chosen if offered on an equal basis with others.*" In opposition to that definition, Hall et al. (1997) proposed instead to name habitat preference the result of selection, i.e. "*the consequence of the process, resulting in the disproportional use of some resources over others.*" Aarts et al. (2008) proposed a compromise between the previous two definitions by stating that habitat preference is "*the ratio of the use of a habitat over its availability, conditional on the availability of all habitats to the study animals*".

Based on these contradictory definitions, I would advice to use Johnson's approach, with the help of the ice cream metaphor: Imagine that we are at the end of a nice dinner, and I would like some dessert[1]. At that moment, I would love to have that ice cream, with melted chocolate on top of it. That is my very preference. However, I do not want too much sugar and fat in my blood because I care about my health, and I also want to stay slim and attractive, in response to a strong social pressure. These are part of my selection process. As a consequence, I will choose an apple instead.

Following that example, preference may be defined in a cognitive way, as the taste for a component *ceteris paribus* (i.e. all other things being equal). This way, it corresponds also very much to the common meaning of a preference. In the case where it is necessary to distinguish between

---

[1] We French people usually do that...

the result of selection and the process of selection itself, the result might be called *habitat choice*. **Thus, preference is the source of motivation, selection is the process involved, and choice is the final result.**

In an ideal free world, i.e. without constraint of any type, animals would choose their favored habitat. However, between preference and choice, animals do experience many constraints, either abiotic (physical characteristics of the landscape, distance, etc.), or biotic (intraspecific interactions: sociality, territoriality; or interspecific interactions: predation, parasitism, etc.). In other words, due to external constraints, an animal might select a component even if it might prefer not to. The response to these constraints involves "*a series of innate and learned behavioral decisions*" (Morrison et al., 2006) during the process of selection. However, because of their mental nature, we can generally not access to (and thus study) these decisions nor can we assess preference. This would be possible only in controlled experiments, with every constraint removed, and the test of different features together, which is almost never possible with wildlife. Instead, we often measure the disproportionate use in comparison to the availability, which is habitat choice. At best, can we evaluate some constraints, and interpret the current choice in the light of them. As the main process is the one of selection, it is thus by no mean a non-sense to encompass the whole succession under the general term *habitat selection*.

1.2.4 The problem of availability

The constraints that prevent animals to merely yield themselves to their preference immediately bring us to the problem of availability. By availability, we mean the subset of conditions bounded by the physical accessible world and limited by all internal and external constraints, and its abundance (in case of resource) or value (otherwise). However, in habitat selection studies, we generally only measure accessibility without constraints. This has been discussed extensively by Hall et al. (1997) who stated that availability was seldom measured to the benefit of abundance (in case of resources) which refers only to their quantity in the habitat. As Johnson (1980) underlined, "*conclusions about whether an individual component is used above, in proportion to, or below its availability are critically dependent upon the array of components the investigator deems available to the animal. This decision is often made somewhat arbitrarily by* *the investigator."* Thus availability is a subjective element. This can be a major problem since habitat selection is studied by comparing usage and availability, and may be biased in this respect.

To avoid such a situation, availability must be checked, and as much as possible rely on previous knowledge on the species. In some situations, it is however possible to simplify the problem when the scale of the study can only lead to coarse results (see e.g. Chapter 7). It might also be easier to detect what is really available in the case of animals establishing home ranges (then the availability for one individual is bounded by the borders of its home range) and/or expressing territoriality (then the availability for one individual is limited by conspecific territories). Another example is given by females brown bear (*Ursus arctos*), which in Sweden only establish in the neighbourhood of their mother (Martin et al., data, see Appendix 1). Thus, in that case, availability is totally constrained by a strong social (maternal) constraint within the accessible world.

1.3 From species to individuals

1.3.1 A hierarchical approach

Habitat selection has been consistently seen as a hierarchical process (Johnson, 1980; Wiens, 1989; Orians and Wittenberger, 1991; Hall et al., 1997). This hierarchy can be expressed at different levels. In particular, there is a hierarchy in an individual's decisions, which can result in a trade-off for his space use (e.g. between food and security, Houston et al., 1993). The hierarchy of selection also occurs in space, from broad scale to fine scale, which is generally related to time scale (Wiens, 1989). Finally, the hierarchy on a ecological level, from species to individuals, through populations.

As a summary for these related hierarchies, Johnson (1980) proposed to approach habitat selection at four selection orders. First-order selection relates to the distribution of the species. Within that range, second-order selection describes the individual habitat choice to establish their home ranges. Third-order selection is given by habitat use within individual home ranges. Finally the choice of food items at a fine scale defines the fourth-order selection. As emphasized by Johnson (1980), it is possible to further divide these selection orders more finely, and to adjust the four levels to a particular system. Senft et al. (1987), for example, adapted these levels to a functional approach of large herbivore foraging.

The next Chapter (Chapter II) discusses the habitat and niche concepts in the context of a species' population. We will later on (Chapter III) consider habitat selection of a single individual following the second and third-order selection *sensu* Johnson (1980).

This simple hierarchy may correspond to different study designs, as discussed by Thomas and Taylor (1990, 2006). The authors distinguished between four designs, one for each of Johnson's selection orders. In Design 1, data are collected at the population level (individuals are not identified), and this corresponds to the first-order selection of Johnson (1980). In Design 2 and 3, use data are collected at the individual level. In Design 2 (second-order selection), availability is measured at the level of the population, and thus is the same for all individuals, whereas in Design 3 (third-order selection), availability is measured for each individual independently. Finally, the Design 4 is intended to deal with micro-habitats selection (fourth-order selection), with availability and use data measured at multiple time for each individual.

1.3.2 Measuring performance

At any level of the hierarchy, it is generally assumed that the actual use of different components of the habitat is related to the performance of the animals (Morris, 2003b; Rosenzweig, 1981). As underlined by Thomas and Taylor (2006), *"resource selection studies are commonly conducted because it is generally assumed that if animals select habitat and food resources disproportionate to their availability, that habitat or food improves their fitness, reproduction, or survival."* In other words, as animals are expected to maximise their fitness through selective use of their environment, the spatial distribution and the abundance of a species should reflect the environmental quality. This postulate may however not always be true, unless animals are omniscient. It has been shown that dynamics including source and sink habitats can be both ecologically and evolutionarily stable (Pulliam, 1988). Sinks can emerge in case of social exclusion of some individuals from the best conditions (Pulliam and Danielson, 1991), or from a maladaptative habitat selection, leading to a choice of unfa-

Figure 1.2: A roe deer (*Capreolus capreolus*) within its habitat in southern Norway. Is the presence of an individual in this particular environmental conditions an evidence of the high suitability of these conditions? Photo credit: John Linnell.

vorable habitats (Delibes et al., 2001). Delibes et al. (2001) furthermore demonstrated that in some demographic conditions, the number of individuals in such attractive sink may even increase for a given time, while the number of individuals in source habitat decreases. In such conditions, a field worker collecting species' occurrence may conclude that the sink habitat is best suited for the species (because the number of individuals within these conditions increases). The existence of sinks demonstrates that the detection (and further, abundance) of a species in an area does not necessarily mean that this area is optimal for the species. In other words, the species' occurrences may delineates its habitat *sensu* Whittaker et al. (1973), rather than fully describe its shape.

It is thus necessary, following Whittaker et al. (1973), to measure performance throughout the conditions encountered by a species to derive a proper definition of its habitat. This spatial measure of performance might take the form of a species' population density (but see Van Horne, 1983 who warned about the systematic use of such a measure), or growth rate, or a combination of both. A better approach would be to use individual based measures of performance, e.g. survival (Pollock et al., 1989; Murray, 2006; McLoughlin et al., 2005) or reproduction (Pidgeon et al., 2006). However, relating fitness measures to multivariate resource selection has received little attention in ecology (McLoughlin et al., 2005).

In the recent years, there has been some progress towards the use of synthetic individual measures of performance related to habitat selection. For instance, the lifetime reproductive success (LRS) can be computed as the number of female young to which a female gave birth and which survived to the age of independence. Computed for every female, this index integrates both survival (since the number of birth events is directly related to longevity) and reproduction (by the numbers of youngs). This measure enabled McLoughlin et al. (2007) to show that a roe deer's incorporation into its home range of habitat components important to food, cover, and edge was significantly related to individual performances (See Fig. 1.2). We should relate more (individual) performance to habitat selection in order to fully describe a species habitat (*sensu* Whittaker et al., 1973).

1.3.3 A dynamic process

This large scale overview of habitat selection could not pass silently over the climate change (Zwiers and Hegerl, 2008; Rosenzweig et al., 2008). As a matter of fact, the world is dynamic! Despite this, habitat selection is often defined static. As underlined by Hutchinson (1957), "*the model refers to a single instant of time*". While he referred to his model of ecological niche, his statement can easily be extended to the conceptualisation of habitat. What Hutchinson meant is that the study of the niche (and/or habitat) should be undertaken in a fixed state of the environment. However, this condition does not necessarily hold, as the environment experienced by animals may change. It can be on a daily basis (nocturnal vs. diurnal species, Hutchinson, 1957), or a seasonal basis (Arthur et al., 1996), or life cycles (e.g. vegetation or parasitism Orians and Wittenberger, 1991).

Eventually, and this might be even more interesting, animals themselves are dynamic. Their habitat selection behavior is unlikely to be the same at different periods of their life cycle (different ages or seasons), of for different activities (e.g. feeding time, resting, reproduction Cowlishaw, 1997). Study Design 4 (Thomas and Taylor, 2006) is especially suited for this approach, through the analysis of steps, i.e. trajectories. A step is defined by the straight-line segment between successive relocations. For example, (Fortin et al., 2005) used a resource selection function (RSF Manly et al., 2002) applied to trajectories (steps) of elk, and was able to show that, as the risk of wolf encounter increased, the preference of elk for aspen stands gradually decreased, and selection became strongest for steps ending in conifer forests in high wolf-use areas. Finally, technologies like activity censors or recent movement analyses might enable the use of activity patterns into habitat selection.

Chapter 2

Further evolution of the habitat and niche concepts through niche differentiation

Mathieu Basille[1,2,*] & Bram Van Moorter[1]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69 622, Villeurbanne, France.

[2] Centre for Conservation Biology, Department of Biology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[*] Corresponding author: `basille@biomserv.univ-lyon1.fr`

In prep. for submission.

"The more life there is within a system, the more niches there are for life."
[Kynes' father, Imperial Planetologist on the planet of Dune]
"Dune", by Frank HERBERT.

COMPETITION IS A CORNERSTONE BETWEEN ECOLOGY AND EVOLUTION: it is grounded in ecology (as early defined by Ernst Haeckel in 1870 as the study of the natural environment and the relations of organisms to each other and their surroundings, Ricklefs 1990), to explain evolutionary mechanisms. More specifically, inter– and intraspecific competition can result in functional specialization (in case of coexistence in the same area) or geographical isolation (Davies et al., 2007; Rosenzweig, 1978). The former occurs in the case of niche differentiation, whereas the latter is generally know as habitat differentiation, justifying concurrent use of both concepts of *niche* and *habitat*. These two processes of exclusion form the starting points of the ultimate process of speciation, through allopatric speciation (Mayr, 1963) and sympatric speciation (Maynard Smith, 1966), respectively.

Udvardy (1959) was amongst the first to approach the *habitat* concept together with the *niche* concept, in an attempt to reconsider both in the same theoretical background. Despite this early focus on theory clarification, we have seen an increasing confusion between habitat and niche concepts in the last decades (Kearney, 2006). As Whittaker et al. (1973) noted 35 years ago, *"it is regrettable that two of the most important terms in ecology,* "niche" *and* "habitat"*, are among the most confused in usage."* While an increasing number of habitat selection studies since the early 80s enhanced the focus onto habitat theory (Morris, 2003b), we believe that the situation is currently not any better. Indeed, several authors recently attempted to provide definitions for these terms (e.g. Guisan and Thuiller, 2005; Mitchell, 2005; Kearney, 2006; Aarts et al., 2008), but to date no definitive one has emerged and usage of these terms remains ambiguous. In this respect Mason and Langenheim (1957) stated that there "[...] *is no right or no wrong definition for a particular word."* However, words can become standardized through our use of them, i.e. through common usage. Mason and Langenheim (1957) continue with the observation that: "[...] *there are very important exceptions to this rule of common usage.*" Words can be used "*so indefinitely that dissatisfaction results from perpetuation of its use.*" This dissatisfaction drove Mitchell (2005) to question the usefulness of the habitat concept. The reason for this "*rule of common usage*" not to work for certain words is that it are often "*inclusive term[s] covering many confusions and hence must be either clarified or avoided.*" Mitchell (2005) provocatively suggested for the later option and abandon the use of the habitat concept altogether. We prefer to keep the baby and merely throw away the bathwater. In our opinion a thorough investigation of the habitat and niche concepts will not kill ecology instead it will just make it stronger (*sensu* Nietzsche, 1889). In this paper we will follow Mason's suggestion: "[...] *it becomes mandatory for the scientist to analyze the words in his field in terms of the phenomena upon which they rest. He is then in a position to construct a precise definition.*"

However, our aim herein is not to provide yet another definition that could probably contribute to the reining chaos, but rather to synthetise the notions of habitat and niche from the current literature. Following Mason and Langenheim (1957), we analysed the habitat and niche concepts in our own field, namely animal ecology in terrestrial ecology; our approach might therefore be somewhat focused in this respect. Our analysis starts from the observation of the origin of the problem in the literature, with the seminal paper of Hutchinson (1957); we will undertake a clarification of the habitat and niche concepts by using a general framework, which borrows heavily from the one presented by Whittaker et al. (1973), and eventually develop a synthetic view.

2.2 A 50 years old problem

Following the work of Grinnell (1917), Hutchinson (1957) defined the niche as the conditions in which a species can survive and reproduce. In its seminal paper, Hutchinson described the niche geometrically as the "*n-dimensional hypervolume* [...], *every point in which corresponds to a state of the environment which would permit* [a species] *to exist indefinitely*" (Fig. 2.1). This hypervolume, which Hutchinson called the *fundamental niche*, was the first attempt to "mathematicize" (*sensu* Haskell,

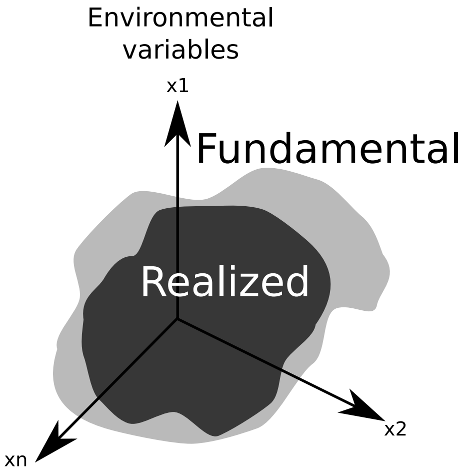


Figure 2.1: The fundamental niche is defined as the *"n-dimensional hypervolume* [...], *every point in which corresponds to a state of the environment which would permit* [a species] *to exist indefinitely"* (Hutchinson, 1957). In the presence of interspecific competition, the niche is reduced and form the realized niche.

1940) ecological theory on the niche. With this point of view, a set of environmental variables —*"any property outside the organisms under consideration"* (*sensu* Hutchinson)— define a hyperspace, i.e. a multidimensional space. Within this hyperspace the conditions suitable for the persistence of a species consist of the niche hypervolume. Even if its model presupposed an equal probability of persistence in every part of the niche, Hutchinson pointed out that there were both an optimum and suboptimal conditions near the boundaries. Whittaker et al. (1973) considered the niche was not as that hypervolume itself but as the species' response within that hypervolume, so that every point of the niche is characterized by a measure of performance of the species, which can be assessed with e.g. density, growth rate, etc. The niche hypervolume therefore defines the boundaries of the niche, whereas its shape arises from the species' response to the environment (see Fig. 2.2).

Within that hypervolume, not every condition is accessible to the species. In particular, competition with other species may exclude the focal species from a part of the niche (or even the whole niche in extreme cases). In the presence of interspecific competitors, the fundamental niche may be reduced, and form what Hutchinson called the *realized niche* (Fig. 2.1). Based on the competition exclusion principle of Volterra & Gause (Volterra, 1926; Gause, 1934), it is stated that competition between two species on a common resource (either by exploitation or interference) should result in the exclusion of one of the species (the least favoured by the competition) in the overlap of the two fundamental niches, so that realized niches do not intersect. It follows that the fundamental niche encloses all potential conditions for a species' persistence, whereas the realized niche is the part of the fundamental niche a species' population is actually able to colonize successfully in a given biotic context.

The multivariate representation described by Hutchinson became widely accepted in both theoretical and applied approaches. Indeed, this formalism is very convenient as it enables the niche to be quantified on many dimensions simultaneously. However, this representation soon overshadowed the theoretical notion behind it: The niche became identified by this representation, and this representation associated with the niche. In other words, the niche concept was so strongly associated with the abstraction of Hutchinson that it precluded any other use of this representation. In particular, the concept of habitat, previously defined using a similar multivariate space (Haskell, 1940), became closely related to the niche concept with its representation. As a consequence, niche and habitat have often been used to describe the same object (see e.g. Pulliam, 2000). It is interesting to note that it was probably not Hutchinson's aim to deal with the habitat concept at all. Rather, the niche concept defined in his paper was designed for synecology. Precisely, Hutchinson meant to deal with competing species within communities, and the habitat was just left out of the picture. Despite being rooted in community ecology, Hutchinson's paper faced a great success in habitat selection theory (it is still widely cited), and introduced a lot of confusion on the habitat concept as a side-effect, much beyond its original scope. However, as we will see, habitat and niche are related though not synonymous concepts, and both can be represented using the same n-dimensional formalism, while representing different concepts.

2.3 Niche and habitat: It is all about scale

Whittaker et al. (1973) were amongst the first to recognize this potential confusion between habitat and niche. In reaction they defined a set of proposals allowing one to deal with the habitat and the niche of a species in a common conceptual framework. The main focus of their advances was related to the scale of the environmental variables. Whittaker et al. (1973) separated all possible environmental variables into two categories, habitat variables and niche variables. Habitat variables are defined as intercommunity variables, which are extensive variables, with an intrinsic spatial component related to the species' distribution. They often consist of spatial gradients defined at a large scale. Conversely, niche variables are defined as intracommunity variables and describe the functional relationships of a species within a community (Whittaker et al., 1975). These are intensive, local variables, with no specific spatial component, described at a fine scale.

Habitat and niche variables are thus both subsets of the whole set of environmental variables and both can be represented geometrically, using Hutchinson's approach. Habitat (resp. niche) variables define axes of a habitat (resp. niche) hyperspace. The part of this hyperspace where a species can persist is its habitat (resp. niche) hypervolume, which shape is described by the species' response confronted to the environmental variables. Only with the species' response are the habitat and the niche fully described. Practically, habitat and niche variables are two poles of a continuum of environmental variables, from extensive to intensive variables, and it may be difficult to distinguish among them. For such cases, Whittaker et al. (1973) proposed to use the term *ecotope*, as a broader concept that encompass both niche and habitat (Fig. 2.2). They furthermore suggested to consider the ecotope as the ultimate evolutionary unit, defining the species' response to the full range of environment variables to which it is exposed, whereas the niche might be seen as the proximal context for a species. However, since (1) the term "ecotope" has to our knowledge not found entrance in the research community, and (2) niche variables are often included in habitat studies, we will not promote the use of the term ecotope. Rather, we suggest to consider habitat in a broad and habitat in a narrow sense; the former includes the niche, whereas the later excludes it. Note that the opposite is not possible: whereas habitat variable are extensive and might encompass all niche variables, the reverse is not true. In cases where it is needed to make a clear distinction between these two forms of habitat, the terms *macrohabitat* and *microhabitat* might be used.

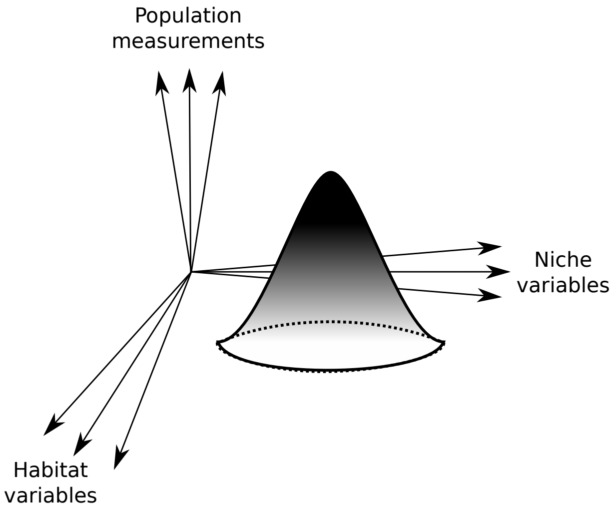


Figure 2.2: The ecotope, *sensu* Whittaker et al., 1973, encompasses both concepts of niche and habitat. The whole set of environmental variables can be divided into habitat variables and niche variables. The part within the ecological hyperspace a species occupies is represented with the circle below, whereas the species' response (measured on a set of population measurements) defined the habitat, and the niche, on the habitat variables, and niche variables, respectively (after Whittaker et al., 1973, Fig. 5).

Macrohabitat and microhabitat were recognized as large spatial scale vs. small spatial scale features (Begon et al., 1986; Hall et al., 1997), in accordance to a precision of Whittaker et al. (1975) after a comment of Kulesza (1975). Accordingly, macrohabitat and microhabitat factors correspond tightly to the definitions of habitat and niche variables by Whittaker et al. (1973). Note that they are usually used in different fields, macrohabitat and microhabitat being more common in ecology, and niche more common in evolution.

Different geostatistical methods have been developed to determine scales of variation of environmental variables (e.g. variogram analysis Mayor et al., 2007 and Fourrier analysis Stein and Weiss, 1971). When the extent of variation of the environmental variable is greater (or with the same order of magnitude) than the geographical range of the population, the variable is said to be a habitat variable; on the other hand, when the scale of variability is at a small scale the variable is said to be a niche variable. From this it follows that an animal may hold a potential larger influence from niche variables, being variables that varies at the corresponding scale.

On the other hand, a single individual does not experience the whole range of variation of a habitat variable. Thus, the direct impact of such a variable is rarely measurable, although this variable may have an effect on the species distribution on the landscape. This is precisely what Whittaker et al. (1973, 1975) emphasized, stating that "*niche refers to the functional relationships of a species within a community; and habitat to its distributional response to environmental factors at different points in the landscape.*" As pointed out by Kearney (2006), a habitat variable that describes association between organisms and features of the landscape may be best studied with a correlative approach. On the other hand, the study of the niche implies knowledge of the behavioural, morphological and physiological properties of an organism, and is best suited for a mechanistic approach. However, we believe that both concepts can be indifferently used in both contexts. The framework defined by Whittaker et al. (1973) affords the possibility to deal with both, which are not mutually exclusive and do not preclude from a more mechanistic approach of the habitat (yet not so intuitive) or a more descriptive approach of the niche (which is consistent with the multivariate formalism of Hutchinson 1957, as admitted by Kearney, 2006). Modern niche theory has already absorbed this dichotomy of the niche, stated in terms of "*requirements*" (correlative approach) and "*impacts*" (mechanistic approach), successfully considering Hutchinson's niche and the purely functional approach of Elton (1927) in the same framework (we return the reader to Leibold, 1995 for more details regarding the elucidation of the niche within the context of modern niche theory).

2.4 Theoretical implications and synthesis

Whittaker et al. (1973) noted that the fundamental niche of a species is not directly measurable; only the realized niche of a species' population is measurable, i.e. the performance of a population *in situ*, bounded by the fundamental niche, and accounting for interspecific relationships. Whenever experimentation, such as competitor removal, is not possible (which is unfortunately often the case in ecology), the fundamental niche may be only approximated by the sum of all measurable realized niches. In other words, the fundamental niche of a *species* can be seen as the generalization of all realized niches of this species' *populations*. Even though it is not often stated explicitly, the same reasoning holds for the habitat concept: Similarly, the *fundamental habitat* canx be considered as the generalization of populations' *realized habitats*. It is important to note here that, even if habitat variables are defined at the level of the community, the habitat itself refers to a population (and is generalized over a species).

A few implications result from the measure of a species's performance as defined by Whittaker et al. (1973, see Fig. 2.3), in particular on the definition of the niche boundaries. Hutchinson (1957) encompassed in his concept of niche every condition that allows theoretically a species to persist indefinitely. In other words, every place of the niche should be associated with a positive global performance, if we can imagine such an synthetic measure. In contrast, Whittaker et al. (1973) defined the boundaries of the niche (resp. habitat) by the limit of the species' occupancy, within which the species expresses a response to environmental conditions. This response is typically bell-shaped, with asymptotic tails at extreme conditions, and defines the shape of the niche (or habitat). How artefactual this approach might be (the limits will often be dictated by the ability to detect a species' presence), this is however much more flexible and powerful. First, a species may face a changing environment, and be confronted from time to time to unfavourable conditions. In such conditions, the population would become extinct unless sustained by immigration from more favourable regions. At this moment, the performance of the species would be very low, probably negative, and should condemn the species to extinction. However, species can often survive in unfavourable conditions in changing environments by "cutting their losses", as highlighted by Whittaker et al. (1973). Thus, the habitat and the niche shall include them as well, to consider the temporal dimension of both habitat and niche.

Van Horne (1983) questioned the assessment of habitat quality by density only. She called for a distinction between low-quality habitats (later on called *sinks*) sustained by immigration, where individuals are unlikely to survive and reproduce, and high-quality habitats (*sources*), with positive demographic outputs. By doing so, she opened the way to the source-sink dynamics theory (Holt, 1985), which states that large sink populations can exist thanks to migration from source populations (Pulliam, 1988). This source-sink habitats approach can easily be extended at the niche level, as

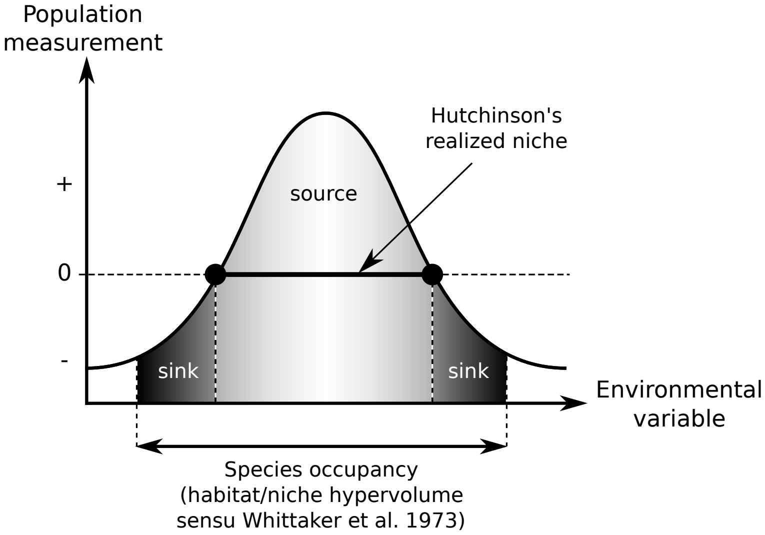


Figure 2.3: A synthetic representation of the performance measure. The performance is approached by one population measurement, and is related to one environmental variable, that can be either a habitat or a niche variable. Hutchinson (1957) only included in his niche concept conditions related to the persistence of the species, i.e. a positive performance. The extended concept of Whittaker et al. (1973) is bounded by the species' occupancy, and thus includes both sources and sinks.

it was defined above. At the turn of the century, Pulliam (2000) advocated for a new definition of Hutchinsonian's niche, in order to integrate the influence of dispersal and sink habitats. Even though Pulliam used the niche as a mere geometrical representation of the habitat, he indeed identified an important limitation of Hutchinson's niche, which does not include sinks by definition. The proposal of Whittaker et al. (1973) then has the great advantage to consider sources and sinks within the species' habitat or niche, answering Pulliam's request. Thus, the niche and the habitat, bounded by the species' occupancy, encompass conditions associated to a positive population performance (sources) and conditions associated to a negative population performance (sinks, Fig. 2.3).

This point of view associated to the problematic of the competition exclusion principle of Volterra & Gause solves one of the major problem of this principle: the coexistence of competing species (Tilman, 1987). Tilman underlined that *"for species to coexist stably, there must be trade-offs in their physiological, morphological, or behavioural traits."* These trade-offs occur when (at least) two species are competing for common resources, and result from a better performance of each species for the use of different resources. McPeek (1996) stated that these trade-offs operate both at the between-community and within-community scales, i.e. at the habitat and niche levels. The competitive advantage of one species for a given resource would consequently reduce the other species' performance regarding to this resource, leading to coexistence (i.e. overlapping habitats or niches on their boundaries) or extinction (i.e. habitat or niche exclusion) depending on the level of decrease. If the performance becomes negative, the species would be in a sink, thus would sustain in these conditions only in case of immigration. In the extreme case, if the performance falls below the limits of the species' tolerance (thus occupancy), the species would become extinct in these specific conditions. This mechanism forces the species to have responses spread over environmental variables, and ultimately leads to habitat or niche segregation. Competition thus does not have a binary impact on the size of the realized niche (*sensu* Hutchinson), but merely changes the shape of populations' performance, and *in fine* the boundaries of their habitat or niche, depending on the level of investigation.

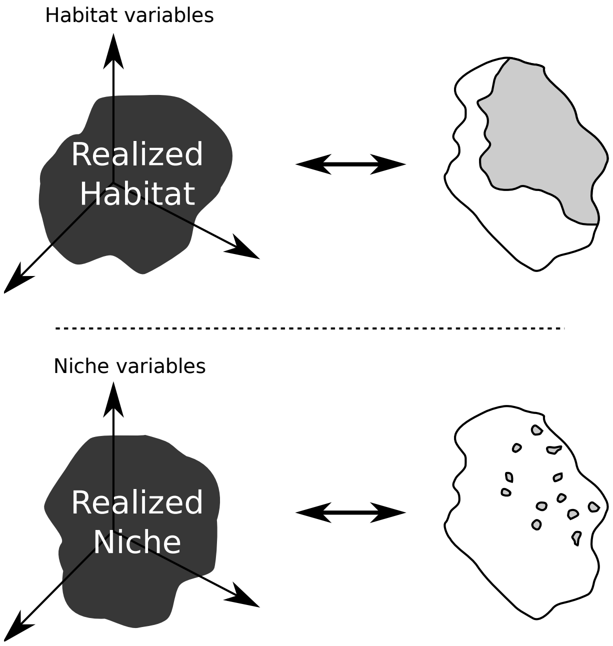


Figure 2.4: Both the habitat and the niche can be projected into the ecological space and mapped in the geographic space. A given habitat generally appears as a continuous area in a given landscape (top) whereas a given niche generally corresponds with a set of patches (below).

2.5 Habitat, niche, and the rule of common usage

Kearney (2006) rightly advocated the use of mechanistic niche modelling to predict a species distribution. The habitat and the niche of a species can be projected simultaneously in the ecological space and the geographical space (Fig. 2.4). Both concepts are indeed defined in the multi-dimensional space formed by environmental variables; they consist of a limited subset of conditions in both cases (bounded by the limits of the habitat or niche hypervolume). Both concepts can in turn be projected in the 2 or 3 dimensions (x, y, and potentially z) physical space, i.e. the geographical space, and thus be used to model species' distribution.

Habitat variables are defined extensively, and often consist of gradients in the geographical space. Consequently, a projection of the habitat in the geographical space generally results in a continuous area that often can be physically delineated (Fig. 2.4, top). By extension, the term habitat itself is often used in reference to the area it delineates in the geographical

Figure 2.5: A gradient of landscapes from central to southern Norway (from upper left to lower right corner). Photo credit: Scandlynx / John Linnell.

space (Morris, 2003b). In this context, the habitat is no longer a species' characteristic, but rather a geographic feature. As pointed out by Kearney (2006), we can describe a given habitat "*without reference to an organism, even though the potential presence of some kind of organism is always implied.*" However, for this very purpose, we would advocate instead the use of the term *landscape*, as is currently used in landscape ecology (see e.g. Turner et al., 1995), as both abiotic and biotic characteristics of a large area (see Fig. 2.5).

On the other hand, niche variables are defined intensively, i.e. locally. That is, the range of variation of a niche variable is many times greater than the size of the area occupied by a species' population. This results in niche variables that are not of spatial nature *per se*. The niche can, however, also be projected into the geographical space, and will generally correspond to a multitude of patches with similar conditions, that are not necessarily contiguous and thus do not define a continuous area (Fig. 2.4, below).

In the early 1980s, following Tilman (1982), with the emergence of what is known as the modern niche theory, focus shifted to the resource part of the niche. After Tilman's work, a resource is defined as *"a factor which* [...] *leads to higher population growth rates as its availability is increased, and which is consumed, in the broad sense, by the population"* (Tilman, 1980). This definition highlights three characteristics of a resource: it is an environmental factor, that is depletable and potentially limiting (Leibold, 1995). As a result, the niche in the modern niche theory is based only on variables that represent resources, through a dimensionality reduction. In other words, because not every environmental variable is included, the modern niche is not equivalent to the ecological niche, but is a subset taken from it. However, these two concepts are by no means incompatible, as they both rely on the same multidimensional approach of environmental variables. For instance, Leibold (1995) showed that the "zero growth net isocline" or "ZGNI" as used by Tilman (1980) (i.e. the set of conditions for which the available resources determines a mortality equal to the reproduction of the population) was the niche of Hutchinson (1957), actually the niche hypervolume as defined by Whittaker et al. (1973). One of the major fields of application of the modern niche theory is competition within communities. Whenever two (resource) niches overlap (that is, two species rely on the same resource range), there is competition, either by interference or by exploitation. Interspecific relationships (for instance competition, but also facilitation or predation) can be seen as a particular resource (and represented by another environmental variable) or explicitly stated as separated niches for each species, allowing one to compare them directly. The framework detailed here thus provides the basis for the inclusion of interspecific relationships (Guisan and Thuiller, 2005), in particular in a given trophic chain. The approach presented here is therefore consistent with modern niche theory, providing that the latter relates to a niche restricted to resources, i.e. the resource niche (Begon et al., 1986).

2.6 Conclusion

The past literature, especially in more recent years, showed an increasing demand for separation of the habitat and niche concepts. Unfortunately, a synthetic view remained largely absent. Here, we summarized the main sources of confusion and resolved them by putting both concepts in a common framework, initially presented by Whittaker et al. (1973) that we further developed.

We believe that Whittaker et al. (1973)'s approach is most useful to define habitat and niche clearly. In particular, it accommodates the main current fields where both terms are in use, like habitat selection theory and modern niche theory. In addition, it is highly congruent with different common usages, and the use of this framework would only require minor adjustments. This would allow for a more unambiguous use of both terms, which are in our opinion still very useful (and anyway actually used). This framework might then also contribute to a increased exchange of ideas between ecology and evolutionary biology. Finally, the unambiguous definition of the niche and habitat could dissipate the current dissatisfaction from the perpetuation of their (mis)use, as underlined by Haskell (1940), without actually avoiding their use altogether, as advocated by Mitchell (2005).

Acknowledgments

The French Ministry of Research, the University of Lyon, the Office National de la Chasse et de la Faune Sauvage and the Norwegian University of Science and Technology provided financial support for this work.

We are indebted to Jean-Michel Gaillard and to the GROS (*Groupe de Recherche sur les Ongulés Sauvages*), for constructive comments at several stages of that work.

Chapter 3

Habitat selection against limiting factors at different spatial scales

Bram Van Moorter[1,*], Mathieu Basille[1,2], Mark S. Boyce[3] &
Jean-Michel Gaillard[1]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69 622, Villeurbanne, France.

[2] Centre for Conservation Biology, Department of Biology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[3] Department of Biological Sciences, University of Alberta; Edmonton, Alberta, Canada T6G 2E9.

[*] Corresponding author: `moorter@biomserv.univ-lyon1.fr`

In prep. for submission.

"Il y a les malins et les pas malins. Et le moins malin des malins n'est pas beaucoup plus malin que le plus malin des pas malins. Mais en cas de nombre impair, il peut être les deux à la fois.."

"L'avenir du Chat", by Philippe GELUCK.

ANIMAL HABITAT SELECTION has potentially profound impacts on population dynamics (e.g. Pulliam and Danielson, 1991), life-history evolution (e.g. Holt, 1997) and animal conservation (e.g. Morris, 2003a). Habitat selection is defined as the process by which an animal chooses which habitat components to use (Hall et al., 1997; Johnson, 1980). The result from this selection process is the disproportionate use of some resources over others (Hall et al., 1997).

Since Johnson's (1980) seminal paper where he discussed the hierarchical nature of habitat selection, an increasing number of studies include multiple spatial scales in their analysis of habitat selection. The scale at which one investigates a phenomenon greatly influences the patterns one might observe (Wiens, 1989); many studies of animal habitat selection have shown different selective behaviors depending on the spatial scale (Bowyer and Kie, 2006). Johnson distinguished four hierarchical levels: species range, home range selection, selection of patches in the home range and finally selection of individual food items within a patch; in the following we will focus on individual habitat selection. However, despite the fact that now almost thirty years have passed since this publication, hardly any theory is available for predicting habitat selection across spatial scales (however, see Rettie and Messier, 2000).

One exception on the above mentioned lack of theory about hierarchical habitat selection is in the paper by Rettie and Messier (2000). They propose: *"that there should be a direct relationship between the importance of specific factors potentially limiting populations and the level at which individuals exhibit selective behaviour to minimize the effects of those factors."* Or in other words: *"Factors with greater potential to reduce individual fitness should be avoided at coarser spatial and temporal* *scales, as avoiding the factors that are most limiting at each successive scale will maximise an individual's fitness."* A limiting factor in population dynamics is an environmental factor that prevents a population from attaining its maximum population growth. Since, population growth is the average reproductive success in a population (note: as common in population biology, we only work on the female part of the population), a limiting factor can then easily be extended to the individual level, as the set of factors that prevent individuals from realizing their full potential fitness.

Since Rettie & Messier's (2000) publication different authors have investigated these ideas in different animal species (including mammals, birds and reptiles; e.g. McLoughlin et al., 2004; Harvey and Weatherhead, 2006). However, different publications reported failures to support Rettie & Messier's prediction (e.g. Morin et al., 2005; Gustine et al., 2006), and didn't find selection against the most limiting factor at the highest hierarchical level. We identified three important ceavats in this current theory of hierarchical habitat selection, which we see as likely candidates to explain the failures to support this theory: spatial variability and temporal variability of limiting factors, and co-variation of different resources.

Morin et al. (2005) and (Harvey and Weatherhead, 2006), did not find selection against the most limiting factor for the home range choice, whereas they found selection at the interior of the home range. In both publications the absence of home range selections was attributed at the relative homogeneity of the environment at the home range scale within the study site. The importance of environmental spatial variability for habitat selection studies has been extensively discussed by Åberg et al. (2000).

Population ecologists have shown that knowledge of the critical demographic parameters —i.e. the parameters for which population growth is most sensitive— does not imply these rates to be most important for explaining temporal changes in population size (Gaillard et al., 1998, 2000b; Coulson et al., 2005). The temporal variability of these rates is at least as important as the potential demographic impact of these rates for them influencing population dynamics. Similarly, it seems unlikely that for habitat selection there would be a direct relationship between the potential limiting factor and selective behavior of the animal (*sensu* Rettie and Messier, 2000), we should expect the environmental spatial variability to play an important role as well.

The spatial variability of environmental features is often scale dependent (Turner et al., 2001). The scale of variability can then be expected to be crucial in our understanding of habitat selection across spatial scales. However, until very recently no attempts have been made to understand changes in spatial variability of resources across scales in habitat selection studies. The use of variograms is one approach that allows the comparison of variability across scales, see Mayor et al. (2007) for an application in habitat selection. Both Morin et al. (2005) and (Harvey and Weatherhead, 2006) discussed the low spatial variability at the level of the home range as a possible explanation for the absence of habitat selection by the animals at this level.

The time frame of habitat selection studies not necessary corresponds with the one used in population ecology to determine limiting factors. Habitat selection is found to change with for instance reproductive status (Loseto et al., 2006; Candolin and Voigt, 2003) such temporal changes in habitat selection have lead Wittmer et al. (2006) to suggest temporal changes in limiting factors. Limiting factors for an animal may show temporal variation, due to either environmental variability or changing requirements of the animal. Seasonal and even diurnal changes can cause previously suitable areas to become unsuitable and *vice versa*.

Not only environmental change can lead to temporal change in limiting factors, also the animal's requirements can change. Reproductive status can alter the energetic requirements or the need for security. For instance McLoughlin et al. (2002) showed for grizzly bears that females with cubs avoided areas preferred by males; whereas this did not occur in females without young. They interpreted these results as a consequence of infanticide avoidance by females. Also at smaller time scales changes in requirements can occur; safety is more important for resting animals, whereas feeding animals will have higher demands for forage quality and availability. Consequently there might be a mismatch between the temporal range of the limiting factor for population growth and the limiting factor for habitat selection.

By assuming that an animal can only select its habitat in relation to one limiting factor at a given spatial scale, Rettie and Messier (2000) implicitly assumed a perfectly negative relationship between the occurrence of limiting factors in space. Negative correlations between different environmental factors potentially influencing an animal's fitness are not uncommon, especially the trade-off between food and cover has received considerable attention in the literature (e.g. Edwards, 1983; Mysterud et al., 1999). However, different studies report simultaneous habitat selection in relation to more then one limiting factor at the highest hierarchical level (e.g. Dussault et al., 2005; Gustine et al., 2006). This indicates that the choice of environmental characteristics in these studies is not exclusive and animals can select for areas providing good conditions with respect to more then one limiting factor simultaneously. Such suggests that spatial correlations between resources are not always perfectly negative, they might even be positive.

The fast acceptance of Rettie & Messier's (2000) proposition demonstrates its intuitive appeal, despite the fact that its initial presentation lacked a thorough development of its underlying logic. However, their theory seems to show three major incompletenesses: the spatial variability and temporal variability of limiting factors, and the co-variability of different limiting factors. In this paper we will further develop a theory of hierarchical habitat selection. Using simple simulated environments we develop additional ideas on ideal habitat selection at different spatial scales in the absence of social interactions. Hierarchical habitat selection will be considered at two levels: home range selection and the selection of patches —represented by pixels in our landscapes— within the home range (respectively, level 2 and 3 of Johnson's [1980] hierarchy); we will refer to them in the following by respectively home range selection and patch selection. The logic used at these levels, can be readily generalized towards finer levels, like food item selection.

3.2 Spatial variation of resources

The effect of spatial variation of the resources on hierarchical habitat selection is illustrated using a simplified system of selection occurring in simulated environments with only one potentially limiting factor, for instance, food or cover. In this paper limiting factors can be either living conditions or resources for an animal —resources are by definition consumable, whereas conditions are not— only spatial heterogeneous limiting factors are considered. This treatment does not deal with depletion; thus, resources and conditions can be treated similarly and are both referred to by the term resources. We will assume the quantity of a resource transformed as to assure a linear relationship with fitness; landscapes are further simplified by the binary resource content of a pixel: a pixel eithser contains the resource or it does not.

Spatial variation can be scale dependent; to demonstrate the role of variability at different spatial scales we simulated two landscapes with different degree of resource clustering (see Fig. 3.1). In one environment each pixel had an equal probability to contain the resource, called the random environment. The other environment was clumped (Moran's I of 0.75), this environment was generated by iterative permutation of the pixels in the random environment. The common geostatistical techniques to represent variation of a spatial environmental variable across scales is by the use of variograms or correlograms (Turner et al., 2001). One important characteristic of a correlogram is its range, this is the lag where the spatial autocorrelation is no longer different from zero. In other words, the range of a correlogram is the range of spatial autocorrelation; beyond this range, the spatial variable is no longer autocorrelated. Fig. 3.1 demonstrates how increased clumping of the environmental variable leads to an increased range in the correlogram.

The potential for selection depends on the available variability —without variation selection can not occur; for home range selection the variation between potential home ranges is important, whereas patch selection within the home range relies on variation within the home range. Thus, before investigating the effect of habitat selection we first investigated the potential for selection by looking at the available spatial variability in both environment with contrasting ranges, hereafter we studied the effect of home range size on this variability, finally we added more realism by investigating the effect of spatial variability at different spatial scales simultaneously.

3.2.1 How does spatial scale affect the potential for habitat selection?

We placed 25 home ranges at random on both landscapes in Fig. 3.2, for each home range we determined the mean resource content and its variation. We then compared —using the F-statistic— the variation of home range quality (i.e. the average resource content of a home range)

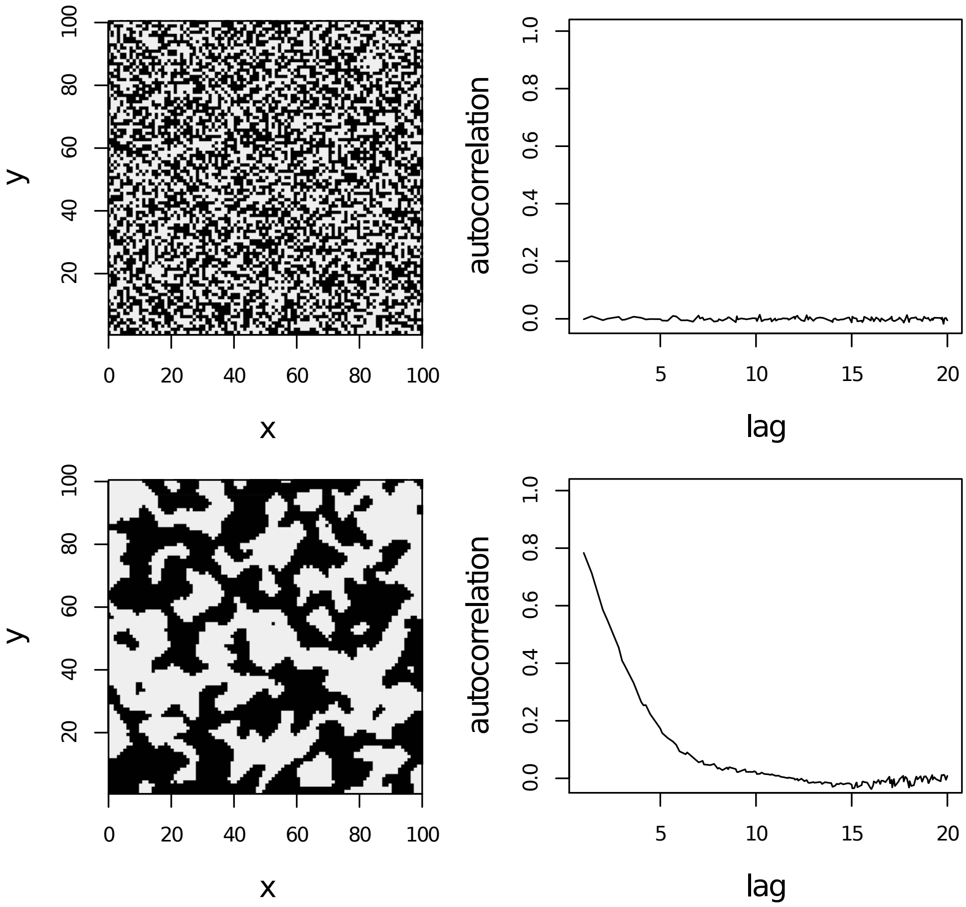


Figure 3.1: On the left hand you see simulated environments, on the right hand the corresponding correlogram is shown. The upper panels are from an environment with a Moran's I of 0, whereas the lower panel has a score of 0.75.

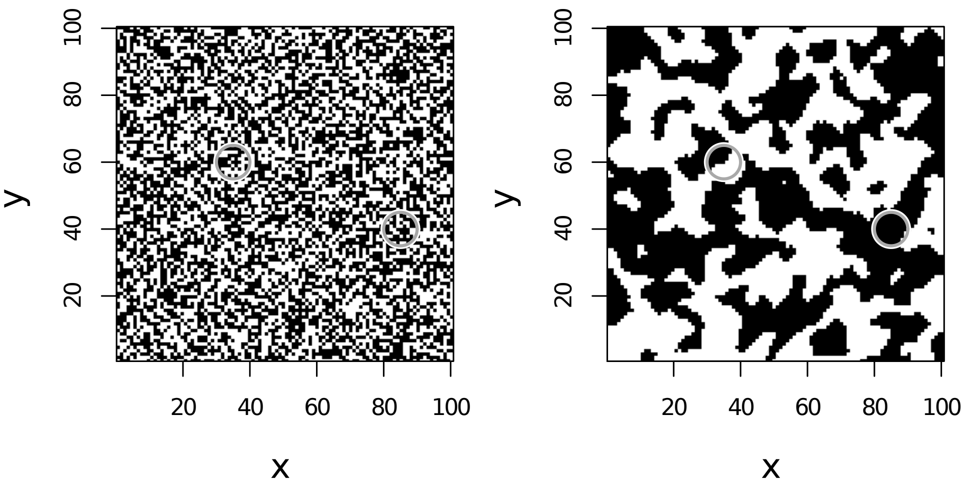


Figure 3.2: On both environments we simulated a home range. At coordinates (90, 40) you see the occupied points, points are only on black pixels. For illustration purposes we depicted a random "home range" of the same dimensions as the real one at a random location in the landscape (40, 60). We simulated this way 25 random home ranges. The variability of resources within the home range is zero (or without selection mean variability is 0.41) in the clumped landscape (similar to results by Mitchell and Powell, 2004), whereas the variability in the mean resource content between potential home ranges is 0.26. For the random landscape a reverse pattern emerges: within the home range the variability is 0.50, the mean between home ranges however is very low 0.05.

between both landscapes to assess the landscape differences in variation at the home range level. The differences in mean within-home-range variability between both landscapes were compared using a t-test, which gives the potential for patch selection.

The variation of home range quality is significantly higher in the clumped than in the random landscape (sd is respectively 0.29 and 0.05; $F = 0.04$, $df = 24$, p-value < 0.001); thus, in the clumped landscape home ranges are more different. Whereas, the within-home-range variability shows the reverse pattern, the variability of the patches within a home range is significantly lower in the clumped landscape than in the random landscape (the mean sd is respectively 0.40 and 0.50; $t = 5$, $df = 24$, p-value < 0.001); in other words, home ranges in the clumped landscape are more homogeneous.

The within-home-range variability of the clumped landscape is lower than for the landscape as a whole (the mean sd of the home ranges and the sd of the landscape are respectively 0.40 and 0.50; $t = -5$, $df = 24$, $p\text{-value} < 0.001$) whereas this is not true for the random landscape (the mean sd of the home ranges and the sd of the landscape are respectively 0.50 and 0.50; $t = 0.13$, $df = 24$, $p\text{-value} = 0.90$). A consequence of these patterns is that in the random landscape selection of the home range is of little to no importance, whereas selection within the home range will be of crucial. Whereas, the reverse is true for the clumped landscape; careful selection of the home range is important and will lead to rather homogeneous home ranges, which makes selection within the home range of less importance.

3.2.2 How does home range size affect the potential for habitat selection?

Whether the landscape is either clumped or random with respect to home range and patch selection will not only depend on the resource distribution, but also on the size of the home range. To investigate the effect of home range size we increased the home range size used previously. We placed again 25 home range randomly on the clumped landscape in figure 2, however this time the home ranges are approximately 4 times as large (using a radius of 10 instead of 5).

In the same clumped environment the variation of home range quality becomes significantly lower for larger home ranges (sd for the small and the large home range quality is respectively 0.29 and 0.17; $F = 2.87$, $df = 24$, $p\text{-value} < 0.05$); thus, in the same environment the larger home ranges will be more similar. On the other hand, the mean within-home-range variation has increased significantly for the larger home ranges (the mean sd for the small and large home range are respectively 0.40 and 0.47; $t = -3.25$, $df = 28$, $p\text{-value} < 0.005$); in other words, larger home ranges are more heterogeneous in the same landscape. Thus, in this landscape animals with the smaller home ranges will have to be more selective in their home range selection, whereas animals with larger home ranges should be more selective in their patch selection within the home range.

Correlogram analysis of the resource —which shows how variability changes with spatial scale— in relation with the spatial extent of the home range shows whether the variability is relatively larger between or within home ranges. When the range of the correlogram from the resource is small in comparison with the size of the home range, the variation between home ranges will be small and there will be important variation within the home range. On the other hand, when the range of the correlogram is large, the home ranges will be more different; if in addition the spatial autocorrelation is high, then the home ranges will be rather homogeneous regarding their resource content. This can be easily seen when we compare the two home range sizes we used in the previous paragraph with the range of the clumped landscape. The range of this landscape is around 7 units (see Fig. 3.1); thus the range of this landscape is larger than the radius of the small home range (5 units), whereas this range is smaller than the radius of the large home range (10 units).

3.2.3 How does large range affect the potential for habitat selection?

Correlograms of environmental variables of real landscapes do not necessary have a limited range, studies have reported spatial heterogeneity over a wide range of spatial scales (Milne, 1988, 1991). Fractal landscapes have been defined as landscapes where the range is unlimited; more precisely, a landscape is called fractal when the log-log variogram is linear. We simulated a fractal landscape using the midpoint displacement procedure described by Fournier et al. (1982), which we then rendered binary using a median split (Gardner, 1999). Fig. 3.3 shows an example of a simulated fractal landscape (r-square of the log-log variogram is 0.99; the fractal dimension is estimated as 2.81); thus there is not a limited range.

To assess the variability between and within home ranges for landscapes with spatial variability across a large range, we simulated 25 home ranges of each size (radius 5 and 10) randomly on the fractal landscape. We found again that due to the larger range of environmental variability in comparison with the home range size (radius of 5 units), the patch variability is smaller than for the environment at large (mean sd of the patches within the home range and the sd for the landscape are respectively 0.37 and 0.5; $t = -4.28$, $df = 24$, p-value < 0.001). The unlimited range of the fractal landscape results in a not significantly different patch variability between both home ranges sizes (mean sd of the patches within the small and large home range are respectively 0.37 and 0.43; $t = -1.84$, $df = 34$, p-value $= 0.08$). The larger home range (with 10 units radius) has still

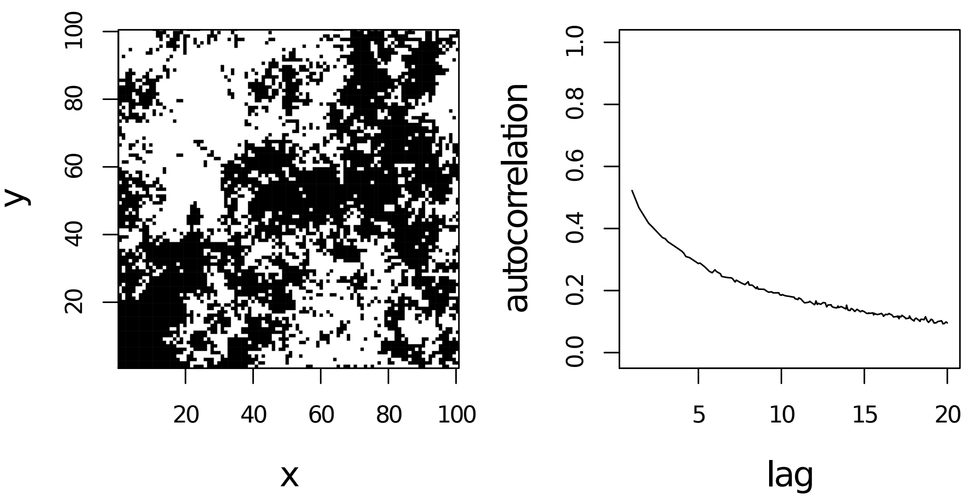


Figure 3.3: A fractal landscape and its correlogram.

a patch variability lower than the environment (mean sd of the patches within the home range and the sd of the landscape are respectively 0.43 and 0.5; t = -4.83, df = 24, p-value < 0.001). Also the variability of the overall home range quality is not significantly different for both home range sizes (sd for the small and large home range quality is respectively 0.31 and 0.24; F = 1.71, df = 24, p-value = 0.19).

In this fractal environment with spatial variation at different spatial scales, the effect of home range size is less pronounced. All home range sizes showed important between home range variation and also significant patch variation within the home range. In these environments, irrespective of home range size an animal should be selective at both home range and patch level.

3.2.4 How does actual habitat selection affect scale-specific spatial variability?

When randomly placed home ranges show high variation in quality, it pays for an animal to be selective in its home range selection. We simulated selection by choosing the 10 percent home ranges (with 5 unit radius) with the highest quality from 250 random ones on three landscapes: random, clumped and fractal (Figs. 3.2 & 3.3).

Both landscapes with a high between home range variability, i.e. the clumped and the fractal landscape, show a significantly reduced within-home-range variability in the selected home ranges versus all random home ranges: for the clumped landscape (mean sd of the patches within the random home ranges and the selected ones is respectively 0.40 and 0.17; t = 10.45, df = 23, p-value < 0.001) and for the fractal landscape (mean sd of the patches within the random home ranges and the selected ones is respectively 0.38 and 0.15; t = 10.51, df = 28, p-value < 0.001). Also the random landscape showed such a reduction although only a small one (mean sd of the patches within the random home ranges and the selected ones is respectively 0.50 and 0.49; t = 6.75, df = 28, p-value < 0.001); the small effect size is a consequence of the small difference between home ranges, hence selection does not have a large effect.

The reduction of patch variability within the home range following home range selection is the simple consequence of the parabolic relationship between the mean and the variance. The highest variance occurs for intermediate mean values; whereas the extreme mean values show low variability. Thus, selection for a resource at the home range level will further reduce the variability of this resource at the lower patch level.

3.2.5 Summary

In summary, resources with a large range in comparison with the home range size will result in home-range variation. When there is home-range variation animals should select for better home ranges. Both the smaller extent of the home range in comparison with the landscape and home-range selection lead to a reduced variability of the resource at the interior of the home range. Following the reduced variability of the patches within the home range as a result of home range selection, it is clear that home-range selection has potentially a larger effect on the animal's performance than patch selection within the home range. The importance of selection within the home range will depend on the homogeneity of the selected home ranges. Thus whether a resource should have a wide range for habitat selection (*sensu* Rettie and Messier, 2000) will depend on the range of this resource in a correlogram.

3.3 Temporal variation of limiting factors

In the Introduction we discussed how limiting factors might change through time due to environmental changes or changes in the requirements of the animal. Above we showed how in general selection for a resource at the home range level leads to decreased spatial variation of this resource at the patch level. When limiting factors change through time this decreased variation limits the opportunities to escape the new limiting factor after a change in limiting factors occurred. When animals are faced with temporal change in limiting factors, two alternatives exist for simply "sweating the bad times": choose a home range that provides a good compromise or move back and forth between two "temporal" home ranges after each change of limiting factor.

An important characteristic of many landscapes is that the scale of spatial variation is not constant. Landscapes often exist of zones with high spatial autocorrelation, and transition zones between these characterized by low spatial autocorrelation. We replicated this characteristic in a simulated landscape consisting of two homogeneous zones with a heterogeneous transition zone between both (see Fig. 3.4). On this landscape we simulated 250 home range pairs, each home range with a radius of 5 units and both home ranges of a pair separated by 50 units. These home range pairs provide an animal with two alternatives: it can choose either for one of both home ranges or it can switch between both "temporal" home ranges when the limiting factor changes. Changes in home range, however, do come at a cost, due to for instance energy expenditure during the movement or increased predation risk during transit. This cost will likely influence the outcome of the choice. It is important to note that when the temporal change of limiting factors occurs at high frequency, home range switching should also occur at high frequency, which results ultimately in a higher cost of movement. From the 250 choice sets the 10 percent sets with the highest outcome were selected, thus animals did not place their home ranges at random they selected the best ones.

In determining the outcome for each home range pair we assumed a perfect negative correlation between resource content of a pixel before and after the change of limiting factor. For example, for an animal, searching for food in a foraging period and for cover during a resting period, this

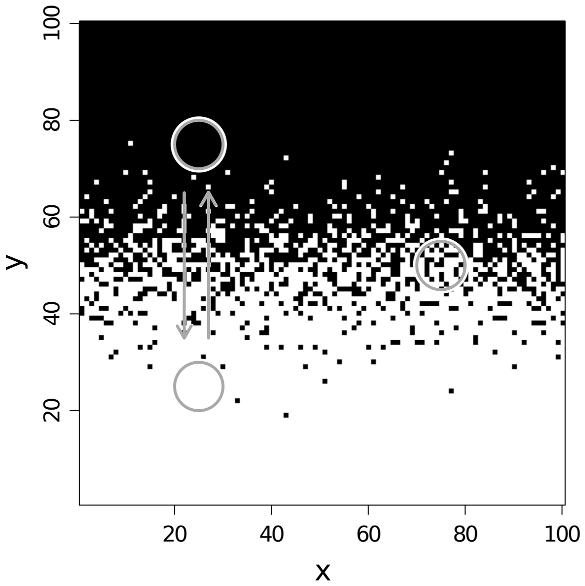


Figure 3.4: Home range placement in two communities with a transitional zone (i.e. ecotone). On the right a typical home range is shown that will be selected when the limiting factors show temporal variation and the movement is constrained (due to for instance high movement costs). On the left two typical "temporal" home ranges are shown that will be selected when the limiting factors change and the movement is not too costly. In this later case, the animal will move between both home ranges depending on which one matches best the animals requirements.

means that a pixel containing food does not contain cover and *vice versa*. The consequence of this simplifying assumption is that the distribution of both resources is the same, they are simply the inverse of each other. The quality of one home range from the pair is then given by the product of both resources. The quality of both home ranges together, as a pair of temporal home ranges, will be determined by the amount of one resource in the first home range and the amount of the other resource in the second home range. The quality difference between the pair of temporal home ranges and the best single home range determines the movement cost the animal should be willingly to pay to occupy two "temporal" home ranges.

As we have discussed in the previous section in the absence of temporal changes of the limiting factors, animals select for the single home range of the highest quality. The selected home ranges have a lower patch variability of the resource within the home range than the available random home ranges ($t = 18.42$, $df = 249$, $p\text{-value} < 0.001$). This replicates the reduction of patch variability due to home range selection in the absence of temporal changes of the limiting factor.

However, when there is temporal change of limiting factor, then the cost of movement will strongly affect the outcome. Without a cost, all selected home ranges consist of a pair of "temporal" home ranges. The animals will shift home ranges according to the changes in limiting factors. The patch variability in both "temporal" home ranges is lower for the 10 percent selected ones ($t = 11.76$, $df = 87$, $p\text{-value} < 0.001$). Thus, temporal changing limiting factors lead to the selection for two different homogeneous home ranges, each containing a different resource.

When the cost of movement outweighs its benefit, animals settle for a single home range that fulfills the need for both limiting factors. The result of this selection is an increased within-home-range variability ($t = -24.01$, $df = 257$, $p\text{-value} < 0.001$). Thus, instead of selecting for a homogeneous home range, animals should select for a heterogeneous home range when confronted with changing limiting factors when the movement is constrained due to for instance high movement costs.

The cost of movement will increase if either the distance that has to be covered is high or the distance has to be covered regularly. If spatial autocorrelation of the resources is high (i.e. as in a highly clumped environment), then an animal will have to move larger distances to reach areas with a different composition of resources. Thus, in such highly spatial autocorrelated landscapes, unless the advantage of changing home ranges is very high, we can expect animals to select for heterogeneous home ranges on transition zones (like ecotones). Similar, when the frequency of temporal changes in limiting factors is high, the animal will have to move very often, thus, unless the gain is high, animals should select in this case also for heterogeneous home ranges. In other words, when the temporal autocorrelation of the limiting factors is low, then there will be selection for regions with lower spatial autocorrelation, and *vice versa*. When movement is costly we expect thus a positive relationship between the temporal autocorrelation of the limiting factors and selection for spatial autocorrelation of the resources related to these factors.

3.4 Co-variation of resources

In the discussion above we assumed a perfect negative correlation between both resources, here we will discuss the consequences of generalizing towards non-perfect correlations. The perfect negative correlation between resources assumed before seems an idealization; different field studies reported habitat selection for more then one resource simultaneously, which suggest a non-perfect negative correlation or even a positive relationship. In the following we will focus on negative correlations, i.e. trade-offs; the logic can however also be applied to positive correlations. Trade-offs are most interesting as they present a conflict in the animal's choice.

Faced with a trade-off between two resources an animal can either select for one of both resources or it could try to optimize both at the same time. If it simply selects for one, then the other resource does play no role at all, a situation we discussed before. An animal can only try to optimize both at the same time when the correlation is not perfect; because, when the correlation is perfectly negative, a pixel containing one resource will be void of the other one. Optimization of a trade-off between two resources can be seen as maximization of a function of these resources. Tilman (1980) discussed different forms of interactions between two resources in the following we only consider perfectly substitutable and interacting resources. The result of two perfectly substitutable resources correspond with adding both resources, whereas the combination of two interacting resources is the product of both. The trade-off between two resources will thus be another variable, which we will call the "trade-off", with its own spatial distribution.

3.4.1 What is the distribution of a trade-off between two similar distributions?

Landscapes with transition zones (Figs. 3.4 & 3.5) show regions with high spatial autocorrelations and regions with low autocorrelations. As we saw in the section on temporal changing limiting factors with costly movement, when the temporal autocorrelation of the liming factors is low the animal should select for a heterogeneous home range in a transition zones, whereas when the this autocorrelation is high it should migrate between homogeneous regions. Unless, the cost of migratory movements become too large, then it should always settle in the transition zone. However, when the animal tries to optimize two resources simultaneously, movement can not solve the problem since there is no time to move as the resources are chosen simultaneously.

Both resource landscapes in Fig. 3.5 are generated by a similar general process, resulting in a transition zone in the same area. There is however a negative correlation between both resources, albeit not a perfect one. As previously, we simulated selection by taking the 10 percent home ranges with the highest quality from 250 home ranges located randomly in the landscape. We tested the effect of home range selection on the distance from the transition zone between both relatively homogeneous areas. Selection for the best compromise based on either the sum ($t = 6.83$, $df = 36$, p-value < 0.001) or the product ($t = 16.16$, $df = 103$, p-value < 0.001) make the animal favoring home ranges closer to the transition zone; this transition zone offers a higher occurrence of pixels showing a beneficial compromise. These chosen home ranges show an increased variation for both perfectly substitutable ($t = -9.17$, $df = 42$, p-value < 0.001) and interacting resources ($t = -24.34$, $df = 266$, p-value < 0.001); thus selection of patches within the home range will further increase the efficiency of home range use. Interesting to note is an important difference between substitutable and interacting resources with respect to the importance of home range selection. For substitutable resources both the worst and the best home ranges occur in the transition zone, with both homogeneous

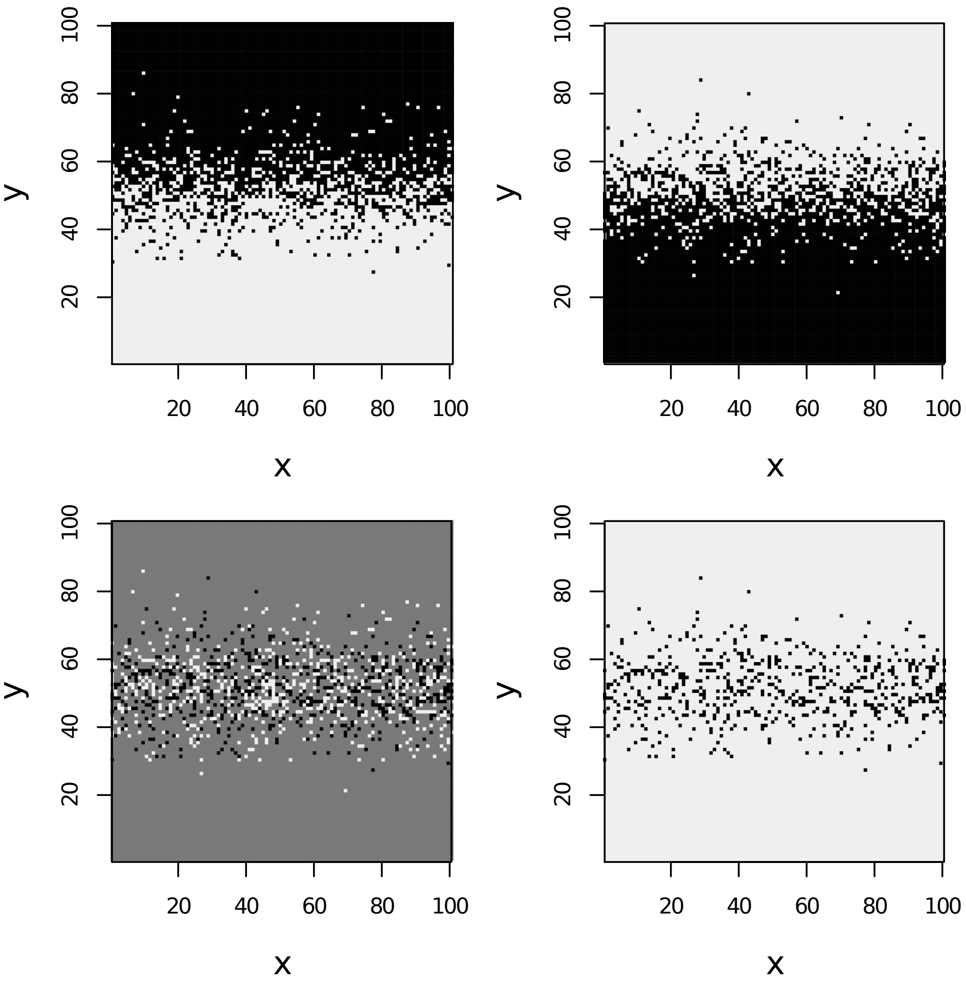


Figure 3.5: The trade-off between environments showing a gradient. The correlation between the environment on the left in the bottom panel and the environments on the top panel is negative (r = -0.84), although not perfect. The lower right panel shows the trade-off between the environment on the left and the environment in the upper panel (determined by multiplication). It can be seen that the transition zone offers the most pixels with beneficial trade-offs.

zones containing intermediate home ranges. Thus, home range selection in the transition zone will be critical. Whereas, for interacting resources, all home ranges in the transition zone are of better quality than in the homogeneous zones. Home range selection in the transition zone will improve the home range quality but is not as crucial as for substitutable resources.

3.4.2 What is the distribution of a trade-off between two contrasting distributions?

Two resource landscapes were simulated with different degrees of spatial clustering (Moran's I 0.25 versus 0.75), and with a negative correlation between both ($r = -0.37$; $t = -40$, $df = 9998$, p-value < 0.001). Fig. 3.6 shows the distribution of a trade-off between both resources (as similar results were obtained for substitutable and interacting resources we only present the results from the later). We discussed above the influence of spatial distribution on habitat selection across scales. Fig. 3.6 demonstrates how the range of the trade-off is determined by the most clumped resource (i.e. the one with the largest range) and that on the other hand the correlation at lag 1 is determined by the lowest correlation at lag 1 (aka the nugget) from both resources. As a result, home range selection (10% with the highest home range quality with respect to the trade-off from 250 random home ranges) will be completely determined by the distribution of the most clumped resource; selection based on the trade-off leads to a significant increase of this highly clumped resource (home range quality of the highly clumped resource in the random and selected home ranges is respectively 0.51 and 0.90; $t = -14.90$, $df = 77$, p-value < 0.001), whereas no such increase, even a decrease occurs in the less clumped resource (home range quality of the lowly clumped resource in the random and selected home ranges is respectively 0.49 and 0.44; $t = 3.77$, $df = 35$, p-value $<$ 0.001). Consequently, there is no reduction in the variability of the patches for this lowly clumped resource (patch sd of the random and selected home ranges is respectively 0.49 versus 0.50; $t = -2.16$, $df = 47$, p-value $=$ 0.04), whereas as expected this reduction did occur for the highly clumped resource (0.40 versus 0.27; $t = 4.88$, $df = 26$, p-value < 0.001). Due to the low autocorrelation of the trade-off at small lags (see Fig. 3.6) there is high within home range variability left (average sd within the selected home ranges = 0.50), thus within home range selection will still be very important. The within-home-range selection should occur mainly based on the resource with the highest variability at this level. The correlation within the selected home ranges between the patches with a beneficial trade-off and the lowly clumped resource is significantly higher than for the highly clumped one (for the lowly and highly clumped variable this correlation is respectively: $r = 0.86$ versus 0.25; $t = -18.65$, $df = 42$, p-value < 0.001); whereas at the landscape level the correlation is for both resources the same ($r = 0.43$). Thus, facing trade-offs animals should select at the larger scale for the large-scale resource, whereas at the smaller scale they should select for the small-scaled resource.

3.4.3 Summary

In summary, the spatial structure of the trade-off will depend on the spatial structure of both resources. Its range will be the largest range available in the resource, whereas its autocorrelation at lag 1 will be largely determined by the smallest autocorrelation in the resources. Animals trying to optimize a trade-off should select at the largest scale for the large-scale resource and for the small-scale resource at the small scale. Alternatively, when the environment contains a gradient, then optimizing a trade-off between two limiting factors that are not perfectly negatively correlated could happen by selecting for transition zones as these are the regions where the most locations will occur with a beneficial outcome for the trade-off.

3.5 Conclusion

Rettie and Messier (2000) proposed that animals should select for the most limiting factor at the highest scale. They, however, did not explicitly take into account the variability of the limiting factors. As reported for the population dynamics of species, factors for which population growth is less sensitive can compensate this by a higher variability and still play an important role in explaining population dynamics. Similarly, higher variability might result in increased habitat selection for a resource despite it being not related to the most limiting factor; variability is, however, scale dependent. Animals should select for the resource that has a combination of high sensitivity and high variability, i.e. with a high selective pressure.

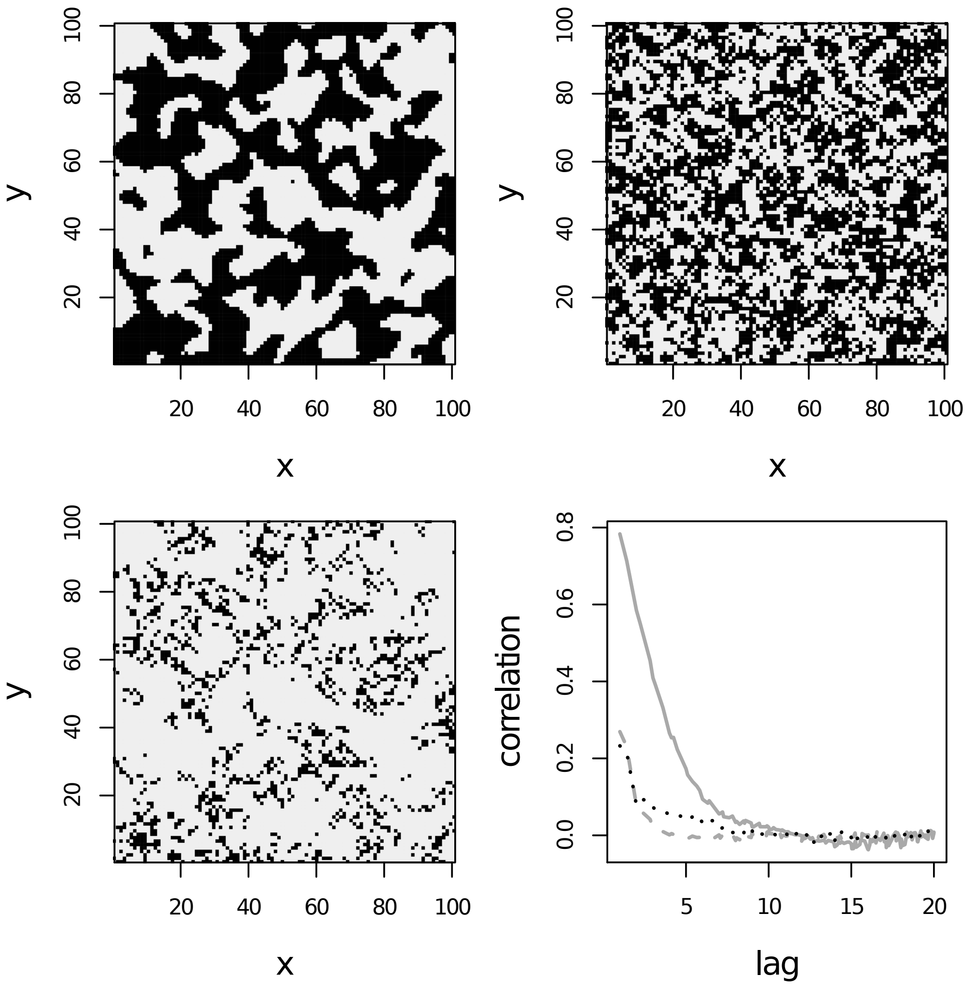


Figure 3.6: The correlogram for a trade-off between two limiting factors with different spatial structure. The two upper panels show the distribution of two limiting factors with different spatial structure (Moran's I of 0.75 and 0.25 for respectively the environment in the left and right panel). The lower left panel shows the distribution of the trade-off, which is the result of multiplying the two upper panels. The lower right panel shows in grey the correlogram for both upper panels, in black the correlogram for the trade-off is shown. Note that the range of the trade-off corresponds with the range of the factor with the longest range and that the correlation at lag 1 is determined by the factor with the lowest correlation at lag 1.

Selection for a resource at one scale in general leads to reduction of variability at the lower scale; the amount of reduction depends on the spatial distribution of the resource. Thus, depending on the distribution of the resource, the variable should also be selected at the smaller scale. Concrete, variables with a high spatial autocorrelation and large range should be narrow and only selected at the home range level, since within the home range this variable will be rather homogeneous. Whereas, variables with a small range will only be selected within the home range. Variables with a large range and lower spatial autocorrelation, i.e. high variability at the small scale, should be selected for at different spatial scales; these resources have a wide range for habitat selection.

Whereas home range selection for a resource in general leads to increased homogeneity of this resource within the home range, we have identified two situations when selection for increased heterogeneity within the home range might occur. First, when animals are faced with limiting factors that are temporally variable then selection for heterogeneous home ranges can occur, especially when these changes in limiting factors show low temporal autocorrelation. Second, when animals are optimizing a trade-off between two resources, home range selection should favor heterogeneous areas as there are more spots providing beneficial solutions for this trade-off.

Previous studies on hierarchical habitat selection focused almost exclusively on the change of habitat selection across scales, without considering the changes in variation of resources and co-variation between these resources across different spatial scales. Tools like variograms and correlograms provide promising avenues to increase our understanding of resource selection across spatial scales by incorporating these changes in scale-dependent variability in the analysis.

Part II

Niche analyses

"The old truth is that most data sets have 2.5 dimensions: those two that you can show in a printed plot, and that half a dimension that you must explain away in the text. Wouldn't that be a sufficient solution?"
Jari OKSANEN.

THIS AXIOM OF THE 2.5 DIMENSIONS, ATTRIBUTED TO KRUSKAL, is a very good summary of this Part of the thesis. During my Master year, I mainly focused on how to reduce a niche dimensionality in order to interpret it more easily. In other words, I was interested in reducing it to a dimensionality small enough to be plotted. My master work dealt with the ENFA ("Ecological-Niche Factor Analysis", Hirzel et al., 2002a; Basille et al., 2008), which is a method based on a Design I, (both use and availability data collected at the population level, Thomas and Taylor, 1990, 2006). At the same time, Clément CALENGE was working on Designs II and III (use data collected at the individual level, availability data collected either at the population level or at the individual level), specifically on the K-select analysis (Calenge et al., 2005). We were both working on how to study and represent niches in the ecological space, he as a biometrician, me as a biologist with some methodological skills. Here started our joint work.

Above all, I have to reveal that this work preceded the previous Part in time, especially Chapter 2, relative to the habitat and niche confusion. While it is more logical to present the tools after the theory, it might present some inconsistencies regarding the niche and habitat concept. Concrete, Clément and I worked on the study of habitat selection with the model presented by Hutchinson (1957) for his niche concept, following other authors (e.g. Doledec et al., 2000; Hirzel et al., 2002a). This apparent contradiction can be quickly resolved, by acknowledging that every method applicable to the niche can be automatically applied to the habitat (and *vice versa*). Thus, for the following three chapters, the niche *sensu* Hutchinson used as a model is to be understood as the conceptual abstraction of the niche. In other words, we focused on the geometrical representation of it instead of its biological meaning. However, whereas it does not solve the probable mistakes in the text, it is important to note that the three analyses detailed hereafter can interchangeably be used for niche or habitat studies.

The premises: the ENFA

When I started my Master work, in 2003, the ENFA was a brand new multivariate method, carrying great promises. The ENFA was actually an attempt to split all niche dimensions into a few relevant dimensions of marginality and specialization. The marginality expresses the departure of the ecological niche from the average available habitat (in other words, the marginality is a multi-dimensional version of the mean), whereas the specialization expresses the narrowness of the niche (that is, a multi-dimensional version of the variance). With an approach similar to a PCA (Principal Component Analysis), the ENFA was able to extract first one axis of marginality, and second, several axes of specialization. The method was then used to compute habitat suitability maps (maps that give the probability of a species to be in every place) based on the position of the niche on these axes.

However, while the ENFA was very similar to a PCA, it did not use one of its strength: factorial maps, and in particular biplots (i.e. factorial maps with the projection of both variables and individuals at the same time). Our first step was thus to enable this kind of plots. As each specialization axis is orthogonal to the marginality axis, it was actually possible to project the scores of both variables and individuals on the orthogonal plane formed by the marginality axis and any of the specialization axes. To anyone familiar with factorial maps, this biplot gives a lot of information at first glance. Details on this are the subject of the next Chapter (Chapter 4).

Beyond this refinement, the method itself was not yet totally satisfactory. By computing first the marginality, the ENFA was actually removing one dimension that could be particularly relevant. This first step was indeed a prerequisite for the specialization extraction; for the specialization to be computed, it needed to be centered twice, both on the niche and on the environment (i.e. use and available data). But the marginality axis was actually carrying a part of specialization by itself. In other words, the niche variance may be smaller on the marginality axis than the available variance. Additionally, the extent of specialization was related to the extent of marginality: the more marginal a niche, the greater potential for specialization (we could say that a very marginal niche has no further room to express a great variance). The specialization could be highly modified depending on the marginality: if the marginality was very high, the marginality axis could carry most of the specialization, and thus lead to specialization axes that could only express a weak residual specialization. On the contrary, if the marginality was very weak, the marginality axis could be oriented in any direction. Thus, the constraint of orthogonality to compute specialization axes may lead them to be not relevant any more. Altogether, the specialization procedure, which is the core of the ENFA (see Chapter 6), was not pure and could just be an artifact of the marginality.

Of course, it was possible to calculate the amount of specialization accounted for in the dimension of the marginality axis, but still this wasn't totally satisfactory as the method was actually extracting axes of different mathematical nature (marginality and specialization).

Development: the MADIFA

This led me to ask Clément the following question: "Can we get rid of the marginality axis?", or in other words: "Is it possible to extract specialization only?" The answer came quite indirectly, but the question as usual led to most interesting trails. While working on the ENFA, we were trying to implement every feature provided by the initial authors (Hirzel et al., 2002a). For instance, we implemented the algorithm of this paper, and while several other ad-hoc algorithms later appeared (Hirzel and Arlettaz, 2003a,b), it became quite clear that the original one was not optimal, and that the authors were looking for one more accurate. We thus started to work on the Mahalanobis distances, a method previously used to derive habitat suitability maps by itself (Clark et al., 1993), as an algorithm to derive such maps on the reduce dimensionality of the ENFA.

This solution did not solve the major problem of the ENFA though: the extracted axes were of different mathematical nature, and it is difficult to combine all these axes into one single index of environmental suitability. We were thus caught between two unsatisfactory alternatives: in order to derive habitat suitability maps, we could either use the ENFA with the Mahalanobis distances algorithm on top of it (or any other ad-hoc algorithm), with the former problem of axes of different nature; or we could use directly the Mahalanobis distances, with the problem of not being able to interpret them regarding the environmental variables included in the analysis. The results of the two approaches were however highly similar, not surprisingly since the core of both methods rely on the same metric.

There was actually so many similarities that Clément started to work on the Mahalanobis distances as a sister-method of the ENFA. The Mahalanobis distances were not yet a classical multivariate analysis, but soon became one. This is the story of the second Chapter of this Part (Chapter 5). Clément was actually able to develop a new method to partition the Mahalanobis distances into a restricted set of biologically meaningful axes. This method, which was called MADIFA ("Mahalanobis distance factor analysis") extracted successive axes that support the greatest Mahalanobis distances, all with the same mathematical nature. This allowed us to compute habitat suitability maps based on a small number of axes, which extracted most of the relevant information.

Synthesis: the GNESFA

The MADIFA was thus a step further towards our aim, but not the final destination. By dint of examining the similarities between the ENFA and the MADIFA, Clément managed to build a common framework that encompasses both methods, the "General niche-environment system factor analysis" (GNESFA). In the process of this unification, a new method appeared, the FANTER ("Factor analysis of the niche, taking the environment as the reference"), which can be seen as the reverse point of view of the MADIFA, applied to the same problematic. Thus, the FANTER had the same advantages as the MADIFA: the FANTER is another multivariate analysis, that was extracting uncorrelated axes of the same mathematical nature. With the three methods together, the GNESFA actually allowed an extensive exploration of the niche within its environment (Fig. 3.7), as is shown in the last Chapter of this Part (Chapter 6).

But the best thing was given by the FANTER properties: this method actually extracted axes that support on one hand most of the marginality, and on the other hand most of the specialization. That is the first axes of the FANTER were axes of marginality, whereas the last ones were axes of specialization. With this method, after more than 3 years of work, we were

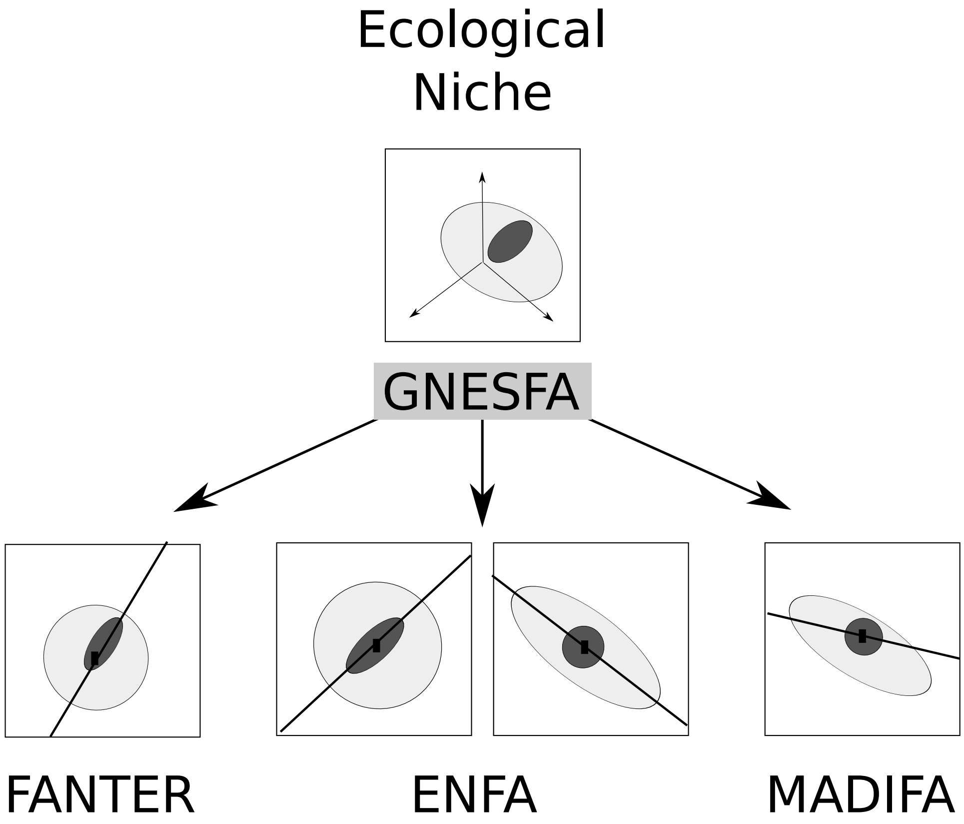


Figure 3.7: A summary of the GNESFA. The GNESFA allows an extensive exploration of the niche within its environment, taking the point of view of the expert of the study area (FANTER), the point of view of the expert of the studied species (MADIFA), or an intermediate approach (ENFA, with two equivalent possibilities). See details in Chapter 6.

finally able to extract specialization axes without worrying about marginality (note that the marginality was still extracted but the interpretation of specialization axes no longer needed to detail the marginality first).

This work on the exploration of the (abstraction of the) niche was typical for a scientific study. I started with a simple question for the biometrician, but we faced new problems which raised in turn new questions, and trying to solve them, we found out new properties and ended up with a global framework that is much more comprehensive than the initial question.

Chapter 4

Assessing habitat selection using multivariate statistics: Some refinements of the Ecological-Niche Factor Analysis

Mathieu Basille[1,*], Clément Calenge[1,2], Éric Marboutin[3], Reidar Andersen[4] & Jean-Michel Gaillard[1]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69 622, Villeurbanne, France.

[2] Office national de la chasse et de la faune sauvage, 95 rue Pierre Flourens, 34 000 Montpellier, France.

[3] Office national de la chasse et de la faune sauvage, Z.I. de Mayencin, 5 allée de Béthléem, 38 610 Gières, France.

[4] Museum of Natural History and Archaeology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.

[*] Corresponding author: `basille@biomserv.univ-lyon1.fr`

Published in Ecological Modeling 211, 2008, pp. 233–240.

"All models are wrong but some are useful."
George E. P. Box.

Abstract

We propose here some refinements of the Ecological-Niche Factor Analysis (ENFA) to describe precisely one organism's habitat selection. The ENFA is based on the concept of the ecological niche, and provides a measure of the realised niche within the available space from the computation of two parameters, the marginality and the specialization. By measuring the departure of the ecological niche from the average available habitat, the marginality identifies the preference of the individual, population, or species for specific conditions of the environment among the whole set of possibilities. The specialization appears as a consequence of the narrowness of the niche on some environmental variables. The ENFA is a factorial analysis that extracts one axis of marginality and several axes of specialization. We present here the use of biplots (i.e. the projection of both the pixels of the map and the environmental variables in the subspace extracted by the ENFA) as a way to identify the key-variables for management, assessing which habitat features are of prime importance and should be preserved or reinforced. With the help of this tool, we are now able to describe much more precisely the habitat selection of the organism under focus. In our application to the lynx in the Vosges mountains, based on sightings as well as other indices of lynx presence, we thus underlined a strong avoidance of agricultural areas by the lynx. We also highlighted the relative indifference of the lynx to the proximity of artificial areas and at the opposite, the sensitivity to the proximity of highways. The ENFA provides a suitable way to measure habitat use/selection under a large range of ecological contexts and should be used to define precisely the ecological niche and therefore identify the characteristics searched for by the organism under study.

Keywords: biplot; ENFA; Lynx; *Lynx lynx*; marginality; presence-only data; specialization; Vosges mountains

ASSESSING THE RELATIONSHIPS BETWEEN INDIVIDUALS in a population and their environment is required in most ecological studies, both from a theoretical and a management viewpoint. In particular the habitat use and the intensity of habitat selection displayed by individuals are likely to influence markedly the response of organisms to density dependence and environmental variation (Gilpin and Hanski, 1991; Tilman and Kareiva, 1997). The increasing availability of advanced tools such as GIS

(Geographic Information System), and the ever increasing power of computers offer the possibility to include much more biological information in the analyses. This allows habitat use/selection to be assessed in a much more precise way (Guisan and Zimmermann, 2000). Consequently, work has been performed to develop new multivariate statistics in order to account for the complexity of the environment.

The lack of absence data in most sampling designs so far applied to study habitat use or selection, is one of the major problems ecologists have to solve (Hirzel et al., 2002a; Soberón and Peterson, 2005). While collecting reliable data on animal presence is straightforward in most case studies, it is difficult to assess the true absence of an animal in a given habitat. Are the animals really absent because the environment is not suitable for the species? Or because the animals did not yet colonize their whole suitable habitat (hunting, history of colonization, demographic stochasticity)? Or do we face an apparent absence because the animals are present, but not detected during the sampling, or temporarily absent (Martin et al., 2005)? For all these reasons, the absence of observation at a given location cannot be reliably interpreted as a true absence, thus we have to rely on the presence data only. The well-known concept of ecological niche (Hutchinson, 1957) provides a suitable way to analyse presence-only data. It is defined as the n-dimensional hypervolume, in which every point corresponds to a state of the environment which would permit the species to exist indefinitely. Each environmental variable then corresponds to a dimension in the so-called ecological space which defines the available habitat for the animals under study. The observed presences are used to assess the utilization of the space by the animals, i.e. the ecological niche (Fig. 4.1). Although originally developed to describe the ecological requirements of a species, the concept of ecological niche can easily be applied to other scales of biological organisation (community for larger scale, see e.g. Doledec et al., 2000, individual for finer scale, see e.g. Calenge et al., 2005). Here, we will focus on the analyses of the distribution of populations of a given species, i.e. corresponding to a second-order selection study according to Johnson's selection order (Johnson, 1980). However, the concept of ecological niche as defined above could be used for the study of habitat selection at all levels (e.g. the selection of the distribution range by a species, the selection of the home range of an animal within a region, the selection of patches

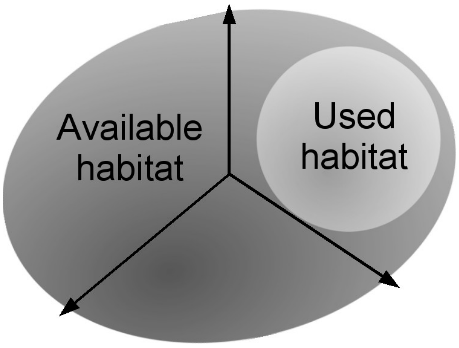


Figure 4.1: Representation of the ecological niche. The arrows identify the environmental variables defining the ecological space. The dark grey cloud stands for the available habitat and the light grey cloud stands for the used habitat, i.e. the ecological niche.

within home range, ...).

Several statistical analyses have been recently developed to assess habitat selection using presence-only data (Elith et al., 2006). These methods can be classified into two complementary approaches (see a review in Pearce and Boyce, 2006): exploratory analyses, which aim at extracting the characteristics of the environment used by a given organism (e.g. Calenge et al., 2005; Doledec et al., 2000) and modelling analyses (Manly et al., 2002). Exploratory analyses can be seen as a required preliminary for modelling analyses as they lead to select the variables of interest to model the habitat. Among these, the ENFA (Ecological-Niche Factor Analysis, Hirzel et al., 2002a) searches for directions in the ecological space so that (i) the difference between the conditions used in average by the species and the conditions available on the study area (i.e. the marginality) is maximised, and (ii) the ratio between the variance of available conditions on the variance of conditions used by the species (specialisation) is maximised. According to the structure of (Austin, 2007), the ENFA belongs to the theory of the ecological niche, relies on presence-only data and takes place in the well-studied family of multivariate analyses. Up to now, biologists have used the ENFA in order to build habitat suitability maps. Such maps rely on the assumption that habitat suitability in a given pixel of a map can be estimated by the probability of presence of the individual, population, or species under study. Several algorithms have been proposed and compared to assess the reliability of the maps (Hirzel and Arlettaz, 2003b,a).

The ENFA has then been used to predict the potential habitat in a large range of animal taxons (insects: e.g. Gallego et al., 2004; cetaceans: e.g. Compton, 2004; birds: e.g. Hirzel et al., 2004; mammals: e.g. Dettki et al., 2003; Zimmermann, 2004), in some plants (e.g. Zaniewski et al., 2002) and rare or endangered species (e.g. Reutter et al., 2003).

However, the usefulness of the ENFA in other ecological contexts have been overlooked. In addition of providing an answer to the where-question ("Where can the organisms establish?"), the ENFA can be used to answer the what-question ("What do the organisms search for?").

The ENFA is indeed suitable to assess the habitat features that are preferred by the individual, population or species under study. Therefore, while valuable, the construction of habitat suitability maps appears to us as a secondary task after having identified the processes behind the habitat use or selection by a given individual, population, or species.

In this paper, we develop the required refinements of the ENFA to reach such a goal, and we show how the use of biplots (i.e. the projection of both the ecological-niche and the environmental variables on the subspace defined by the axes of the ENFA) is an essential step in that direction. As an illustration of the usefulness of our approach, we use these extensions of the ENFA in the study of the habitat selection by the lynx (*Lynx lynx*) in the Vosges mountains (France). The data used corresponds to sightings as well as other indices (scats, hairs, carcasses, ...) of lynx presence, the kind of presence-only data that perfectly fulfills the requirements of the ENFA. All the statistical procedures are implemented in the R-software (R Development Core Team, 2008) within the R-package "adehabitat" (Calenge, 2006).

4.2 The Ecological-Niche Factor Analysis

4.2.1 Description of the design

The available habitat is described by a set of raster maps of the study area, giving the values of P environmental variables in N pixels. Let $\mathbf{Z}$ be the $N \times P$ matrix with the values of the P variables in the N pixels; $\mathbf{Z}$ defines a cloud of points (the available space) in the P-dimensional ecological space (Fig. 4.2A). The $\mathbf{Z}$ matrix is column-centered and scaled so that its variance is equal to 1 and the centroid (barycenter) of the scatterpoint corresponds

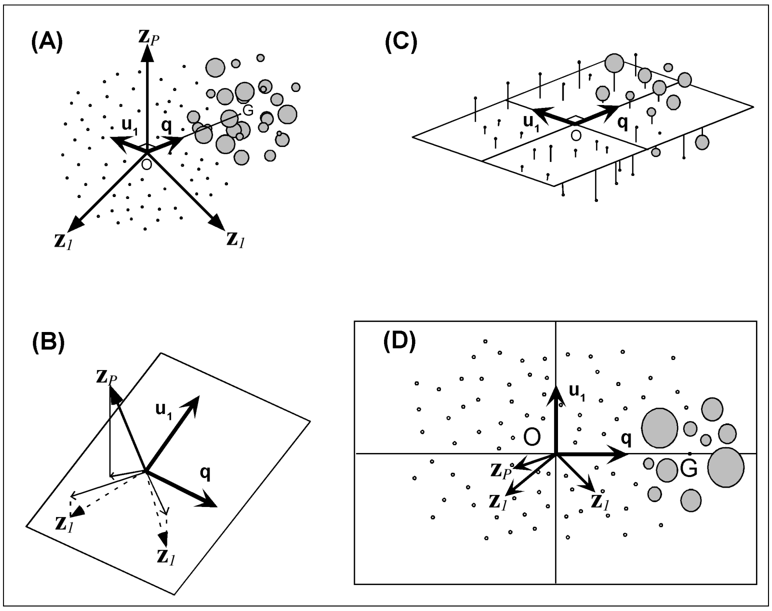


Figure 4.2: Construction of the biplot built from the ENFA. (A) The ecological space is defined here by 3 environmental variables $\mathbf{z}_1$, $\mathbf{z}_j$ and $\mathbf{z}_\mathbf{p}$. The marginality vector $\mathbf{m}$ connects the centroid of the available space O to the centroid of the used space G. The vector $\mathbf{q}$ corresponds to the marginality vector normed to 1. The vector $\mathbf{u}_1$ corresponds to the first vector of specialization. (B) The vectors $\mathbf{q}$ and $\mathbf{u}_1$ are orthogonal and define the plane on which the points are projected. (C) The variables are projected in the same plane. (D) The simultaneous representation of coordinates of the points and the variables on the same plot leads to the best representation of the ecological niche, and an easy interpretation of it.

to the origin O of the ecological space and represents the average available habitat. To each available pixel is associated an "availability weight" describing the availability of the pixel to the species, population or individual (with weights summing to 1; defaulting to $1/N$ for all pixels). Let $\mathbf{D}$ be the $N \times N$ matrix containing these weights on the diagonal (defaulting to $\mathbf{D} = Diag(1/N)$).

The locations of the individual, population, or species sampled define the used habitat. The vector $\mathbf{p}$ of length N provides the proportion of locations in each pixel and defines the utilization weights (with weights summing to 1). Let $\mathbf{D}_\mathbf{p}$ be the $N \times N$ matrix containing these utilisation weights $\mathbf{p}$ on the diagonal: $\mathbf{D}_\mathbf{p} = Diag(\mathbf{p})$.

The points in the available space for which the corresponding utilization weights are upper than zero define the used space, i.e. the ecological niche. Therefore the centroid G of the niche corresponds to the average used habitat (Fig. 4.2A).

4.2.2 Concept of marginality

The marginality is measured as the squared distance from the mean available space to the mean used space. It is geometrically defined as the squared norm of the vector from the origin O of the ecological space to the centroid G of the niche (Fig. 4.2A). Let **m** be this vector:

$$\mathbf{m} = \mathbf{Z}^t \mathbf{D_p} \mathbf{1}_N \tag{4.1}$$

where $\mathbf{Z}^t$ is the transpose of $\mathbf{Z}$, $\mathbf{1}_N$ is a N–vector of 1.

The marginality is then the squared norm of the vector **m**:

$$M = \mathbf{m}^t \mathbf{m} \tag{4.2}$$

The vector of marginality is normed for further analyses:

$$\mathbf{q} = \frac{\mathbf{m}}{\sqrt{\mathbf{m}^t \mathbf{m}}} \tag{4.3}$$

The marginality measures a position and expresses the magnitude of the deviation of the niche relative to the available space (Fig. 4.2A). The higher the marginality, the more the niche deviates from the average conditions of the available habitat. When marginality is high, the individual, population, or species are present in areas displaying quite different habitat features compared to what is available.

4.2.3 Concept of specialization

The specialization measures the narrowness of the niche. The specialization corresponds to the axes on which the ratio of variance of the available habitat to the used habitat is the highest. We are thus looking for the $P - 1$ vectors orthogonal to the vector of marginality that lead to extract most of the specialization. Such a procedure is equivalent to find a vector **u** matching the following conditions:

$$\mathbf{u}^t \mathbf{u} = 1 \tag{4.4}$$

$$\mathbf{u}^t\mathbf{m} = 0 \tag{4.5}$$

$$R = \frac{\mathbf{y}^t\mathbf{D}\mathbf{y}}{\mathbf{y}^t\mathbf{D_p}\mathbf{y}} \quad \text{Max} \tag{4.6}$$

where $\mathbf{Zu} = \mathbf{y}$.

In other words, the vector $\mathbf{u}$ is of length 1, and is orthogonal to the marginality vector $\mathbf{m}$. Because the vector $\mathbf{u}$ is normed, the vector $\mathbf{y}$ is the projection of the rows of $\mathbf{Z}$ on the vector $\mathbf{u}$. The vector $\mathbf{y}$ contains the scores of the pixels projected on $\mathbf{u}$. The condition (4.6) therefore implies that the specialization is maximized on the vector $\mathbf{u}$.

We compute the covariance matrices $\mathbf{S} = \mathbf{Z}^t\mathbf{D_p}\mathbf{Z}$ and $\mathbf{G} = \mathbf{Z}^t\mathbf{D}\mathbf{Z}$ and define $\mathbf{x} = \mathbf{S}^{-\frac{1}{2}}\mathbf{m}$, $\mathbf{b} = \frac{\mathbf{x}}{\sqrt{\mathbf{x}^t\mathbf{x}}}$ and $\mathbf{W} = \mathbf{S}^{-\frac{1}{2}}\mathbf{G}\mathbf{S}^{-\frac{1}{2}}$.

Let the matrix $\mathbf{H}$:

$$\mathbf{H} = (\mathbf{I}_v - \mathbf{b}\mathbf{b}^t)\mathbf{W}(\mathbf{I}_v - \mathbf{b}\mathbf{b}^t) \tag{4.7}$$

If $\mathbf{H}$ is of rank v, then this matrix has v non-null eigenvalues, associated to v eigenvectors $\mathbf{v}_i$. Hirzel et al. (2002a) have shown that the eigenvectors $\mathbf{v}_i$ of the matrix $\mathbf{H}$ are related to the vectors $\mathbf{u}_i$ by the equation:

$$\mathbf{u}_i = \frac{\mathbf{S}^{-\frac{1}{2}}\mathbf{v}_i}{\sqrt{\mathbf{v}_i^t\mathbf{S}^{-1}\mathbf{v}_i}} \tag{4.8}$$

The vector $\mathbf{u}_i$ is the i^{th} vector of the analysis. The eigenvalues λ_i are the values of the specialization on the vectors $\mathbf{u}_i$.

Note that:

$$\mathbf{u}_i^t\mathbf{u}_j = \frac{\mathbf{v}_i^t\mathbf{S}^{-1}\mathbf{v}_j}{\sqrt{\mathbf{v}_i^t\mathbf{S}^{-1}\mathbf{v}_i}\cdot\sqrt{\mathbf{v}_j^t\mathbf{S}^{-1}\mathbf{v}_j}} \neq 0 \tag{4.9}$$

The axes of specialization are therefore not orthogonal.

The specialization measures the dispersion of the ecological niche and expresses the restriction of the ecological niche on some particular directions (Fig. 4.2A). The higher the specialization, the more restricted is the niche in that dimension. A high specialization on a given dimension indicates that the individual, population, or species does not tolerate large variation of the habitat features that mostly determine that dimension.

4.2.4 Identifying the ecological niche

The vectors $\mathbf{q}$ and $\mathbf{u}_i$ provide the scores of the environmental variables on respectively the marginality axis and the specialization axes (Fig. 4.2B). The coordinates of the pixels are defined with $\mathbf{f} = \mathbf{Zq}$ on the marginality axis, and $\mathbf{y}_i = \mathbf{Zu}_i$ on the specialization axes since the vectors $\mathbf{q}$ and $\mathbf{u}_i$ are of length 1. Since the vectors of specialization $\mathbf{u}_i$ are orthogonal to the vector of marginality $\mathbf{m}$ (or $\mathbf{q}$), the plot of $\mathbf{f}$ and $\mathbf{y}_i$ displays the projection of the rows of $\mathbf{Z}$ on the plane $\mathbf{q} - \mathbf{u}_i$, which exactly (i.e. not altered, Fig. 4.2C) corresponds to the best possible 'photograph' of the ecological niche *sensu* Hutchinson (1957).

We can project the used and available points in the ecological space on the plane defined by the marginality axis and one specialisation axis to obtain a biplot (Fig. 4.2D) in the sense of Gabriel (1971). This biplot is of primary help to assess the habitat selection, with respect to the marginality and the specialization. On the biplot, the environmental variables are represented by an arrow with 2 components of importance: the length and the direction. The length of the arrow identifies the contribution of a given environmental variable to the definition of the axes of the ENFA, i.e. their influence on the position and volume of the ecological niche within the available habitat. The direction measures how this contribution is decomposed on the marginality or specialization axes. The first step is then to identify the variables which correspond to the longest arrows. These are the critical variables in terms of habitat selection. In the second place, the relative contribution of the marginality or specialization is assessed by the coordinates of the arrow on the corresponding axis.

4.2.5 The relationship between the marginality axis and the specialization axes

It is noteworthy that the specialization is constrained by the marginality: all the specialization axes are orthogonal to the marginality axis, but not to each other. To conserve the distances and angles of the projection of both variables and pixels in the biplot, the use of an orthogonal base is required. Therefore, we can only use the plan formed by the marginality axis and any specialization axis to compute the biplot with a representation of the niche not altered, i.e. the distances and angles between points are exact.

Conversely, the specialization axes are not necessarily orthogonal so that the projection of the niche in the plan formed by two specialization axes will be twisted because of the straightening of the specialization axes in the biplot. Note that the marginality axis in itself expresses some specialization (the higher the marginality, the higher is the specialization, due to the departure of the ecological niche from the centroid of the ecological space). If the ecological niche is most narrow in the dimension of the marginality axis, the main part of the specialization will already be taken on this axis. The first axis of specialization which is next extracted is constrained to be orthogonal to the marginality axis and will not explain the main part of specialization but the remaining part of it, thus resulting in a meaningless analysis of specialization. However, it is possible to estimate the specialization accounted for in the dimension of the marginality axis: it is given by the ratio of variances projected on the marginality axis of the available habitat to the used habitat. This ratio is computed in the same way as the eigenvalues of specialization and can be compared to them.

4.3 Application to a case study: The lynx in the Vosges mountains (France)

4.3.1 Study area

From 1983 to 1993, 21 lynx have been reintroduced to the Vosges mountains (Vandel et al., 2006). The issuing population later colonized the whole southern part of the massif. The study area is about 16 500 km^2 (Fig. 4.4A) and is bordered by an intensive human-used area on the eastern part (along the Strasbourg-Mulhouse connection) connected by highways to the northern and southern directions. The lynx mainly colonized the central part of the area that includes a large patch of high-elevation deciduous forests (from 500 to 1 400 m a.s.l.), with almost no agriculture and urbanized areas. The surrounding area has lower elevation (less than 500 m a.s.l.) and is more used for agriculture and human activities.

4.3.2 Data

The French Lynx network organised the collection of all signs of presence of the lynx, including sightings, carcasses (of both lynx and preys), hairs,

Figure 4.3: Signs of presence collected by the French Lynx network. From left to right, top to bottom: a sighting of lynx in a garden; a lynx found dead; tracks of lynx; scats of lynx; a roe deer killed by a lynx. Photo credit: Réseau Lynx – Oncfs.

tracks, and scats (Vandel and Stahl, 2005, , see Fig. 4.3). During the study period (1998-2002), 292 indices were collected. A minimum convex polygon was drawn from these locations, and a buffer of 5 km was added to define the available habitat for the lynx. The choice of a 5 km buffer corresponded to the average radius of a female lynx home-range in the area (around 80 km^2, see Vandel et al., 2006). We selected some environmental variables that could a priori affect the use of space by the lynx within the available habitat. We included characteristics of the physical environment and the vegetation, as well as the influence of humans (Table 4.1).

4.3.3 Assessing the ecological niche of the lynx: results and interpretation

We first normalized through a square root transformation all the environmental variables that deviated from normality. Indeed, although the ENFA is quite robust to departure from normality, it is optimal when the environmental variables are unimodal and roughly symmetric. We then performed the first step of the ENFA that involves the selection of the number of specialization axes to retain. The diagram of the eigenvalues clearly indicated that only one axis accounted for the main part of special-

Name	Description
agri	Proportion of agricultural areas within a radius of 5 km
artif	Distance to artificial areas
dem	Digital Elevation Model: altitude
forest	Proportion of forests within a radius of 5 km
highway	Distance to highways
quickway	Distance to main roads (without highways)
railway	Distance to railways
rivers	Distance to rivers
roads	Density of all kind of roads within a radius of 5 km
slope	Slope

Table 4.1: Environmental variables used in the analysis.

ization (Fig. 4.4B). Thus, in the present case, only two axes (i.e. the axis of marginality and the first axis of specialization) accounted for most of the information.

The biplot of the ENFA provided us much information. The distance between the centroid of the ecological niche and the centroid of the available habitat was quite high, resulting in a pronounced marginality (X-axis, Fig. 4.4B), i.e. the optimum of the species was rather different from the mean available conditions. On the other hand the specialization (Y-axis, Fig. 4.4B) corresponded to an eigenvalue of 8, which means that the variance of the available habitat was 8 times higher than the variance of the ecological niche in this dimension, thus the ecological niche was much narrower than the available habitat. The significance of both the marginality and the first eigenvalue of specialization was assessed with a Monte-Carlo test. One thousand sets of 292 localizations were randomly distributed over the area. For each one, the marginality and the specialization were computed, and the actual values were compared with these random distributions. Both statistics were highly significant ($P < 0.001$). The most relevant information was provided by the projection of the environmental variables in this new space (Fig. 4.4B). Five variables were of prime importance for the analysis: the elevation, the slope, the proportion of deciduous forest, the distance to highways and the proportion of agricultural areas. The elevation and the slope contributed the most to the marginality, followed by the proportion of agricultural areas, the proportion of deciduous forest, and to a lesser extent the distance to highways. The lynx searched

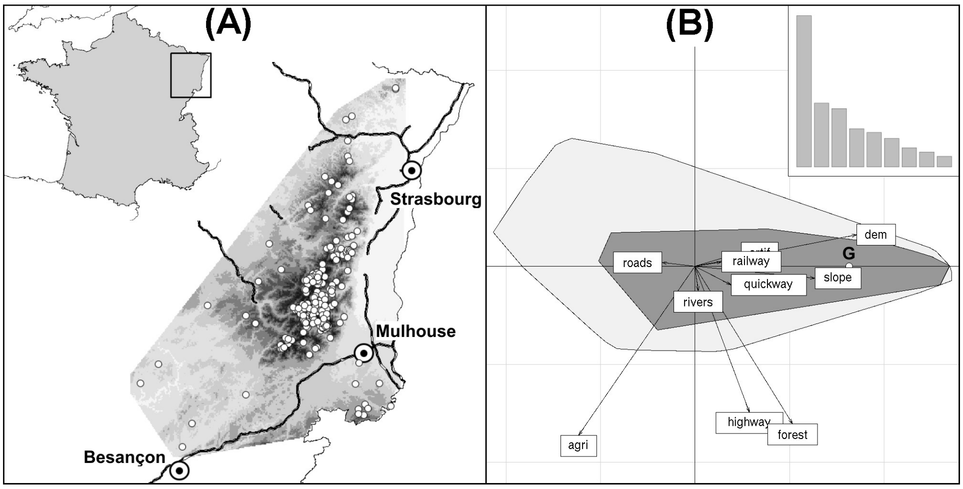


Figure 4.4: Results of the illustration. (A) The study area is situated in the eastern part of France, in the southern part of the Vosges mountains. The lines represent the highways and the dots are the locations of lynx used in the analysis. The elevation is represented in the background. (B) Biplot of the ENFA, in the plane formed by the marginality axis (X-axis) and the first specialization axis (Y-axis). The light and dark areas correspond to the minimum convex polygon enclosing all the projections of the available and used points respectively. The white dot G corresponds to the centroid of the used habitat. The arrows are the projections of the environmental variables. The insert gives the eigenvalues of specialization. One axis of specialization explains most of the specialization and is kept for the analysis.

for high values of both elevation and slope, for low use of agricultural area, for high proportions of forest, and avoided highways. The proportion of deciduous forest, the distance to highways and the proportion of agricultural areas contributed the most to the specialization axis. However, the proportions of deciduous forest and agricultural areas were strongly negatively correlated (correlation of -0.95) so that their contribution at the same level on the specialization axis vanishes. The lynx was thus not tolerant to the variation of distance to highways (i.e. the lynx was restricted on a limited range on this variable), with a mean shifted toward high distances.

The lynx was reintroduced to the central part of the study area characterized by a high elevation, a dense forest cover, and a low human use, however, later colonizing the whole central part, but avoiding the proximity of the surrounding area and particularly the eastern part, characterized by the Strasbourg-Mulhouse highway. Based on the interpretation of the marginality using the biplot, we can assess that the lynx was actually searching for a high elevation (and therefore high slopes), a dense forest cover and was avoiding highways and high agricultural use. Even more interestingly, the ENFA allowed us to assess the high specialization (i.e. low tolerance of variations) on the distance from highways. We thus found that the lynx was restricted to areas with low values of agricultural use, far from highways, and with a high proportion of forest, and was really sensitive to departure from a high distance to highways.

Another important result highlighted by the ENFA was the weak influence of artificial areas on lynx habitat use. While apparently counter-intuitive, such a result is actually not really surprising. The lynx can be seen really close to houses (Bunnefeld et al., 2006); additionally, Sunde et al. (1998) already showed that the lynx could endure a high human activity, provided that there is a high density of forested plots. From a human point of view, the lynx was just restricted to low agricultural-use areas far from the highways, thus, the eastern part was avoided due to a high proportion of agricultural areas and the presence of a highway, but not because of the presence of two big cities. Finally, the critical habitat features for the lynx include the proportion of forest and agricultural areas, and the distance from highways. The apparent selection for a high elevation and steep slopes could just be a byproduct of the proportion of forest, agricultural areas, and highways in the area. High elevation and steep slopes indeed occur in areas which are not suitable for agriculture and highways.

4.4 Discussion

In habitat selection studies, the where-question ("Where can the organisms establish?") has been under focus for many years. With the assumption that this probability of occurrence is proportional to the quality of habitat, many methods have been developed to compute habitat suitability maps (Clark et al., 1993; Guisan and Zimmermann, 2000). Computing habitat suitability maps allows the identification of suitable areas not yet, or not anymore colonized, and critical areas that need to be preserved, such as faunistic corridors (Gibson et al., 2004; Chefaoui et al., 2005). It can also be used to assess the habitat loss and fragmentation (Ciarniello et al., 2003), to estimate the population size, and to simulate spatial population dynamics (Mladenoff et al., 1995; Fielding and Bell, 1997).

Although answering the where-question is of first importance, Rushton et al. (2004) pointed out the need to understand the factors determining the distribution of the population or species. This answer to the what-question ("What do the organisms search for?") is needed to know as well as possible the ecology of the individual, population or species under study. Before a conservation plan is set, any decision should be taken with a lot of care, based on the knowledge of the processes that drive the species distribution. Soberón and Peterson (2005) underlined the lack of effective tools for exploring, analyzing, and visualizing ecological niches in many-dimensional environmental space. We present here such a tool with a new development of the ENFA (Ecological-Niche Factor Analysis).

The ENFA is based on the concept of the ecological niche, and provides a measure of the realised niche within the available space from the computation of two parameters with a clear biological meaning, as first described by Perrin (1984). By measuring the departure of the ecological niche from the average available habitat, the marginality identifies the preference of the individual, population, or species for specific conditions of the environment (e.g. high proportion of forests, high altitude, … in our case study) among the whole set of possibilities. When the niche is unimodal, the position of the centroid of the niche defines the optimum of the individual, population, or species, i.e. the conditions of the environment associated to the highest probability of presence. The specialization appears as a consequence of the narrowness of the ecological niche that involves the restriction of the occurrence on some environmental variables. It can also be interpreted as the sensitivity of the individual, population or species to variations around its optimum, highlighting limiting factors for the use of the space.

The ENFA presents several advantages. First, being fundamentally a descriptive analysis, it does not rely on any underlying hypothesis for the data, in particular autocorrelation is not a problem as such. However, for the sake of interpretation, the niche is supposed to be normal multivariate. Second, the ENFA relies on the concept of ecological niche and is therefore especially suited to a presence-only design (Hirzel et al., 2002a). The ENFA was first implemented in the Biomapper software (Hirzel et al., 2002b) which is aimed at computing habitat suitability maps, i.e. answering the where-question. The widespread use of this software resulted in biologists computing such maps, without looking carefully at the factors that are responsible for this map. Moreover, the accuracy of such maps has been questioned (Calenge et al., 2008) and as it depends on the ad-hoc algorithm used (Hirzel and Arlettaz, 2003b,a), it can be less accurate than classical linear modelling techniques in some cases (see Olivier and Wotherspoon, 2006, for an example). The ENFA, however, provides a way to identify precisely the ecological niche and therefore to answer the what-question. For this task the biplot we proposed here is probably one of the best tools (Gabriel, 1971). Marginality and specialization can be used to identify key-variables for management, assessing which habitat features are of prime importance and should be preserved or reinforced. In our application to the lynx in the Vosges mountains, we thus underlined the importance of the proportion of agricultural areas, on which the lynx had a clear preference for low values. Highlighting the relative indifference of the lynx to the proximity of artificial areas and at the opposite, the sensitivity to the proximity of highways, we got a precise picture of the influence of the human use of land for the ecology of the lynx. Thus, balancing the development of human activities and the conservation of viable lynx populations, such information will be crucial.

The ENFA provides a suitable way to measure habitat use/selection under a large range of ecological contexts. The ENFA allows us to compute uncorrelated axes from correlated variables. The method presented here uses the same core procedure as in Hirzel et al. (2002a) but we incorporated the utilization weights so that the method can handle the case where several occurrences of the species fall in the same pixel. Additionally, although we only used quantitative variables as an illustration, Calenge (2005) generalized the theory to show that qualitative variables can be included in the analysis as well. Consequently, the ENFA is probably the only analysis based on the concept of ecological niche that describes precisely the specialization, in addition to the marginality. As the marginality and the specialization are two complementary measures of the niche with different status, future work would need to tease apart the analyses of marginality and specialization. By proceeding step by step we could have a proper representation of the specialization of the organism under study, in addition to its marginality.

Software availability

The ENFA is implemented in the R-package "adehabitat" (Calenge, 2006), which collects many tools for the analysis of habitat selection by animals and trajectories of individuals. The R-software itself (R Development Core Team, 2008) is freely available on the Internet at the URL http://www.r-project.org/.

Acknowledgments

We warmly thank everyone involved in the French lynx network in countless hours of data collection. The French Ministry of Research, the University of Lyon and the Office National de la Chasse et de la Faune Sauvage (ONCFS) provided financial support of this work. We are indebted to Alexandre Hirzel for constructive comments on earlier drafts of the manuscript.

Chapter 5

The factorial decomposition of the Mahalanobis distances in habitat selection studies.

Clément Calenge[1,2,*], Gaëlle Darmon[1], Mathieu Basille[1] & Jean-Michel Jullien[2]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69 622, Villeurbanne, France.

[2] Office national de la chasse et de la faune sauvage, 95 Rue Pierre Flourens, 34 000 Montpellier, France.

[*] Corresponding author: `calenge@biomserv.univ-lyon1.fr`

Published in Ecology, 89, 2008, pp. 555–566.

"There are three kinds of lies: lies, damned lies, and statistics."

Leonard H. COURTNEY.

Abstract

The Mahalanobis distances have been introduced in habitat selection studies for the estimation of environmental suitability maps (ESMs). The pixels of raster maps of a given area correspond to points in the multidimensional space defined by the mapped environmental variables (ecological space). The Mahalanobis distances measure the distances in this space between these points and the mean of the ecological niche (i.e. the hypothesized optimum for the species) regarding the structure of the niche. The map of these distances over the area of interest is an estimated ESM. Several authors recently noted that the use of a single optimum for the niche of a species may lead to biased predictions of animal occurrence. They proposed to use instead a minimum set of basic habitat requirements, found by partitioning the Mahalanobis distances into a restricted set of biologically meaningful axes. However, the statistical approach they proposed does not take into account the environmental conditions on the area where the niche was sampled (i.e. the environmental availability), and we show that including this availability is necessary. We used their approach as a basis to develop a new exploratory tool, the Mahalanobis distance factor analysis (MADIFA), which performs an additive partitioning of the Mahalanobis distances taking into account this availability. The basic habitat requirements of a species can be derived from the axes of the MADIFA. This method can also be used to compute ESMs using only this small number of basic requirements, therefore including only the biologically relevant information. We also prove that the MADIFA is complementary to the commonly used ecological-niche factor analysis (ENFA). We used the MADIFA method to analyze the niche of the chamois Rupicapra rupicapra in a mountainous area. This method adds to the existing set of tools for the description of the niche.

Key words: chamois; ecological-niche factor analysis; environmental suitability maps; exploration; French Alps; habitat selection; Mahalanobis distances factor analysis; niche; *Rupicapra rupicapra*.

THE DETAILED KNOWLEDGE OF SPECIES DISTRIBUTION is of major concern for a large range of ecological topics. Among the tools available to improve this knowledge, environmental suitability maps (ESMs) occupy the first place (Guisan and Zimmermann, 2000; Manly et al., 2002;

Elith et al., 2006). Such maps are essential for decision making in wildlife management (Knick and Rotenberry, 1998) and for building conservation plans (Araújo and Williams, 2000).

Most methods developed to build ESMs rely on the concept of ecological niche (Guisan and Zimmermann, 2000). These maps are generally estimated using a sample of species occurrences on an area mapped for several environmental variables (e.g. elevation, slope, vegetation). Each environmental variable defines a dimension of a multidimensional space, hereafter termed "ecological space." The values of these variables can be determined for each species occurrence, so that the whole set of occurrences defines a cloud of points in the ecological space, the species niche. Environmental suitability mapping implies the computation of one environmental suitability index for each pixel of the map, based on the position of the corresponding point in the ecological space relative to the species niche. These indices are then mapped in the geographical space to provide an ESM.

The commonly used Mahalanobis distance between the available point and the mean of the niche is such an index (Mahalanobis, 1948; Clark et al., 1993; Knick and Dyer, 1997; Knick and Rotenberry, 1998; Corsi et al., 1999; Farber and Kadmon, 2003; Cayuela, 2004; Thompson et al., 2006). The mean of the niche is supposed to reflect the environmental conditions optimal for the studied species. The Mahalanobis distance for a given point expresses the distance between this point and the species optimum in the ecological space, regarding the niche structure (see Appendix A for a precise graphical description of these distances). If we assume that smaller distances correspond to areas that are more likely to be occupied by the species, the Mahalanobis distances can be mapped over the study area to provide a reliable ESM.

Recently, several authors noted that the mean of the niche of a species on a given study area can be a poor proxy for its optimum (Dunn and Duncan, 2000; Rotenberry et al., 2002, 2006; Browning et al., 2005). More suitable characteristics of the environment found in another area, but not in the original one, will be characterized by large Mahalanobis distances, and therefore low estimated suitability. The Mahalanobis distances may therefore lead to biased predictions of animal occurrence under different environmental conditions. These authors proposed to use, instead of this optimum, a minimum set of basic habitat requirements. They advocated that the variables that maintain a consistent value where the species occur (i.e. the variables with a low "used" variance) are those most likely to be associated with basic habitat requirements. For this reason, they argued that the last axes of a principal component analysis (PCA) of the niche, on which the variance is the smallest, can be used to define this basic set. Moreover, they demonstrated that this PCA is a natural way to partition the Mahalanobis distances. Therefore, these authors recommended estimating ESMs by computing a reduced-rank Mahalanobis distance for each pixel of the map of the study area, by considering only this restricted set of principal components. They consider this statistic as the distance from the pixel to this minimum set of basic requirements.

However, although this linear partitioning of the Mahalanobis distance relies on both solid mathematical bases and sound biological issues, it is also problematic. The PCA recommended by these authors is performed on the table giving the value of the environmental variables (columns) in the sites used by the species (rows), without consideration of the availability of the environmental variables. Note that this table is standardized before the PCA is applied, so that all the environmental variables have a unit variance. This preliminary operation is necessary, as the variables may not be measured on the same scale (e.g. the elevation measured in meters and slope measured in percent). However, this scaling has an unexpected consequence: maximizing the variance of the standardized niche on the first axes of the PCA is just a way of maximizing the sum of the squared correlations between the environmental variables and the first axis (Legendre and Legendre, 1998).

However, the fact that some environmental variables are strongly correlated among each other does not imply that these variables cannot be used to define a basic set of required habitats. For example, hydrobiologists often measure the velocity, the depth, and the flow of a stream when they want to study the niche of a fish species (e.g. Mäki-Petäys et al., 1997). These variables are often strongly correlated among each other, even when the correlations are computed only with the sites used by the species. These variables are therefore likely to define the first axis of the PCA of the niche. However, they are strong limiting factors for many species, in the sense that the range of variation actually experienced by the species is very small relative to the range that could be potentially encountered by the species.

The crucial point here is that the identification of variables with a "low" variance implies that we know what a "normal" variance is for these variables: a reference value is needed. Actually, the used sites are generally sampled on a given area, which defines the context in which the niche takes place. The whole set of pixels of this area defines a cloud of "available points" in the ecological space, of which the niche is a subset. The shape of the niche in the ecological space is partly the result of the influence of this context. Actually, we defend the idea that the identification of the required habitat for a species distribution from a sample of used sites should also take into account the environmental availability at the time of sampling in some way.

However, the biological issue raised by Rotenberry et al. (2002, 2006) is important. The definition of a restricted set of basic habitat requirements could improve the predictive capabilities of ESMs based on the Mahalanobis distances. In this paper, we used the work of Rotenberry et al. (2002, 2006) as a basis to solve the problem of the identification of the basic habitat requirements. We therefore developed a new exploratory approach to tackle the problem, which we called the "Mahalanobis Distances Factor Analysis" (MADIFA). This approach also performs an additive partitioning of the Mahalanobis distances, but the first components of the analysis now explain most of the Mahalanobis distances for the set of available points on a given area. The factorial maps of these axes allow both the exploration of the niche in the ecological space and the identification of the environmental variables corresponding to basic habitat requirements. The factorial axes can also be used to compute ESMs on a lower number of dimensions (and therefore with increased generality) that take into account a large part of the niche restriction. We illustrate how this analysis may find its place among other exploratory tools of the niche with the analysis of the niche of the chamois (*Rupicapra rupicapra*) in a mountainous environment.

5.2 The computation of the Mahalanobis distances

We assume that the values of P environmental variables are known for N pixels (where N can be a random sample or the whole set of pixels of a map). We consider here that the N available pixels have the same weight in the analysis, contained in the $N \times N$ (rows × columns) diagonal matrix $\mathbf{D} = \mathrm{Diag}(1/N)$. Moreover, we consider a set of N utilization weights, summing to one, which reflects the use of the N pixels by the focus species. For example, these weights may correspond to the proportion of locations of the studied species in the pixels of the map. These weights are stored in an $N \times N$ diagonal matrix $\mathbf{D}_p$. In the rest of this paper, we will term "available pixels" the whole set of N pixels of the analysis, and "used pixels" or "niche" the set of pixels for which the utilization weights are greater than zero.

Let the matrix $\mathbf{Z}$ contain the value of the P environmental variables (columns) in each one of the N available pixels (rows). The matrix $\mathbf{Z}$ is centered and scaled for the weighting $\mathbf{D}_p$ (i.e. respectively, the origin of the space defined by the columns of $\mathbf{Z}$ is located at the mean and the variance is 1 for all columns of $\mathbf{Z}$). Finally, let $\mathbf{\Sigma} = \mathbf{Z}^{\mathbf{T}}\mathbf{D}_p\mathbf{Z}$ be the correlation matrix as the columns of $\mathbf{Z}$ have a unit variance (where $\mathbf{Z}^{\mathbf{T}}$ is the transpose of $\mathbf{Z}$).

The squared Mahalanobis distance D_i^2 between any available point i (associated to a pixel in the geograph- ical space) and the mean of the niche provides an index of the environmental suitability at this place. Let $\mathbf{Z}_{i\bullet}$ be the row vector containing the values of the P environmental variables for the ith pixel (that is, the ith row of the matrix $\mathbf{Z}$). In these conditions, the squared Mahalanobis distance between the point i and the mean of the niche can be computed with

$$D_i^2 = \mathbf{Z}_{i\bullet}\mathbf{\Sigma}^{-1}\mathbf{Z}^{\mathbf{T}}{}_{i\bullet} \qquad (5.1)$$

5.3 Linear partitioning of the Mahalanobis distances: The point of view of Rotenberry et al. (2002, 2006)

Rotenberry et al. (2002, 2006) noted that the computation of these distances relies on the computation of the inverse of the matrix $\mathbf{\Sigma}$ (Eq. 5.1). This computation may be performed by its diagonalization (i.e. the computation of its eigenvectors and eigenvalues). More formally,

$$\mathbf{\Sigma} = \mathrm{A}\delta\mathrm{A}^{\mathrm{T}}$$

where the matrix $\boldsymbol{\delta}$ is the diagonal matrix containing the P eigenvalues λ_j of the matrix $\boldsymbol{\Sigma}$, i.e. $\mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$, and $\mathbf{A}$ is the matrix containing the P eigenvectors α_j of the matrix $\boldsymbol{\Sigma}$ concatenated by columns, i.e. $[\alpha_1, \alpha_2, \dots, \alpha_P]$. The inverse of the matrix $\boldsymbol{\Sigma}$ is given by the following (Harville, 1997):

$$\boldsymbol{\Sigma}^{-1} = \mathbf{A}\boldsymbol{\delta}^{-1}\mathbf{A}^{\mathrm{T}}$$

Consequently, the Mahalanobis distance between the point i and the mean of the niche can be computed using

$$D_i^2 = \mathbf{Z}_{i\bullet}\mathbf{A}\boldsymbol{\delta}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}}{}_{i\bullet} \tag{5.2}$$

Rotenberry et al. (2002, 2006) noted that this formula provides a natural way of partitioning the Mahalanobis distances, as it is related to the principal components analysis (PCA) of the niche (i.e. a PCA of the table $\mathbf{Z}$ using the matrix $\mathbf{D}_p$ as row weights; as in Fig. 5.1B). The axes of this PCA correspond to the eigenvectors of $\boldsymbol{\Sigma}$ (i.e. α_1, α_2, etc.). The first axes represent the directions in the ecological space for which the niche width is maximal. The variance of the niche projected onto a given axis j of this PCA is the corresponding eigenvalue λ_j. Note that because $\mathbf{Z}$ has been scaled, this maximized variance is just the sum of squared correlations between the environmental variables and the axis j of the PCA (Legendre and Legendre, 1998). The vector $\mathbf{Z}_{i\bullet}$ contains the coordinates of the available point i in the ecological space. Therefore the coordinate of the available point i projected onto the jth axis of the PCA is computed by $\mathbf{Z}_{i\bullet}\alpha_j$. The normed coordinate b_{ij} of the point i on the jth factorial axis corresponds to the raw coordinate divided by the standard deviation of the niche on this axis. Then, using Eq. 5.2, it is straightforward to show that the Mahalanobis distances can be computed by the sum of the squared b_{ij}:

$$D_i^2 = \sum_{j=1}^{p} b_{ij}^2 = \sum_{j=1}^{p} \left(\frac{\mathbf{Z}_{i\bullet}\alpha_j}{\sqrt{\lambda_j}} \right) \tag{5.3}$$

Rotenberry et al. (2002, 2006) advocated the use of a limited set of PCA axes to compute reduced-rank Mahalanobis distances. They noted that the first axes of the PCA are unlikely to describe required habitats, precisely because they thought that the large variance on these axes indicated that

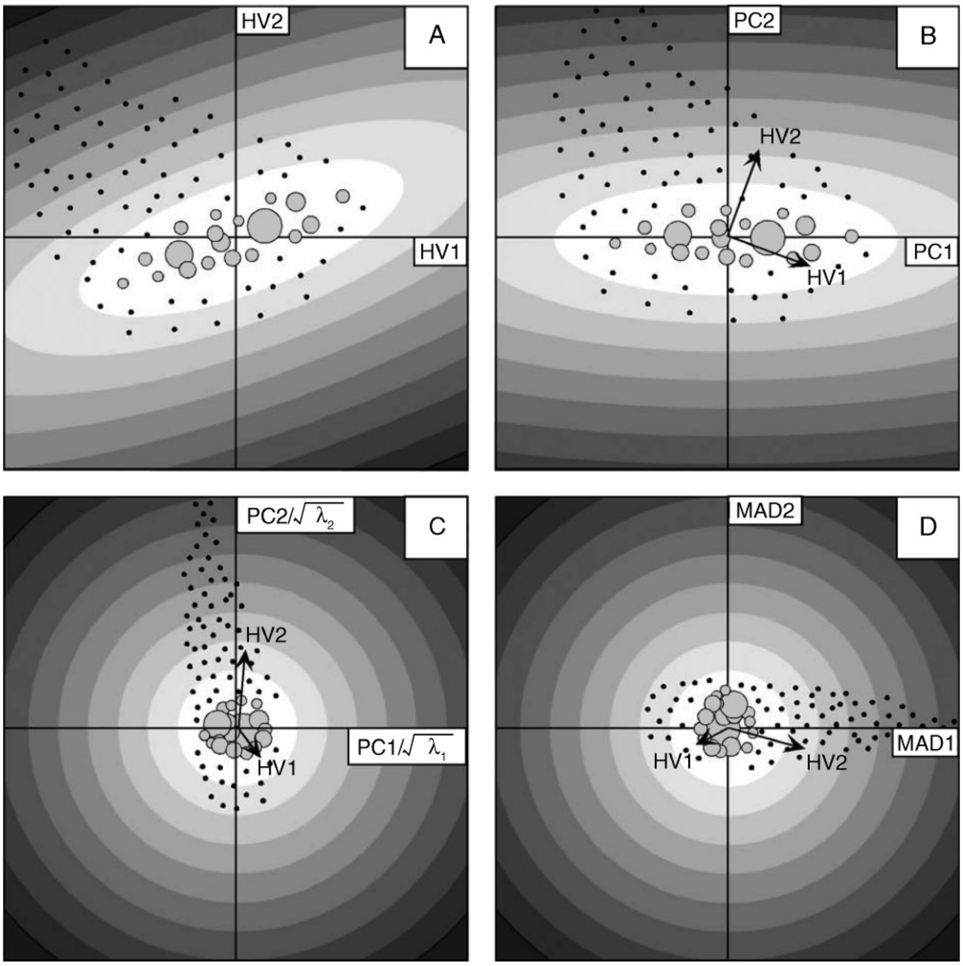


Figure 5.1: The MADIFA procedure (see *The three steps to perform the MADIFA*). The black dots are points considered as available to the species. To each available point is associated one utilization weight proportional to its use by the species (indicated by a gray circle with an area proportional to this weight). The values of the Mahalanobis distance to the optimum of the niche are indicated by gray levels (i.e. the darker the shade, the farther from optimum). (A) The ecological space is defined by two environmental variables, HV1 and HV2, and is centered on the mean of the niche. (B) The first principal component analysis (PCA) of the niche (PC1 and PC2 are the principal components). (C) The scores of the points on the two principal components are divided by the square root of their respective eigenvalues. (D) The second PCA (not centered) maximizes the mean-squared Mahalanobis distances between the available points and the mean of the niche on the first axes, MAD1 and MAD2.

the ecological variation experienced by the species was large (whereas this variance is just the sum of squared correlation with the environmental variables). They proposed instead to compute the reduced-rank Mahalanobis distances using the last eigenvectors of the PCA, arguing that the dimensions on which the niche is the narrowest are likely to describe required habitats. For example, using the last R axes of the PCA, the reduced-rank squared Mahalanobis distances $\tilde{D}_i^2$ is computed using

$$\tilde{D}_i^2 = \sum_{j=P-R}^{P} b_{ij}^2$$

5.4 Some refinements of this point of view: The MADIFA

5.4.1 The three steps to perform the MADIFA

We develop here a new partitioning of the Mahalanobis distances, which identifies the directions in the ecological space for which the niche is the narrowest in comparison to the width of the cloud of available points (see Fig.5.1). We call it the "Mahalanobis Distances Factor Analysis" (MADIFA). This analysis is performed in three steps. The first two steps of this analysis are exactly the approach proposed by Rotenberry et al. (2002, 2006).

A PCA is first performed on the table $\mathbf{Z}$ using the matrix $\mathbf{D}_p$ as row weights, which returns the directions partitioning the variance of the standardized niche into orthogonal components (Fig. 5.1B), i.e. the set of eigenvectors α_j and of eigenvalues λ_j $(j = 1, \ldots, P)$ of the matrix $\mathbf{\Sigma}$ as defined in Eq. 5.2. Second, the ecological space is distorted: the correlation structure is removed by rescaling the variance of all axes to one (Fig. 5.1C). The scores of the available pixels in this distorted space are stored in the $N \times P$ matrix $\mathbf{B}$:

$$\mathbf{B} = \mathbf{ZA}\delta^{-1/2} \tag{5.4}$$

The matrix $\mathbf{B}$ contains the normed scores b_{ij} as defined in Eq. 5.3 (Rotenberry et al., 2002, 2006).

Thereafter, we add another step to this approach: we perform a PCA on matrix $\mathbf{B}$ using the uniform row weights stored in $\mathbf{D}$. This second PCA is the core of the MADIFA, and we show hereafter that it returns linear combinations of the environmental variables so that the width of the niche is the smallest in comparison to the width of the cloud of available points (Fig. 5.1D).

5.4.2 Mathematical properties of the second PCA

The matrix being diagonalized is $\mathbf{G} = \mathbf{B}^{\mathbf{T}}\mathbf{D}\mathbf{B}$. This analysis returns a set of P orthogonal eigenvectors $\mathbf{v}_k$ stored in a matrix $\mathbf{V}$, and P corresponding eigenvalues θ_k stored on the diagonal of the matrix $\boldsymbol{\beta}$, so that $\mathbf{G} = \mathbf{V}\boldsymbol{\beta}\mathbf{V}^{\mathbf{T}}$. The pixel scores are computed by $\mathbf{L} = \mathbf{B}\mathbf{V}$:

$$\mathbf{L} = \mathbf{Z}\mathbf{A}\boldsymbol{\delta}^{-1/2}\mathbf{V} \tag{5.5}$$

This formula summarizes the three steps of the MADIFA (Fig. 5.1): the factorial axes of this analysis are found after a rotation (matrix $\mathbf{A}$), a distortion (matrix $\boldsymbol{\delta}^{-1/2}$), and another rotation (matrix $\mathbf{V}$) of the cloud of available points in the ecological space (matrix $\mathbf{Z}$). All these transformations of $\mathbf{Z}$ can be summarized in a matrix $C = \mathbf{A}\boldsymbol{\delta}^{-1/2}\mathbf{V}$. The pixels scores are the linear combinations of the environmental variables (i.e. $\mathbf{L} = \mathbf{Z}\mathbf{C}$).

The value maximized on the first axes of the MADIFA is equal to the following:

$$\theta_j = \frac{1}{N}\sum_{i=1}^{N} l_{ij}^2 = \frac{\sum_{i=1}^{N}\frac{1}{N}\left(l_{ij} - \bar{l}_j^u\right)^2}{\sum_{i=1}^{N} u_i\left(l_{ij} - \bar{l}_j^u\right)^2} \tag{5.6}$$

where l_{ij} is the score of the pixel i on the jth axis of the MADIFA, $\bar{l}_j^u$ is the mean of the scores of the used pixels on the jth axis of the analysis, and u_i is the utilization weight associated with the pixel i. This result derives from the observation that the used variance (denominator of θ_j) is equal to 1 on the axes of the MADIFA, and that $\bar{l}_j^u = 0$ (as the origin of the ecological space is the mean of the niche).

Thus, the denominator of θ_j is the variance of the niche on the first axis of the MADIFA. However, the numerator is not a variance: it is the mean of the squared deviations of the available points from the mean of the scores of used points. Consequently, the MADIFA indicates the directions where the niche is the narrowest (low variance) compared to the width of the distribution of available points. This direction is likely to define a basic habitat requirement.

We show in Appendix B that

$$D_i^2 = \sum_{j=1}^{P} l_{ij}^2 \tag{5.7}$$

Note that this result implies that the sum of the eigenvalues θ_j over all the axes j of the analysis is equal to the mean of the squared Mahalanobis distances for the available pixels. It is therefore possible to compute the proportion of the mean-squared Mahalanobis distances explained by each axis.

Now, like Rotenberry et al. (2002, 2006), we can compute reduced-rank squared Mahalanobis distances with the set of R first axes (chosen so that the variance of the niche is the smallest as compared to the variance of the available points), reflecting the distance between the available points and the set of basic habitat requirements. From Eq. 5.7, one can derive the reduced-rank squared Mahalanobis distance:

$$\tilde{D}_i^2 = \sum_{j=1}^{R} l_{ij}^2 \tag{5.8}$$

The scores of the pixels on the axes of the MADIFA can be used to draw factorial maps to identify the structures of the niche in the ecological space (as in Fig. 5.1D). Alternatively these scores can be used to map reduced-rank Mahalanobis distances over the area, to provide clearer and sharpened environmental suitability maps (ESMs; using Eq. 5.8). The biological meaning of the factorial axes can be found either by using the coefficients in **C** or the correlations with the original environmental variables.

The MADIFA is programmed in the function `madifa()` of the free package adehabitat (Calenge, 2006) for the R software (R Development Core Team, 2008). It can be used as a classical exploratory tool (Legendre and Legendre, 1998) to draw a conceptual model of the studied biological system.

5.4.3 The MADIFA and the ecological-niche factor analysis

The MADIFA is closely related to the ecological- niche factor analysis (ENFA) developed by Hirzel et al. (2002a). Indeed, these authors noted that basic habitat requirements are likely to be associated with the directions of the ecological space where the variance of the niche is very small in comparison to the variance of the available points. The ratio of these two variances computed for a given variable is an index of the specialization of the species on this variable. The ENFA is a factor analysis of the niche maximizing this index on the first axis. More formally, for a given axis j, the specialization ratio S is equal to

$$S(w_j) = \frac{\sum_{i=1}^{N} \frac{1}{N} \left(w_{ij} - \overline{w}_j^a \right)^2}{\sum_{i=1}^{N} u_i \left(w_{ij} - \overline{w}_j^u \right)^2} \tag{5.9}$$

where w_{ij} is the score of the ith pixel on the jth axis of the ENFA, $\overline{w}_j^a$ is the mean of the scores of available points on the jth axis of the ENFA, and $\overline{w}_j^u$ is the mean of the scores of the used points on the same axis. Note that $S(w_j)$ is very similar to θ_j (compare Eq. 5.6 and Eq. 5.9). The only difference is that the former uses the variance of available points as a measure of the width of the distribution of available points, while the latter uses the mean of the squared deviation of available points from the mean of the scores of used points.

Maximizing the ratio $S(w_j)$ is possible only if the marginality vector has first been extracted from the data (i.e. the vector connecting the mean of the cloud of available points to the mean of the cloud of used points; Hirzel et al., 2002a). However, the marginality vector is often biologically important, and several authors stressed the need to take into account this vector in the interpretation of the results (e.g. Hirzel et al., 2002a). Consequently, the available and used points are projected onto this vector to define a marginality axis as a first step. The interpretation of the results of the ENFA includes the interpretation of the scores of used and available points on this marginality axis.

Note that the ratio θ_j maximized by the axes of the MADIFA can be rewritten:

$$\theta_j = \frac{m_j^2}{v_j^2} + S(w_j)$$

where m_j^2 is the squared difference between the mean of the scores of used points and the mean of the scores of available points on the jth axis of the analysis (i.e. the marginality), and v_j^2 is the variance of the niche on the jth axis of the analysis. The MADIFA therefore combines the marginality and the specialization into one single measure of niche restriction.

Thus, the ENFA may be used to complement the results of the MADIFA as it allows identification of the part of the Mahalanobis distances corresponding to the specialization and to the marginality, respectively. Used jointly, these two approaches lead to a more precise conceptual model elaborated for the niche of the focus species. The ENFA can also be used to draw factorial maps of the niche (Basille et al., 2008).

On the other hand, as the marginality axis does not have the same mathematical status as the specialization axes of the ENFA (the marginality axis is orthogonal to the specialization axes, but the specialization axes are not orthogonal among each other; Hirzel et al., 2002a), it is often difficult to combine all these axes into one single index of environmental suitability. So far, existing methods trying to combine the marginality and specialization axes use ad hoc algorithms (Hirzel et al., 2002a; Hirzel and Arlettaz, 2003b). Although these ENFA-based methods have proven to return biologically consistent environmental suitability maps (ESMs; e.g. Bryan and Metaxas, 2007), the MADIFA is probably a better way to build environmental suitability maps: it returns axes, all with the same mathematical status, which can be combined into ESMs in a consistent manner.

5.5 Application: Exploration of habitat selection by the chamois

We explored the habitat component of the niche of the chamois (*Rupicapra rupicapra*; see Fig. 5.2) in open areas of the wildlife reserve of Les Bauges (French Alps, 45°25′ N, 6°5′ E; Fig. 5.3A). The data were collected during censuses carried out every year from 1994 to 2004 in June using the same protocol (flash counts; see e.g. Houssin et al., 1994). Volunteers and professionals working in various French wildlife and forest management organizations walked along 24 transects and looked around two fixed points, which were distributed over the reserve so that all open areas (i.e. nonforested areas) were visible to the observers. All transects were traveled simultaneously at dawn by teams of two observers, and each detected chamois group was located on a map of the reserve (precision of ~10 m). At the end of the census, hours and locations of observations were compared in order to delete the double counts. Because the study of habitat selection requires a homogeneous sampling effort, we used the upper elevation limit of the forests to delimit our study area (6430 ha dominated by open meadows located at an elevation >1200 m). Preliminary analysis showed that the number and the spatial distribution of the detected groups did not vary greatly among years (C. Calenge and G. Darmon, unpublished data). We therefore considered the pooled data set here to reduce these sampling fluctuations. During the seven years of the study, 650 chamois groups were detected (Fig. 5.3B). We split the data set in two, one for calibration (from 1994 to 2000; 400 groups detected), and one for validation (from 2001 to 2004; 250 groups detected). Seventeen environmental variables were included in the analysis of the chamois habitat (Table 5.1, Fig. 5.3C). These variables were supposed to reflect the chamois distribution, either because they reflect the location of secure areas (e.g. distance to trails, visibility, slope; von Elsner-Shak, 1985), or because they represent vegetal associations in which the chamois may search for food (Ferrari et al., 1988; Garcia-Gonzalez and Cuartas, 1996). Note that although we focused only on the chamois distribution in the open areas, we also included in the analysis the distance to forested areas, because these surrounding habitats may also influence the habitat use by the chamois in open areas (Hamr, 1985).

We first investigated habitat selection using the calibration data set. Before the application of the MADIFA, we explored the structure of the environmental composition over the study area, using a principal component analysis of the table giving the values of the environmental variables (columns) in the pixels of the maps of the area (rows). One main pattern is highlighted (see Appendix C): the elevation, which is the variable best correlated with the first axis, affects the value of several environmental variables. Such an altitudinal structure was expected in this mountainous area. Areas close to the screes, to the meadows dominated by *Sesleria* and *Carex sempervirens*, and to the meadows dominated by *Carex ferruginea* are generally found at high elevations (Rameau et al., 2001).

We also performed a PCA restricted to the pixels where chamois were

Figure 5.2: A chamois (*Rupricapra rupricapra*) photographed in the Bauges mountains (French Alps). Photo credit: Marc Cornillon.

Abbreviation	Variable name
Elev	elevation
D.Alder	distance to alder woods
D.Screes	distance to screes
D.Forest	distance to forested areas
D.Fodder	distance to fodders
D.Brachy	distance to meadows dominated by *Brachipodium*
D.CarexF	distance to meadows dominated by *Carex ferruginea*
D.TallHe	distance to meadows dominated by tall herbs
D.Nardus	distance to meadows dominated by *Nardus* sp.
D.SeCarS	distance to meadows dominated by *Sesleria* and *Carex sempervirens*
D.Rhodo	distance to moors dominated by *Rhododendron*
D.Trail	distance to recreational trails
Hydro	hydrography
Slope	slope
Sunshine	sunshine
Visib	visibility (area seen from each pixel, computed using Elev)
Visib1000	visibility computed within a radius of 1000 m

Table 5.1: Variables included in the "Mahalanobis Distances Factor Analysis" (MADIFA).

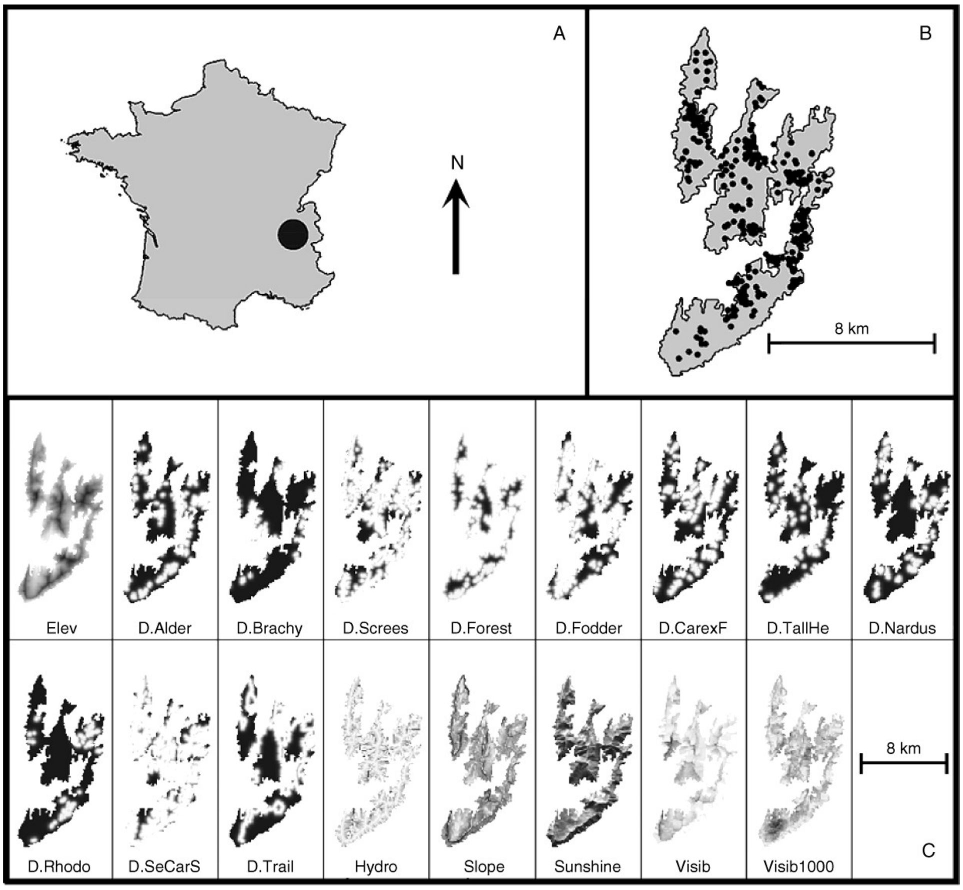


Figure 5.3: (A) Location of the wildlife reserve of Les Bauges in France; (B) distribution of the chamois detected on the area from 1994 to 2000; and (C) maps of the 17 environmental variables over the area, where levels of each environmental variable increase from light to dark gray (see Table 5.1 for the full names of the variables).

located (i.e. on its habitat). The altitudinal structure highlighted on the study area was also the main structure of the chamois habitat (Appendix C). The correlation between the first axis of the PCA of the available points and the first axis of the PCA of the habitat is very strong ($R = -0.87$). Actually, the altitudinal structure is so strong in the study area that it also affects the shape of the cloud of used points in the ecological space. However, the fact that the variance of used pixels is maximal on this direction does not imply that it does not describe a habitat required by the chamois, as shown next.

We then studied habitat selection of the chamois with the MADIFA. We first performed a preliminary Monte Carlo test to determine whether the habitat selection is significant in at least one direction of the ecological space. At each step of the process, we simulated a random habitat use by the chamois by generating a uniform distribution of 400 points over the study area, and we computed the first eigenvalue of the MADIFA of this simulated data set. We repeated this simulation 500 times to derive a distribution of eigenvalues under the hypothesis of random habitat use. We finally compared the first eigenvalue of the MADIFA of the observed 400 chamois groups to this simulated distribution to derive a P value. There is actually a highly significant habitat selection value ($\theta_1 = 3.7, P < 0.002$).

The proportion of the mean of the squared Mahalanobis distances explained by each axis j is measured by the corresponding eigenvalue θ_j. The exploration of these eigenvalues helps in choosing a number of axes to interpret (Fig. 5.4A). The MADIFA returned one main eigenvalue (15% of the mean of the squared Mahalanobis distances are explained on the first axis). The percentage of the mean of the squared Mahalanobis distances explained by the following axes is much lower (10.5%, 9.5%, and 7.6% for the second, third, and fourth axis, respectively). We therefore focused our interpretation on the first axis of the MADIFA.

The biological meaning of this axis can be deduced from the correlation coefficients between the first axis of the MADIFA and the environmental variables (Fig. 5.4E). The positive scores on this axis correspond to areas located at high elevations (correlation between elevation and the first axis: $R = 0.59$), close to the screes (D.Screes, $R = -0.67$), and, above all, close to the meadows dominated by *Sesleria* and *Carex sempervirens* (D.SeCarS, $R = -0.78$). The negative scores correspond to areas with the

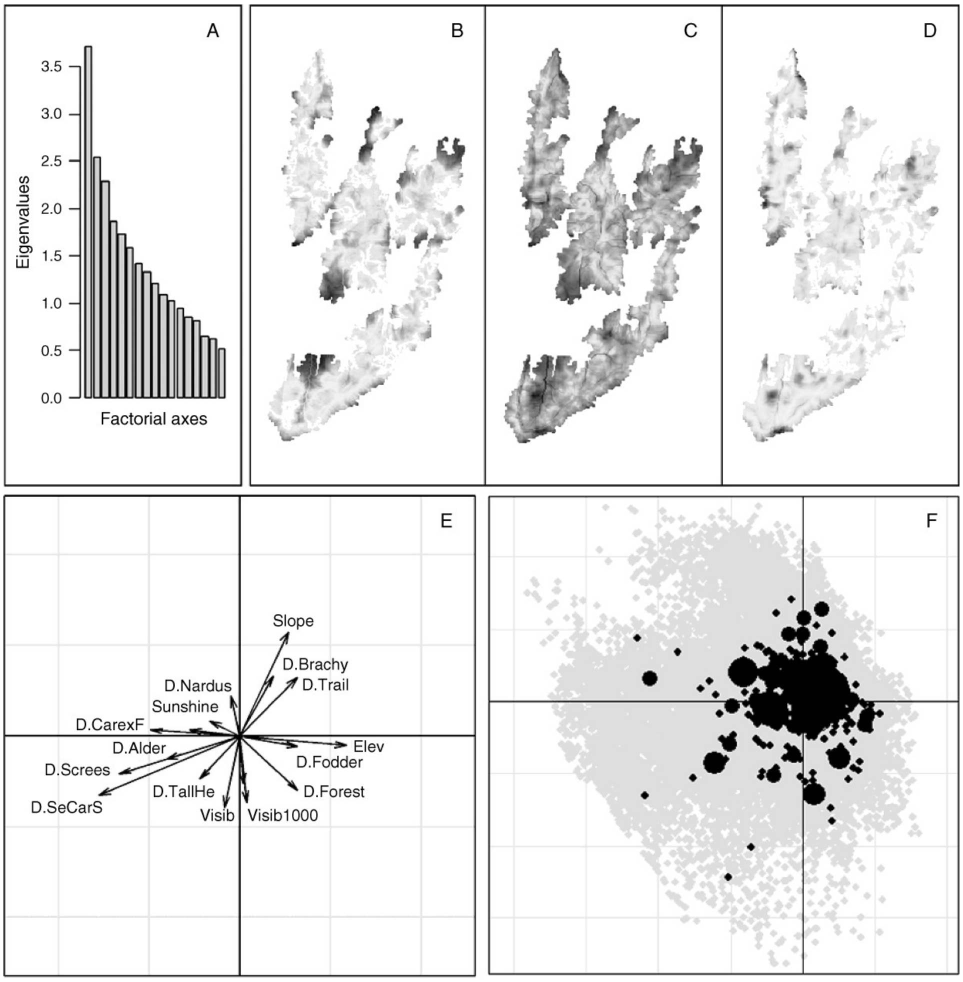


Figure 5.4: Results of the MADIFA performed to analyze the chamois distribution with respect to the 17 environmental variables. Despite the fact that only one axis is highlighted by the analysis, results are presented for the first two axes in panels (E) and (F). For panels (B)-(D), levels of each environmental variable increase from light to dark gray. (A) Bar plot of the eigenvalues; (B) environmental suitability map of the area computed using the first axis of the MADIFA; (C) environmental suitability map of the area computed using the 17 environmental variables (full-rank Mahalanobis distances); (D) environmental suitability map of the area computed using the last seven axes of the PCA of the niche (method of Rotenberry et al., 2002, 2006); (E) graph of the correlations between the environmental variables and the first (x-axis) and second (y-axis) axes of the MADIFA (see Table 5.1); and (F) factorial map of the ecological niche of the chamois on the first (x-axis) and second (y-axis) axes of the MADIFA. The gray points correspond to the available points (pixels of the maps), and their intensity of use is proportional to the area of the black points. The whole set of black circles defines the niche of the species.

opposite characteristics. The chamois habitat is the narrowest on this dimension of the ecological space, regarding the width of the distribution of available points. The factorial map of the ecological space indicates that the distribution of the available environment is shifted to the negative values of the first axis (whereas the used points are still centered on zero; see Fig. 5.4F). Within the studied context, it seems that the chamois select the areas close to the screes (50% of the detections within 111 m of this environment type) and, above all, close to meadows dominated by *Sesleria* and *Carex sempervirens* (75% of the detection within 70 m of this vegetation type).

The environmental suitability maps (ESMs) built using the first axis confirmed these results (Fig. 5.4B). The comparison of the ESMs with the maps of environmental variables showed that the most suitable areas are found close to meadows dominated by *Sesleria* and *Carex sempervirens*, and close to screes (Fig. 5.3C). The effect of the elevation here seems indirect: the most suitable areas are found at high elevation, which correspond to low distances to meadows dominated by *Sesleria* and *Carex sempervirens* (this environment type is on average located at an elevation of 1588 ± 183 m [mean ± SD]) and to screes (which were, on average, located at an elevation of 1748 ± 163 m). The indirect effect is consistent with the sharp aspect of the map that indicates a clear frontier between suitable and unsuitable environments, whereas the elevation map is more continuous. Note that the main spatial structures of the map of the full-rank Mahalanobis distances (Fig. 5.4C) are clearer on the ESMs built from the analysis (Fig. 5.4B): the increased precision (reduced generality) of the full-rank Mahalanobis distances is manifest in the identification of less area as potentially suitable (more noise is included in this measure).

Female chamois give birth to young in May and need a lot of resources to feed them (Hamr, 1985; Ferrari et al., 1988). The prolific regrowth of the vegetation results in many energetic shoots in the meadows dominated by *Sesleria* and *Carex sempervirens*, which may therefore explain the abundance of the chamois in such environments at this time of the year. The distance to screes is also well-correlated with the first axis of the MADIFA, but this probably results from a confounding effect, as the screes are close to such meadows. This proximity of the screes probably increases the chamois preference for these meadows, as the screes may provide both an escape in case of predators (Bleich et al., 1997) and saline resource.

We then measured the goodness of fit with the validation data set. Following, Knick and Dyer (1997), we computed the cumulative frequency of the reduced-rank Mahalanobis distances (Fig. 5.4B) for (i) the pixels of the study area, (ii) the pixels containing chamois detections of the calibration set, and (iii) the pixels containing detections of the validation set (Fig. 5.5). We used the curves of both the study area and the validation set to derive a measure of the predictive capabilities of the analysis. The area located above the curve of the study area and below the curve of the validation set on this graph measures the quality of the prediction. Indeed, this area would be maximum in the case of a perfect prediction, because the value of the cumulative frequency of distances for the validation set would be equal to one whatever the value of distance (indicating that these distances are equal to zero for all the detections of the validation set). Therefore, dividing the quality of prediction of the validation set by the area located above the curve of the study area and below the line $Y = 1$ (theoretical perfect prediction) on this graph gives a standardized measure Q of quality of prediction. We also computed this ratio for the calibration data set, to give a measure G for the goodness of fit.

The goodness of fit of the MADIFA is rather high ($G = 74\%$; Fig. 5.5A). The curve of cumulative frequency distribution for the validation set is similar to the curve of the calibration set, indicating good predictive capabilities ($Q = 73\%$). Indeed, 94% of the detections of these sets are in the top 75% of the reduced-rank Mahalanobis distances of the pixels of the study area.

Finally, we compared the results of the MADIFA with those of the PCA of the used points advocated by Rotenberry et al. (2002, 2006). We computed an ESM using the last seven axes of the PCA of the used points (Fig. 5.4D). The goodness of fit was lower than for the MADIFA ($G = 66\%$), and the predictive capabilities of this ESM were even lower ($Q = 59\%$, Fig. 5.5B). In fact, the main factor limiting the chamois distribution is closely related to the elevation, which is the main pattern on the study area. Therefore, this basic habitat requirement is unlikely to define the last axes of the PCA of the used points. Using the last axes of the PCA to build an ESM is likely to keep only the "noisy part" of the Mahalanobis

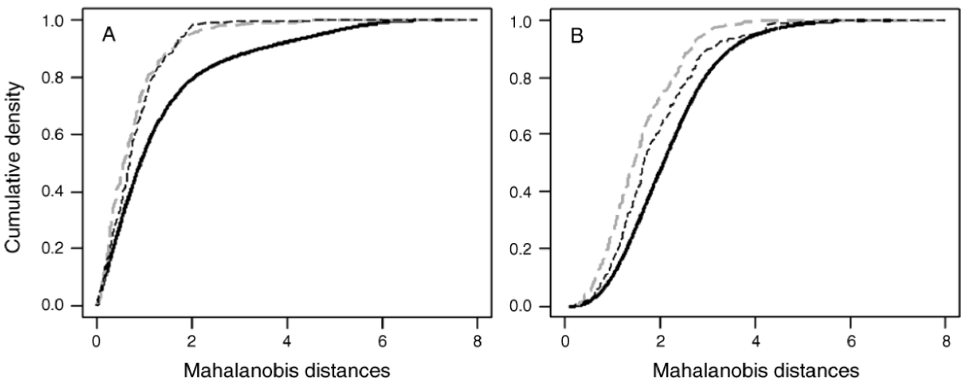


Figure 5.5: Cumulative frequency distribution of the reduced-rank Mahalanobis distances computed for the pixels of the study area in the Bauges mountains (French Alps, solid black line), for the calibration data set (pixels where chamois groups were detected from 1994 to 2000, dashed gray line), and for the validation data set (pixels where chamois groups were detected from 2001 to 2004, dashed black line). (A) Reduced-rank Mahalanobis distances computed using the first axis of the MADIFA; and (B) reduced-rank Mahalanobis distances computed using the last seven axes of the PCA of the niche.

distances. This again stresses the need to take into account the availability when one wants to identify habitat requirements.

5.6 Discussion

We developed Mahalanobis distance factor analysis (MADIFA) to explore, analyze, and visualize the niche in the ecological space. Furthermore, these results can be used to derive environmental suitability maps (ESMs) to visualize the patterns of the niche in the geographical space. This method led us to identify the main characteristics of the environment selected by the chamois, and provided an ESM of the area. We pointed out that the elevation is correlated to all the environmental variables included in the analysis (e.g. screes, meadows dominated by *Sesleria* and *Carex sempervirens* are generally found at high elevation) and is also the main structure of the chamois habitat: the variance of the species habitat is maximal for the elevation. However, although the chamois habitat is wider on this dimension, it is narrow relative to the range of available environment, indicating that this dimension contributes to the definition of a basic habitat requirement for this species (although indirectly, through its effect on the vegetation). This example clearly illustrates the need to take into account the availability in the partitioning of the Mahalanobis distances.

Accounting for the environmental availability at the time of sampling is also important for the "classical" Mahalanobis distances method. In most papers using this method, the environmental suitability is estimated on the area where the sample of used site has been collected (e.g. Clark et al., 1993). However, the environmental conditions may vary beyond the limits of this area. If the limits of the area on which the Mahalanobis distances are mapped are not carefully checked, the environmental conditions on the mapped area may not be representative of what was actually available to the species at the time of sampling. In such a case, the Mahalanobis distances may indicate an unsuitable environment in areas where the environmental conditions vary in a biologically positive direction (Knick and Rotenberry, 1998). Consequently, even if the Mahalanobis distances method is a powerful method for ESM modeling, it does not circumvent the problem of the definition of availability.

5.6.1 Hypotheses underlying the MADIFA

The main assumption underlying the MADIFA is that the maximized statistic θ_j is relevant to capture the patterns of the niche in its environment. Because this statistic is a ratio between two sums of squared deviations from the mean of the niche, this assumption will be met if the mean of the niche is close to its mode (i.e. unimodal and symmetric niche). This hypothesis is also required by all factorial methods relying on the concept of ecological niche (ter Braak, 1985, 1986; Knick and Rotenberry, 1998; Hirzel et al., 2002a). It ensures that the sum of squared deviations from the mean of the niche is a measure of the distance from the conditions most frequently used by the species.

This sum of squared deviations is very sensitive to outliers (Cleveland, 1993), and so is the optimality criterion θ_j. Although this criterion allows MADIFA to be placed in a consistent theoretical framework (including the ecological-niche factor analysis [ENFA] and the Mahalanobis distances), further research needs to be done on factor analyses relying on more robust criteria, for example based on the median of absolute deviations from the median of the niche (Cleveland, 1993).

Finally, one of the main issues regarding the statistical analysis of this type of data (therefore including MADIFA) is that most of the time the sample is not obtained using proper sampling designs that lead to unbiased estimation (e.g. random sampling or systematic sampling). The data concerning the chamois in the mountains of Les Bauges were obtained after a complete, therefore unbiased, census of the population in open areas, so that we did not meet this kind of problem. However, such sources of bias should be carefully checked in studies carried out at very large scale, especially in biogeography, where proper sampling is not possible (e.g. Spichiger et al., 2004).

5.6.2 Conclusions

The MADIFA is to be used jointly with other exploratory methods to visualize the structures of the niche. Classical PCAs can be used to identify correlates between environmental variables both in the species niche and on the study area. The MADIFA returns an image of the ecological space, and also allows visualization of the niche patterns in the geographical space, through the computation of an environmental suitability map (ESM). The ENFA may, in addition, be used to distinguish the parts of the Mahalanobis distances caused by the specialization and the marginality of the species. By matching all these results and the results of simpler descriptive statistics (e.g. histograms), the researcher can build a conceptual model of the biological system under study. The understanding of this system may be of major use for the estimation of more complex predictive models.

Acknowledgments

We warmly thank the Office national de la chasse et de la faune sauvage for their financial support, and the Parc Naturel Régional des Bauges for providing the maps of environmental variables. We also thank two anonymous referees for their very constructive comments, which greatly improved the quality of this paper.

Appendix A

Mahalanobis distances (Ecological Archives E089-030-A1).

Appendix B

Demonstration: the sum of the squared scores of the pixels on the factorial axes of the MADIFA is equal to the Mahalanobis distances (Ecological Archives E089-030-A2).

Appendix C

Results of the principal component analyses performed to identify the correlations on the study area, and in the chamois niche (Ecological Archives E089-030-A3).

Chapter 6

A General Framework for the Statistical Exploration of the Ecological Niche

Clément Calenge[1,2,*] & Mathieu Basille[1,3]

---

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69 622, Villeurbanne, France.

[2] Office national de la chasse et de la faune sauvage, 5, rue de Saint Thibault, Saint Benoist, 78 610, Auffargis, France.

[3] Centre for Conservation Biology, Department of Biology, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway.

[*] Corresponding author: `calenge@biomserv.univ-lyon1.fr`

Accepted in the Journal of Theoretical Biology, 2008.

"If you can't win by reason, go for volume."
"Calvin and Hobbes", by Bill WATTERSON.

Abstract

We propose a new statistical framework for the exploratory analysis of the ecological niche, the "General niche-environment system factor analysis" (GNESFA). The data required for this analysis are (i) a table giving the values of the environmental variables in each environment unit (EU, e.g. the patches of habitat on a vector map), (ii) a set of weights measuring the availability of the EUs to the species (e.g. the proportion of the study area covered by a given patch), and (iii) a set of utilization weights describing the use of the EUs by the focal species (e.g. the proportion of detections of the species in each patch). Each row of the table corresponds to a point in the multidimensional space defined by the environmental variables, and each point is associated with two weights. The GNESFA searches the directions in this space where the two weight distributions differ the most, choosing one distribution as the reference, and the other one as the focus. The choice of the utilization as the reference corresponds to the MADIFA (Mahalanobis distances factor analysis), which identifies the directions on which the available EUs are in average the furthest from the optimum of the niche, allowing habitat suitability modelling. The choice of the availability as the reference corresponds to the FANTER (Factor analysis of the niche, taking the environment as the reference), which identifies the directions on which the niche is the furthest from the average environment (marginality) and those on which the niche is the narrowest compared with the environment (specialization). The commonly used ENFA (Ecological niche factor analysis) is at the middle point between the MADIFA and the FANTER, considering both distributions as the reference and the focus simultaneously. When used concurrently, these three analyses allow an extensive exploration of the system.

Keywords: Multivariate analysis; Factor analysis; inertia ratio; habitat selection; General niche-environment system factor analysis

TWO KINDS OF STUDIES CAN BE CARRIED OUT to study the relationships between a species and its environment (Morrison et al., 2006; Calenge et al., 2005). On one hand, hindcasting studies aim to emphasize among a large set of environmental variables those that are of interest for the species. On the other hand, forecasting modelling is used to predict suitable environments for the species in new unsampled areas, and/or under different environmental conditions (Guisan et al., 2006; Knick and

Rotenberry, 1998). With the predicted global warming of the climate, it is of increasing importance to predict the behavior of keystone species in response to various scenarios of future climate (e.g. Araújo et al., 2005). For this reason, the social and scientific demand for predictive models is presently very strong (Elith et al., 2006). However, statistical methods allowing forecasting modelling, such as general linear model or related methods, can deal only with a limited number of environmental variables (Guisan and Zimmermann, 2000). For this reason, it is generally supposed in such studies that "the modeller knows the limiting factors that influence the distribution and abundance of the study organism" (Boyce and McDonald, 1999). All these methods rely on the hypothesis that a large amount of biological knowledge concerning the species is available prior to the study (Burnham and Anderson, 1998). In other words, hindcasting studies should necessarily precede a forecasting approach (Soberón and Peterson, 2005).

Visualization is an essential step of hindcasting studies. As noted by Cleveland (1993), "visualization is critical to data analysis. It provides a front line of attack, revealing intricate structure in data that cannot be absorbed in any other way. We discover unimagined effects, and we challenge imagined ones". In particular, only the use of a visualization approach would reveal hidden structures and other "surprises" in the data (Cleveland, 1994).

The graphical exploration of the relationships between a species and its environment may rely on the formal concept of ecological niche (Hutchinson, 1957). Each environmental variable can define a dimension of a multidimensional space, namely the ecological space. In that space, the distribution of the species occurrences represents the niche, which can be compared to the environment defined as available to the species (e.g. pixels of a raster map). This concept allows both a graphical and a quantitative exploratory analysis, in order to identify the directions in the ecological space where the distribution of the species is most different from the distribution of points describing the environment available to the species. However, the present "lack of effective tools for exploring, analysing, and visualizing ecological niches in many-dimensional environmental space" (Soberón and Peterson, 2005) may render this task difficult.

Factor analyses have numerous desirable qualities for data exploration in highly multidimensional spaces, especially for visualizing structures in the data (e.g. Hill, 1974; Blondel et al., 1988; Thioulouse and Chessel, 1992; Dray et al., 2003). For this reason, factor analyses may have a key role in hindcasting studies of species-environment relationships. They could be used to sort factorial axes which support most of the difference between the niche and its environment. The patterns of the niche-environment system identified by the analysis can then be related to the choices (habitat selection) or the requirements (niche characteristics) of the species concerning its habitat.

The Ecological-niche factor analysis (ENFA, Hirzel et al., 2002a) and the Mahalanobis distances factor analysis (MADIFA, Calenge et al., 2008) are two such methods. On one hand, the ENFA distinguishes two kinds of information measured in the niche-environment system, marginality and specialization. The marginality is a measure of the eccentricity of the niche relative to the distribution of available points in the ecological space, whereas the specialization is a measure of the niche restriction relative to the distribution of available points. The ENFA comes up with the directions in the ecological space where first the marginality, and then the specialization are maximised. On the other hand, the MADIFA relies only on one measure of habitat suitability, Mahalanobis distance, computed in the ecological space from the centroid of the distribution of the species occurrences to all available points. The Mahalanobis distance gives an index of the environmental suitability at this place, as it indicates the departure from the species' optimum (a low distance value is supposed to indicate a strong suitability, see Clark et al., 1993; Knick and Dyer, 1997). The MADIFA returns the directions in the ecological space where the Mahalanobis distances of the available sites are, on average, the largest. Both the ENFA and the MADIFA have been proposed as appropriate tools for drawing factorial maps of the niche in the ecological space, or for building reduced-rank environmental suitability maps (Hirzel et al., 2002a; Calenge et al., 2008).

From a formal point of view, the ENFA and MADIFA are actually closely related (Calenge et al., 2008). In this paper, we extend the mathematical relationships between the ENFA and the MADIFA to develop a more general framework encompassing these two methods, the General

Niche-Environment System Factor Analysis (GNESFA). This framework also includes another factor analysis of the niche-environment system, the Factor analysis of the niche, taking the environment as the reference (FANTER), which offers a third point of view regarding this system. These three exploratory methods return complementary results, as illustrated by two examples. When used concurrently, they provide an extensive summary of the patterns in the data.

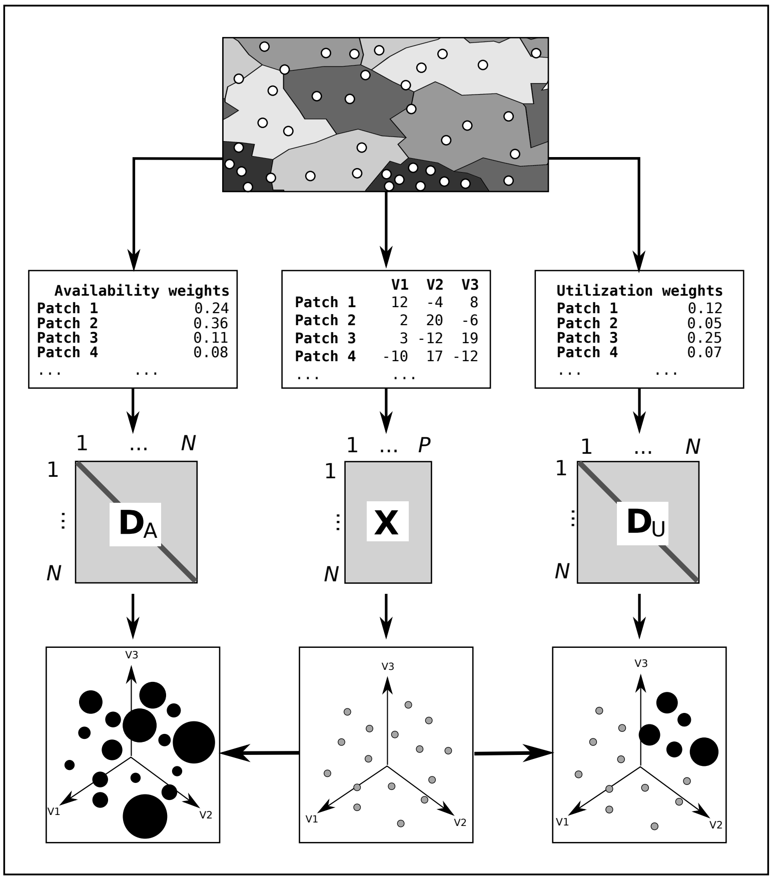
6.2 The algorithm

6.2.1 Notations and Definitions

We suppose that the study area is made of a set of N discrete environment units (EU), on which P environmental variables are measured (Fig. 6.1). These EUs may be, for example, the pixels of a raster map, or the patches of environment on a vector map. These values are stored in the $(N \times P)$ matrix $\mathbf{X}$. Because the environmental variables may not be measured using the same units (e.g. the elevation is measured in meters, and the slope in percent), we suppose that the columns of $\mathbf{X}$ have been standardised prior to the analysis (with zero mean and unit variance).

A weight describes the availability of each EU to the focal species. This "availability weight" could be, for example, the proportion of the study area covered by a patch of environment (in the case of raster maps, all these availability weights may be equal). This set of weights —defining the "availability distribution"— is stored on the diagonal of the $(N \times N)$ matrix $\mathbf{D}_A$ (the values of the non-diagonal elements of $\mathbf{D}_A$ are set to 0).

Additionally, an "utilization weight" describes the intensity of use of each EU by the focal species. This weight could be, for example, the proportion of detections of the species located in the patch during a census operation. The set of utilization weights —defining the "utilization distribution"— is stored on the diagonal of the $(N \times N)$ matrix $\mathbf{D}_U$ (the values of the non-diagonal elements of $\mathbf{D}_U$ are set to 0). Note that both the utilization and availability weights sum to 1 by construction. The aim of the GNESFA is to identify the directions in the ecological space where the two weight distributions differ most, which we call "niche patterns".



	Availability weights
Patch 1	0.24
Patch 2	0.36
Patch 3	0.11
Patch 4	0.08
...	...

	V1	V2	V3
Patch 1	12	-4	8
Patch 2	2	20	-6
Patch 3	3	-12	19
Patch 4	-10	17	-12
...	...		

	Utilization weights
Patch 1	0.12
Patch 2	0.05
Patch 3	0.25
Patch 4	0.07
...	...

Figure 6.1: The data design required by the GNESFA: the study area is partitioned into a set of N discrete environment units (here, patches of environments). For each patch, P environmental variables are measured (here, V1, V2, and V3 and stored in the matrix X). Each variable defines a dimension of the ecological space. For each environment unit, an "availability weight" (stored in the diagonal matrix $\mathbf{D}_A$) defines its availability to the species, and an "utilization weight" (stored in the diagonal matrix $\mathbf{D}_U$) measures its intensity of use by the species.

6.2.2 Choice of a reference

The GNESFA implies a choice of one of these two weight distributions as a reference distribution, and the other as a focus distribution (Fig. 6.2). The cloud of points defined by the rows of X will be distorted so that this cloud, considered from the point of view of the reference distribution, will take a standard spherical shape (i.e. with a variance of the available EUs equal to one for all directions of the ecological space). Then, the shape of the cloud of points considered from the point of view of the focus distribution will be investigated in this standard space, and any deviation from this spherical shape will indicate a pattern. The choice of a reference distribution depends on the needs of the analyst.

On one hand, when the main interest of the analysis is related to the identification of the variables affecting the shape (unimodal vs multimodal niches), the central tendency (marginal species or not), and the spread of the niche relative to the environment (specialized species or not), the availability distribution should be chosen as the reference, and the utilization distribution as the focus. Such a choice implies that the patterns of the available EUs are known (i.e. the correlation structure of the environmental variables on the study area), and would correspond to the point of view of the expert of the study area. Therefore, the choice of the availability as a reference allows a detailed exploratory analysis of the patterns displayed by the niche in the ecological space. This is the case of the Factor analysis of the niche, taking the environment as the reference (FANTER, Fig. 6.2).

On the other hand, in some cases, we are more interested in the patterns of the environment relative to the niche. For example, the suitability of the available EUs can be measured by the distance between these EUs and the utilization distribution as a whole (Clark et al., 1993). It follows that, if one wants to know the most suitable conditions of an area considered as available to the species, the utilization distribution should be chosen as the reference. The distribution of used EUs will then take a standard shape, and the GNESFA will indicate the direction of the ecological space in which the available EUs are the most different from this distribution. Such a choice implies that the patterns of the utilization distribution are known, and corresponds to the point of view of the expert of the studied species. In that context, we are mainly interested in whether the species

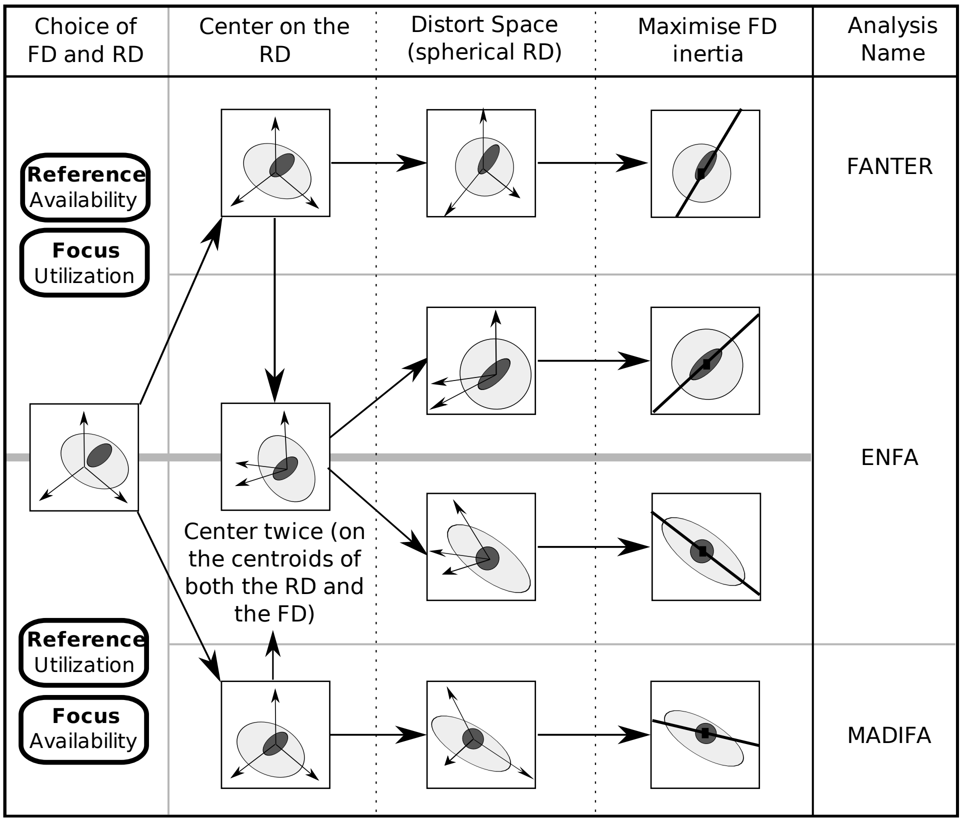


Figure 6.2: The three possible analyses encompassed by the GNESFA. The light grey ellipse represents the distribution of availability weights in the ecological space and the dark grey ellipse represents the distribution of utilization weights in the ecological space. The FANTER uses the availability distribution as reference and the utilization distribution as focus. The MADIFA uses the utilization distribution as reference and the availability distribution as focus. The ENFA can use both approaches (RD = Reference distribution; FD = Focus distribution).

"considers" the proposed environment (available EUs) as suitable (within the niche) or not (far from the niche). This is the approach used by the Mahalanobis distances factor analysis (MADIFA, Fig. 6.2).

Finally, another point of view is possible, for which each of the two distributions are both the reference and the focus distribution. This symmetrical point of view has the advantage of not relying on the choice of one distribution as the reference. However, we will later prove that this choice also implies the loss of one dimension of the ecological space, a dimension which may carry important biological information. This special case is the basis of the Ecological-niche factor analysis (ENFA, Fig. 6.2).

These three approaches define the framework of the GNESFA that we describe now more formally. The reference distribution is described by the weight matrix $\mathbf{R}$, and the focus distribution is described by the weight matrix $\mathbf{F}$. For example, if the availability distribution is chosen as reference, then $\mathbf{R} = \mathbf{D}_A$ and $\mathbf{F} = \mathbf{D}_U$. In this paper, we will refer to the "reference mean", "reference variance", "focus mean" or "focus variance" of a variable, depending on the computed statistic (mean or variance) and the weight distribution chosen to compute this statistic ($\mathbf{R}$ and $\mathbf{F}$, respectively). We call the "centroid of the reference distribution" the point in the ecological space defined by the vector $\mathbf{X}^t\mathbf{R}\mathbf{1}_N$ (i.e. the P-vector of reference means computed for all the environmental variables). Similarly, we will refer to the "centroid of the focus distribution" if the chosen weight matrix is the matrix $\mathbf{F}$.

6.2.3 The centering

The first step of the GNESFA is the centering of the table $\mathbf{X}$. Actually, "mathematically and geometrically, centering involves the specification of the origin (...). It is the 'point of zero information'; anything that is at it, is trivial and uninteresting; anything that deviates from it is information" (Noy-Meir, 1973). Therefore, it seems logical to choose the centroid of the reference distribution as the origin of the ecological space to perform the GNESFA. The centered table $\mathbf{Z}$ is thus computed by (Seber, 1984):

$$\mathbf{Z} = \left(\mathbf{I}_N - \mathbf{1}_N\mathbf{1}_N^t\mathbf{R}\right)\mathbf{X}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix, and $\mathbf{1}_N$ is a N-vector of 1s.

6.2.4 First Principal Component Analysis

The second step of the GNESFA consists of a principal component analysis (PCA) of the table **Z**, using the matrix **R** as row weights. Actually, the PCA of **Z** consists of the search for the eigenstructure of the variance-covariance matrix **V**:

$$\mathbf{V} = \mathbf{Z}^t\mathbf{R}\mathbf{Z}$$

Let **U** be the $(N \times P)$ matrix containing the P eigenvectors $\mathbf{u}_j$ of **V** concatenated by column, and let $\boldsymbol{\delta}$ be the diagonal matrix containing the corresponding eigenvalues λ_j, on the diagonal. In other words:

$$\mathbf{V}\mathbf{U} = \mathbf{U}\boldsymbol{\delta}$$

The $(P \times P)$ matrix **U** contains the scores of environmental variables (rows) on each principal axis of the analysis (columns). Moreover, the $(N \times P)$ matrix $\mathbf{L}^\bullet = \mathbf{Z}\mathbf{U}$ contains the coordinates of the EUs (rows) on the principal components of the analysis (columns) (Legendre and Legendre, 1998).

Because the table **Z** is centered for the weight matrix **R**, this analysis is just a classical PCA, i.e. a rotation of the cloud of points so that: (i) the reference variance of the EU coordinates on the first principal components is maximised (it is equal to the corresponding eigenvalues), and (ii) the correlation between the coordinates of the EUs on different principal components is equal to 0 (Legendre and Legendre, 1998). Let the $(N \times P)$ matrix **L** be equal to:

$$\mathbf{L} = \mathbf{L}^\bullet\boldsymbol{\delta}^{-1/2} = \mathbf{Z}\mathbf{U}\boldsymbol{\delta}^{-1/2}$$

This matrix contains the normed coordinates of the EUs (rows) on the principal components (columns) of the analysis. This matrix has the following property:

$$\mathbf{L}^t\mathbf{R}\mathbf{L} = \mathbf{I}_P \qquad (6.1)$$

where $\mathbf{I}_P$ is the $P \times P$ identify matrix. This equation shows that the product of the EU coordinates and $\boldsymbol{\delta}^{-1/2}$ results in a distortion of the cloud of EUs in the ecological space, so that the reference variance of these coordinates after distortion is equal to 1 for all the principal components. As these components are still uncorrelated, it follows that the cloud of points described by X has been “sphericized”, from the point of view of the reference distribution.

6.2.5 Second Principal Component Analysis: the core of the GNESFA

The last step of the GNESFA is the analysis of the focus distribution in this distorted ecological space. This analysis is done using a non-centered PCA of the table **L**, with the matrix **F** as row weights.

Since the cloud of points has been “sphericized” with respect to the reference distribution, it should also be spherical from the point of view of the focus distribution, if it is identical to the reference distribution. In other words, all the eigenvalues of this PCA should be equal, which would indicate that: (i) the centroid of the focus distribution is the same as the centroid of the availability distribution (because the analysis is not centered), and (ii) the variance of the focus distribution is the same in all the directions of the ecological space. More formal justifications will be given in the next section.

This PCA is performed by computing the eigenstructure of the matrix **H**:

$$\mathbf{H} = \mathbf{L}^t\mathbf{F}\mathbf{L}$$

This PCA is non-centered because **L** is not centered for the weight matrix **F**. Let **W** be the matrix $P \times P$ containing the eigenvectors $\mathbf{w}_j$ of **H** concatenated by column, and the diagonal matrix $\boldsymbol{\alpha}$ containing the eigenvalues γ_j of **H** on the diagonal:

$$\mathbf{H}\mathbf{W} = \mathbf{W}\boldsymbol{\alpha}$$

The coordinates of the EUs on the principal components of the GNESFA are contained in the $N \times P$ matrix **P**:

$$\mathbf{P} = \mathbf{L}\mathbf{W} = \mathbf{Z}\mathbf{U}\boldsymbol{\delta}^{-1/2}\mathbf{W}$$

This equation summarizes the GNESFA: it consists of a first rotation (matrix **U**), a distortion (matrix $\boldsymbol{\delta}^{-1/2}$) and a second rotation (matrix **W**) of the cloud of points in the ecological space (matrix **Z**).

Note that the cloud of points is still spherical with respect to the reference distribution on the components of the GNESFA:

$$\mathbf{P}^t\mathbf{R}\mathbf{P} = \mathbf{W}^t\mathbf{L}^t\mathbf{R}\mathbf{L}\mathbf{W} = \mathbf{I}_P \tag{6.2}$$

This arises from equation 6.1 and from the observation that $\mathbf{W}$ is an orthogonal matrix (i.e. $\mathbf{W}^t\mathbf{W} = \mathbf{W}\mathbf{W}^t = \mathbf{I}_P$).

Factorial maps of the niche in the ecological space can be obtained by plotting the coordinates of the EUs on a restricted number of principal components, as in classical PCA. The biological meaning of the principal components can be derived from the correlations between the environmental variables and the principal components of the analysis. Note that some school of thought prefer to interpret the meaning of the principal components from the scores of the environmental variables on the principal axes of the GNESFA, contained in the matrix $\mathbf{A}$ (Rotenberry et al., 2006):

$$\mathbf{A} = \mathbf{U}\delta^{-1/2}\mathbf{W}$$

We advocate the use of the correlations to interpret the meaning of the principal components. Indeed, the coefficient associated with a given environmental variable may be misleading when this variable is correlated with other environmental variables in $\mathbf{X}$ (Basille et al., 2008).

To choose the number of components to keep for the interpretation, we can look for a break in the decrease of the eigenvalues (broken-stick method, Barton and David, 1956; Frontier, 1976). The biological meaning of these eigenvalues depends on the weight matrix chosen as reference distribution, as developed in the next section.

6.2.6 The inertia ratio

We now investigate the mathematical meaning of the statistic maximised by the GNESFA on the first components of the analysis. Because the second step of the GNESFA is a PCA, the j^{th} eigenvalue γ_j of this analysis is equal to:

$$\gamma_j = \mathbf{p}_j^t\mathbf{F}\mathbf{p}_j \tag{6.3}$$

where $\mathbf{p}_j$ is the j^{th} column of $\mathbf{P}$ (i.e. the j^{th} component of the analysis). Note that, as in classical PCA, this statistic is at its maximum on the first axis of the analysis. For this reason, we focus on this first axis to make explicit the properties of this statistic. The first axis of the GNESFA $\mathbf{a}_1$ (i.e. the first column of $\mathbf{A}$) fulfills the following conditions:

$$\begin{aligned}
\mathbf{Z}\mathbf{a}_1 &= \mathbf{p}_1 && (6.4)\\
\mathbf{p}_1^t\mathbf{R}\mathbf{p}_1 &= 1 && (6.5)\\
\gamma_1 &= \mathbf{p}_1^t\mathbf{F}\mathbf{p}_1 \quad Max && (6.6)
\end{aligned}$$

The condition 6.4 means that the vector $\mathbf{p}_1$ contains a linear combination of the environmental variables, using the values stored in $\mathbf{a}_1$ as coefficients. This linear combination gives the coordinates of the EUs on the principal components of the analysis, such that: (i) the reference variance of these coordinates is equal to 1 (condition 6.5, arising from equation 6.2), and (ii) the focus mean of squared coordinates is maximised (condition 6.6, arising from equation 6.3).

Actually, we can show that the conditions 6.4, 6.5, and 6.6 define a problem mathematically equivalent to the search for a P-vector $\mathbf{g}_1$, fulfilling the following conditions (see appendix A):

$$\begin{aligned}
\mathbf{g}_1^t\mathbf{g}_1 &= 1 && (6.7)\\
\mathbf{Z}\mathbf{g}_1 &= \mathbf{y} && (6.8)\\
\gamma_1 &= \frac{\mathbf{y}^t\mathbf{F}\mathbf{y}}{\mathbf{y}^t\mathbf{R}\mathbf{y}} \quad Max && (6.9)
\end{aligned}$$

The vector $\mathbf{g}_1$ is normed to 1 (condition 6.7). Consequently, the vector $\mathbf{y}$ contains the coordinates of the orthogonal projections of the undistorted cloud of EUs in the ecological space on the vector $\mathbf{g}_1$ (condition 6.8). The coordinates of these projections are such that the ratio γ_1 is maximised (condition 6.9). This equivalence between the two problems is demonstrated in appendix A, with

$$\mathbf{g}_1 = \frac{\mathbf{a}_1}{\sqrt{\mathbf{a}_1^t\mathbf{a}_1}}$$

Therefore, the first eigenvalue γ_1 of the GNESFA is equal to:

$$\gamma_1 = \frac{\sum_{i=1}^{n} f_i (y_i - \bar{y}_r)^2}{\sum_{i=1}^{n} r_i (y_i - \bar{y}_r)^2}$$

where y_i is the i^{th} element of the vector $\mathbf{y}$, $\bar{y}_r$ is the reference mean of $\mathbf{y}$, r_i is the i^{th} reference weight, and f_i is the i^{th} focus weight.

The denominator of γ_j is the reference variance of $\mathbf{y}$. On the contrary, the numerator is generally not a variance, except if the focus mean of $\mathbf{y}$ is equal to its reference mean: it is the focus mean of the squared distances between the EUs and the centroid of the reference distribution. Such a sum of squared distances is sometimes called "inertia" (Dray et al., 2003). For this reason, we named γ_j the "inertia ratio" of the j^{th} component of the GNESFA. The biological meaning of this ratio depends on the distribution chosen as the reference (see below).

6.3 Three special cases

As explained before, the GNESFA encompasses three methods, depending on the choice made by the analyst for the focus and reference distributions (Fig. 6.2). When the utilization distribution is chosen as reference, the resulting analysis is the MADIFA (Calenge et al., 2008). When the availability distribution is chosen as reference, the resulting analysis is a new analysis, which we called the Factor analysis of the niche, taking the environment as the reference (FANTER). Finally, the ENFA (Hirzel et al., 2002a) is also a special case of the GNESFA, provided that a special transformation of the table X has been carried out prior to the analysis. Because the point of view of the ENFA is central to the understanding of the differences between these three analyses, we first describe how the ENFA takes place within the framework of the GNESFA.

6.3.1 The ENFA: a prior transformation of the table

A classical approach for characterizing a statistical distribution is to provide a measure of its central tendency (e.g. mean, median) and a measure of its spread (e.g. variance, interquartile range). The ENFA relies on this approach, in a multidimensional space, to characterize the niche-environment system.

On one hand, the "central tendency" of the utilization distribution relative to the availability distribution is measured by the vector connecting the centroids of the two distributions, named the "marginality vector". Its computation can be performed in the following way: first the table $\mathbf{X}$ is centered for the weight matrix $\mathbf{D}_A$:

$$\mathbf{C} = \left(\mathbf{I}_N - \mathbf{1}_N \mathbf{1}_N^t \mathbf{D}_A\right) \mathrm{X}$$

Then the marginality vector $\mathbf{m}$ can be computed as the P-vector containing the utilization means of the columns of $\mathbf{C}$:

$$\mathbf{m} = \mathbf{C}^t \mathbf{D}_U \mathbf{1}_N$$

This vector contains the differences between the utilization and availability means of all the environmental variables. Its squared length (equal to $m^2 = \mathbf{m}^t\mathbf{m}$), named "marginality", is a measure of the eccentricity of the utilization distribution relative to the availability distribution.

On the other hand, the spread of the utilization relative to the availability distribution may vary according to the considered direction of the ecological space. Hirzel et al. (2002a) proposed the use of the specialization ratio S:

$$S = \frac{\text{availability variance}}{\text{utilization variance}}$$

For a given environmental variable, a large S indicates that the environmental variability experienced by the species is much smaller than the range of variability that is actually available, and therefore that the species is highly specialized on this variable. The aim of the ENFA is to identify the directions, in the ecological space, where the specialization ratio S is maximised.

The equivalence of the ENFA and the GNESFA can be proved by noting that the specialization ratio S is a special case of the inertia ratio γ_j. We noted previously that the numerator of γ_j is a variance only if the centroid of the reference distribution is the same as the centroid of the focus distribution. Therefore, the only way to compute the specialization ratio is to consider only those directions of the ecological space orthogonal to the marginality vector. For these directions, both the availability means and the utilization means are equal to zero. The projection of the EUs on the hyperplane orthogonal to $\mathbf{m}$ is carried out by the following operation (Harville, 1997):

$$\mathbf{C}_c = \mathbf{C}(\mathbf{I}_v - \frac{\mathbf{m}\mathbf{m}^t}{\mathbf{m}^t\mathbf{m}}) \tag{6.10}$$

The GNESFA of the table $\mathbf{C}_c$, using $\mathbf{D}_U$ as the reference distribution, and $\mathbf{D}_A$ as a focus distribution, is the solution to the following problem:

$$\begin{aligned}
\mathbf{g}_1^t\mathbf{m} &= 0 && (6.11)\\
\mathbf{g}_1^t\mathbf{g}_1 &= 1 && (6.12)\\
\mathbf{Z}\mathbf{g}_1 &= \mathbf{y} && (6.13)\\
S &= \frac{\mathbf{y}^t\mathbf{D}_A\mathbf{y}}{\mathbf{y}^t\mathbf{D}_U\mathbf{y}} \quad Max && (6.14)
\end{aligned}$$

The conditions 6.12 and 6.14 are identical to the conditions 6.7 and 6.9 previously defined. The condition 6.11 derives from equation 6.10 (as the EUs are all located on the hyperplane orthogonal to the marginality vector). The condition 6.13 derives from the condition 6.8 (i.e. in this case $\mathbf{C}_c\mathbf{g}_1 = \mathbf{y}$) and the condition 6.11. This problem is exactly the problem of the ENFA defined by Hirzel et al. (2002a). The ENFA is therefore a special case of the GNESFA.

We can see that the first component of the GNESFA of the table $\mathbf{C}_c$ using $\mathbf{D}_A$ as reference distribution maximises $1/S$. Consequently, this component is the same as the last component of the GNESFA of the table $\mathbf{C}_c$, using $\mathbf{D}_U$ as reference distribution. Thus, the GNESFA of the table $\mathbf{C}_c$ can be seen as an ENFA, whatever the weight distribution chosen as reference. When $\mathbf{D}_U$ is chosen as the reference, the GNESFA is the classical ENFA. When $\mathbf{D}_A$ is chosen as the reference, the GNESFA is a "reversed" ENFA (with the first axes of the classical ENFA being the last axes of the "reversed" ENFA). Therefore, the roles of the utilization and availability distributions are symmetric in the ENFA (see Fig. 6.2). This corresponds to the symmetrical point of view described previously.

Whatever the importance of the specialization, the dimension of the ecological space defined by the marginality vector may carry important information about the niche-environment pattern, and the biological meaning of this dimension should also be interpreted (Hirzel et al., 2002a). Thus, the EUs are also projected on this vector, to define the marginality component $\mathbf{b}$:

$$\mathbf{b} = \mathbf{C}\frac{\mathbf{m}}{\sqrt{\mathbf{m}^t\mathbf{m}}}$$

and the values in $\mathbf{b}$ can be plotted vs. values in $\mathbf{y}$ (row coordinates on the specialization axes) to build factorial maps of the niche-environment system. Basille et al. (2008) have proved that such maps give an optimal image of the niche from the point of view of the ENFA (distinction between marginality and specialization), and are undistorted because the marginality vector is orthogonal to the specialization axes of the ENFA.

However, the ENFA may give problematic results in three cases: (i) for some datasets, the marginality is not biologically significant. In such cases, imposing the constraint of orthogonality of the specialization axes to the marginality vector may lead to meaningless results. Indeed, even if the marginality is not strong, the specialization may be important for the marginality component, and the orthogonality constraint may obscure this characteristic of the data; (ii) although this method has been widely used to build habitat suitability maps (e.g. Zaniewski et al., 2002; Reutter et al., 2003), we believe that this method should not be used to reach such a goal. Because the marginality component does not have the same mathematical status as the successive specialization components, it is very peculiar to combine all of them into a single value of habitat suitability. Even though the *ad hoc* methods existing in the literature (Hirzel and Arlettaz, 2003b) have returned biologically consistent results, we believe that the ENFA is not designed to build such maps, and that better methods exist for this objective (see below); (iii) finally, the ENFA relies on the hypothesis that both the utilization and the availability distributions are symmetric and unimodal (Hirzel et al., 2002a). In the case of a multimodal niche, the ENFA is not recommended (see examples below).

6.3.2 The MADIFA: utilization distribution as reference

The specialization and the marginality are clear measures of the niche patterns in the ecological space. We can express the inertia ratio γ_j as a function of the marginality and the specialization, in order to give insight into the differences between the ENFA, the MADIFA and the FANTER.

When the reference is the utilization distribution (i.e. $\mathbf{R} = \mathbf{D}_U$ and $\mathbf{F} = \mathbf{D}_A$), the ratio $\gamma_{(m)j}$ (subscript "m" stands for "MADIFA"), maximised by the analysis can be reformulated:

$$\gamma_{(m)j} = \frac{m_j^2}{s_{(u)j}^2} + S_j$$

where $s_{(u)j}^2$ is the utilization variance of the j^{th} component, m_j^2 measures the marginality on this component (the squared difference between the availability mean and the utilization mean of this component), and S_j is the specialization on this component. Thus, this analysis combines the marginality and the specialization into one single value: the larger the inertia ratio, the higher the marginality and/or the specialization.

Calenge et al. (2008) has already described this special case of the GNESFA, and called it MADIFA (Mahalanobis distances factor analysis). The authors noted an interesting property of this analysis: the sum of squared scores of an available EU over all the components of the analysis is equal to the Mahalanobis distance between this EU and the centroid of the utilization distribution. More formally,

$$D_i^2 = \mathbf{P}_{i\bullet}\mathbf{P}_{i\bullet}^t = \sum_{j=1}^{P} p_{ij}^2$$

where the P-vector $\mathbf{P}_{i\bullet}$ is the i^{th} row of the matrix $\mathbf{P}$, and p_{ij} is the coordinate of the i^{th} EU on the j^{th} component of the GNESFA. Thus, the squared coordinate of a EU on a given component of the GNESFA can be considered as the contribution of this component to the Mahalanobis distance between this EU and the centroid of the reference distribution. This property is interesting because the Mahalanobis distances have been used in many studies as indices of environmental suitability for species (Clark et al., 1993; Farber and Kadmon, 2003; Knick and Dyer, 1997; Cayuela, 2004; Browning et al., 2005; Corsi et al., 1999), especially to build environment suitability maps.

Recalling equation 6.3, the inertia ratio on the j^{th} component of the GNESFA is equal to:

$$\gamma_j = \mathbf{p}_j^t \mathbf{D}_A \mathbf{p}_j = \sum_{i=1}^{N} a_i p_{ij}^2$$

Where a_i is the availability weight associated with the i^{th} EU. It follows that the j^{th} eigenvalue of the analysis corresponds to the availability mean of the contributions of the j^{th} component to the Mahalanobis distances between the EUs and the centroid of the utilization distribution.

Therefore, the MADIFA finds the directions of the ecological space where these distances are the largest, corresponding to the environmental conditions that are scarcely used by the species. Because they all have the same mathematical status, these components can be easily combined to build reduced-rank environment suitability maps with increased generality (see Calenge et al., 2008, for details). These directions are those where the marginality and/or the specialization are the largest. Thus, this analysis identifies all the patterns of the niche-environment system on the first principal components.

Note that the MADIFA relies on the hypothesis that the utilization distribution is both unimodal and symmetric (Calenge et al., 2008) and is therefore not recommended for the study of multimodal niches (as the ENFA). However, this analysis does not rely on any assumption concerning the shape of the availability distribution.

6.3.3 The FANTER: availability distribution as reference

When the reference is the availability distribution (i.e. $\mathbf{R} = \mathbf{D}_A$ and $\mathbf{F} = \mathbf{D}_U$), the ratio $\gamma_{(f)j}$ (subscript "f" stands for "FANTER") maximised by the GNESFA can be reformulated:

$$\gamma_{(f)j} = \frac{1}{S_j} + \frac{m_j^2}{s_{(a)j}^2}$$

where $s_{(a)j}^2$ is the availability variance of the j^{th} component. Consequently, a large $\gamma_{(r)j}$ may indicate that the marginality is large, but also that the specialization is low. On the other hand, a low $\gamma_{(r)j}$ indicates a strong specialization and/or a low marginality. Therefore, the first components of this analysis are those for which the marginality is maximised, whereas the last components are those on which the specialization is maximised. As such, both the first and the last components are of interest. Thus, the FANTER could be used as a preliminary to the ENFA, because it assesses whether it is of interest to partition the niche inertia into marginality and specialization components.

As for the MADIFA, it is straightforward to show that the j^{th} eigenvalue of the analysis corresponds to the utilization mean of the contributions of the j^{th} component to the Mahalanobis distances between the EUs and the centroid of the availability distribution. The first components are those on which the utilization distribution is the furthest from the availability distribution (i.e. the most marginal), whereas the last components are those on which the used EUs are the most concentrated around the availability mean (the most specialized).

Although the FANTER supposes that the availability distribution is symmetric and unimodal, it does not make the same hypotheses about the niche (contrary to the ENFA and the MADIFA). Thus, this analysis is suitable for the exploration of multimodal niches.

6.4 Examples

We illustrate here the framework of the GNESFA with the concurrent use of the ENFA, the MADIFA and the FANTER, based on two datasets collected on the chamois (*Rupicapra rupicapra*). For these two datasets, we performed these analyses to develop a conceptual model of the niche-environment system under study. For each analysis, we also tested the significance of the first (and last, for the FANTER) eigenvalue of the analysis using a randomisation test. At each of the 500 steps of the randomisation process, and for a given analysis, we randomly distributed the chamois locations on the area considered as available to it, and we computed again the eigenvalue of the analysis. We finally compared the observed eigenvalue with the distribution of eigenvalues simulated under the hypothesis of random habitat use, to derive a P-value. We also tested the significance of the marginality value (i.e. m^2), using similar randomization tests (see Basille et al., 2008), to assess the importance of this dimension in the ENFA. Because all the pixels of the raster maps cover the same area, we gave to them an equal availability weight (i.e. $\frac{1}{N}$) in the analyses.

6.4.1 The chamois population in the Chartreuse mountains

The first dataset was collected during censuses of the chamois population carried out in November 1997 in the Chartreuse mountains (French Alps, N. 45.33°, E. 5.80°) by the Departmental association of hunters of Isere.

Abbreviation	Description
Slope	Slope
Deciduous	Distance to deciduous woodland
Coniferous	Distance to coniferous woodland
Mixed	Distance to mixed woodland
Open	Distance to open land
Ecotone	Distance to the ecotone between open and forested areas (takes a positive value in open areas and a negative value in closed areas)

Table 6.1: Variables included in the analyses of habitat selection by 239 chamois groups detected during a census in the Chartreuse mountains (French Alps).

During the census, 239 chamois groups were detected in the study area. More details about the sampling protocol can be found in Michallet (2003). The raster maps of six environmental variables describing the habitat were used to define the ecological space (Table 6.1). We used the GNESFA to compare the distribution of the locations of the animals (utilization) vs. the distribution of the pixels of the entire area (available) in the ecological space.

The eigenvalue diagram of the FANTER indicates two patterns in the data, on the first and on the last components of the analysis (Fig. 6.3A). Indeed, the tests of the first ($\gamma_1 = 1.61$, $P < 0.002$) and of the last eigenvalue ($\gamma_6 = 0.53$, $P < 0.002$) were both significant. A clear "break" is apparent on this diagram before the last eigenvalue, and a slight break appeared after the first one. The first component was mainly correlated with the slope ($R = 0.84$), while the last was defined by the distance to the ecotone open/forested areas ($R = 0.62$), and to a lesser extent, by the distance to the mixed woodland ($R = 0.51$, fig 6.4A). The factorial map of the niche revealed that the niche was rather marginal on the first axis (fig. 6.4B): chamois were rarely located on flat terrain (only 16% of the chamois were detected on slopes lower than 25%, while this habitat type represented 40% of the study area). On the other hand, the specialization of the chamois population was maximum for the distance to the ecotone open/forested areas (25% percent of the chamois were located at an absolute distance value less than 100 meters, while this distance class represented only 13% of the study area) and, to a lesser extent, for the distance to the mixed

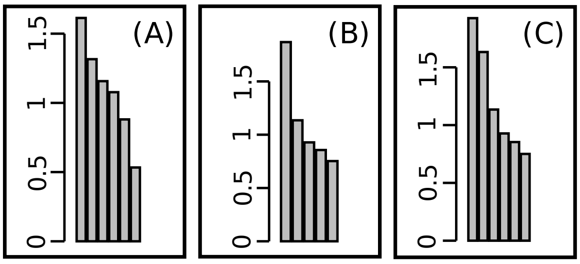


Figure 6.3: Barplots showing the eigenvalues of the GNESFA performed to identify correlates between the distribution of chamois detections in the Chartreuse mountain (French Alps) and 6 environmental variables (listed in Table 6.1) of the study area: (A) eigenvalues of the FANTER, the first and the last are kept for the analysis, (B) eigenvalues of the ENFA, only the first is kept, (C) eigenvalues of the MADIFA, the first two are kept.

woodland (40% percent of the chamois were located between 10 and 150 meters from this vegetation type, while this distance class represented only 30% of the study area).

The ENFA confirmed these results. A randomisation test of the first eigenvalue of this analysis indicated a significant specialization on at least one component ($S_1 = 1.87$, $P < 0.002$). There was a clear break after the first eigenvalue (Fig. 6.3B), so that we kept only the first specialization component —and the marginality component as it is a prerequisite of the analysis— for the interpretation. Note that the marginality value in the ENFA was also highly significant ($m^2 = 0.56$, $P < 0.002$). As expected, there was a very strong correlation between the marginality component of the ENFA and the first component of the FANTER ($R = 0.92$), and between the first specialization component of the ENFA and the last component of the FANTER ($R = -0.99$). Thus, in this example, the two analyses highlighted the same patterns. The position of the niche relative to the availability distribution was mainly determined by the slope, while its narrowness was determined by the distance to the ecotone open/forested areas, and to a lesser extent, by the distance to the mixed woodlands.

The test of the first eigenvalue of the MADIFA was also highly significant ($\gamma_1 = 1.93$, $P = 0.008$). The eigenvalue barplot indicated a clear break after the second eigenvalue (Fig 6.3C), and we therefore focused our interpretation on the first two components. Actually, the first component of the MADIFA was correlated with both the first specialization component of

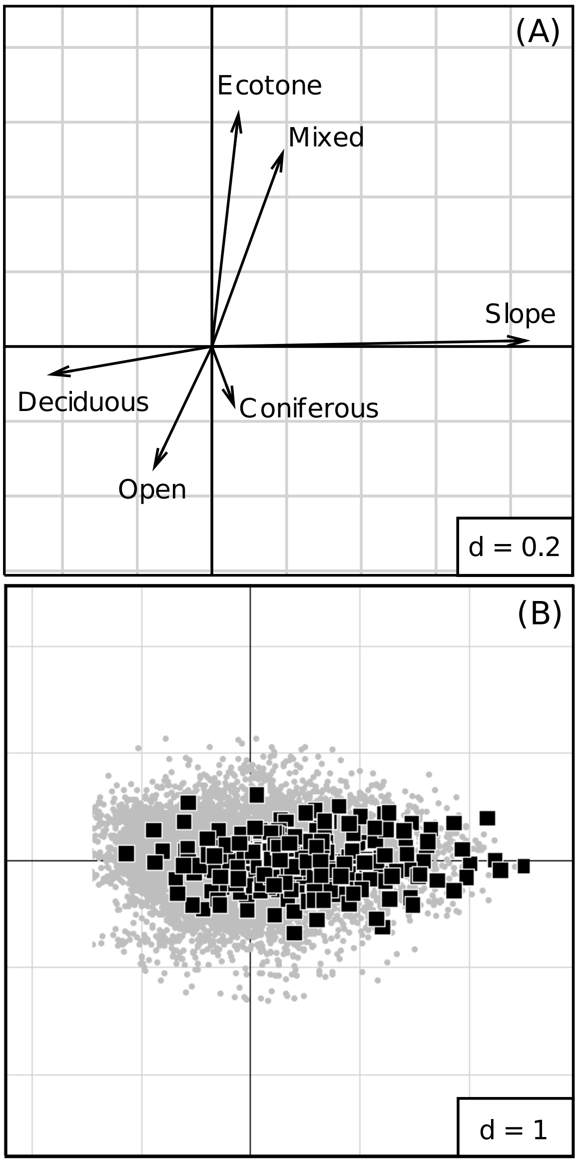


Figure 6.4: Results of the FANTER performed to identify correlates between the distribution of chamois detections in the Chartreuse mountain (French Alps) and 6 environmental variables (listed in Table 6.1) of the study area. (A) the correlations between the environmental variables and both the first (abscissa) and the last (ordinate) component of the analysis are indicated by arrows. For each variable, the length of an arrow on a given axis gives the value of the correlation between the variable and this component. Grid lines (separated by a distance of 0.2) can be used to measure these correlations on the graph; (B) scatterplot diagram of the cloud of available (grey circle) and used (black squares) points on the first two axes of the MADIFA.

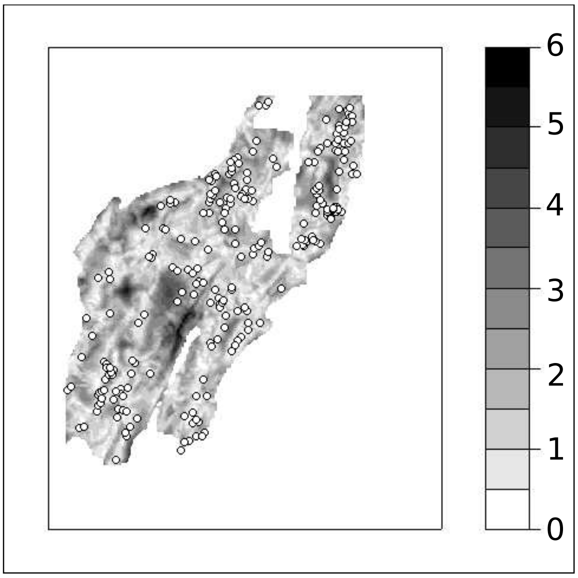


Figure 6.5: Environment suitability map for the chamois in the Chartreuse mountain (French Alps), computed by summing the squared coordinates of the pixels on the first two components of the MADIFA (approximate Mahalanobis distance between the pixels and the centroid of the niche). Lighter pixels correspond to suitable areas (low Mahalanobis distance) whereas darker pixels correspond to unsuitable areas (high Mahalanobis distance). The distribution of the chamois detections are also displayed.

the ENFA ($R = 0.94$) and the last component of the FANTER ($R = -0.96$). The second component of the MADIFA was strongly correlated with both the marginality component of the ENFA ($R = 0.90$), and the first component of the FANTER ($R = -0.96$). The coordinates of every pixel on the principal components of the MADIFA were combined to compute a reduced-rank environment suitability map (Fig. 6.5). This is done, for each pixel, by summing its squared coordinates on the first two components of the analysis (Calenge et al., 2008). This map can then be examined to identify the areas where rarely used environmental combinations are found, an additional information which helps to interpret the results.

The three analyses identified similar and consistent patterns in the data, and helped to draw a conceptual model of the niche. Chamois avoided flat terrain and preferred the ecotone between open land and mixed woodland. However, the three analyses were used to deliver different outputs. The FANTER was used as a preliminary analysis to identify the patterns in-

Abbreviation	Description
Elev	Elevation
D.FR	Distance to fallen rocks
D.SeCarS	Distance to meadows made of *Sesleria* and *Carex sempervirens*
D.Trail	Distance to recreational trails
Hydro	Hydrography
Slope	Slope
Sunshine	Sunshine
Visib1000	Visibility computed within a radius of 1000 m

Table 6.2: Variables included in the analyses of habitat selection by one chamois monitored using a GPS collar in the Bauges mountains (French Alps).

volved in the determination of the niche-environment system, whereas the ENFA distinguished precisely between the patterns determining the position and the spread of the niche. On the contrary, the MADIFA was unable to explicitly disentangle between marginality and specialization. Rather, it was used to combine these results to compute an environment suitability map. The three analyses returned the same pattern here, but they can highlight dramatically different results in some cases, as demonstrated by the next example.

6.4.2 The chamois radio-tracking in the Bauges mountain

The second dataset describes 56 daily relocations of one chamois in the Bauges mountains (French Alps, N. 45.63°,E. 6.23°). These data were collected from October 1st to November 27th 2003, using a GPS collar. We studied the habitat selection by this individual within its home range (third level of selection according to Johnson, 1980). We therefore computed the home-range limits of this chamois using the minimum convex polygon (Mohr, 1947), and defined the environment by eight environmental variables describing different characteristics potentially important for the chamois, measured for each pixel of a raster map (Table 6.2). We used the GNESFA to compare the distribution of the relocations of the monitored chamois (utilization) vs. the distribution of the pixels of its home range (available) in the ecological space.

The FANTER indicated a very clear structure driven by the first component (Fig. 6.6), confirmed by a randomisation test of the first eigenvalue ($\gamma_1 = 2.85$, $P = 0.02$). This component contrasted the areas located at high elevations, far from recreational trails, and located on the crests of the mountains (low hydrography), with areas with opposite characteristics. The niche of the individual was bimodal on this component: there was a first mode around the origin, and a second one located on the positive side of this component. Such a pattern can be understood by considering that the shape of the "niche" resulted from both a function giving the probability of selection by the chamois, and by the distribution of the available points. Because the distribution of available points showed a high peak at coordinates around zero, a moderate probability of selection for zero coordinates resulted in a peak at these coordinates (the proportion of used points was high at zero because the proportion of available points was high at zero). However, the peak on the positive side of this component revealed a strong selection of the mountain crest by this chamois (Fig. 6.6). In other words, the utilization distribution is a mix between a random use of space by this chamois (same shape as the available distribution) and a search for mountain crest (with a peak on the positive side). Note that the last eigenvalue of the FANTER, on which the specialization was maximised, did not differ significantly from the hypothesis of random habitat use ($\gamma_8 = 0.33$, $P = 0.73$).

The MADIFA did not highlight any particular pattern since the randomisation test of the first eigenvalue was not significant ($\gamma_1 = 3.29$, $P = 0.28$). Similarly, the test of the first eigenvalue of the ENFA did not reject the hypothesis of random habitat use ($\gamma_1 = 2.842$, $P = 0.24$). Note that the marginality was significant in the ENFA ($m^2 = 0.44$, $P = 0.01$). Actually, the marginality component of the ENFA was related to the first axis of the FANTER ($R = 0.76$). However, it does not make sense to use the ENFA or the MADIFA in such situations, as they both rely on the hypothesis of unimodal and symmetric niches. In such situations, the FANTER proves to be very useful, allowing one to describe the shape of the niche under study.

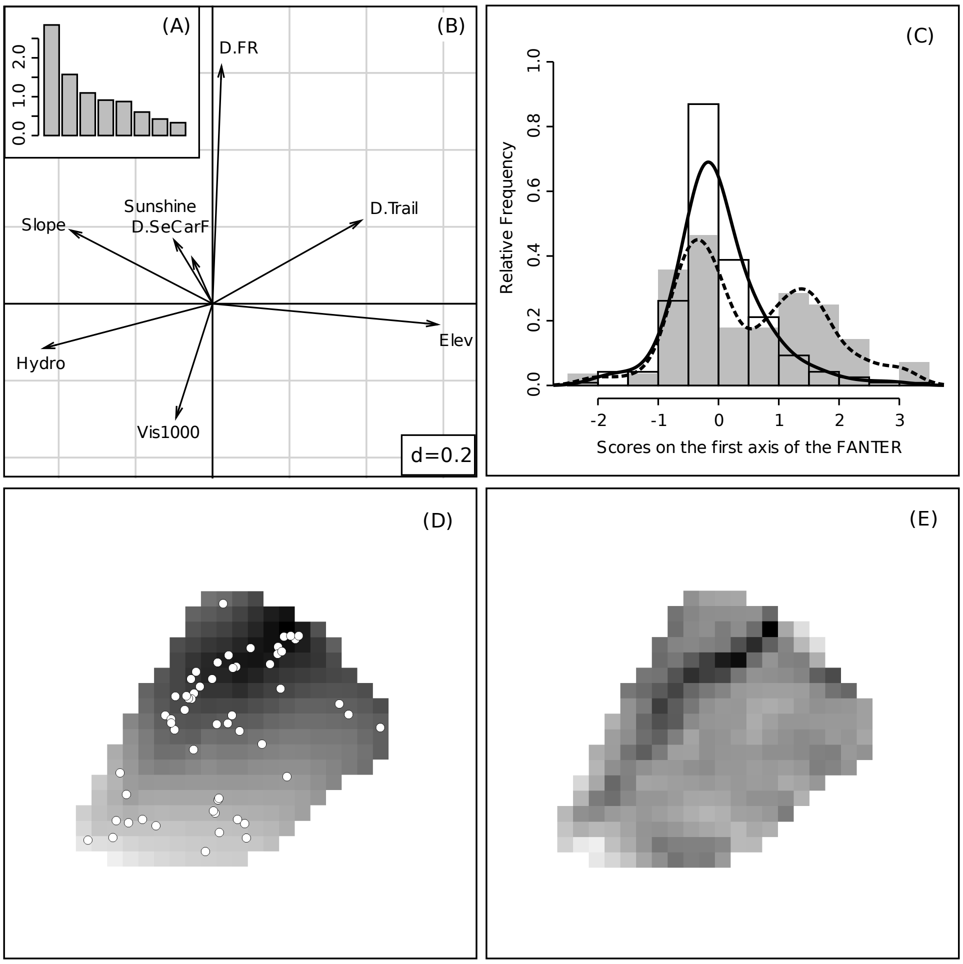


Figure 6.6: Results of the FANTER carried out to identify correlates between the distribution of the chamois relocations in the Bauges mountain (French Alps) according to 8 environmental variables (listed in Table 6.2) mapped in its home range. (A) eigenvalue diagram of the analysis. Only the first axis is kept for the analysis; (B) the correlations between the environmental variables and the first (abscissa) and the second (ordinate) component of the analysis. Grid lines (separated by a distance of 0.2) can be used to measure these correlations on the graph; (C) histogram and smoothed density (using a kernel smoothing with smoothing parameter equal to 0.3) of the coordinates of the available points (white bars and continuous curve) and of the used points (grey bars and dashed curve) on the first component of the FANTER; (D) map of the chamois relocations on a map of the elevation in its home range (darker areas are higher); (E) map of the scores of the pixels of the home range on the first component of the FANTER (darker areas are higher).

6.5 Discussion

We introduced a new general framework for the analysis of the niche-environment system. Because the GNESFA is by its very nature exploratory, it does not rely on many constraining hypotheses (e.g. no minimum sample size required), which renders it appealing. This framework encompasses three consistent factor analyses. Among these analyses, the FANTER offers a new point of view on the niche-environment system. We also proved that the ENFA and the MADIFA can be viewed as special cases of the GNESFA. Finally we showed that the application of these three methods to biological data may give different outputs, as they are based on different biological points of view of the niche-environment system.

On one hand, the MADIFA corresponds to the point of view of the expert of the studied species: when the biologist has a good prior knowledge of the kind of environment required by the species, an examination of the niche itself is not of prime interest. In this kind of studies, the aim is frequently to determine whether the environment in the study area (the available environment) is similar to the environment the species usually occupies (the reference). This is typically the point of view used for environmental suitability modelling.

On the other hand, the FANTER corresponds to the point of view of the expert of the studied area: when the biologist has a good prior knowledge of the environmental structure in the study area (e.g. correlates between environmental variables), an examination of the availability distribution in the ecological space may not be of prime interest. In this kind of studies, the aim is frequently to identify the patterns of the niche itself, and in what it differs from the study area (classical point of view in habitat selection studies). The patterns identified by the analysis may be due to a particularly strong or low inertia of the niche within the cloud of available points. A strong inertia is likely due to the marginality of the niche, but may also be the result of a multimodal niche. A low inertia indicates that the niche restriction is high on some directions of the ecological space. However, this method will fail to identify the directions of the ecological space where the niche is both marginal and very restricted, as the first characteristic counterbalances the second one (Calenge, pers. obs.). Fortunately, the other analyses belonging to the framework of the GNESFA can be used to detect such cases. The main advantage of the FANTER is that it does not assume the unimodality of the niche, contrary to the other analyses of the framework.

Finally, the ENFA is at the middle point between the FANTER and the MADIFA. Both distributions are used simultaneously as reference and focus distributions. This symmetric role of the two distributions is gained to the detriment of the ecological space, with the loss of one of its dimensions through the marginality vector. The ENFA can be used to distinguish between the position and the narrowness of the distributions relative to each other. Eventually, only the concurrent use of the three methods with graphical displays of the niche within its environment would lead to a consistent knowledge of the system. This statement underlines the interest of the GNESFA as a general framework for the statistical exploration of the ecological niche. In most cases, the two points of view described previously (species and study area) are of interest for the biologist (as in the examples presented in this paper), so that the three complementary analyses may be used concurrently to build a conceptual model of the niche-environment system under study.

The GNESFA is easy to perform with any statistical software, as it relies on a succession of two principal component analyses, which are widely available in most standard statistical packages. In particular, the package adehabitat (Calenge, 2006) for the R software (R Development Core Team, 2008) contains a set of functions allowing the application of the GNESFA (function `gnesfa()`) and several graphical displays of the results. This package also contains numerous functions for managing raster maps, and is especially well-designed for the exploration of the niche-environment system.

We presented here the GNESFA for the exploratory analysis of one single ecological niche. However, radio-tracking studies (involving numerous animals) and multi-species designs are frequent among ecological studies, and most of them aim to identify the common characteristics of the environment affecting the distribution of the organisms under study (whether animals or species). Preliminary results indicated that the framework of the GNESFA can be extended to cover more complex study designs: in particular, canonical OMI analysis (Chessel and Gimaret, 1997) and the eigenanalysis of selection ratios (Calenge and Dufour, 2006) can be reformulated as special cases of the FANTER (Calenge et al., in prep). These first results also allow for relating the GNESFA to the large family of methods belonging to the duality diagram framework (Escoufier, 1987), a family also containing most factor analysis (principal component analysis, discriminant analysis, etc.). This family has very interesting properties for the exploration of multidimensional spaces, especially in ecology (see Calenge and Dufour, 2006, for a deeper discussion). Further studies are required to clarify the relationships between these analyses and the GNESFA. With the increasing concern of the ecological community for the study of ecological niches subject to climate change, there is an urgent need to have more than one string to our bow, to ensure the reliability of our conclusions. A multi-niche generalization of the GNESFA would allow for a more effective exploration of a species niche within a community in the ecological space, and would allow for the building of habitat suitability maps for several species at once.

Acknowledgments

We warmly thank the Office national de la chasse et de la faune sauvage (ONCFS) for their financial support. We are also grateful to Gaëlle Darmon (University of Lyon), Sonia Saïd (ONCFS) and Jean-Michel Jullien (ONCFS) for providing the GPS data collected on the chamois in the Bauges mountain, and to the Fédération Départementale des chasseurs de l'Isere for the data on the chamois in the Chartreuse mountain.

Appendix A: Equivalence between the two problems of the GNESFA

In this appendix, we demonstrate that the first problem of the GNESFA:

$$\mathbf{Z}\mathbf{a}_1 = \mathbf{p}_1 \tag{A.1}$$

$$\mathbf{p}_1^t\mathbf{R}\mathbf{p}_1 = 1 \tag{A.2}$$

$$\gamma_1 = \mathbf{p}_1^t\mathbf{F}\mathbf{p}_1 \quad Max \tag{A.3}$$

is equivalent to the second problem of the GNESFA:

$$\begin{aligned} \mathbf{g}_1^t\mathbf{g}_1 &= 1 \qquad (A.4) \\ \mathbf{Z}\mathbf{g}_1 &= \mathbf{y} \\ \beta_1 &= \frac{\mathbf{y}^t\mathbf{F}\mathbf{y}}{\mathbf{y}^t\mathbf{R}\mathbf{y}} \quad Max \qquad (A.5) \end{aligned}$$

We demonstrate this equivalence, with:

$$\mathbf{g}_1 = \frac{\mathbf{a}_1}{\sqrt{\mathbf{a}_1^t\mathbf{a}_1}} \qquad (A.6)$$

and

$$\mathbf{a} = \frac{\mathbf{g}_1}{\sqrt{\mathbf{g}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{g}_1}} \qquad (A.7)$$

And finally, we prove that $\beta_1 = \gamma_1$.

First, we demonstrate that if $\mathbf{a}_1$ is a solution to the first problem, then the use of equation A.6 gives the vector $\mathbf{g}_1$ as a solution to the second problem. Note that the equation A.6 implies that the condition A.4 is fulfilled. Moreover,

$$\beta_1 = \frac{\mathbf{g}_1^t\mathbf{Z}\mathbf{F}\mathbf{Z}\mathbf{g}_1}{\mathbf{g}_1^t\mathbf{Z}\mathbf{R}\mathbf{Z}\mathbf{g}_1} = \frac{\mathbf{a}_1^t\mathbf{Z}^t\mathbf{F}\mathbf{Z}\mathbf{a}_1}{\mathbf{a}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{a}_1} = \mathbf{a}_1^t\mathbf{Z}^t\mathbf{F}\mathbf{Z}\mathbf{a}_1 = \gamma_1$$

It follows that the condition A.5 is fulfilled. Consequently, if the vector $\mathbf{a}_1$ is a solution to the second problem, the solution $\mathbf{g}_1$ to the first problem can be found using equation A.6.

Now, we prove that if $\mathbf{g}_1$ is a solution to the second problem, then the use of equation A.7 gives the vector $\mathbf{a}_1$ as a solution to the first problem. First note that

$$\mathbf{a}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{a}_1 = \frac{\mathbf{g}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{g}_1}{\mathbf{g}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{g}_1} = 1$$

The condition A.1 is fulfilled. Moreover,

$$\gamma_1 = \mathbf{a}_1^t\mathbf{Z}^t\mathbf{F}\mathbf{Z}\mathbf{a}_1 = \frac{\mathbf{g}_1^t\mathbf{Z}^t\mathbf{F}\mathbf{Z}\mathbf{g}_1}{\mathbf{g}_1^t\mathbf{Z}^t\mathbf{R}\mathbf{Z}\mathbf{g}_1} = \beta_1$$

and the condition A.2 is fulfilled.

This completes the demonstration: the two problems are mathematically equivalent, and the relationships between the solutions of the two problems are described in equations A.7 and A.6. Furthermore, $\beta_1 = \gamma_1$.

Part III

Lynx predation in a human dominated landscape

> *"In theory, there is no difference between theory and practice. But, in practice, there is."*
>
> Jan L. A. van de Snepscheut.

In this last part, we finally reach the ultimate function of the thesis, the question that is at the origin of this work: *"How can we describe the spatial relationships between lynx, roe deer, and human?"* Now that we know the necessary theory and tools developed in the previous Parts (Parts I & II), we need some information about the system under study, to show why it is particular in southern Norway. I will then present the next two chapters, with some additional elements that need to be carefully checked before publication.

A three components system in southern Norway

There are very few systems that can be considered as closed in terrestrial ecosystems. Among the exceptions, some islands present predator-free systems of ungulate populations, and allow thus intensive studies on the focal species. It is the case for example on the Isle of Rum (Scotland, McLoughlin et al., 2006), with a large population of red deer or on Storfosna Island (Norway, Andersen and Linnell, 1998) with an increasing roe deer population, both in a natural state (free range, predator-free, unharvested). Compared to these ideal areas, the system lynx–roe deer–human in southern Norway is surely far from perfect, but presents indeed a few interesting characteristics.

First of all, predator-prey relationships between lynx and roe deer are predominant in southern Norway. Roe deer are available in most parts of southern Norway, but at very low densities (0.3 per km^2, Odden et al., 2006). Despite this relative scarcity, roe deer constitutes the main part of lynx diet. Indeed, whereas around 20 species are found in the diet of lynx in Norway, roe deer represent up to 83% of ingested biomass by lynx in winter (Odden et al., 2006). It is more contrasted in summer, while sheep (*Ovis aries*), when available, can constitute a consequent alternative prey, with up to 26% of the ingested biomass (Odden et al., 2006). Conversely, lynx are also the main predator for roe deer throughout most of their southern range, with low numbers of alternate predators such as wolves (*Canis lupus*, Wabakken et al., 2001) or wolverines (*Gulo gulo*, Flagstad et al., 2004). Lynx mainly prey on adults but fawns are also commonly killed: 68% of lynx-killed roe deer in a study conducted in the area were adults (among which 65% of females), whereas only 32% were fawns (note that this proportion would be much lower in terms of ingested biomass). Lynx are however not the main cause of death for roe deer fawns, since another study showed that red foxes were responsible for 62% of all fawns mortality in a neighbouring area (Panzacchi et al., 2008).

On the other hand, both roe deer and lynx are heavily hunted by humans. As such, humans may be regarded as top predators in the system, like would be hyenas (*Crocuta crocuta*) and lions (*Panthera leo*) compared to cheetahs (*Acinonyx jubatus*) in the Serengeti (Durant, 1998). Both lynx and cheetahs are subject to intra-guild predation and present low competitive ability in this respect. Contrary to lynx, humans mainly harvest on male roe deer (Andersen et al., 2007, reported 44% of adult males vs. 28% of adult females). So far, no lynx have been reported to kill humans in Norway, and the species is not regarded as being threatening (Røskaft et al., 2003). Thus lynx are hunted as a game species, using quotas since 1994. Legal hunting and poaching are by far the main causes of death: a study conducted in southern Norway reported that these were the only causes of mortality for yearling males and adult males and females (Andrén et al., 2006). To sum up, the system may be idealized as a three component system, with humans on top, being the only cause of mortality of lynx, together accounting for most of roe deer mortality (Melis et al., data, see Fig. 6.7).

What shapes Eurasian lynx distribution in human dominated landscapes: selecting prey or avoiding people?

The next Chapter (Chapter 7) present the analysis of lynx distribution in southern Norway. In this study we examined the factors that were responsible for the large-scale distribution of lynx. This study actually raised the general concern about the spatial scale defined in terms of "*extent*" and "*grain*" (Wiens, 1989). Extent is the overall area encompassed

Figure 6.7: Illustration of the three components system in Norway. Upper left corner: lynx hunted by humans. Upper right corner: roe deer bucks hunted by humans. Lower left corner: a roe deer fawn killed by a lynx. Lower right corner: a roe deer buck killed by a lynx. Photo credit: Scandlynx & John Linnell.

by a study; grain is the size of the individual units of observation. In our study case, the extent was thus around 100,000 km^2, while the grain was approximately 50 km^2. Extent and grain together delineate the resolution of a study: as explained by Wiens (1989), they defined the upper and lower limits of resolution. Follows from this a logical rule of thumbs: we cannot detect any process below the grain, whereas generalization beyond the extent would be at the cost of assumptions about the independence of the studied process and the scale. It is of outermost importance that the data used and collected correspond to the hypotheses investigated. Otherwise we may not detect patterns at a given scale just because they were studied at a wrong resolution (Coulson et al., 1997). We believe that in our case the conclusions were drawn at the appropriate resolution.

Large-scale data may show large imprecision due to their broad nature. However, they are likely to express a strong biological signal for the same reason. In other words, the signal-to-noise ratio for large scale data may not be so unfavourable as generally expected. On the contrary, fine-scale data allow one to investigate processes more deeply, and to test more refined hypotheses, but at the cost of a cautious treatment of the noise. It is not unlikely that the signal-to-noise ratio becomes weaker at finer scales. As underlined by Wiens (1989), while *"fine-scale studies may reveal greater detail about the biological mechanisms underlying patterns"*, *"generalizations are more likely to emerge at broader scales."*

With a large-scale approach of lynx distribution, we were able to highlight a clear trade-off between abundance of prey and avoidance of human activity. This trade-off is however a result of the species distribution at this specific scale and may thus not hold at finer scale. This should be indeed carefully investigated in order to acknowledge the hierarchical nature of lynx habitat selection (see Chapter 1).

Predation by lynx has the greatest impact on roe deer population growth at lower environmental productivity.

After demonstrating that lynx select areas with high prey abundance, we addressed the question of the impact of predation (i.e. lynx) on roe deer dynamics, in the following chapter. While carnivores are expected to limit herbivore biomass, it seemed from our preliminary results that the impact of lynx was even higher than expected. As a matter of fact, in the absence of lynx, all roe deer populations were stable along a wide range on the productivity gradient. In the presence of lynx, however, only in those areas with high productivity populations were able to sustain this additional predation, whereas in unproductive areas, populations were decreasing.

Two concerns should be addressed on this preliminary study. First, we used relatively high correlated variables in our regression model. For instance, there was a positive correlation between annual net primary productivity and mean annual temperature ($r_p = 0\ 57$, df = 176, $t = 9\ 292$, $P < 0.001$). The multicollinearity between environmental variables may be misleading and lead to confounding results (Graham, 2003). Two approaches could be developed to circumvent the problem. First, we might perform a PCA on the whole set of environmental variables (3 climatic variables and 3 phenology variables, all highly correlated) in order to extract uncorrelated axes with most of the variance for the regression (Graham, 2003). This approach seems quite promising and the preliminary results showed a strong effect of lynx, as well as an interaction between environmental conditions (scores of the PCA) and lynx presence. Another approach would be to reduce the data set to temperature and productivity only (in order to improve the interpretation), and regress one on the other (Graham, 2003). As we are mainly interested in environmental productivity, we could use productivity and the residuals from temperature.

The second issue follows from the data we used. The trends in roe deer growth rate were estimated using hunting statistics from every municipality. We assumed for the analysis that these statistics reflects actual changes in the population's abundance (see details in the text). However, we could expect a direct effect of hunting on the dynamics of these populations: if a high number of roe deer is harvested one year, it is likely that the harvest the year after will be reduced. This possible occurrence of density dependence may be handled by different means (Freckleton et al., 2006). For example, let us assume that the number of roe deer harvested at the year t is given by N_t. We might expect a negative relationship between N_{t+1} and N_t (that is: a year with a high harvest is followed by a year with a low harvest, see Fig. 6.8). A covariance analysis with all municipality together would assess a common trend (common slope), and we could thus remove this trend and use the residuals by municipality instead in the regression model.

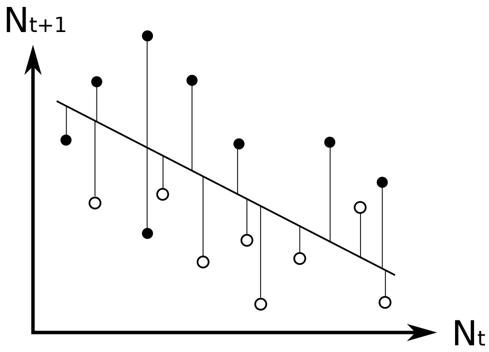


Figure 6.8: A proposition to correct for a possible hunting effect. Because hunting has a direct impact on roe deer numbers, we would expect a negative relationship, per municipality, between two successive years. A covariance analysis might remove the common trend, if needed. Here dark points and empty points represent two different municipalities.

However, these two concerns need further developments and validation before being used. Unfortunately, I ran out of time before finalizing the analyses. Consequently the results presented in Chapter 8 should be considered cautiously.

Chapter 7

What shapes Eurasian lynx distribution in human dominated landscapes: selecting prey or avoiding people?

Mathieu Basille[1,2,*], Ivar Herfindal[2], Hugues Santin-Janin[1], John D. C. Linnell[3], John Odden[3], Reidar Andersen[3,4], Kjell Arild Høgda[5] & Jean-Michel Gaillard[1]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, F-69 622, Villeurbanne, France.

[2] Centre for Conservation Biology, Department of Biology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[3] Norwegian Institute for Nature Research; Tungasletta 2, NO-7485 Trondheim, Norway.

[4] Vitenskapsmuseet; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[5] Northern Research Institute Tromsø, Postboks 6434 Forskningsparken, NO-9294 Tromsø, Norway.

[*] Corresponding author: `basille@biomserv.univ-lyon1.fr`

Submitted to Ecography.

> *"There's another way to phrase that and that is that the absence of evidence is not the evidence of absence. It is basically saying the same thing in a different way. Simply because you do not have evidence that something does exist does not mean that you have evidence that it doesn't exist."*
>
> Donald RUMSFELD.

Abstract

In the multi-use landscape of southern Norway, the distribution of lynx *Lynx lynx* is likely to be determined both by the availability of their favoured prey —the roe deer *Capreolus capreolus*— and the presence of humans because hunter harvest of lynx is used as a tool to manage conflicts with depredation on livestock. We described the distribution of the reproductive part of the lynx population based on snow-track observations of females with dependent kittens collected over 10 years (1997–2006) in southern Norway. We used the Ecological-Niche Factor Analysis to examine how lynx distribution was influenced by roe deer, human activity, habitat type, environmental productivity and elevation. Our first prediction that lynx should be found in areas of relatively high roe deer abundance was supported. However, our second prediction that lynx should avoid human activity was rejected, and lynx instead occupied areas more disturbed than available. Finally, our third prediction of a trade-off between abundance of prey and avoidance of human activity was supported. While roe deer benefit to a large extent from current human land use practices, allowing them to escape predation from lynx, the situation seems not so favourable for the predators who are restricted in competition refuges at medium to low prey density. At the same time, lynx are very efficient predators and can survive on numerous alternative prey species. Management plans for the conservation of the lynx should therefore focus primarily on regulating lynx hunting (both legal and poaching) to ensure that the most favourable part of the habitat are freely available.

Keywords: attractive sink, competition refuge, Ecological-Niche Factor Analysis, Eurasian lynx, habitat selection, human activity, roe deer

LARGE MAMMALIAN CARNIVORES typically occupy the upper trophic level in terrestrial ecosystems. They often have a keystone role, as they are strongly interactive species. Their absence usually coincides with marked changes in structure and composition, resilience to disturbance, or species diversity of ecosystems (Jędrzejewska and Jędrzejewski, 2005;

Soulé et al., 2005). Large carnivores also cause severe conflicts with human interests, through competition for prey and depredation on livestock, which in the past often led to extermination campaigns that have seriously reduced their distribution and numbers (Breitenmoser, 1998). In the absence of large carnivores, humans have usurped the role of top-predator, in addition to modifying the original habitat through conversion and infrastructure development. The last three decades have seen a global reversal of carnivore policy with the passing of conservation legislation. As a result, and through both natural expansion and reintroduction, a widespread increase in the abundance and distribution of large carnivores is currently taking place in most European countries. However, after a long period of continuous human development, the landscape which large carnivores are returning to is very different from the landscape they left one or two centuries ago. There has been much debate about both the ability of large carnivores to persist in human dominated landscapes (Woodroffe, 2000; Linnell et al., 2001b) and the ecological role that they will play in these new, highly modified ecosystems (Linnell et al., 2005). For instance Woodroffe (2000) found a positive relationship between historical patterns of large carnivore extinction probability and human population density. However, in the presence of favourable legislation, large carnivores can usually recover much of their past range despite the presence of high human densities (Linnell et al., 2001b).

Eurasian lynx *Lynx lynx* in Norway have followed the general temporal pattern of large carnivore abundance and distribution in Europe. Subject to a state sponsored bounty program from 1846, they were nearly exterminated by the mid 20^{th} century, persisting in one or two remnant populations in southeastern and central Norway. However, under increasingly restrictive hunting legislation lynx expanded during the late 20^{th} century and are now widespread throughout the whole country with the exception of the southwest. In response to conflicts with sheep farmers (depredation on domestic sheep; Odden et al., 2002) and hunters (competition for game species; Odden et al., 2006), Norwegian lynx populations are managed as a game species and objectives have been set for limiting their density and distribution (Ministry of the Environment, 2004). When planning for the conservation of large carnivores in human-dominated landscapes, one needs reliable information about their range of habitat tolerance (Linnell et al., 2005). Lynx habitat tolerance is likely to be shaped by two main factors: access to food and mortality risk. In Europe, the major prey item for Eurasian lynx are roe deer (Odden et al., 2006), although they can survive on alternative prey such as hares *Lepus* sp., tetraonids and other small ungulates. In most cases, lynx mortality is human-caused (hunting, poaching, vehicle collisions; Andrén et al., 2006), and lynx, especially for females with kitten, avoid human-dominated areas (Bunnefeld et al., 2006). However, high roe deer densities often occur in fragmented and disturbed areas associated with high human activity. Lynx may thus have to balance selection for prey density against mortality risk from humans. Habitat selection should then reflect the response of animals to the trade-off between food and mortality.

Previous studies at the landscape scale have focused on modelling lynx tolerance to habitat fragmentation and human infrastructure in Central and Eastern Europe. In historical areas of lynx presence (e.g. Niedziałkowska et al., 2006), the habitat of native lynx populations was characterized by a higher proportion of forest and a lower fragmentation than observed nowadays. Conversely, lynx occurrence was negatively associated with human settlements and transportation infrastructure. Zimmermann and Breitenmoser (2002) provided support for the importance of forest and roads from a reintroduced population in Switzerland. However, the main factors were here elevation and slope, which were likely to include both the presence of forest and roads as a result of human activities in this area. Guisan and Zimmermann (2000) stated that such models should only be applied to regions and situations similar to those where the basic data were originally gathered. Schadt et al. (2002) however extended these findings to design a Germany-wide conservation plan, and found that the critical factor was the connectivity of forested and non-forested semi-natural areas. Using the natural recovery of lynx in southern Norway, where both the level of fragmentation and the degree of human infrastructure development are much lower than elsewhere in Europe, we aimed to assess how lynx distribution at the population scale (i.e. southern Norway) relates to a range of environmental characteristics including an index of prey abundance, different habitat types, and human impact using recent developments of the Ecological-Niche Factor Analysis approach (ENFA, Basille et al., 2008). We expected that (1) lynx should select areas with high prey availability and therefore the habitat of the lynx should be associated with relatively high roe deer availability; (2) lynx should avoid areas of high human activity and therefore the habitat of lynx should be associated with a relatively low level of human activity; (3) as roe deer abundance and human activities should co-vary positively to some extent because roe deer take advantages of human transformation of the landscape, lynx should trade searching for food for avoiding humans, so that some optimal combination between roe deer availability and human disturbance should occur.

7.2 Material and methods

7.2.1 Study area

The study took place in southern Norway, between approximately 58°N and 63°N (Fig. 7.1). The study area (c. 100,000 km^2) was defined as the 8 southern counties with a permanent lynx presence and where roe deer are the main prey. We did not include the county of Oslo that is too densely populated to provide reliable estimates of both lynx and roe deer.

The study area covers a gradient from highlands covered with alpine tundra in the northwest to lowlands covered with a matrix of boreal forest and farmland in the southeast. The proportion of forest that has been converted to farmland or given over to human infrastructure increases close to the coast.

7.2.2 Lynx distribution data

Since 1996, lynx have been monitored using a standard methodology based on non-replicated counts of family group (i.e. a female with dependent young of the year). Records of tracks in the snow (see Fig. 7.2) are collected by hunters, game wardens and the public, and checked by game wardens (Linnell et al., 2007). Outside the mating season (late March and April), observations of 2 or more lynx together are attributed to a family group. We used these data to measure lynx distribution: the occurrence of a family group indicates that lynx are resident in the area. We pooled 10 years of data, from 1997 to 2006 (911 observations).

As adult females lynx are territorial (Breitenmoser et al., 1993), several observations in the same pixel from one year are likely to be from the same female. We therefore used as an index of distribution the number of

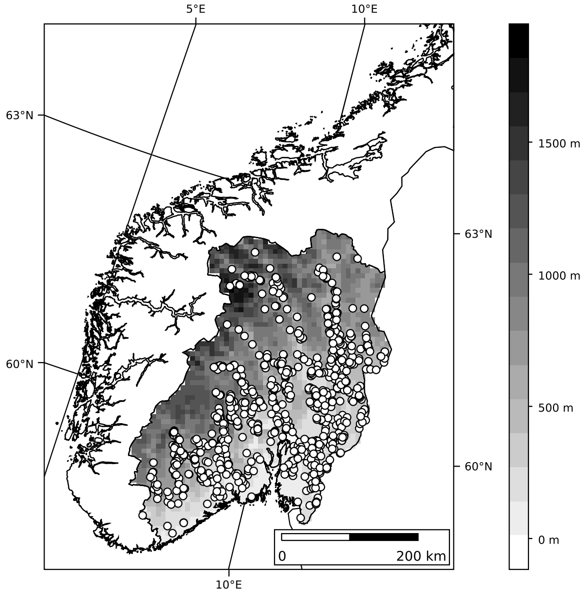


Figure 7.1: The study area situated in southern Norway. Family group observations are represented by white dots. Elevation is represented by the gray scale.

Figure 7.2: Records of lynx tracks in the snow in southern Norway. Photo credit: Lars Gangås / Scandlynx

Name	Description
Agri	Proportion of agricultural areas
Alpine	Proportion of alpine areas
Bare	Proportion of bare ground
DEM	Mean elevation
DEMslope	Mean slope
DEMstd	Standard deviation of the elevation
Forest	Proportion of forest
Human	Mean value of human density
NDVIderived	Derived of the annual NDVI-curve in spring
NDVIint	Integrated NDVI between spring and autumn
NVDIpeak	Peak of the annual NDVI-curve
NDVIspring	The week number in spring when NDVI-values reach levels corresponding to leaf burst on birch
PublR	Total length (km) of public roads (paved roads, ranging from municipality roads to national highways) per km^2
PrivR	Total length (km) of private roads (roads in connection to farming and logging or some recreational resorts) per km^2
Roe	Mean number of roe deer shoot by legal hunting from 1997 to 2005
Simpson	Simpson's index of habitat type diversity
Urban	Proportion of urban areas

Table 7.1: Environmental variables used in the analysis. See details in the text.

years of monitoring during which lynx reproduction was observed within a given pixel. Thus, the distribution variable varied from 0 (pixels with no observation of lynx during the ten years) to 10 (pixels with at least 1 observation of a family group every year).

7.2.3 Environmental variables

A range of environmental variables were used as potential habitat variables. We included information on habitat type, elevation and environmental productivity, as well as roe deer density and human activity (Table 7.1, see below).

Environmental phenology derived from NDVI. The Normalised Difference Vegetation Index is an index based on the difference in reflectance between the red and near-infrared wavelengths. NDVI is closely related to photosynthetic activity and plant productivity (Reed et al., 1994). Thus, by using annual curves of NDVI for a specific area, it is possible to calculate several indices that can be interpreted as the change in plant productivity, or environmental phenology, through the year (Pettorelli et al., 2005). We used the GIMMS-dataset 1982–2002 to calculate four parameters of environmental phenology: The onset of spring measured as the date when the green-up of the vegetation starts, the derived NDVI corresponding to the rate of change in NDVI at this date, the peak value representing the highest NDVI-value reached during the season, and the integrated NDVI calculated as the sum of NDVI-values throughout the growing season. The NDVI dataset we used in this study had a spatial resolution of approximately 7×7 km (Karlsen et al., 2006), and this resolution therefore determined the spatial resolution of our analyses, henceforth referred to as the grand pixel resolution. The GIMMS dataset was available from summer 1981–2003 (Tucker et al., 2005). To avoid calibration problems, we used the GIMMS dataseries only based on average phenological values. We justify the use of environmental phenology covering only 60% of our study period (1997–2006) with the rather small change in climatic conditions during the study period.

Habitat type. We used the habitat typology based on the Global Land Cover 2000 database (Bartholomé et al., 2002). This dataset has a spatial resolution of approximately 1×1 km in our study area and consists of 23 different habitats types. We redistributed these habitats types into five main classes: Urban areas, agricultural areas, forest, alpine tundra, and bare rock-gravel. We then calculated the proportion of each of these classes per grand pixel (7×7 km). Based on the composition of habitat types inside each grand pixel, we calculated the Simpson's index of diversity as $L = \sum_{i=1}^{S} p_i^2$, where p_i is the proportion of habitat type i inside the grand pixel. This index goes from low values when there are many different habitat types in low proportion, to high values when there are a few dominating habitat types.

Elevation measures. The elevation data were obtained from the Norwegian Mapping Authority as a raster digital elevation model (DEM) with a spatial resolution of 100×100 m. We then calculated, for each grand pixel, the mean elevation and slope, as well as the standard deviation in elevations in order to get a measure of the variation within each grand pixel.

Measure of roe deer abundance. An index of roe deer abundance was calculated from the Norwegian hunting statistics on a municipality basis (Herfindal et al., 2005). The data from 1997 to 2005 were averaged to match the data of lynx family group. All but 3 municipalities ($N = 157$) had data for at least 7 years. For the three missing municipalities, we used the mean of roe deer abundance of the neighboring municipalities.

Human impact. We measured human impact as road density and human density. Data on road density were obtained from the Norwegian Mapping Authority. Road density was calculated as the total length of roads (km) within each grand pixel for private and public roads separately. The human density map was obtained from Statistics Norway (Takle, 2002), and corresponded to the number of inhabitants per square kilometre, with a spatial resolution of 1×1 km. Such data were re-sampled to the grand pixels by summarising the values of the $1\times$ km pixels within each grand pixel.

7.2.4 Habitat selection

We had data on observations of lynx throughout the study area, based on a presence-only design (Pearce and Boyce, 2006). The Ecological-Niche Factor Analysis (ENFA; Hirzel et al., 2002a) is a factorial analysis designed for presence-only data that summarizes the habitat selection into two components, the marginality and the specialization, to distinguish the habitat used from the availability. The marginality measures the position of the habitat within the environment (i.e. deviation of the average conditions in the habitat used from the average conditions available in the environment). The specialization measures the dispersion of the habitat within the available environment (i.e. tolerance of the species according to characteristics of its environment). The ENFA is like a Principal Component Analysis

(PCA); however, it is based on two components with well-defined *a priori* biological meaning to assess the position of the habitat within the environment. Using the ENFA we first extracted one axis of marginality, followed by several uncorrelated axes of specialization which successively account for the maximum of the specialization until the number of initial variables have been exhausted. The number of axes of specialization to keep was defined using the broken-stick method (Jackson, 1993). The ENFA can be used to build factorial maps of the habitat (Basille et al., 2008) with a projection of the environmental variables on the marginality and specialization axes. Such factorial maps provide an optimal (in terms of distinction between marginality and specialization) and undistorted representation of the habitat within the environment. A Monte-Carlo procedure was used to assess the significance of both marginality and specialization axes. The same number of localisations as observed ware randomly distributed 1,000 times over the study area and an ENFA was run at each step. A comparison of the 1,000 sets of marginality and first eigenvalues of specialization with the observed values provides the significance of each component, expressed as the proportion of random values higher than the observed value.

The ENFA is an exploratory tool that allows identifying the variables responsible for the position and the shape of the habitat. These analyses allowed us to test our two first predictions. For the last prediction, we confronted lynx space use, roe deer abundance and human-related variables. This was done in two steps, first from a lynx perspective and second from a roe deer perspective. To test the use of roe deer abundance and human-related variables by the lynx over their whole range, we divided each of the variables into a few classes in order to get approximately equal numbers of pixels in each, and no empty classes. Roe deer abundance and public road density were square root transformed, and human density was log-transformed to get approximately equal bin sizes. When the ratios of density of used pixels over density of available pixels per class was greater than 1, the class was more used than expected by chance (selection), whereas when it was less than 1, the class was less used than expected by chance (avoidance). We then created 1,000 random distributions with the same number of used pixels and constrained by the observed available distribution per class. At each step, we again computed the ratios so that we obtained a distribution of 1,000 random ratios. The comparison of the observed value to the 95% confidence interval of the random distribution gave us an assessment of the significance of the selection for each class.

Finally, to assess the relationship between roe deer abundance and human-related variables, we fitted a generalized linear model with a negative binomial error and a log link function to handle overdispersed count data (Ver Hoef and Boveng, 2007). The roe deer abundance at the municipality level was the response variable with the human-related variables as explanatory variable, and the municipality area as offset variable (Venables and Ripley, 2002). We tested the influence of the different explanatory variables using likelihood ratio tests. We computed Monte-Carlo simulation to derive confidence intervals of the conditional mean and the optimal value.

All statistical analyses were conducted using R 2.4 (R Development Core Team, 2008) and the R-package "adehabitat" (Calenge, 2006).

7.3 Results

Using the ENFA we found a clear structure in the selection of the habitat by lynx (Fig. 7.3); both marginality and specialization were highly significant ($P < 0.001$). The specialization was very strong on the first axis of specialization (variance ratio of 8.9, which means that the used habitat was almost 9 times narrower compared to what was available), and to a lesser extent on the following three axes (variance ratios of 4.8, 3.5 and 2.7 respectively), which were retained from our analysis (Fig. 7.3A). The position of the used habitat was clearly distinct from the available habitat on the scatterplot (Fig. 7.3A) enabling us to interpret the correlations of the variables with the axes of the analysis (Table 7.2). Using the marginality axis we identified the preferred habitat of the lynx as including a higher integrated NDVI, a higher NDVI peak value, a larger proportion of forest, a higher Simpson's index, a lower proportion of alpine areas, an earlier onset of spring and a lower elevation than what was available on average, with absolute values of correlations ranging from 0.75 to 0.87 (Table 7.2; Fig. 7.3B). The habitat of lynx also included higher roe deer abundance than what was generally available (correlation of 0.37), therefore validating our first prediction. However, contrary to our second prediction, the density of private and public roads were positively associated with the lynx distribution (correlations of 0.68 and 0.35, respectively). Only agricultural

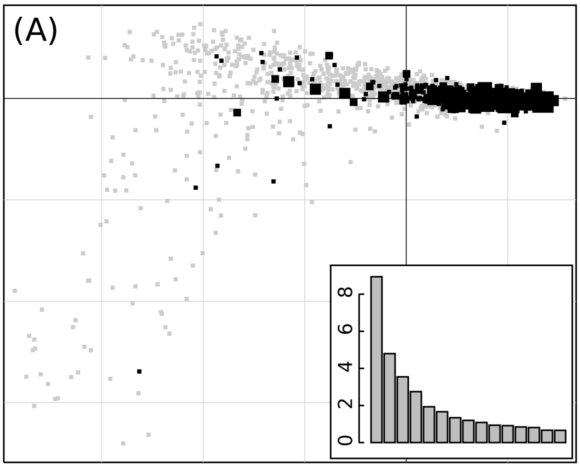


Figure 7.3: Results of the ENFA. (A) Display of the habitat in the plane formed by the marginality axis (x axis) and the first specialization axis (y axis). The grey squares represent the available pixels, whereas the black squares represent the used pixels, with a side proportional to the frequency of lynx observations. This graph shows both the position and the shape of lynx habitat within the environment. The lower-right insert presents the eigenvalues of specialization from the analysis. (...)

areas seemed to be slightly avoided in comparison to what was available on average (correlation of -0.05). A strong specialization of lynx on the proportion of urban areas and human density on the second specialization axis (correlations of -0.77 and -0.38, respectively, Table 7.2; Fig. 7.3C), however, indicated that lynx distribution was negatively influenced by both extremely low and extremely high values of human occupancy. Note that the high specialization on the first axis against the proportion of bare ground (correlation of -0.92, Fig. 7.3B) simply meant that lynx were almost never found at medium to high proportion of bare ground and were therefore strongly restricted on areas with a very small proportion of bare ground.

Additionally, there was a marked selection of areas characterized by intermediate values of roe deer abundance, whereas areas with very low or very high abundance were used less than expected (Fig. 7.4A). The same pattern also occurred for the public road density (with an optimum between 0.39 and 0.6 km of roads per km^2, Fig. 7.4B) and human density (with an optimum between 2.3 and 6.5 inhabitants per km^2, Fig. 7.4C), therefore lending strong support for our third prediction. Lynx used ar-

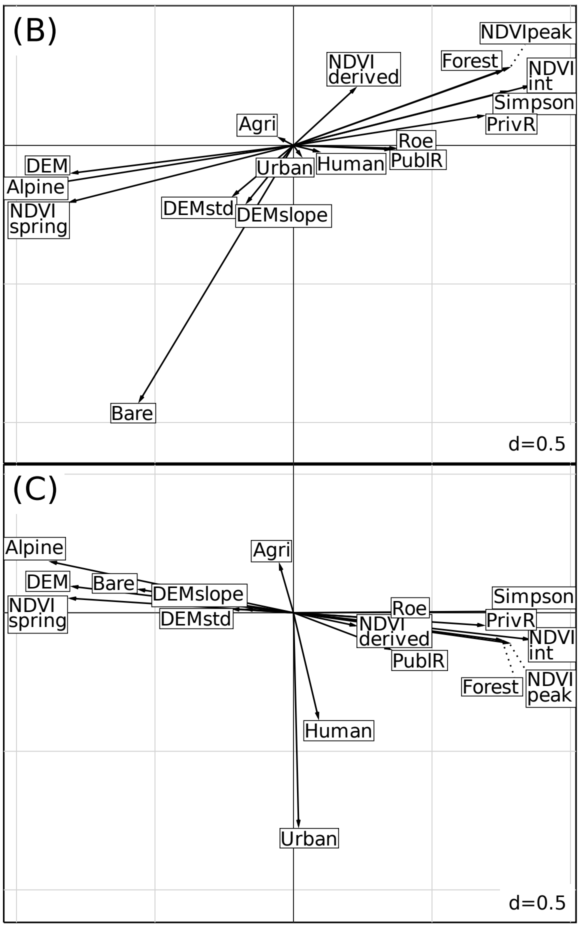


Figure 7.3: Results of the ENFA (continued). (B) Graph of the correlations between the environmental variables and the marginality axis (x axis) and the first specialization axis (y axis). The projection of the arrows on the axes give the contribution of each variable to the definition of the axes. Compared to the horizontal axis, long arrows towards the right (respectively the left) indicate that the habitat is characterized by high (respectively low) values on these variables. Compared to the vertical axis, long arrows indicate that the habitat is restricted on some characteristics of the variable, regardless to the direction (the sign of the specialization is not important). (C) Graph of the correlations between the environmental variables and the marginality axis (x axis) and the second specialization axis (y axis).

Name	Mar	Spe1	Spe2
Agri	-0.050	0.026	0.173
Alpine	-0.874	-0.139	0.184
Bare	-0.555	-0.922	0.084
DEM	-0.796	-0.100	0.095
DEMslope	-0.167	-0.205	0.013
DEMstd	-0.215	-0.180	0.013
Forest	0.751	0.268	-0.099
Human	0.090	-0.022	-0.378
NDVIderived	0.222	0.205	-0.046
NDVIint	0.848	0.214	-0.096
NDVIpeak	0.775	0.280	-0.109
NDVIspring	-0.804	-0.204	0.052
PublR	0.348	-0.014	-0.131
PrivR	0.683	0.107	-0.045
Roe	0.366	-0.011	-0.025
Simpson	0.766	0.195	0.005
Urban	0.018	-0.026	-0.767

Table 7.2: Correlations of the environmental variables with the axes of the ENFA, given for the marginality axis (Mar) and the first two axes of specialization (respectively Spe1 and Spe2).

eas with lowest and highest roe deer abundance and human disturbance less than expected, whereas areas with intermediate values were selected for, indicating that a trade-off occurred between these two variables. This trade-off was furthermore supported by similar findings for roe deer, with however higher optima. Indeed, roe deer abundance was strongly related to both public road density and human density, both simple and quadratic terms being highly significant in both cases ($\chi^2_{Roads} = 1941.6$, $df = 1$, $p < 0.001$; $\chi^2_{Roads^2} = 490.9$, $df = 1$, $p < 0.001$; $\chi^2_{Human} = 1474.6$, $df = 1$, $p < 0.001$; $\chi^2_{Human^2} = 369.6$, $df = 1$, $p < 0.001$). There was an optimum of 1 km of public roads per km^2 (with 95% confidence interval ranging from 0.97 to 1.02, Fig. 7.5A), and an optimum of 136 inhabitants per km^2 (with 95% confidence interval ranging from 115 to 163, Fig. 7.5B), both optima being higher than for lynx (about twice and 30 times, respectively).

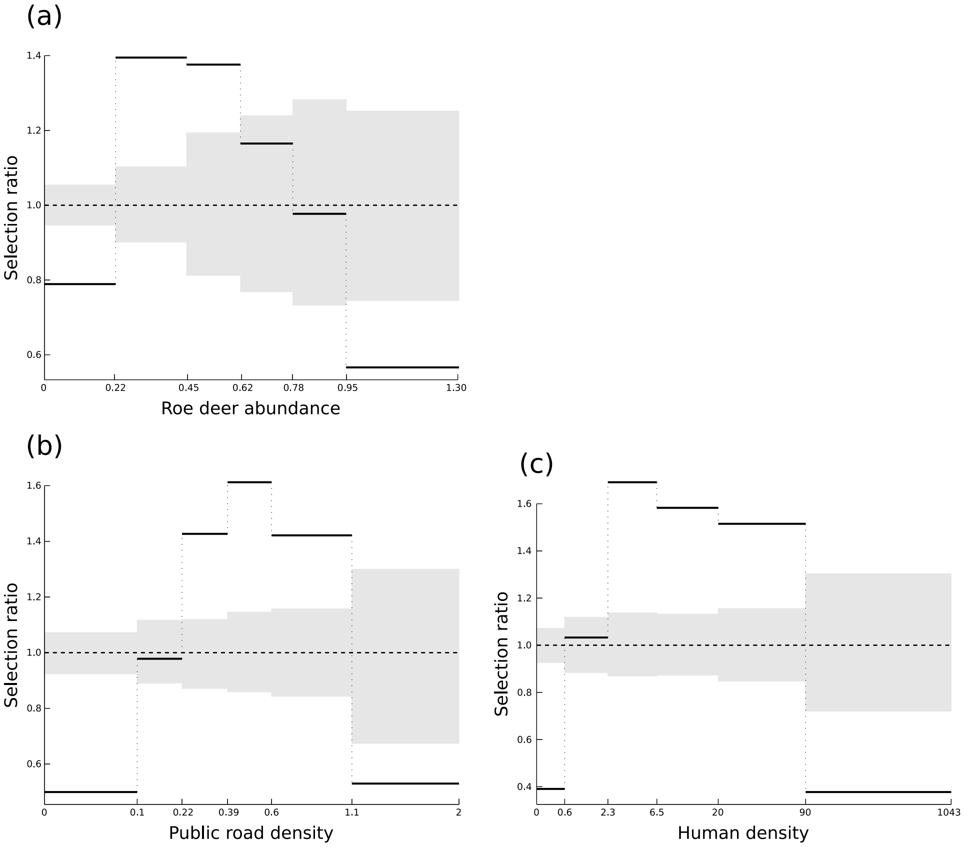


Figure 7.4: Ratios of density of used pixels over density of available pixels for different environmental variables. The observed ratio is represented by a thick line and should be compared to the random distribution with its 95% confidence interval in grey. The horizontal line at $y = 1$ indicates a null selection. (a) Roe deer abundance. (b) Public road density (c) Human density. For reading convenience, the x axis has been square root transformed for (a) and (b) and log-transformed for (c).

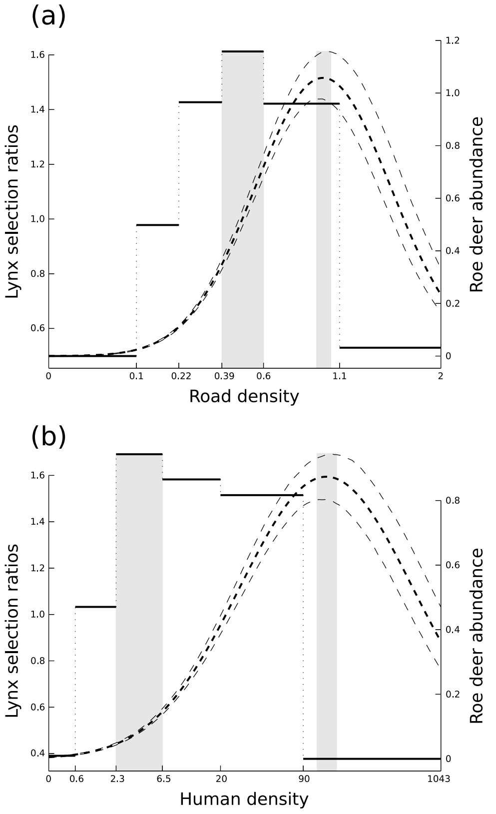


Figure 7.5: Lynx selection ratios (thick lines) and roe deer abundance (dashed lines) for (a) public road density, and (b) human density. For roe deer abundance, 95% confidence intervals of the predictions are provided. The grey bins indicate lynx and roe deer optima. The x axis has been transformed prior the analysis to match the scale transformation of Fig. 7.4, i.e. with a square root function for (a) and a log function for (b).

7.4 Discussion

We found that (1) lynx were found in areas with a higher roe deer density than generally available in the study area, therefore validating our first prediction; (2) lynx occupied areas with more human activity (relatively to roads primarily) than generally available, therefore rejecting our second prediction; (3) there was a clear trade-off in lynx habitat selection so that an optimal combination of intermediate roe deer abundance and intermediate human disturbance occurred.

The distribution range of lynx in southern Norway is characterized by a high proportion of forests, which correspond to productive areas with an early green-up of the vegetation and a low elevation, and by a relatively high roe deer abundance. This supports previous studies in Poland (Niedziałkowska et al., 2006), in Switzerland (Zimmermann and Breitenmoser, 2002) and in France (Basille et al., 2008). Forest can indeed be considered as the main habitat for lynx, with roe deer as the main prey species.

On the other hand, lynx habitat was characterized by very low values of proportion of bare ground and alpine tundra. This selection against mountainous areas is potentially a side effect of the absence of a suitable main prey species at high elevation in Norway. For instance, in the Alps of central Europe, chamois *Rupicapra rupicapra* can constitute a significant proportion of lynx diet (up to 22% of lynx kills) with a higher preference of lynx to prey upon this species (e.g. Molinari-Jobin et al., 2004).

Contrary to our second prediction, lynx were found in areas with slightly more human activity than generally available. Such findings support previous results reported by Bunnefeld et al. (2006) who found that even if their mortality risk is higher close to roads, houses and fields, lynx are attracted by human-dominated areas. Additionally, lynx could tolerate human activity provided that there is a high density of forested areas (Sunde et al., 1998). Family group data are collected by hunters, game wardens and the public, and are therefore potentially influenced by the accessibility. If this was the case in our study, we would have expected to find more signs of lynx presence in areas close to roads and the results would have been biased towards areas of high disturbance. However, there was a minimum effort in the sampling with 1,819 transects (3 km long) distributed among the different counties. For instance, 80% of the collared lynx were detected during 2 trials in Hedmark (16 collared lynx), among which all females with kittens were detected (Odden et al., 2001). Additionally, females with young take less risk than males and females without young, and stay further from human activities (Bunnefeld et al., 2006). If there is a bias in our data set, it should be towards less disturbed areas, therefore reinforcing our findings regarding to human impact.

Finally, lynx avoided areas with the highest human and road densities, which matched closely with an avoidance of the highest roe deer abundance. At the same time, roe deer in Norway are well known to occur in fragmented and disturbed areas (Mysterud, 1999). Indeed, in our study area, roe deer abundance positively correlated with human density and roads density up to very high thresholds. Concerning road density, the optimum for roe deer was about twice higher than for lynx (corresponding to 0.4–0.6 and 1 km of roads per km^2 for lynx and roe deer, respectively). Even more important, the human density optimum was about 30 times higher for roe deer than for lynx (corresponding to 2.3–6.5 and 136 inhabitants per km^2 for lynx and roe deer, respectively). Roe deer were therefore much more tolerant to human disturbance than lynx, and could therefore sustain themselves in most converted areas. These results supported our third prediction and suggest that some areas used by lynx could act like "attractive sinks" (sensu Delibes et al., 2001); attractive due to the presence of abundant roe deer, but sinks because of mortality risks caused by the proximity of people. Despite these mortality risks, maladaptive behaviors tend to the selection of these areas where the species is unable to replace itself without immigration. While the lynx population is still expanding in southern Norway, such attractive sinks could lead to local extinction. However, lynx seemed to select against these risky areas and therefore perceive them as unsuitable. Further analyses are needed on the exact locations where lynx are killed to test this hypothesis. From a prey perspective, we can also question whether roe deer are associated with humans as a way to escape lynx predation.

Optimal Foraging Theory (Stephens and Krebs, 1986) predicts that animals should balance both energy intake and predation risk, which often grow concurrently (Houston et al., 1993), leading to a necessary trade-off. Whereas the theory is well developed for herbivores, with a particular emphasis on modelling (e.g. Lima and Bednekoff, 1999), there has been little empirical evidence of this mechanism for large carnivores. Cheetahs *Acinonyx jubatus* in Africa have low competitive ability compared with their principal competitors, hyenas and lions, which are directly responsible for their low density (Durant, 1998). In response to that, cheetahs actively avoid lions and hyenas, and are therefore restricted in areas with lower prey density, which become "competition refuges" (Durant, 2000). Our results suggest that the same mechanism could occur in southern Norway where lynx face competition with humans. In this context, their low "competitive" ability compared with humans created both "competition" refuges at low to moderate roe deer density and potential attractive sinks at high prey density.

7.4.1 Conclusion

Lynx and roe deer share a common pattern of response to human disturbance. However, while roe deer benefit to a large extent from current human land use practices, potentially allowing them to escape predation from lynx, the situation seems not so favourable for the predators who are restricted in competition refuges at medium to low prey density. At the same time, lynx are very efficient predators even at low prey density and in addition can survive on numerous alternative prey species. Management plans for the conservation of lynx should therefore focus primarily on regulating lynx hunting (both legal and poaching) to ensure that the most favourable part of the habitat are freely available.

Acknowledgments

The data collection was financed by the Norwegian Directorate for Nature Management as part of the National Large Carnivore Monitoring Program that is operated by the Norwegian Institute for Nature Research. The analysis was conducted within the framework of Scandlynx, the Scandinavian Lynx Project, which receives funding from the Directorate for Nature Management, the Research Council of Norway and the county governor's offices of Akershus, Østfold, Oppland and Buskerud counties.

The French Ministry of Research, the University of Lyon, the Office National de la Chasse et de la Faune Sauvage and the Norwegian Institute for Nature Research provided financial support for this work. We are indebted to Clément Calenge, Jim Casaer, Anne Viallefont, Vincent Tollon and the GROS (Groupe de Recherche sur

les Ongulés Sauvages), for constructive comments at several stages of that work.

Chapter 8

Predation by lynx has the greatest impact on roe deer population growth at lower environmental productivity.

Claudia Melis[1,*], Mathieu Basille[1,2], Ivar Herfindal[1], John D. C. Linnell[3], John Odden[4], Reidar Andersen[3,4], Kjell Arild Høgda[5] & Jean-Michel Gaillard[1]

[1] Centre for Conservation Biology, Department of Biology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[2] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, F-69 622, Villeurbanne, France.

[3] Norwegian Institute for Nature Research (NINA); Tungasletta 2, NO-7485 Trondheim, Norway.

[4] Norwegian Institute for Nature Research (NINA); Gaustadalléen 21, NO-0349 Oslo, Norway.

[5] Museum of Natural History and Archaeology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[6] Norut Information Technology AS, NO-9291, Tromsø, Norway.

[*] Corresponding author: `claudia.melis@bio.ntnu.no`

In prep. for submission.

"A new statistic proves that 73% of all statistics are pure inventions."

J.J.A. WEBER.

FOR DECADES ECOLOGISTS HAVE DEBATED the relative importance of top-down and bottom-up forces in structuring populations and ecosystems (e.g. Murdoch, 1966; Ehrlich and Birch, 1967; Hunter and Price, 1992). Various studies conducted on a wide range of species in a variety of environments have produced a range of conflicting results, which support both mechanisms. In an attempt to unify the existing data, Oksanen et al. (1981) developed the exploitation ecosystems hypothesis (EEH), which combines these two views, predicting that the relative strength of top-down and bottom-up limitation will vary along productivity gradients. For instance, in a two-links system (i.e. a system with plants and herbivores), the EEH predicts that an increasing productivity would result in an increasing herbivore biomass which would regulate plant biomass at a stable level. In a three-links system (i.e. with the presence of carnivores), the EEH indicates that increasing productivity would result in an increasing carnivores biomass (up to a given threshold) that would limit herbivore biomass at a stable level. This in turn would result into an increasing plant biomass, no longer limited by herbivores (Oksanen et al., 1981; Oksanen and Oksanen, 2000). While some studies have found support for various components of the EEH (e.g. Crête, 1999; Aunapuu et al., 2008; Bråthen et al., 2007) others have found contradictory results (Kuijper and Bakker, 2005). Regardless of whether the EEH in its present form is correct, it has helped to structure the debate and inspired research to examine how the effect of predation varies with productivity (Meserve et al., 2003).

For example, a recent longitudinal study from Białowieża Primeval Forest (Poland) (Jędrzejewska and Jędrzejewski, 2005) suggested that predation exerted a stronger effect on ungulate populations when climate was most unfavourable and ungulates occurred at lowest densities. This result implies that in harsh conditions predation can not only limit (*sensu* Sinclair, 1989) herbivore biomass as predicted by the EEH but even cause a decrease in the herbivore biomass that would be sustained by the biomass of vegetation. In order to contribute to the ongoing discussion about the relationship between on one hand environmental productivity and on the other hand top-down and bottom-up processes, we have conducted the present study to investigate if it is possible to observe an effect of predation on wild ungulates from a recovering population of a large carnivore, and to determine how this effect varies along a productivity gradient, using Eurasian lynx *Lynx lynx* L. and European roe deer *Capreolus capreolus* L. in Norway as a case study.

The European roe deer is a widespread small ungulate that has successfully recolonised Norway during the 20th century (Andersen et al., 1998, 2004). Roe deer in Norway are mainly subjected to predation by Eurasian lynx, red foxes *Vulpes vulpes* L., wolves *Canis lupus* L., and to hunter harvest. In addition to predation, other important factors known to shaping their distribution and dynamics are climate (e.g. Danilkin, 1989; Grøtan et al., 2005, see Fig. 8.1) and productivity of vegetation (e.g. Pettorelli et al., 2006).

Lynx in Norway have followed the general pattern of large carnivore abundance and distribution in Europe (Linnell et al., ress). After being almost exterminated by the mid 20th century as consequence of a sponsored bounty program that started in 1846, they persisted in two remnant populations in south-eastern and central Norway. However, as a result of a restrictive hunting legislation lynx expanded during the late 20th century and are now present throughout the whole country with the exception of the south-west. As a consequence of conflicts with humans because of depredation on domestic sheep (Odden et al., 2002), semi-domestic reindeer (Pedersen et al., 1999) and competition for game species (Odden et al., 2006), Norwegian lynx populations are managed as a game species with the aim of limiting their density and distribution (Ministry of the Environment, 2004). The result has been a considerable debate about the impact of lynx populations on roe deer populations, and the extent to which hunters must adjust their hunting quotas. Determining the extent of this impact is important in both a management context and as a step in understanding predator-prey interactions. Our goal in this study has been to avail of existing monitoring data from both lynx and roe deer to dexamine how population trends of roe deer are influenced by presence of reproductive

Figure 8.1: Roe deer in the snow in Norway. Photo credit: Svein Ulvund.

lynx while controlling for the effects of climate, environmental productivity and roe deer density. Accordingly, based on existing literature, we formulated and tested the following predictions: 1) roe deer population growth rate varies along a gradient of climatic and vegetation productivity; 2) the trend in roe deer populations is negatively influenced by the presence of reproductive lynx, 3) the effect of lynx presence on population trend is greater in areas of low environmental productivity and / or areas of low roe deer density.

8.2 Materials and methods

8.2.1 Roe deer data

Roe deer are present in the whole southern part of Norway between 58°N and 65°N, with the exception of western areas (Fig. 8.2). The number of roe deer harvested annually in each municipality (H. Brøseth, pers. comm.) was divided by the square km of suitable area (i.e. excluding open water and alpine areas) and was used as an index of roe deer density. This type of index has been previously used for roe deer in Norway (Herfindal et al., 2005; Mysterud and Østbye, 2006) and its validity for some part of the study areas has been controlled by means of other indices (Grøtan et al., 2005; Mysterud and Østbye, 2006). The hunting success of roe deer in Norway is low (Mysterud and Østbye, 2006), during the years 1995-2002 only 33% ($\pm 17.6\%$) of quotas were filled on average in Norway ($n = 1883$), therefore the numbers of harvested roe deer are likely to reflect changes in population abundance rather than being artefacts of quotas (Grøtan et al., 2005; Mysterud and Østbye, 2006). We conducted our analysis on the population growth rate of roe deer instead than on roe deer density because we believed that the growth rate as calculated by hunting statistics would be more consistent throughout the country, independently from possible local differences in hunting effort. Moreover, we wanted to detect the impact of lynx after its reestablishment in many areas of southern Norway during the last decade and average roe deer density was more likely to be shaped by the underlying environmental conditions before lynx recolonisation. The population growth rate of roe deer between the years 1997–2005 was calculated for each municipality by fitting a linear regression to the log of the density index over the years, and then by

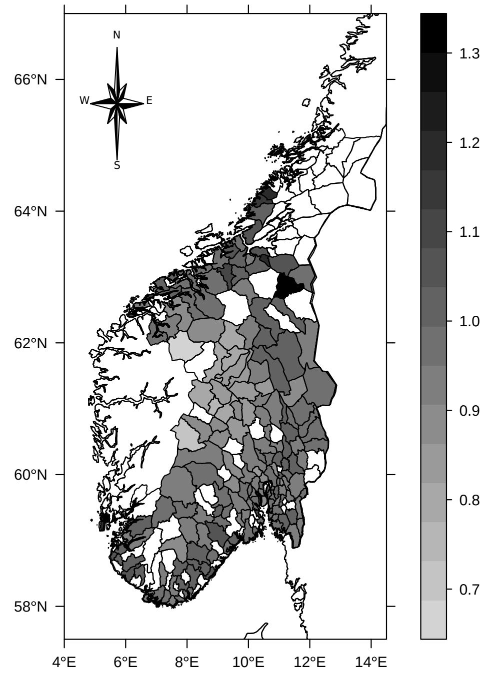


Figure 8.2: Roe deer growth rate between 1997-2005 in 178 municipalities in Norway, white = no data.

computing the exponential of the slope of that curve (n = 178) (Royama, 1992).

8.2.2 Lynx data

Data on the presence of reproductive lynx in each municipality are available since 1996 based on non-replicated counts of family groups (i.e. a female with dependent young of the year) and records of tracks in the snow collected by hunters, game wardens and the public, and checked by game wardens (Linnell et al., 2007) within the framework of a national large carnivore monitoring program. We used these data to obtain a index of the presence of absence of a resident lynx population based on absence (no reproductive lynx were present during any of the 9 years of the study) vs. presence (the municipality hosted a reproductive lynx at least in one year between 1997–2005) of reproductive lynx. On the 178 municipalities, 70 were categorized with lynx absence and 108 with lynx presence.

8.2.3 Environmental data

Climatic data (annual temperature, length of the growing season and temperature of the growing season) were obtained by Norwegian Meteorological Institute on a municipality base as normal for the years 1961–1991. Variables describing the plant phenology were extracted from the annual NDVI-curves based on the Global Inventory Monitoring and Modelling System (GIMMS) dataset (Karlsen et al., 2006). The GIMMS dataset consist of the maximum values of NDVI for 15-day periods with a spatial resolution of approximately 8×8 km^2, covering the world and available from 1982 until the present (Karlsen et al., 2006). The NDVI is based on the relationship between reflected red and near-infrared radiation from the ground. NDVI is closely related with the photosynthetic activity, plant biomass and net primary productivity (Myneni et al., 1995). The GIMMS dataset allows the calculation of annual NDVI-curves and the extraction of variables that describe the annual plant phenology (Pettorelli et al., 2005, 2006; Garel et al., 2006; Herfindal et al., 2006b). Plant phenology variables were calculated for each pixel in the GIMMS and averaged annually within each municipality for the years 1997–2002, excluding pixels representing large areas of open water and alpine areas. The fact that plant phenology data cover only two thirds of our study period (1997–2005) is justified by the rather small changes in climatic conditions during the study period.

Since plant phenology and climatic variables were highly intercorrelated ($r_p > 0.8$ in each group) we chose to use one for each group as explanatory variable in our model, based on their biological meaning and on their Pearson correlation coefficient in order to reduce multicollinearity (Graham, 2003). We therefore retained in our dataset integrated NDVI (annual net primary productivity) and mean annual temperature.

8.2.4 Statistical analyses

We used multiple linear regression to analyze the relationship between roe deer population growth rate, mean annual temperature, annual net primary productivity, the presence of reproductive lynx, an index of roe deer density (n harvested roe deer / km^2 excluding unsuitable areas) and the interactions between roe deer density and lynx presence and between annual net primary productivity and lynx presence. For this purpose, the data on density were log-transformed to improve normality and reduce skewness. We used AICc (Akaike Information Criterion corrected for small sample sizes, Burnham and Anderson, 2002) to select the most parsimonious model. To investigate for lack of spatial independence that could possibly lead to a Type I error (Legendre, 1993), a Moran test was performed on the residuals of the best model as selected by AICc. Since the test was significant for positive spatial correlation, a spatial simultaneous autoregressive model based on generalized least squares method was run as suggested by Diniz-Filho et al. (2003) and the estimates were compared to the ones of the ordinary least squares model. However, since the estimates of the two models were highly correlated (see Results) and the P-values were very consistent, we retained the ordinary least squares model according to the principle of parsimony. The analyses were conducted using R 2.6.2 Software (R Development Core Team, 2008). Spatial autocorrelation was investigated by means of the R package `spdep` (Bivand, 2007).

8.3 Results

Roe deer population growth rate in 1997–2005 averaged 0.96 (±0.08) across all municipalities. It averaged 1.00 (±0.07) in the absence of lynx, and 0.94 (±0.08) in the presence of lynx and was normally distributed. There were positive correlations between annual net primary productivity and mean annual temperature (r_p = 0.57, df = 176, t = 9.292, $P < 0.001$), annual temperature and roe deer density (r_p = 0.63, df = 176, t = 10.932, $P < 0.001$) and between roe deer density and annual net primary productivity (r_p = 0.29, df = 176, t = 4.017, $P < 0.001$).

The most parsimonious model as selected by AICc for roe deer population growth rate included all the explanatory variables: net primary productivity, mean annual temperature, lynx, roe deer density and the interaction between lynx and productivity and lynx and roe deer density (Table 8.1; df = 171, R^2 = 0.35). This model explained 35% of the observed variation in roe deer population growth rate. Mean annual temperature had a positive effect on the population growth rate, whereas presence of lynx and roe deer density had a negative effect (Table 8.2, Fig 8.3). Net primary productivity did not influence the growth rate in a consistent di-

		n = 178 municipalities			
Model	K	R^2	AIC$_c$	ΔAIC$_c$	ω_i
Full model: NDVI + Temp + log (Density) + Lynx + Lynx* log (Density) + Lynx*NDVI	8	0.35	-431.47	0.00	0.380
NDVI + Temp + log (Density)+ Lynx + Lynx*NDVI	7	0.34	-429.82	1.43	0.185
Temp + log (Density)	4	0.32	-429.59	1.90	0.148
Temp + log (Density)+ Lynx	5	0.32	-428.64	2.83	0.092
NDVI + Temp + log (Density)+ Lynx	6	0.33	-428.10	5.72	0.070

Table 8.1: Set of linear regression models with roe deer population growth rate (n = 178 municipalities in Norway, 1997–2005) as dependent variable and annual net primary productivity (NDVI), presence of lynx (Lynx), mean annual temperature (Temp), roe deer density (log-transformed) as explanatory variables. The models were ranked by the corrected Aikake Information Criterion (AIC$_c$). (K = number of parameters; ΔAIC$_c$ = difference in AICc between the best and the actual model; ω_i = Akaike's weights, i.e. normalized likelihoods of the models). All models contain the intercept. The most parsimonious model is on the top of the list and the first 5 ranked models are shown.

rection. The assessment of the effect of lynx presence / absence with varying temperature showed a consistently lower growth rate of roe deer in presence of lynx throughout the climatic gradient (Fig 8.3a). On the contrary the visual inspection of the effect of interaction between lynx and net primary productivity (Fig 8.3b) suggested that the effect of predation was higher at low productivity and that in the absence of lynx, productivity had little influence on the population growth rate. Regarding the interaction between lynx and roe deer density (Fig 8.3c) there was a weak higher effect of predation at lower roe deer density, however this relationship was not significant (Table 8.2).

The inspection of the residuals revealed that there was a source of spatial correlation not taken into account by our data set (Moran's I = 0.26, $P < 0.001$). An autoregressive model including the same variables of the ordinary least squares full model gave a very similar result in term of direction of the effects and significance of all variables. The estimates of the autoregressive model and of the ordinary least squares model were highly positively correlated ($r_p = 0.99$, df = 5, $t = 19.108$, $P < 0.001$),

Coefficients	Estimate	S.E.	*t*	*P*
Intercept	-0.147	0.059	-2.496	*
Lynx	-0.278	0.094	-2.498	**
NDVI	-0.004	0.008	-0.539	0.590
Temp	0.028	0.005	5.824	***
log (Density)	-0.026	0.009	-2.784	**
Lynx*NDVI	0.034	0.013	2.722	**
Lynx*log(Density)	-0.021	0.011	-1.876	0.062

Table 8.2: Estimates for the most parsimonious model of roe deer population growth rate in Norway (1997–2005), as selected by AICc. * < 0.05, ** < 0.01, *** < 0.001, for further details see Table 8.1.

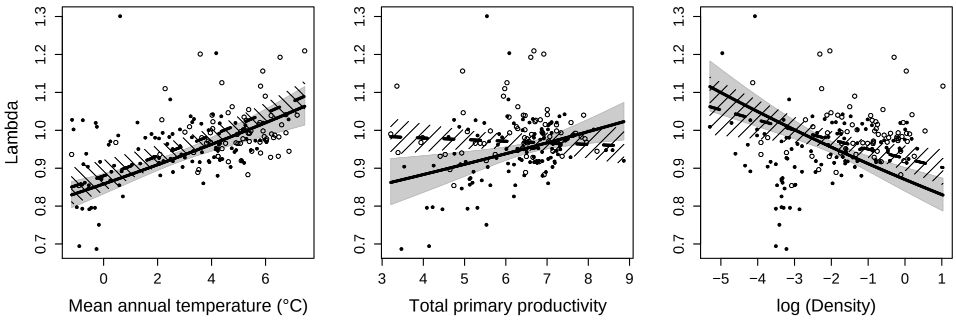


Figure 8.3: Prediction lines according to the most parsimonious model on population growth rate of roe deer in Norway (n = 178 municipalities) with varying (a) mean annual temperature, (b) net primary productivity, (c) roe deer density (number of individuals harvested per square km) in absence (black line) and presence (grey line) of reproductive lynx. The other explanatory variables are hold constant at their mean value. The dotted curves represent the 95% confidence intervals.

therefore minimizing the risk of making a Type I error.

8.4 Discussion

This study has shown that the presence of lynx had a negative impact on the growth rate of roe deer, and this impact was greatest in environments that had a low productivity. In other words, the strength of the top-down effect of this predator is highly context dependent.

8.4.1 Influence of large scale patterns (prediction 1)

Among our variables describing environmental conditions, mean annual temperature was positively associated with higher population grow rate of roe deer, whereas the main effect of net primary productivity was not significant. In roe deer populations climate have been considered one of the major causes of density independent mortality at the biogeographic scale (Danilkin, 1989, 1996). Deep snow restricts mobility and increases the energetic cost of movement, as well as reducing access to forage and increasing the hunting success of predators (Holand et al., 1998). Among northern ungulates roe deer may be among the most vulnerable to snow conditions (Grøtan et al., 2005), despite having a range of physiological and behavioural adaptations to survive during winter (Holand et al., 1998; Lamberti et al., 2004). Mysterud and Østbye (2006) found that population growth rate was negatively affected by increasing snow depth in southern Norway. Ideally we would have included snow conditions in our dataset, but these data were not available on a municipality scale and they would anyway have increased the level of intercorrelation between explanatory variables. We thus included mean annual temperature as a variable that would give us a more general picture of weather conditions. However it is likely that the strong positive effect of mean annual temperature on roe deer population growth rate operates also through snow conditions which are highly correlated with temperature at this study scale.

NDVI-related indices have been successfully related to herbivore performance (e.g. Pettorelli et al., 2005; Garel et al., 2006; Herfindal et al., 2006a; Pettorelli et al., 2006). At a small spatial scale, in the roe deer population at Chizé (western France), density-dependence has also been shown to rely on the interaction between population density and habitat quality

(Pettorelli et al., 2003). Roe deer are selective browsers (Tixier and Duncan, 1996) and they are small compared to other cervids, thus they are less dependent on quantity and more on quality of vegetation. Moreover, roe deer prefer early stage succession forests to aged ones (Gill et al., 1996) and avoid habitats like mountain tundra and coniferous taiga (Danilkin, 1996). NDVI, which measure the quantity of radiation adsorbed by plants, including coniferous and old stages of forests, might not be the best measure of food availability to roe deer. This, together with the collinearity between mean annual temperature, might contribute to explain why the main effect of net primary productivity on roe deer population growth rate was not significant.

On the other hand, according to the EEH, in the absence of lynx roe deer biomass should increase along the productivity gradient. Whereas in the presence of lynx, roe deer biomass would be predicted to be stable along the productivity gradient. We would therefore expect the roe deer population growth rates to be stable in the absence of lynx (although each population will stabilize at a different level according to local environmental conditions) if climate, productivity of vegetation and hunting pressure are relatively stable. In this respect our results are consistent with the EEH.

8.4.2 Influence of predators (predictions 2 and 3)

The presence of lynx had a negative effect on roe deer population growth rate across Norway. In accordance with the predictions of Jędrzejewska and Jędrzejewski (2005) the impact of predation was greatest in areas with low environmental productivity. The fact that this appears to contradict some of the predictions of the EEH needs explanation. Along a gradient of productivity, herbivore biomass would be predicted to reach a threshold in absence of carnivores, and then stabilize at this level; at the same time carnivores are increasing, indirectly exploiting the higher level of productivity through the herbivores. Our results suggest a top-down limitation of lynx on roe deer, but this limitation goes further than expected according to the EEH, since roe deer populations actually appeared to be decreasing in the presence of the predator. The implication is that lynx have a disproportionate impact on prey at low prey densities and in areas of low productivity. This in turn requires that predator kill rates remain high even at quite low prey density. In other words, lynx kill roe deer with the same rate, regardless of their availability, therefore will only have a limiting effect which corresponds to a type II functional response (Holling, 1959). Eurasian lynx vary their kill-rates only by a factor of two across a range of prey densities spanning two orders of magnitude (Breitenmoser and Haller, 1993; Linnell et al., 1996; Okarma et al., 1997; Molinari-Jobin et al., 2002; Nilsen et al., data). This, even in the absence of numerical responses on the part of the predator, will automatically imply a greater impact on low density prey populations.

We did not formulate any prediction about the influence of roe deer density on population growth rate. Reduced population growth rate at high densities have been found in many studies (e.g. for a review Fowler, 1987; Gaillard et al., 2000a). A weak density dependence in body weight has been documented in the absence of predation in the population on Storfosna island (central Norway), a real two-link system *sensu* Oksanen and Oksanen (2000), where the population density was up to 10 times higher than inland (Andersen and Linnell, 2000). In Sweden roe deer generally occur at higher densities and density dependence has been found both in body weight and reproduction (Kjellander and Nordström, 2003). A recent longitudinal study in southern Norway did not find any evidence of density dependence in the population growth rate of roe deer (Mysterud and Østbye, 2006). Conversely, we did find a negative effect of density on the population growth rate of roe deer. On the other hand, the analysis of the interaction (albeit not significant) between lynx and roe deer density suggested that in the presence of predation the relationship was reversed and the population growth rate was higher at higher roe deer densities. This might be due to the fact that at low densities (and productivity) roe deer are not able to recover the loss caused by predation.

However, when interpreting the results of this type of analysis it is important to consider the quality of the data and the wide range of confounding factors that can be present. The source of positive spatial correlation that was found in the residuals might be due to the presence of small clustered municipalities in our dataset, each of which cannot be attributed to one distinct roe deer population. Also, lynx home ranges are so large that they are rarely, if ever, embraced by a single municipality (Herfindal et al., 2005). Nevertheless, grouping the municipalities by county would have been rather artificial, since there is no biological or human-related background, relying on hunting traditions or management administration, for using county as a grouping variable. In addition, the size of counties is so large that they would embrace too wide a range of environmental conditions. Migration of yearlings between neighbouring populations (e.g. Wahlström and Liberg, 1995) might also explain the spatial autocorrelation in our data. Finally, we could not account for the impact of red fox predation that exerts a strong impact on roe deer at fawn stage (e.g. Cederlund and Lindström, 1983; Aanes and Andersen, 1996; Panzacchi et al., 2008; Elmhagen and Rushton, 2007).

These uncertainties underline the limitations intrinsic to this type of broad scale study. There was a great deal of variation that we were not able to explain. A wide range of human and non-human factors (biotic and abiotic) clearly influence roe deer populations across Norway, and it is unlikely that any model will be able to account for more than a fraction of them. However, despite this, our model was able to explain a significant proportion of the variation in terms of climate, productivity and predation. It is also highly unlikely that any of the aforementioned limitations will induce any bias into the analysis that is likely to produce the results we have documented. Therefore, we have a fairly high degree of confidence that our results do reflect real biological processes, They are also consistent with our results on lynx kill rates, that are consistently high even at low prey densities (Nilsen et al., data) and roe deer hunting success which is higher on roe deer with increasing snow depth (Odden et al., data). The results therefore support the pattern of environmentally conditional predation impacts found by Jędrzejewska and Jędrzejewski (2005) and indicate that roe deer managers have an urgent need to adjust hunting quotas in low productivity areas where lynx are present.

Acknowledgments

We are indebted to Mathieu Garel, Stéphane Dray and Christophe Bonenfant for fruitful discussions on the approach developed in this study. The study was funded by the Norwegian Directorate for Nature Management, The Norwegian Institute for Nature Research, the Research Council of Norway and the offices of Environmental Affairs from Hedmark, Oppland, Akershus, Buskerud and Østfold counties. CM was funded by a post-doctoral scholarship from the Norwegian University of Science and Technology.

Part IV

General Synthesis

"Seul le hasard peut nous apparaître comme un message.
Ce qui arrive par nécessité, ce qui est attendu et se répète
quotidiennement n'est que chose muette. Seul le hasard est
parlant. On tente d'y lire comme les gitanes lisent au fond
d'une tasse dans les figures qu'a dessinées le marc de café."
"L'insoutenable légereté de l'être", by Milan KUNDERA.

DURING THESE ALMOST FOUR YEARS OF PHD, I think I have been confronted to the typical scientific process. I was faced first with a very applied problematic: *"How can we describe the spatial relationships between lynx, roe deer, and human?"* Lynx are, as a in Norway (see Fig. 8.4). While roe deer is the main prey of lynx throughout southern Norway, and both species are heavily hunted, this question has concrete impacts in terms of conservation and management (see Part III). But this question in turn raised several theoretical concerns that soon required both answers (see Part I) and tools (see Part II) to be applied. In a sense, this work may not fulfill his initial aim regarding the system lynx–roe deer–human in Norway, but probably laid the foundations of a deeper and more fundamental knowledge of this system.

From theory to application

The first pitfall to circumvent was about terminology. One can not rely on concepts and terms without a clear and unambiguous definition of them. If it is not possible, there's a need to explain what is really meant every time a term is used. That was the case for the habitat and niche concepts (**Chapter 2**). Habitat and niche are central terms in ecology, used since many decades and at the basis for many theories (competition, speciation, ideal free distribution, etc.) but still lack a general agreement about what is meant by the two terms. We had to disentangle first the meanings and the field of application of both, following the approach first suggested by Whittaker et al. (1973). The habitat is often considered in a classical multivariate approach: *n* environmental variables are measured and define a

Figure 8.4: A lynx radio-monitored in southern Norway. Photo credit: Scandlynx.

n-dimensional geometrical space called *ecological space*. The whole set of variables can however be split in two categories: habitat variables at large scale, and niche variables at fine scale, each defining its own sub-space. In the space of niche (resp. habitat) variables, the ecological niche (resp. habitat) is defined as the hypervolume in which a species can potentially persist. This formalization can be extended with a projection of a species population performance that would define completely the niche or habitat shape, depending on the variables considered.

While it was clear that there was no short and straightforward definitions of these two terms together, the operational clarification of both enabled us to embrace a general theory of habitat selection (**Chapter 1**). In habitat selection studies, the term habitat is generally used in an inclusive way, including both niche variables and habitat variables *sensu stricto*. We gave some definitions regarding availability, use and preference. Habitat selections studies should integrate scales, species' performance and dynamics as tightly as possible.

The scientific process being what it is, answering one question usually does not allow the scientist to go on at the point he stopped while asking this very question. Instead, one answer opened new areas of investigation. In our case, it opened both exciting theoretical and methodological concerns. First, in the context of a multi-scale hierarchical habitat selection, we investigated individual habitat selection at different spatial scales (**Chapter 3**). It is generally assumed that factors with greater potential to reduce individual fitness should be avoided at coarser spatial and temporal scales, as a strategy to maximise an individual's fitness (Rettie and Messier, 2000). However, we demonstrated, with the help of simple simulations, that the variability of a factor should be accounted for first, as it defines the potential of selection for a factor. Thus, a factor most variable may eventually have a great effect on individual fitness even though the species is less sensitive to that factor. As variability of a factor is scale dependant, a species should express the strongest habitat selection at the scale where the combination of sensitivity and variability for a given factor is the highest.

Methodological issues

The initial work was a refinement of a recent method, the Ecological-niche factor analysis (ENFA, Hirzel et al., 2002a). This method is able to consider at the same time directions in the ecological space on which the difference between what is available and what is used is the greatest (marginality), and those on which the ratio of available to used variance is the greatest (specialization). The previous part allowed us to assess precisely the model on which the method rely, the ecological niche, and to note that the method may be applied indifferently on formalizations of the habitat and the niche. Following Cleveland's philosophy (Cleveland, 1993), we provided graphical tools to easily interpret the results of the method, for instance with the help of factorial maps (**Chapter 4**).

We then extended this approach of the study of the formalization of the niche by defining a general framework of analysis, based on a similar multivariate approach (**Chapter 6**). This framework called GNESFA (General niche-environment system factor analysis) encompasses three complementary methods, the ENFA, the MADIFA (Mahalanobis distance factor analysis) and the FANTER (Factor analysis of the niche, taking the environment as the reference), and allows an extensive description of a species' niche or habitat. The MADIFA can furthermore compute habitat suitability maps, that gives the probability of encountering a species' individual at any place (**Chapter 5**). The FANTER has the advantage of handling niches (or habitats) of various shapes, for instance bimodal.

Management issues

Both theoretical and methodological developments enabled us to study habitat selection of lynx from an applied point of view. Applied to the case study of lynx in the Vosges mountains (France), this framework underlined the influence of urbanisation corridors to structure lynx space use, as they are confined in forested areas at high altitude (**Chapter 4**).

Most of the work was related to Norway though. In most part of Europe, the main prey of lynx are roe deer. At the same time, lynx are heavily hunted in Norway, in order to maintain limited densities and range. In these conditions, it was primordial to study the whole lynx–roe deer-human system. First, we focused on the large scale determinants of lynx distribution in southern Norway (**Chapter 7**). The habitat of lynx is mainly characterized with productive areas with high prey density. Surprisingly, it was also characterized by relatively high human disturbances. While roe deer benefit to a large extent from current human land use practices, it is likely that this resulted from a trade-off between food (abundance of prey) and security (avoidance of human activity).

In turn, these results called for new concerns. As lynx distribution is primarily driven by roe deer distribution, we can thus question the impact of lynx on roe deer populations. The hypothesis of exploitation ecosystems (EEH, Oksanen and Oksanen, 2000) predicted that herbivore biomass, in the absence of carnivores, should increase along a gradient of productivity until the system can sustain the presence of carnivores. The latter should then limit herbivore biomass and eventually stabilize in case of intra-guild predation. Preliminary results of roe deer dynamics at a large scale showed indeed the strong limiting pressure of lynx along a gradient of productivity (**Chapter 8**). Our results were even stronger than predicted since roe deer populations seem to decrease.

Future perspectives

We believe that this work synthesized a global approach of a case study, with theoretical, methodological, and applied developments. However, a lot remains to be answered. In our system in Norway, fine scale investigations should be considered. In particular, we still have to define habitat selection behaviours of lynx at a fine scale. Since lynx are somewhat attracted to human activities (**Chapter 7**), this may result in a maladaptative habitat selection with lynx using preferentially areas associated to a major risk of mortality. From the roe deer perspective, this would allow the study of a fine-scale model of lynx impact on roe deer. Here again, humans may be the source of maladaptative behaviours: roe deer are often found in disturbed areas, especially in relation to agricultural fields. While this is probably a solution to escape lynx predation, this should result in a higher human-caused mortality. We thus expect roe deer to experience difficult conditions everywhere in southern Norway along a gradient of human disturbance: at a very low level of human activity, roe deer would be freed from predation but confronted with harsh conditions; at intermediate human pressure, lynx would probably kill the roe deer surplus; and in the proximity of human, roe deer would face the highest hunting pressure while escaping lynx.

From a broader perspective, our work highlighted promising fields in habitat selection studies. We showed in **Chapter 1 & 2** that a species' habitat was fully defined by a measure of population performance, describing its quality. In particular, this performance could be approached by a measure of fitness of individuals in every conditions encountered in the habitat. However, there is at least one case where the quality of a given habitat does not correspond to the actual selection, thus highlighting a form of maladaptative selection. This has been described under the name of "attractive sinks" (Delibes et al., 2001), or equivalent "ecological traps" (Battin, 2004), that correspond to low quality habitats (sinks), that are not recognized as such by animals.

This concern should introduce a shift in the current habitat selection paradigm. We should study in parallel habitat quality (as used in source-sink theory) and actual habitat selection. The latter defining attractive habitats (that are positively selected) vs. what may be called *repulsive* habi-

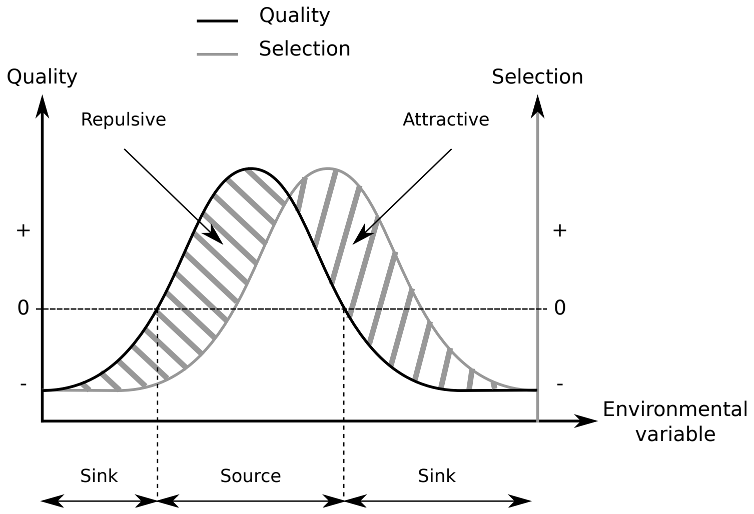


Figure 8.5: A synthetic approach of the habitat. On the left axis (black line), the quality of the habitat is given, measured for instance with the fitness of individuals along a gradient of habitat (environmental variable). On the right, the selection is represented (grey line). Both curves are standardized (the area under each sums to 1) so that they can be compared. A sink is defined with a negative quality, whereas a source corresponds to a positive quality. On the other hand, a repulsive habitat is less selected than expected by the quality only, and an attractive habitat is more selected than expected by the quality only. The figure presents on purpose an extreme case with very little optimal habitat selection.

tats (that are actually avoided), four cases may be considered. The first two relate to adaptative habitat selection (Pulliam, 1988): attractive sources and repulsive sinks; whereas the alternative cases relate to maladaptative habitat selection (Gilroy and Sutherland, 2007): attractive sinks and repulsive sources (the latter being similar to the concept of "undervalued resource", *sensu* Gilroy and Sutherland (2007) (Fig. 8.5). Further exploration of this basic framework is much needed in order to synthesize these different approaches.

In this respect, the integration of animals activity would improve the grain of habitat selection studies, focusing more into individuals behaviours (i.e. processes) than the static result (see **Chapter 1**). Design 4 studies based on trajectories are in this respect a very much needed field of progress for the coming years.

Bibliography

Åberg, J., Jansson, G., Swenson, J., and Mikusinski, G. (2000). Difficulties in detecting habitat selection by animals in generally suitable areas. *Wildlife Biology*, 6:89–99.

Aanes, R. and Andersen, R. (1996). The effects of sex, time of birth, and habitat on the vulnerability of roe deer fawns to red fox predation. *Canadian Journal of Zoology*, 74:1857–1865.

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M., and Matthiopoulos, J. (2008). Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, 31:140–160.

Andersen, R., Duncan, P., and Linnell, J. (1998). *The European roe deer: the biology of success.* Scandinavian University Press, Oslo.

Andersen, R., Herfindel, I., Saether, B. E., Linnell, J. D. C., Odden, J., and Liberg, O. (2004). When range expansion rate is faster in marginal habitats. *Oikos*, 107:210–214.

Andersen, R., Karlsen, J., Austmo, L. B., Odden, J., Linnell, J. D. C., and Gaillard, J. M. (2007). Selectivity of eurasian lynx lynx lynx and recreational hunters for age, sex and body condition in roe deer capreolus capreolus. *Wildlife Biology*, 13:467–474.

Andersen, R. and Linnell, J. (2000). Irruptive potential in roe deer: density-dependent effects on body mass and fertility. *Journal of Wildlife Management*, 64:698–706.

Andersen, R. and Linnell, J. D. C. (1998). Ecological correlates of mortality of roe deer fawns in a predator-free environment. *Canadian Journal of Zoology*, 76:1217–1225.

Andrén, H., Linnell, J. D. C., Liberg, O., Andersen, R., Danell, A., Karlsson, J., Odden, J., Moa, P. F., Ahlqvist, P., Kvam, T., Franzén, R., and Segerström, P. (2006). Survival rates and causes of mortality in Eurasian lynx (*Lynx lynx*) in multi-use landscapes. *Biological Conservation*, 131:23–32.

Apps, C. D., McLellan, B. N., Woods, J. G., and Proctor, M. F. (2004). Estimating grizzly bear distribution and abundance relative to habitat and human influence. *Journal of Wildlife Management*, 68:138–152.

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

Araújo, M. and Williams, P. (2000). Selecting areas for species persistence using occurence data. *Biological Conservation*, 96:331–345.

Arthur, S. M., Manly, B. F. J., McDonald, L. L., and Garner, G. W. (1996). Assessing habitat selection when availability changes. *Ecology*, 77:215–227.

Aunapuu, M., Dahlgren, J., Oksanen, T., Grellmann, D., Oksanen, L., Olofsson, J., Rammul, U., Schneider, M., Johansen, B., and Hygen, H. O. (2008). Spatial patterns and dynamic responses of arctic food webs corroborate the exploitation ecosystems hypothesis (EEH). *American Naturalist*, 171:249–262.

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

Bartholomé, E., Belward, A. S., Achard, F., Bartalev, S., Carmona Moreno, C., Eva, H., Fritz, S., Grégoire, J.-M., Mayaux, P., and Stibig, H.-J. (2002). GLC 2000 Global Land Cover mapping for the year 2000. Technical report, European Commission, Joint Research Centre, Institute for Environment and Sustainability.

Barton, D. and David, F. (1956). Some notes on ordered random intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 18:79–94.

Basille, M., Calenge, C., Marboutin, É., Andersen, R., and Gailllard, J.-M. (2008). Assessing habitat selection using multivariate statistics: Some refinements of the Ecological-Niche Factor Analysis. *Ecological Modelling*, 211:233–240.

Battin, J. (2004). When good animals love bad habitats: Ecological traps and the conservation of animal populations. *Conservation Biology*, 18:1482–1491.

Beckmann, J. P. and Berger, J. (2003). Using black bears to test ideal-free distribution models experimentally. *Journal of Mammalogy*, 84:594–606.

Begon, M., Harper, J. L., and Townsend, C. R. (1986). *Ecology - Individuals, Populations and Communities*. Blackwell Scientific Publications.

Bivand, R. (2007). *Spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.4-20.

Bleich, V., Bowyer, R., and Wehausen, J. (1997). Sexual segregation in mountain sheep: Resources or predation? *Wildlife Monographs*, 134:1–50.

Blondel, J., Chessel, D., and Frochot, B. (1988). Bird species impoverishment, niche expansion and density inflation in mediterranean island habitats. *Ecology*, 69(6):1899–1917.

Bowyer, R. T. and Kie, J. G. (2006). Effects of scale on interpreting life-history characteristics of ungulates and carnivores. *Diversity and Distributions*, 12:244–257.

Boyce, M. S., Mao, J. S., Merrill, E. H., Fortin, D., Turner, M. G., Fryxell, J., and Turchin, P. (2003). Scale and heterogeneity in habitat selection by elk in yellowstone national park. *Ecoscience*, 10:421–431.

Boyce, M. S. and McDonald, L. (1999). Relating populations to habitats using resource selection functions. *Trends in Ecology & Evolution*, 14:268–272.

Breitenmoser, U. (1998). Large predators in the Alps: The fall and rise of man's competitor. *Biological Conservation*, 83:279–289.

Breitenmoser, U. and Haller, H. (1993). Patterns of predation by reintroduced European lynx in the Swiss Alps. *Journal of Wildlife Management*, 57:135–144.

Breitenmoser, U., Kaczensky, P., Dötterer, M., Breitenmoser-Würsten, C., Capt, S., Bernhart, F., and Liberek, M. (1993). Spatial organization and recruitment of lynx (*Lynx lynx*) in a re-introduced population in the Swiss Jura Mountains. *Journal of Zoology (London)*, 231:449–464.

Browning, D., Beaupré, S., and Duncan, L. (2005). Using partitioned Mahalanobis $d^2(k)$ to formulate a GIS-based model of timber rattlesnake hibernacula. *Journal of Wildlife Management*, 69:33–44.

Bryan, T. and Metaxas, A. (2007). Predicting suitable habitat for deep-water gorgonian corals on the atlantic and pacific continental margins of North America. *Marine Ecology Progress Series*, 330:113–126.

Bråthen, K. A., Ims, R. A., Yoccoz, N. G., Fauchald, P., Tveraa, T., and Hausner, V. H. (2007). Induced shift in ecosystem productivity? Extensive scale effects of abundant large herbivores. *Ecosystems*, 10:773–789.

Bunnefeld, N., Börger, L., Nilsen, E. B., Basille, M., Hall, R., Ezard, T. H., Trierweiler, C., Minderman, J., Mangel, M., Gaillard, J.-M., Milner-Gulland, E., and attendees of the *Populations under Pressure* symposium (2007). Coming out of the ivory tower: How to ensure that ecological modelling research remains practical and applied. *Bulletin of the British Ecological Society*, 38:64–66.

Bunnefeld, N., Linnell, J. D. C., Odden, J., van Duijn, M. A. J., and Andersen, R. (2006). Risk taking by Eurasian lynx (*Lynx lynx*) in a human-dominated landscape: effects of sex and reproductive status. *Journal of Zoology (London)*, 270:31–39.

Burnham, K. and Anderson, D. (1998). *Model selection and inference.* Springer, Berlin.

Burnham, K. and Anderson, D. (2002). *Model selection and multi-model inference: A practical information-theoretic approach.* Springer-Verlag, Berlin.

Calenge, C. (2005). *Des outils statistiques pour l'analyse des semis de points dans l'espace écologique.* PhD thesis, Université Claude Bernard Lyon 1.

Calenge, C. (2006). The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling,* 197:516–519.

Calenge, C., Darmon, G., Basille, M., Loison, A., and Jullien, J.-M. (2008). The factorial decomposition of the mahalanobis distances in habitat selection studies. *Ecology,* 89:555–566.

Calenge, C., Dufour, A., and Maillard, D. (2005). K-select analysis: A new method to analyse habitat selection in radio-tracking studies. *Ecological Modelling,* 186:143–153.

Calenge, C. and Dufour, A. B. (2006). Eigenanalysis of selection ratios from animal radio-tracking data. *Ecology,* 87:2349–2355.

Candolin, U. and Voigt, H.-R. (2003). Do changes in risk-taking affect habitat shifts of sticklebacks? *Behavioral Ecology and Sociobiology,* 55:42–49.

Cassing, G., Greenberg, L. A., and Mikusinski, G. (2006). Moose (alces alces) browsing in young forest stands in central sweden: A multiscale perspective. *Scandinavian Journal of Forest Research,* 21:221–230.

Cayuela, L. (2004). Habitat evaluation for the Iberian wolf *Canis lupus* in Picos de Europa National Park, Spain. *Applied Geography,* 24:199–215.

Cederlund, G. and Lindström, E. (1983). Effects of severe winters and fox predation on roe deer mortality. *Acta Theriologica,* 287:129–145.

Chefaoui, R. M., Hortal, J., and Lobo, J. M. (2005). Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: A case study of Iberian Copris species. *Biological Conservation,* 122:327–338.

Chessel, D. and Gimaret, C. (1997). *Niche - ADE-4 modules documentation.* Université Lyon 1.

Ciarniello, L. M., Boyce, M. S., and Beyer, H. (2003). Resource Selection Function for the plateau landscape of the Parsnip grizzly bear project (An update for 2003). Technical report, British Columbia Ministry of Forests.

Clark, J. D., Dunn, J. E., and Smith, K. G. (1993). A multivariate model of female black bear habitat use for a geographic information system. *Journal of Wildlife Management*, 57:519–526.

Cleveland, W. S. (1993). *Visualizing Data*. Hobart Press, Summit, New Jersey. 360 pp.

Cleveland, W. S. (1994). *The elements of graphing data*. Hobart Press, Summit, New Jersey.

Compton, R. C. (2004). Predicting key habitat and potential distribution of northern bottlenose whales (*Hyperoodon ampullatus*) in the northwest Atlantic ocean. Master's thesis, University of Plymouth.

Corsi, F., Dupré, E., and Boitani, L. (1999). A large scale model of wolf distribution in Italy for conservation planning. *Conservation Biology*, 13:150–159.

Coulson, T., Albon, S., Guinness, F., Pemberton, J., and CluttonBrock, T. (1997). Population substructure, local density, and calf winter survival in red deer (*Cervus elaphus*). *Ecology*, 78:852–863.

Coulson, T., Gaillard, J.-M., and Festa-Bianchet, M. (2005). Decomposing the variation in population growth into contributions from multiple demographic rates. *Journal of Animal Ecology*, 74:789–801.

Cowlishaw, G. (1997). Trade-offs between foraging and predation risk determine habitat use in a desert baboon population. *Animal Behaviour*, 53:667–686.

Crête, M. (1999). The distribution of deer biomass in north america supports the hypothesis of exploitation ecosystems. *Ecology Letters*, 2:223–227.

Dahle, B., Sorensen, O. J., Wedul, E. H., Swenson, J. E., and Sandegren, F. (1998). The diet of brown bears ursus arctos in central scandinavia: Effect of access to free-ranging domestic sheep ovis aries. *Wildlife Biology*, 4:147–158.

Danilkin, A. (1989). Structure of the geographic range of the roe deer in the ussr and factors determining its boundaries. *Soviet Journal of Ecology*, 20:233–237.

Danilkin, A. (1996). *Behavioural ecology of the Siberian and European roe deer*. Chapman & Hall, London.

Davies, T. J., Meiri, S., Barraclough, T. G., and Gittleman, J. L. (2007). Species co-existence and character divergence across carnivores. *Ecology Letters*, 10:146–152.

Delibes, M., Gaona, P., and Ferreras, P. (2001). Effects of an attractive sink leading into maladaptive habitat selection. *American Naturalist*, 158:277–285.

Dettki, H., Lofstrand, R., and Edenius, L. (2003). Modeling habitat suitability for moose in coastal northern Sweden: Empirical vs process-oriented approaches. *Ambio*, 32:549–556.

Diniz-Filho, J. A. F., Bini, L. M., and Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12:53–64.

Doledec, S., Chessel, D., and Gimaret Carpentier, C. (2000). Niche separation in community analysis: A new method. *Ecology*, 81:2914–2927.

Dray, S., Chessel, D., and Thioulouse, J. (2003). Co-inertia analysis and the linking of ecological data tables. *Ecology*, 84:3078–3089.

Dunn, J. and Duncan, L. (2000). Partitioning Mahalanobis d^2 to sharpen gis classification. In Brebbia, C. and Pascolo, P., editors, *Management Information Systems 2000: GIS and Remote sensing*, pages 195–204. WIT Press, Southampton, U.K.

Durant, S. M. (1998). Competition refuges and coexistence: an example from serengeti carnivores. *Journal of Animal Ecology*, 67:370–386.

Durant, S. M. (2000). Living with the enemy: avoidance of hyenas and lions by cheetahs in the serengeti. *Behavioral Ecology*, 11:624–632.

Dussault, C., Ouellet, J.-P., Courtois, R., Huot, J., Breton, L., and Jolicoeur, H. (2005). Linking moose habitat selection to limiting factors. *Ecography*, 28:619–628.

Edenius, L., Bergman, M., Ericsson, G., and Danell, K. (2002). The role of moose as a disturbance factor in managed boreal forests. *Silva Fennica*, 36:57–67.

Edwards, J. (1983). Diet shifts in moose due to predator avoidance. *Oecologia*, 60:185–189.

Ehrlich, P. and Birch, L. (1967). The "balance of nature" and "population control". *American Naturalist*, 101:97–107.

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

Elmhagen, B. and Rushton, S. P. (2007). Trophic control of mesopredators in terrestrial ecosystems: top-down or bottom-up? *Ecology Letters*, 10:197–206.

Elton, C. (1927). *Animal ecology*. Sidgwick and Jackson, London, England.

Enserink, M. and Vogel, G. (2006). Wildlife conservation: The carnivore comeback. *Science*, 314:746–749.

Escoufier, Y. (1987). The duality diagram: a means of better practical applications. In Legendre, P. and Legendre, L., editors, *Development in numerical ecology*, Series G, pages 139–156. Springer Verlag, Berlin.

Farber, O. and Kadmon, R. (2003). Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, 160:115–130.

Fernandez, N., Delibes, M., Palomares, F., and Mladenoff, D. J. (2003). Identifying breeding habitat for the Iberian lynx: Inferences from a fine-scale spatial analysis. *Ecological Applications*, 13:1310–1324.

Ferrari, C., Rossi, G., and Cavani, C. (1988). Summer food habits and quality of female, kid and subadult apennine chamois, *Rupicapra pyrenaica ornata* neumann, 1899 (*Artiodactyla, Bovidae*). *Z Säugetier*, 53:170–177.

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

Flagstad, O., Hedmark, E., Landa, A., Broseth, H., Persson, J., Andersen, R., Segerstrom, P., and Ellegren, H. (2004). Colonization history and noninvasive monitoring of a reestablished wolverine population. *Conservation Biology*, 18:676–688.

Fortin, D., Beyer, H., Boyce, M., Smith, D., Duchesne, T., and Mao, J. (2005). Wolves influence elk movements: Behavior shapes a trophic cascade in Yellowstone National Park. *Ecology*, 86:1320–1330.

Fournier, A., Fussell, D., and Carpenter, L. (1982). Computer rendering of stochastic models. In *Communications of the ACM, June 1982*.

Fowler, C. (1987). A review of density-dependence in populations of large mammals. In Genoways, H., editor, *Current mammalogy, Vol 1.*, pages 401–441. Plenum Press, New York.

Freckleton, R. P., Watkinson, A. R., Green, R. E., and Sutherland, W. J. (2006). Census error and th detection of density dependence. *Journal of Animal Ecology*, 75:837–851.

Frontier, S. (1976). Étude de la décroissance des valeurs propres dans une analyse en composantes principales : comparaison avec le modele du baton brisé. *Journal of Experimental Marine Biology and Ecology*, 25:67–75.

Gabriel, K. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58:453–467.

Gaillard, J.-M., Festa-Bianchet, M., Delorme, D., and Jorgenson, J. (2000a). Body mass and individual fitness in female ungulates: bigger is not always better. *Proceedings of the Royal Society of London B Biological Sciences*, 267:471–477.

Gaillard, J.-M., Festa-Bianchet, M., and Yoccoz, N. (1998). Population dynamics of large herbivores: variable recruitment with constant adult survival. *Trends in Ecology & Evolution*, 13:58–63.

Gaillard, J.-M., Festa-Bianchet, M., Yoccoz, N., Loison, A., and Toïgo, C. (2000b). Temporal variation in fitness components and population dynamics of large herbivores. *Annual Review of Ecology and Systematics*, 31:367–393.

Gallego, D., Canovas, F., Esteve, M. A., and Galián, J. (2004). Descriptive biogeography of *Tomicus* (Coleoptera: Scolytidae) species in Spain. *Journal of Biogeography*, 31:2011–2024.

Garcia-Gonzalez, R. and Cuartas, P. (1996). Trophic utilization of a montane/subalpine forest by chamois (*Rupicapra pyrenaica*) in the central pyrenees. *Forest Ecology and Management*, 88:15–23.

Gardner, R. (1999). Rule: A program for the generation of random maps and the analysis of spatial patterns. In Klopatek, J. and Gardner, R., editors, *Landscape Ecological Analysis: Issues and Applications*, pages 280–303. Springer-Verlag, New York, USA.

Garel, M., Solberg, E., Sæther, B.-E., Herfindal, I., and Høgda, K.-A. (2006). The length of growing season and adult sex ratio affect sexual size dimorphism in moose. *Ecology*, 87:745–758.

Gause, G. (1934). *The struggle for existence.* Williams & Wilkins, Baltimore, USA.

George, S. L. and Crooks, K. R. (2006). Recreation and large mammal activity in an urban nature reserve. *Biological Conservation*, 133:107–117.

Gibeau, M. and Herrero, S., editors (1998). *Roads, rails and grizzly bears in the Bow River Valley, Alberta.*, Proceedings International Conference on Ecology and Transportation.

Gibson, L. A., Wilson, B. A., Cahill, D. M., and Hill, J. (2004). Spatial prediction of rufous bristlebird habitat in a coastal heathland: A GIS-based approach. *Journal of Applied Ecology*, 41:213–223.

Gill, R., Johnson, A., Francis, A., Hiscocks, K., and Peace, A. (1996). Changes in roe deer (*Capreolus capreolus* l.) population density in response to forest habitat succession. *Forest Ecology and Management*, 88:31–41.

Gilpin, M. and Hanski, I., editors (1991). *Metapopulation Dynamics: Empirical and Theoretical Investigations.* Academic Press, London. 336 pp.

Gilroy, J. J. and Sutherland, W. J. (2007). Beyond ecological traps: perceptual errors and undervalued resources. *Trends in Ecology & Evolution*, 22:351–356.

Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology*, 84:2809–2815.

Grinnell, J. (1917). The niche-relations of the california thrasher. *Auk*, 34:427–433.

Grøtan, V., Sæther, B. E., Engen, S., Solberg, E. J., Linnell, J. D. C., Andersen, R., Brøseth, H., and Lund, E. (2005). Climate causes large-scale spatial synchrony in population fluctuations of a temperate herbivore. *Ecology*, 86:1472–1482.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M., McC. Overton, J., Aspinall, R., and Hastie, T. (2006). Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, 43:386–392.

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

Guisan, A. and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135:147–186.

Gustine, D., Parker, K., Lay, R., Gillingham, M., and Heard, D. (2006). Interpreting resource selection at different scales for woodland caribou in winter. *Journal of Wildlife Management*, 70:1601–1614.

Hall, L. S., Krausman, P. R., and Morrison, M. L. (1997). The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, 25:173–182.

Hamr, J. (1985). Seasonal home range size and utilisation by female chamois (*Rupicapra rupicapra*) in northern tyrol. In Lovari, S., editor, *The biology and management of mountain ungulates*, pages 106–116. Croom Helm, London, UK.

Harvey, D. and Weatherhead, P. (2006). A test of the hierarchical model of habitat selection using eastern massasauga rattlesnakes (*Sistrurus c. catenatus*). *Biological Conservation*, 130:206–216.

Harville, D. (1997). *Matrix algebra from a statistician's perspective*. Springer, New York, USA.

Haskell, E. F. (1940). Mathematical systematization of "environment," "organism" and "habitat". *Ecology*, 21:1–16.

Herfindal, I., Linnell, J. D. C., Odden, J., Nilsen, E. B., and Andersen, R. (2005). Prey density, environmental productivity and home-range size in the Eurasian lynx (*Lynx lynx*). *Journal of Zoology (London)*, 265:63–71.

Herfindal, I., Solberg, E. J., Sæther, B.-E., Høgda, K. A., and Andersen, R. (2006a). Environmental phenology and geographical gradients in moose body mass. *Oecologia*, 150:213–224.

Herfindal, I., Sæther, B.-E., Solberg, E. J., Andersen, R., and Høgda, K.-A. (2006b). Population characteristics predict responses in moose body mass to temporal variation in the environment. *Journal of Animal Ecology*, 75:1100–1118.

Hill, M. (1974). Correspondence analysis: A neglected multivariate method. *Applied Statistics*, 23:340–354.

Hill, M. and Smith, A. (1976). Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon*, 25:249–255.

Hirzel, A. H. and Arlettaz, R. (2003a). Environmental-envelope based habitat-suitability models. In Manly, B. F. J., editor, *1st Conference on Resource Selection by Animals*, pages 67–76. Omnipress, Laramie, USA.

Hirzel, A. H. and Arlettaz, R. (2003b). Modeling habitat suitability for complex species distributions by environmental-distance geometric mean. *Environmental Management*, 32:614–623.

Hirzel, A. H., Hausser, J., Chessel, D., and Perrin, N. (2002a). Ecological-Niche Factor Analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83:2027–2036.

Hirzel, A. H., Hausser, J., and Perrin, N. (2002b). *Biomapper 3.1*. Lab. for Conservation Biology, Department of Ecology and Evolution, University of Lausanne.

Hirzel, A. H., Posse, B., Oggier, P. A., Crettenand, Y., Glenz, C., and Arlettaz, R. (2004). Ecological requirements of reintroduced species and the implications for release policy: The case of the bearded vulture. *Journal of Applied Ecology*, 41:1103–1116.

Holand, Ø., Mysterud, A., Wannag, A., and Linnell, J. (1998). Roe deer in northern environments: Physiology and behaviour. In Andersen, R., P., D., and Linnell, J., editors, *The European roe deer: The biology of success.*, pages 117–137. Scandinavian University Press, Oslo, Norway.

Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism. *Canadian Entomologist*, 91:285–298.

Holt, R. (1997). On the evolutionary stability of sink populations. *Evolutionary Ecology*, 11:723–731.

Holt, R. D. (1985). Population dynamics in two-patch environments: Some anomalous consequences of an optimal habitat distribution. *Theoretical Population Biology*, 28:181–208.

Houssin, H., Loison, A., Gaillard, J.-M., and Jullien, J.-M. (1994). Validité d'une méthode d'estimation des effectifs de chamois dans un massif des préalpes du nord. *Gibier Faune Sauvage*, 11:287–298.

Houston, A. I., McNamara, J. M., and Hutchinson, J. M. C. (1993). General results concerning the trade-off between gaining energy and avoiding predation. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 341:375–397.

Hunter, M. and Price, P. (1992). Playing chutes and ladders: heterogeneity and the relative roles of bottom-up and top-down forces in natural communities. *Ecology*, 73:724–732.

Hutchinson, G. E. (1957). Concluding remarks. In *Cold Spring Harbour Symposium on Quantitative Biology*, volume 22, pages 415–427.

Hutchinson, G. E. (1959). Homage to Santa Rosalia *or* Why are there so many kinds of animals? *American Naturalist*, 93:145–159.

Jackson, D. A. (1993). Stopping rules in Principal Components Analysis: A comparison of heuristical and statistical approaches. *Ecology*, 74:2204–2214.

Johnson, C. J., Boyce, M. S., Case, R. L., Cluff, H. D., Gau, R. J., Gunn, A., and Mulders, R. (2005). Cumulative effects of human developments on arctic wildlife. *Wildlife Monographs*, 160:1–36.

Johnson, D. H. (1980). The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, 61:65–71.

Jędrzejewska, B. and Jędrzejewski, W. (2005). Large carnivores and ungulates in European temperate forest ecosystems: Bottom-up and top-down control. In Ray, J. C., Redford, K. H., Steneck, R. S., and Berger, J., editors, *Large carnivores and the conservation of biodiversity.*, pages 230–245. Island Press, Washington.

Kalen, C. and Bergquist, J. (2004). Forage availability for moose of young silver birch and scots pine. *Forest Ecology and Management*, 187:149–158.

Karlsen, S. R., Elvebakk, A., Høgda, K. A., and Johansen, B. (2006). Satellite-based mapping of the growing season and bioclimatic zones in fennoscandia. *Global Ecology and Biogeography*, 15:416–430.

Katajisto, J. (2006). *Habitat use and population dynamics of brown bears (Ursus arctos) in Scandinavia.* PhD thesis, University of Helsinki, Department of Biological and Environmental Sciences, University of Helsinki, Finland.

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

Kind, A. (2001). Blåbærlyng i ulike skogmiljø i sørøst-norge (in Norwegian). Master's thesis, Agricultural University of Norway, Ås.

Kindberg, J., Swenson, J., Brunberg, S., and Ericsson, G. (2006). Preliminär rapport om populationsutveckling och -storlek av brunbjörn i sverige (in Swedish). Technical report, Naturvårdsverket.

Kjellander, P. and Nordström, J. (2003). Cyclic voles, prey switching in red fox, and roe deer dynamics – a test of the alternative prey hypothesis. *Oikos*, 101:338–344.

Knick, S. and Dyer, D. (1997). Distribution of black-tailed jackrabbit habitat determined by gis in southwestern idaho. *Journal of Wildlife Management*, 61:75–85.

Knick, S. and Rotenberry, J. (1998). Limitations to mapping habitat use areas in changing landscapes using the Mahalanobis distance statistic. *Journal of Agricultural Biological and Environmental Statistics*, 3:311–322.

Kuijper, D. and Bakker, J. (2005). Top-down control of small herbivores on salt-marsh vegetation along a productivity gradient. *Ecology*, 86:914–923.

Kulesza, G. (1975). Comment on "niche, habitat, and ecotope". *American Naturalist*, 109:476–479.

Lamberti, P., Mauri, L., and Apollonio, M. (2004). Two distinct patterns of spatial behaviour of females roe deer (*Capreolus capreolus*) in a mountainous habitat. *Ethology Ecology & Evolution*, 16:41–53.

Legendre, P. (1993). Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74:1659–1673.

Legendre, P. and Legendre, L. (1998). *Numerical ecology, 2nd Edition.* Elsevier Science BV, Amsterdam, The Netherlands.

Leibold, M. A. (1995). The niche concept revisited — Mechanistic models and community context. *Ecology*, 76:1371–1382.

Lewis, J. S., Rachlow, J. L., Garton, E. O., and Vierling, L. A. (2007). Effects of habitat on GPS collar performance: using data screening to reduce location error. *Journal of Applied Ecology*, 44:663–671.

Lima, S. L. and Bednekoff, P. A. (1999). Temporal variation in danger drives antipredator behavior: The predation risk allocation hypothesis. *American Naturalist*, 153:649–659.

Linnell, J. D., Odden, J., Andersen, R., Brøseth, H., Andrén, H., Liberg, O., Ahlqvist, P., Moa, P., Kvam, T., Segerström, P., Schmidt, K., Jędrzejewski, W., and Okarma, H. (2007). Distance rules for minimum counts of eurasian lynx *Lynx lynx* family groups under different ecological conditions. *Wildlife Biology*, 13:447–455.

Linnell, J. D. C., Andersen, R., Kvam, T., Andren, H., Liberg, O., Odden, J., and Moa, P. F. (2001a). Home range size and choice of management strategy for lynx in Scandinavia. *Environmental Management*, 27:869–879.

Linnell, J. D. C., Breitenmoser, U., Breitenmoser-Würsten, C., Odden, J., and von Arx, M. (in press). Recovery of Eurasian lynx in Europe: What part has reintroduction played? In Sommers, M. and Hayward, M., editors, *Reintroduction of top predators*. Blackwell Press.

Linnell, J. D. C., Promberger, C., Boitani, L., Swenson, J. E., Breitenmoser, U., and Andersen, R. (2005). The linkage between conservation strategies for large carnivores and biodiversity: The view from the "half-full" forests of Europe. In Ray, J. C., Redford, K. H., Steneck, R. S., and Berger, J., editors, *Large carnivores and the conservation of biodiversity.*, pages 381–398. Island Press, Washington.

Linnell, J. D. C., Støen, O. G., Odden, J., Ness, E., Gangås, L., Karlsen, J., Eide, N., and Andersen, R. (1996). Gaupe og rådyr i østre deler av Hedmark. *NINA Oppdragsmelding*, 443:1–83.

Linnell, J. D. C., Swenson, J. E., and Andersen, R. (2001b). Predators and people: Conservation of large carnivores is possible at high human densities if management policy is favourable. *Animal Conservation*, 4:345–349.

Linnell, J. D. C., Swenson, J. E., Andersen, R., and Barnes, B. (2000). How vulnerable are denning bears to disturbance? *Wildlife Society Bulletin*, 28:400–413.

Loseto, L., Richard, P., Stern, G., Orr, J., and Ferguson, S. (2006). Segregation of Beaufort Sea beluga whales during the open-water season. *Canadian Journal of Zoology*, 84:1743–1751.

Mahalanobis, P. C. (1948). Historic note on the d^2 statistic. *Sankhya*, 9:237–240.

Mallik, A. U. (2003). Conifer regeneration problems in boreal and temperate forests with ericaceous understory: Role of disturbance, seedbed limitation, and keytsone species change. *Critical Reviews in Plant Sciences*, 22:341–366.

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., and Erickson, W. P. (2002). *Resource selection by animals. Statistical design and analysis for field studies. Second edition*. Kluwer Academic, Dordrecht, The Netherlands.

Martin, J., Basille, M., Kindberg, J., Allainé, D., and Swenson, J. E. (unpublished data). Hierarchical habitat selection of brown bears: low order does matter.

Martin, J., Calenge, C., Quenette, P.-Y., and Allainé, D. (2008). Importance of movement constraints in habitat selection studies. *Ecological Modelling*, 213:257–262.

Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., and Possingham, H. P. (2005). Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. *Ecology Letters*, 8:1235–1246.

Mason, H. L. and Langenheim, J. H. (1957). Language analysis and the concept *Environment*. *Ecology*, 38:325–340.

May, R., Landa, A., van Dijk, J., Linnell, J., and Andersen, R. (2006). Impact of infrastructure on habitat selection of wolverines *Gulo gulo*. *Wildlife Biology*, 12:285–295.

Maynard Smith, J. (1966). Sympatric speciation. *American Naturalist*, 100:637–650.

Mayor, S., Schaefer, J., Schneider, D., and Mahoney, S. (2007). Spectrum of selection: new approaches to detecting the scale-dependent response to habitat. *Ecology*, 88:1634–1640.

Mayr, E. W. (1963). *Animal Species and Evolution*. Belknap Press (Harvard University Press).

McDougall, W. B. (1927). *Plant ecology*. Lee & Febiger, New York.

McKinney, M. L. (2002). Urbanization, biodiversity, and conservation. *Bioscience*, 52:883–890.

McLoughlin, P. D., Boyce, M. S., Coulson, T., and Clutton-Brock, T. (2006). Lifetime reproductive success and density-dependent, multi-variable resource selection. *Proceedings of the Royal Society of London B Biological Sciences*, 273:1449–1454.

McLoughlin, P. D., Case, R. L., Gau, R. J., Cluff, H. D., Mulders, R., and Messier, F. (2002). Hierarchical habitat selection by barren-ground grizzly bears in the central Canadian Arctic. *Oecologia*, 132:102–108.

McLoughlin, P. D., Dunford, J. S., and Boutin, S. (2005). Relating predation mortality to broad-scale habitat selection. *Journal of Animal Ecology*, 74:701–707.

McLoughlin, P. D., Gaillard, J.-M., Boyce, M. S., Bonenfant, C., Messier, F., Duncan, P., Delorme, D., Moorter, B. V., Saïd, S., and Klein, F. (2007). Lifetime reproductive success and composition of the home range in a large herbivore. *Ecology*, 88:3192–3201.

McLoughlin, P. D., Walton, L. R., Cluff, H. D., Paquet, P. C., and Ramsay, M. A. (2004). Hierarchical habitat selection by tundra wolves. *Journal of Mammalogy*, 85:576–580.

McPeek, M. A. (1996). Trade-offs, food web structure, and the coexistence of habitat specialists and generalists. *American Naturalist*, 148:S124–S138.

Melis, C., Linnell, J. D., Odden, J., and Andersen, R. (unpublished data). Survival rates and causes of mortality of roe deer *Capreolus capreolus* in a multipredator system.

Meserve, P., Kelt, D., Milstead, B., and Gutiérrez, J. (2003). Thirteen years of shifting top-down and bottom-up control. *Bioscience*, 53:633–646.

Michallet, D. (2003). Modélisation de la distribution spatiale du chamois de Chartreuse (*Rupicapra rupicapra cartusiana*) par l'analyse des facteurs environnementaux. Master's thesis, Licence Professionnelle de Traitement de l'Information Géographique, IUT de Perpignan.

Milne, B. (1988). Measuring the fractal geometry of landscapes. *Applied Mathematics and Computation*, 27:67–79.

Milne, B. (1991). Heterogeneity as a multi-scale characteristic of landscapes. In Kolasa, J. and Waters, W. E., editors, *Ecological heterogeneity.*, pages 69–84. Springer, New York, USA.

Ministry of the Environment (2004). Carnivores in Norwegian nature [rovvilt i norsk natur] st.meld nr 15 (2003-2004). Technical report.

Mitchell, M. S. and Powell, R. A. (2004). A mechanistic home range model for optimal use of spatially distributed resources. *Ecological Modelling*, 177:209–232.

Mitchell, S. C. (2005). How useful is the concept of habitat? A critique. *Oikos*, 110:634–638.

Mladenoff, D. J., Sickley, T. A., Haight, R. G., and Wydeven, A. P. (1995). A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. *Conservation Biology*, 9:279–294.

Mladenoff, D. J., Sickley, T. A., and Wydeven, A. P. (1999). Predicting gray wolf landscape recolonization: Logistic regression models vs. new field data. *Ecological Applications*, 9:37–44.

Moe, T. F., Kindberg, J., Jansson, I., and Swenson, J. E. (2007). Importance of diel behaviour when studying habitat selection: Examples from female Scandinavian brown bears (*Ursus arctos*). *Canadian Journal of Zoology*, 85:518–525.

Mohr, C. O. (1947). Table of equivalent populations of North American small mammals. *American Midland Naturalist*, 37:223–249.

Molinari-Jobin, A., Molinari, P., Breitenmoser-Würsten, C., and Breitenmoser, U. (2002). Significance of lynx *Lynx lynx* predation for roe deer and chamois mortality in the Swiss Jura Mountains. *Wildlife Biology*, 8:109–116.

Molinari-Jobin, A., Molinari, P., Loison, A., Gaillard, J.-M., and Breitenmoser, U. (2004). Life cycle period and activity of prey influence their susceptibility to predators. *Ecography*, 27:323–329.

Morin, P., Berteaux, D., and Klvana, I. (2005). Hierarchical habitat selection by north american porcupines in southern boreal forest. *Canadian Journal of Zoology*, 83:1333–1342.

Morris, D. K. (2003a). How can we apply theories of habitat selection to wildlife conservation and management? *Wildlife Research*, 30:303–319.

Morris, D. W. (1987). Ecological scale and habitat use. *Ecology*, 68:362–369.

Morris, D. W. (2003b). Toward an ecological synthesis: A case for habitat selection. *Oecologia*, 136:1–13.

Morrison, M. L., Marcot, B. G., and Mannan, R. W. (2006). *Wildlife-Habitat Relationships: Concepts and Applications.* Island Press, Washington, Covelo, London, 3rd edition.

Murdoch, W. W. (1966). Community structure, population control, and competition — a critique. *American Naturalist*, 100:219–226.

Murray, C. M., Mane, S. V., and Pusey, A. E. (2007). Dominance rank influences female space use in wild chimpanzees, pan troglodytes: towards an ideal despotic distribution. *Animal Behaviour*, 74:1795–1804.

Murray, D. L. (2006). On improving telemetry-based survival estimation. *Journal of Wildlife Management*, 70:1530–1543.

Myneni, R. B., Hall, F. G., Sellers, P. J., and Marshak, A. L. (1995). The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing*, 33:481–486.

Mysterud, A. (1999). Seasonal migration pattern and home range of roe deer (*Capreolus capreolus*) in an altitudinal gradient in southern Norway. *Journal of Zoology (London)*, 247:479–486.

Mysterud, A., Lian, L. B., and Hjermann, D. Ø. (1999). Scale-dependent trade-offs in foraging by European roe deer (*Capreolus capreolus*) during winter. *Canadian Journal of Zoology*, 77:1486–1493.

Mysterud, A. and Østbye, E. (2006). Effect of climate and density on individual and population growth of roe deer *Capreolus capreolus* at northern latitudes: The Lier valley, Norway. *Wildlife Biology*, 12:321–329.

Mäki-Petäys, A., Muotka, T., Huusko, A., Tikkanen, P., and Kreivi, P. (1997). Seasonal changes in habitat use and preference by juvenile brown trout, *Salmo trutta*, in a northern boreal river. *Canadian Journal of Fisheries and Aquatic Sciences*, 54:520–530.

Nellemann, C., Støen, O.-G., Kindberg, J., Swenson, J. E., Vistnes, I., Ericsson, G., Katajisto, J., Kaltenborn, B. P., Martin, J., and Ordiz, A. (2007). Terrain use by an expanding brown bear population in relation to age, recreational resorts and human settlements. *Biological Conservation*, 138:157–165.

Nellemann, C. and Thomsen, M. G. (1994). Terrain ruggedness and caribou forage availability during snowmelt on the arctic coastal-plain, Alaska. *Arctic*, 47:361–367.

Niedziałkowska, M., Jędrzejewski, W., Mysłajek, R. W., Nowak, S., Jędrzejewska, B., and Schmidt, K. (2006). Environmental correlates of Eurasian lynx occurrence in poland – Large scale census and GIS mapping. *Biological Conservation*, 133:63–69.

Nielsen, S. E., Boyce, M. S., and Stenhouse, G. B. (2004a). Grizzly bears and forestry: I. Selection of clearcuts by grizzly bears in west-central Alberta, Canada. *Forest Ecology and Management*, 199:51–65.

Nielsen, S. E., Munro, R. H. M., Bainbridge, E. L., Stenhouse, G. B., and Boyce, M. S. (2004b). Grizzly bears and forestry: II. Distribution of grizzly bear foods in clearcuts of west-central Alberta, Canada. *Forest Ecology and Management*, 199:67–82.

Nietzsche, F. W. (1889). *Götzen-Dämmerung* oder *Wie man mit dem Hammer philosophiert*. KSA 6.

Nikula, A., Heikkinen, S., and Helle, E. (2004). Habitat selection of adult moose *Alces alces* at two spatial scales in central Finland. *Wildlife Biology*, 10:121–135.

Nilsen, E. B., Odden, J., Linnell, J. D. C., and Andersen, R. (unpublished data). Climate, season, and social status modulate the functional response of an efficient stalking predator: The Eurasian lynx.

Noss, R., Quigley, H., Hornocker, M., Merrill, T., and Paquet, P. (1996). Conservation biology and carnivore conservation in the Rocky Mountains. *Conservation Biology*, 10:949–963.

Noy-Meir, I. (1973). Data transformation in ecological ordination. I. Some advantages of non-centering. *Journal of Ecology*, 61:329–341.

Odden, J., Andersen, R., and Linnell, J. D. C. (unpublished data). Hunting success of lynx with respect to prey species and snow depth.

Odden, J., Linnell, J. D. C., and Andersen, R. (2006). Diet of Eurasian lynx, *Lynx lynx*, in the boreal forest of southeastern Norway: The relative importance of livestock and hares at low roe deer density. *European Journal of Wildlife Research*, 52:237–244.

Odden, J., Linnell, J. D. C., Moa, P. F., Herfindal, I., Kvam, T., and Andersen, R. (2002). Lynx depredation on domestic sheep in Norway. *Journal of Wildlife Management*, 66:98–105.

Odden, J., Solvang, H., Maartmann, E., Wabakken, P., Linnell, J., Andersen, R., Haagenrud, H., Lundqvist, O., and Solberg, H. O. (2001). Lynx and wolf census in Hedmark 2001 [registrering av ulv og gaupe i Hedmark 2001: Rapport fra registrering 13. januar 2001]. Technical Report Rapport 11/2001, 1–26, Fylkesmannen i Hedmark Miljøvernavdelingen.

Okarma, H., Jędrzejewski, W., Schmidt, K., Kowalczyk, R., and Jędrzejewska, B. (1997). Predation of Eurasian lynx on roe deer and red deer in Bialowieza Primeval Forest, Poland. *Acta Theriologica*, 42:203–224.

Oksanen, L., Fretwell, S., Arruda, J., and Niemelä, P. (1981). Exploitation ecosystems in gradients of primary productivity. *American Naturalist*, 118:240–261.

Oksanen, L. and Oksanen, T. (2000). The logic and realism of the hypothesis of exploitation ecosystems. *American Naturalist*, 155:703–723.

Olivier, F. and Wotherspoon, S. J. (2006). Modelling habitat selection using presence-only data: Case study of a colonial hollow nesting bird, the snow petrel. *Ecological Modelling*, 195:187–204.

Orians, G. H. and Wittenberger, J. F. (1991). Spatial and temporal scales in habitat selection. *American Naturalist*, 137:S29–S49.

Panzacchi, M., Linnell, J. D. C., Odden, J., Odden, M., and Andersen, R. (2008). When a generalist becomes a specialist: Patterns of red fox predation on roe deer fawns under contrasting conditions. *Canadian Journal of Zoology*, 86:116–126.

Pearce, J. L. and Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43:405–412.

Pedersen, V. A., Linnell, J. D. C., Andersen, R., Andrén, H., Lindén, M., and Segerström, P. (1999). Winter lynx *Lynx lynx* predation on semi-domestic reindeer *Rangifer tarandus* in northern Sweden. *Wildlife Biology*, 5:203–211.

Perrin, N. (1984). *Contribution à l'écologie du genre* Cepaea *(Gastropoda): approche descriptive et expérimentale de l'habitat et de la* *niche écologique*. PhD thesis, Faculté des Sciences de l'Université de Lausanne.

Pettorelli, N., Gaillard, J.-M., Duncan, P., Maillard, D., Van Laere, G., and Delorme, D. (2003). Age and density modify the effects of habitat quality on survival and movements of roe deer. *Ecology*, 84:3307–3316.

Pettorelli, N., Gaillard, J.-M., Mysterud, A., Duncan, P., Stenseth, N. C., Delorme, D., Van Laere, G., Toigo, C., and Klein, F. (2006). Using a proxy of plant productivity (NDVI) to find key periods for animal performance: the case of roe deer. *Oikos*, 112:565–572.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., and Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20:503–510.

Pidgeon, A. M., Radeloff, V. C., and Mathews, N. E. (2006). Contrasting measures of fitness to classify habitat quality for the black-throated sparrow (*Amphispiza bilineata*). *Biological Conservation*, 132:199–210.

Pollock, K., Winterstein, S., Bunck, C., and Curtis, P. (1989). Survival analysis in telemetry studies – The staggered entry design. *Journal of Wildlife Management*, 53:7–15.

Provine, W. B. (2005). Ernst mayr, a retrospective. *Trends in Ecology & Evolution*, 20:411–413.

Pulliam, H. R. (1988). Sources, sinks, and population regulation. *American Naturalist*, 132:652–661.

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3:349–361.

Pulliam, H. R. and Danielson, B. J. (1991). Sources, sinks, and habitat selection: A landscape perspective on population dynamics. *American Naturalist*, 137:S50–S66.

R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rameau, J., Mansion, D., and Dume, G. (2001). *Flore forestière française, guide écologique, tome 2: montagnes*. Institut Pour Le Developpement Forestier, Paris, France.

Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., and Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5:703–714.

Rettie, W. J. and Messier, F. (2000). Hierarchical habitat selection by woodland caribou: Its relationship to limiting factors. *Ecography*, 23:466–478.

Reutter, B. A., Helfer, V., Hirzel, A. H., and Vogel, P. (2003). Modelling habitat-suitability using museum collections: An example with three sympatric *Apodemus* species from the Alps. *Journal of Biogeography*, 30:581–590.

Ricklefs, R. E. (1990). *Ecology*. W. H. Freeman and Company, New York, third edition.

Riley, S. P. D., Sauvajot, R. M., Fuller, T. K., York, E. C., Kamradt, D. A., Bromley, C., and Wayne, R. K. (2003). Effects of urbanization and habitat fragmentation on bobcats and coyotes in southern California. *Conservation Biology*, 17:566–576.

Rolstad, J., Løken, B., and Rolstad, E. (2000). Habitat selection as a hierarchical spatial process: The green woodpecker at the northern edge of its distribution range. *Oecologia*, 124:116–129.

Rolstad, J., Majewski, P., and Rolstad, E. (1998). Black woodpecker use of habitats and feeding substrates in a managed Scandinavian forest. *Journal of Wildlife Management*, 62:11–23.

Rosenzweig, C., Karoly, D., Vicarelli, M., Neofotis, P., Wu, Q., Casassa, G., Menzel, A., Root, T. L., Estrella, N., Seguin, B., Tryjanowski, P., Liu, C., Rawlins, S., and Imeson, A. (2008). Attributing physical and biological impacts to anthropogenic climate change. *Nature*, 453:353–357.

Rosenzweig, M. L. (1978). Competitive speciation. *Biological Journal of the Linnean Society*, 10:275–289.

Rosenzweig, M. L. (1981). A theory of habitat selection. *Ecology*, 62:327–335.

Rotenberry, J. T., Knick, S. T., and Dunn, J. E. (2002). A minimalist approach to mapping species' habitat: Pearson's planes of closest fit. In Scott, J. M., Heglund, P. J., Morrison, M. L., Haufler, J. B., Raphael, M. G., Wall, W. A., and Samson, F. B., editors, *Predicting Species Occurences – Issues of accuracy and scale*, pages 281–289. Island Press, Washington, Covelo, London.

Rotenberry, J. T., Preston, K. L., and Knick, S. T. (2006). GIS-based niche modeling for mapping species' habitat. *Ecology*, 87:1458–1464.

Royama, T. (1992). *Analytical population dynamics*. Chapman and Hall.

Rushton, S. P., Ormerod, S. J., and Kerby, G. (2004). New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41:193–200.

Røskaft, E., Bjerke, T., Kaltenborn, B., Linnell, J. D. C., and Andersen, R. (2003). Patterns of self-reported fear towards large carnivores among the Norwegian public. *Evolution and Human Behavior*, 24:184–198.

Saunders, D. A., Hobbs, R. J., and Margules, C. R. (1991). Biological consequences of ecosystem fragmentation: A review. *Conservation Biology*, 5:18–32.

Schadt, S., Revilla, E., Wiegand, T., Knauer, F., Kaczensky, P., Breitenmoser, U., Bufka, L., Cerveny, J., Koubek, P., Huber, T., Stanisa, C., and Trepl, L. (2002). Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. *Journal of Applied Ecology*, 39:189–203.

Schaefer, J. A. and Messier, F. (1995). Habitat selection as a hierarchy: The spatial scales of winter foraging by muskoxen. *Ecography*, 18:333–344.

Seber, G. (1984). *Multivariate observations*. Wiley Series in probability and mathematical statistics, New York.

Senft, R. L., Coughenour, M. B., Bailey, D. W., Rittenhouse, L. R., Sala, O. E., and Swift, D. M. (1987). Large herbivore foraging and ecological hierarchies. *Bioscience*, 37:789–799.

Sinclair, A. (1989). Population regulation in animals. In Cherrett, J., editor, *Ecological concepts.*, pages 197–241. Blackwell, Oxford, UK.

Soberón, J. and Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2:1–10.

Soulé, M. E., Estes, J. A., Miller, B., and Honnold, D. L. (2005). Strongly interacting species. conservation policy, management, and ethics. *Bioscience*, 55:168–176.

Spichiger, R., Calenge, C., and Bise, B. (2004). The geographical zonation in the neotropics of tree species characteristic of the paraguay-paraná basin. *Journal of Biogeography*, 31:1489–1501.

Stein, E. M. and Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press.

Stephens, D. and Krebs, J. (1986). *Foraging Theory*. Princeton University Press, Princeton, NJ.

Støen, O.-G., Bellemain, E., Sæbø, S., and Swenson, J. (2005). Kin-related spatial structure in brown bears *Ursus arctos*. *Behavioral Ecology and Sociobiology*, 59:191–197.

Støen, O.-G., Zedrosser, A., Sæbø, S., and Swenson, J. (2006). Inversely density-dependent natal dispersal in brown bears *Ursus arctos*. *Oecologia*, 148:356–364.

Sunde, P., Stener, S. Ø., and Kvam, T. (1998). Tolerance to humans of resting lynxes *Lynx lynx* in a hunted population. *Wildlife Biology*, 4:177–183.

Sweden, S. (2003). Assessment of real estate in 2003, part 2. bo 38 sm 0401. Technical report, Statistics Sweden, Stockholm, Sweden.

Swenson, J. E., Dahle, B., Busk, H., Opseth, O., Johansen, T., Soderberg, A., Wallin, K., and Cederlund, G. (2007). Predation on moose calves by European brown bears. *Journal of Wildlife Management*, 71:1993–1997.

Swenson, J. E., Jansson, A., Riig, R., and Sandegren, F. (1999). Bears and ants: Myrmecophagy by brown bears in central Scandinavia. *Canadian Journal of Zoology*, 77:551–561.

Swenson, J. E., Sandegren, F., Bjärvall, A., Söderberg, A., Wabakken, P., and Franzén, R. (1994). Size, trend, distribution and conservation of the brown bear *Ursus arctos* population in Sweden. *Biological Conservation*, 70:9–17.

Swenson, J. E., Sandegren, F., and SO-Derberg, A. (1998). Geographic expansion of an increasing brown bear population: Evidence for presaturation dispersal. *Journal of Animal Ecology*, 67:819–826.

Swenson, J. E., Wabakken, P., Sandegren, F., Bjarvall, A., Franzen, R., and Soderberg, A. (1995). The near extinction and recovery of brown bears in Scandinavia in relation to the bear management policies of Norway and Sweden. *Wildlife Biology*, 1:11–25.

Takle, M. . (2002). Befolkningsstatistikk på rutenett - dokumentasjon. Technical report, Notat 2002/22. Statistics Norway (In Norwegian). 34 pp.

ter Braak, C. (1985). Correspondence analysis of incidence and abundance data : properties in terms of a unimodal reponse model. *Biometrics*, 41:859–873.

ter Braak, C. (1986). Canonical correspondence analysis : a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, 67:1167–1179.

Thioulouse, J. and Chessel, D. (1992). A method for reciprocal scaling of species tolerance and sample diversity. *Ecology*, 73:670–680.

Thomas, D. L. and Taylor, E. J. (1990). Study designs and tests for comparing resource use and availability. *Journal of Wildlife Management*, 54:322–330.

Thomas, D. L. and Taylor, E. J. (2006). Study designs and tests for comparing resource use and availability II. *Journal of Wildlife Management*, 70:324–336.

Thompson, L., van Manen, F., Schlarbaum, S., and DePoy, M. (2006). A spatial modeling approach to identify potential butternut restoration sites in mammoth cave national park. *Restoration Ecology*, 14:298–296.

Thorpe, W. H. (1945). The evolutionary significance of habitat selection. *Journal of Animal Ecology*, 14:67–70.

Tilman, D. (1980). Resources: A graphical-mechanistic approach to competition and predation. *American Naturalist*, 116:362–393.

Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton University Press.

Tilman, D. (1987). The importance of the mechanisms of interspecific competition. *American Naturalist*, 129:769–774.

Tilman, D. and Kareiva, P. M., editors (1997). *Spatial ecology: The role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton. 368 pp.

Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, 20:410–433.

Tixier, H. and Duncan, P. (1996). Are European roe deer browsers? A review of variations in the composition of their diets. *Rev. Ecol. Terre-Vie*, 51:3–17.

Tucker, C., Pinzon, J., Brown, M., Slayback, D., Pak, E., Mahoney, R., Vermote, E., and El Saleous, N. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, 26:4485–4498.

Tufto, J. and Cavallini, P. (2005). Should wildlife biologists use free software? *Wildlife Biology*, 11:67–76.

Turner, M., Gardner, R., and O'Neil, R. (2001). *Landscape Ecology in Theory and Practice: Pattern and process*. Springer, New York, USA.

Turner, M. G., Arthaud, G. J., Engstrom, R. T., Hejl, S. J., Liu, J., Loeb, S., and McKelvey, K. (1995). Usefullness of spatially explicit population management. *Ecological Applications*, 5:12–17.

Udvardy, M. F. D. (1959). Notes on the ecological concepts of habitat, biotope and niche. *Ecology*, 40:725–728.

Van Horne, B. (1983). Density as a misleading indicator of habitat quality. *Journal of Wildlife Management*, 47:893–901.

Vandel, J.-M. and Stahl, P. (2005). Distribution trend of the Eurasian lynx (*Lynx lynx*) populations in France. *Mammalia*, 69:145–158.

Vandel, J.-M., Stahl, P., Herrenschmidt, V., and Marboutin, É. (2006). Reintroduction of the lynx into the Vosges mountain massif: From animal survival and movements to population development. *Biological Conservation*, 131:370–385.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, fourth edition.

Ver Hoef, J. M. and Boveng, P. L. (2007). Quasi-poisson vs. negative binomial regression: How should we model overdispersed count data? *Ecology*, 88:2766–2772.

Volterra, V. (1926). Variazioni e fluttuazioni del numero di individui in. specie animali conviventi. *Memorie della Regia Accademia Nazionale dei Lincei, Ser. 6*, 2:31–113.

von Elsner-Shak, I. (1985). What is good chamois habitat. In Lovari, S., editor, *The biology and management of mountain ungulates*, pages 71–76. Croom Helm., London, UK.

Wabakken, P., Sand, H., Liberg, O., and Bjärvall, A. (2001). The recovery, distribution, and population dynamics of wolves on the Scandinavian peninsula, 1978-1998. *Canadian Journal of Zoology*, 79:710–725.

Wahlström, L. and Liberg, O. (1995). Patterns of dispersal and seasonal migration in roe deer (*Capreolus capreolus*). *Journal of Zoology (London)*, 235:455–467.

Whittaker, R. H., Levin, S. A., and Root, R. B. (1973). Niche, habitat, and ecotope. *American Naturalist*, 107:321–338.

Whittaker, R. H., Levin, S. A., and Root, R. B. (1975). On the reasons for distinguishing "niche, habitat, and ecotope". *American Naturalist*, 109:479–482.

Wiens, J. A. (1989). Spatial scaling in ecology. *Functional Ecology*, 3:385–397.

Wittmer, H., McLellan, B., and Hovey, F. (2006). Factors influencing variation in site fidelity of woodland caribou (*Rangifer tarandus* caribou) in southeastern British Columbia. *Canadian Journal of Zoology*, 537–545.

Woodroffe, R. (2000). Predators and people: Using human densities to interpret declines of large carnivores. *Animal Conservation*, 3:165–173.

Zaniewski, A. E., Lehmann, A., and Overton, J. M. (2002). Predicting species spatial distributions using presence-only data: A case study of native New Zealand ferns. *Ecological Modelling*, 157:261–280.

Zedrosser, A., Støen, O.-G., Sæbø, S., and Swenson, J. E. (2007). Should i stay or should i go? Natal dispersal in the brown bear. *Animal Behaviour*, 74:369–376.

Zimmermann, F. (2004). *Conservation of the Eurasian lynx (*Lynx lynx*) in a fragmented landscape – Habitat models, dispersal and potential distribution.* PhD thesis, Faculté de biologie et de médecine de l'Université de Lausanne, Département d'Écologie et Évolution, Lausanne.

Zimmermann, F. and Breitenmoser, U. (2002). A distribution model for the Eurasian lynx (*Lynx lynx*) in the Jura mountains, Switzerland Mountains, Switzerland. In Scott, J., Heglund, P., Samson, F., Haufler, J., Morrison, M., Raphael, M., and Wall, B., editors, *Predicting species occurrences: Issues of accuracy and scale*, pages 653–659. Island Press, Covelo, California.

Zwiers, F. and Hegerl, G. (2008). Climate change: Attributing cause and effect. *Nature*, 453:296–297.

Appendixes

Appendix A

Hierarchical habitat selection of brown bears: low order does matter

Jodie Martin[1,2,3,*], Mathieu Basille[1,4], Jonas Kindberg[5], Dominique Allainé[1], Jon E. Swenson[2,6]

[1] Université de Lyon, F-69 000, Lyon ; Université Lyon 1 ; CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, F-69 622, Villeurbanne, France.

[2] Department of Ecology and Natural Resource Management, Norwegian University of Life Science, P.O. Box 5003, NO-1432 Ås, Norway.

[3] Office national de la chasse et de la faune sauvage, F-31 800 Villeneuve de Rivière, France.

[4] Centre for Conservation Biology, Department of Biology; Norwegian University of Science and Technology; NO-7491 Trondheim, Norway.

[5] Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden.

[6] Norwegian Institute for Nature Research; Tungasletta 2, NO-7485 Trondheim, Norway.

[*] Corresponding author: martin@biomserv.univ-lyon1.fr

Submitted to the Journal of Applied Ecology.

"I believe in God, only I spell it Nature."
Frank Lloyd WRIGHT.

Abstract

1. The management and conservation of brown bears Ursus arctos in order to maintain viable populations is challenging due to their large spatial requirements. The Scandinavian brown bear population is now expanding, requiring a better understanding of its requirements at multiple spatial scales.
2. To take into account the inherent hierarchical nature of habitat selection, we analysed habitat selection of female brown bears at two spatial scales: (1) establishment of home ranges and (2) use of habitat within home ranges, using Global Positioning System (GPS) data and a K-select analysis. It has been suggested that the most important limiting factors should be selected at coarser scales.
3. Our hypothesis that females should avoid human structures when they establish their home range was rejected. No consistent pattern of habitat selection or avoidance of human structures was found at this spatial scale.
4. Our hypothesis that females should select habitats within their home range that provide food resources and minimize human-caused disturbance was supported, documenting important habitat selection at this low order of scale.
5. *Synthesis and applications.* This study provided useful results for the future management of brown bear habitats, particularly by revealing a trade-off between use of food-rich open habitats and avoidance of human-caused disturbance at a fine spatial scale. Our results also suggested that human-caused disturbance is more important in affecting brown bear habitat in Sweden than present forestry management alone.

Keywords: *Ursus arctos*, selection orders, spatial scales, human disturbance, trade-off, K-select, forestry, Sweden.

MOST HABITAT USE STUDIES OF LARGE CARNIVORES are conducted at the landscape scale (Mladenoff et al., 1995; Schadt et al., 2002), as recommended by some authors (Noss et al., 1996). However, habitat selection is a hierarchical process and is not necessarily congruent across scales (e.g. Schaefer and Messier, 1995; McLoughlin et al., 2002, 2004; Boyce et al., 2003). Indeed, different key factors may be involved according to the scale considered (Orians and Wittenberger, 1991) and it is particularly important to determine the differences at multiple scales (Thomas and Taylor, 1990; Boyce et al., 2003; Fernandez et al., 2003). Johnson (1980) defined 4 orders of selection conveniently ordered through spatial scales. First-order selection corresponds to the geographic distribution of the species. Second-order selection is the process of home range establishment. Third-order selection concerns usage of various habitats within home ranges. Forth-order selection relates to the selection of items (e.g. food items) within these habitats. Rettie and Messier (2000) suggested that the most important limiting factors should be selected at coarser scales. Then, a limiting factor dominating at a large scale should dominate selective behaviour until it becomes less important that the next most important limiting factor (Rettie and Messier, 2000). Often, decisions at large scale reveal avoidance of predation or disturbance, whereas occurrence of food resources drives selection at a finer scale (see e.g. Rettie and Messier, 2000; May et al., 2006). A complete understanding of the relationship between animals and their environment only can be appreciated after a multiple spatial scale analysis, which can greatly improve the effectiveness of management or conservation strategies.

Because of their large requirements for space (Noss et al., 1996; Schadt et al., 2002), large carnivores are particularly affected by expanding human populations (Breitenmoser, 1998) and their impacts on habitat. Although the principal cause of this decline was persecution by humans, the expansion of anthropogenic infrastructure also has contributed to the reduction and fragmentation of their prime habitats (Saunders et al., 1991; Noss et al., 1996; Breitenmoser, 1998; McKinney, 2002). Several studies have reported the effects of human disturbance on animal behaviour, highlighting a shift in use of space or activity patterns (e.g. George and Crooks, 2006; Gibeau and Herrero, 1998; Johnson et al., 2005; May et al., 2006; Mladenoff et al., 1999; Riley et al., 2003). Populations of large carnivores are now recovering in some areas of Europe, but new problems for management arise, due to interactions between humans and wildlife (Enserink and Vogel, 2006).

After almost becoming extinct due partly to overharvesting and overexploitation of habitat by humans, the Scandinavian brown bear Ursus arctos population started to increase in the early 1900s as bounties were removed and protection policies were adopted (Swenson et al., 1994, 1995, 1998; Linnell et al., 2001a). Today, the population is quite large (2000–3000 individuals throughout Scandinavia Kindberg et al., 2006) and is still expanding. As a consequence, bears are colonizing more human-dominated landscapes, with large networks of roads, higher human densities, and concentrations of recreational cabins (Sweden, 2003). Understanding how Scandinavian brown bears use a human-dominated landscape at different spatial scales is a key component for the management or conservation of this population. Until now, habitat selection by Scandinavian brown bears has been studied only at the population level (1^{st} order selection). In this study, we explored individual habitat selection of Scandinavian female brown bears at finer scales. We used a hierarchical approach to investigate factors affecting (1) the choice of the home range location, and (2) the habitats used within the home range (corresponding to 2^{nd} and 3^{rd} order selection).

McLoughlin et al. (2002) showed that food abundance, which is selected at the scale of home range establishment, was the most limiting factor for brown bears in northern Canada. However, because the environment in the boreal forest of Scandinavia is quite homogeneous regarding to vegetation type composition, we assume that food abundance should be allocated relatively evenly in space and should not be as critical at this level. Moreover, previous studies on Scandinavian brown bears at the landscape level highlighted the avoidance of human structures and a selection towards forested areas and rugged terrain (Katajisto, 2006; Nellemann et al., 2007). We thus hypothesize that (i) bears should primarily avoid human structures when establishing their home range and (ii) decisions regarding use of space at a finer scale should correspond to habitat components providing food resources in abundance. However, as fine-scale avoidance of human-caused disturbance could also occur, we explored in more detail the trade-offs that could rise between habitat selection and the avoidance of anthropogenic structures.

A.2 Material and methods

A.2.1 Study area and bear data

The study was conducted in the southernmost reproductive core area of the Scandinavian brown bear population in Dalarna and Gävleborg coun-

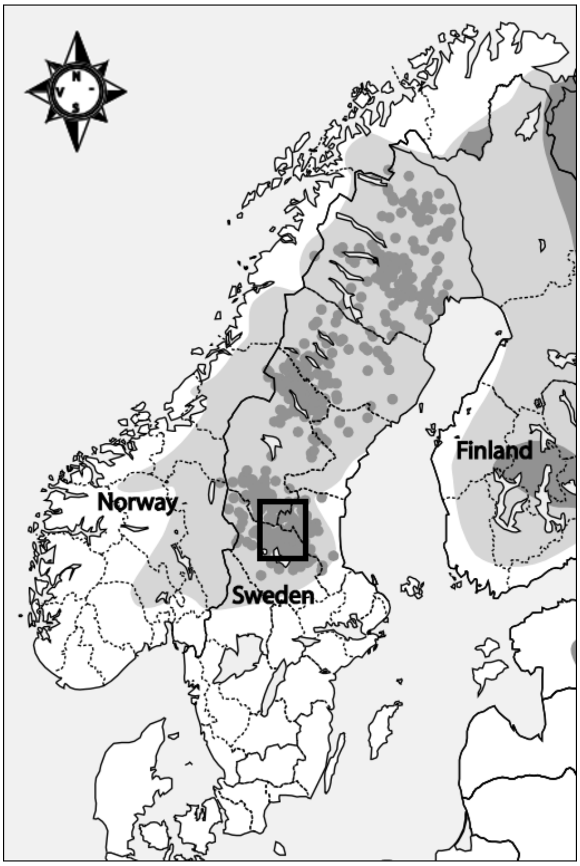


Figure A.1: Distribution of bears (light grey) and location of shot female brown bears (grey dots) in Scandinavia. The black rectangle represents the study area.

ties, south-central Sweden (61°N, 15°E, Fig. A.1). The study area surrounds the home ranges of the GPS-collared females and consists primarily of intensively managed coniferous forest (80%) in patches of different age stands, ranging from clear-cuts to 90-100 years old (Swenson et al., 1999). The remaining area is composed of lakes and bogs. The terrain is hilly and the altitude ranges from 175 m to 725 m, with a southeast-northwest gradient. Settlements are concentrated in the north and in the south and few paved roads with high traffic volumes cross the study area. However, isolated houses and paved and gravel roads with low traffic volumes were distributed throughout the study area.

We used GPS relocations of 11 solitary adult female brown bears (3 to 14 years old) from May to August in 2006 (2 females) and 2007 (9 females). GPS collars were scheduled to take a fix every 30 minutes (i.e. 48 relocations per day). Location errors are inherent with this kind of data and can induce bias in habitat analysis. We therefore eliminated potentially large location errors by data screening based on two-dimensional (2D) and three-dimensional (3D) fixes in relation to the positional dilution of precision (PDOP, see Lewis et al., 2007); 3D positions having a PDOP $>$ 15 and 2D positions having a PDOP $>$ 5 were removed. Due to missing data and large error positions, we obtained 80% of the theoretical number of fixes in average.

A.2.2 Environmental data

The study area was divided into a grid of square pixels (200 × 200 m), which were characterized for 7 variables related to topography, vegetation and human disturbance (Table A.1). A Digital Elevation Model was available for the whole study area (GSD–Höjdkurvor, 25 m ekvidistans Lantmäteriet, Sweden) and used to derive slope. The CORINE Land Cover map (CLC00) was used to define 6 vegetation types. Maps of anthropogenic structures (roads, houses and settlements) were obtained from digital data of Sweden (GSD–Översiktskartan, Lantmäteriet, Sweden) and used to derive distance maps from these structures. ArcView version 3.2a (ESRI Inc., Redlands, California, USA) was used for preliminary preparation of environmental data.

For further analyses, the categorical variable of vegetation classes was converted into 6 binary variables, assessing the presence or absence of each vegetation type using a Hill and Smith (1976) transformation; a weight equal to the proportion of each vegetation type in the area was allocated to these binary variables, so that the overall weight of all vegetation type variables summed to 1, i.e. the weight of a single variable.

A.2.3 Data analyses

Hierarchical habitat selection

We used the K-select analysis (Calenge et al., 2005) to study habitat selection by female brown bears at both spatial scales (establishment of home

Habitat variables		Description	Label
Vegetation classes	Urban (0.5%)	Artificial surfaces like buildings, airports, sport and leisure facilities, green urban areas…	Urban
	Pastures (<1%) Agriculture (1%)	Non-irrigated arable land, pastures, land principally occupied by agriculture…	Agri
	Coniferous forest (66%)	Mainly Scot pines (*Pinus sylvestris*) and Norway Spruce (*Picea abies*)	Conif_F
	Deciduous forest (<1%)	Mainly common birch (*Betula pubescens*)	Decid_F
	Mixed forest (1%)	Mix of coniferous and deciduous forest	Mixed_F
	Regenerating forest (19%)	Young aged stands forests from clear-cut to young forest	Reg_forest
	Wetlands (7%)	Mainly peat bogs (99.5%)	Wet
	Water (5%)	Mainly water bodies (98%)	Lake
Elevation		Digital elevation data in meter	Elev
Slope		Slope in degrees, derived from Digital Elevation model	Slope
Distance to high traffic roads		Linear distance to public roads in km	D_high_traffic
Distance to low traffic roads		Linear distance to private roads in km	D_low_traffic
Distance to houses		Linear distance to houses in km	D_houses
Distance to settlements		Linear distance to human settlements (small villages) in km	D_settle

Table A.1: Description, proportion and label of the different habitat variables used in the analyses of habitat selection of female brown bears in south-central Sweden.

range and within home range). The K-select is based on Hutchinson's concept of ecological niche (Hutchinson, 1957), with a particular emphasis on the marginality (i.e. the deviation of the mean environmental conditions used by individuals from the mean environmental conditions available to them). This is a multivariate analysis similar to a Principal Component Analysis (PCA) based on marginality vectors computed for each animal. It searches for the existence of a general pattern by returning a linear combination of the environmental variables that maximises the mean marginality, i.e. the squared distance between the barycenter of available environment units (EU) and those of used EU, averaged across animals. If all animals have the same pattern of habitat requirements, all their marginality vectors will be oriented in the same direction and the mean marginality explained on the first axis will be the largest. The total inertia explained by the first axis decreases as the variability in individuals' habitat use increases.

At the scale of home range establishment, we considered all the pixels in the study area as available EU to all animals, and pixels within the home ranges were considered as a measure of the utilization (used EU). Home ranges were estimated using the classical method of Minimum Convex Polygon (Mohr, 1947) with the 5% outermost relocations excluded. At the scale of home range use, we considered pixels of each home range as available EU for each corresponding animal. Thus, the availability differed among individuals, and we considered the pixels with GPS relocations as used EU. See Calenge et al. (2005) for details on mathematical procedures of K-select.

Influence of human disturbance

To assess the potential influence of human disturbance on habitat selection variability among individuals, we investigated the relationship between human disturbance and the strength of selection on the variables of interest. An index of human disturbance based on anthropogenic variables (distances to the 4 types of anthropogenic structures) was computed. As the influence of distances to these structures on bears probably may not be linear, we assumed that the potential disturbance was the same above a given threshold. Although bears express a relative tolerance for human-caused disturbance, a review by Linnell et al. (2000) revealed an avoidance of human activity at 1–2 km. Moreover, Swenson et al. (1998) found that brown bears prefer den sites >3 km from villages. We therefore chose a maximum threshold value of 2000 m for distances to roads and houses, and 3000 m for distances to settlements. Above these thresholds, distance values are equal, meaning that the potential disturbance is the same. We finally added all the distance maps to create the human disturbance index. The resulting map provided an index ranging from 200 to 8000, with low values corresponding to high disturbance. For reading convenience, we inverted and then multiplied this index by 1000, so that a low index corresponded to low disturbance. Finally, we calculated the mean disturbance index for each animal's home range by averaging the human disturbance index within each home range.

Selection ratios (Manly et al., 2002) were computed for the key habitat types revealed by the K-select analysis to evaluate the strength of habitat selection on a particular variable. These ratios have been developed for categorical variables defined by several classes, but it can be used as well on continuous variables by defining categories. The selection ratio corresponds to the ratio between the used proportion of a habitat type and the available proportion of this habitat type. It is therefore computed for all individuals for the habitat variables we investigated and were only examined at the fine spatial scale (habitat use within home ranges).

The relationships between the disturbance indices and selection ratios were then investigated using a regression. First, we excluded outliers using a bisquare regression. It is based on an "iterative re-weighted least squares", which gives weights to each observation according to residuals. Therefore, each individual is weighted according to its departure from the model (the more it is likely to be an outlier, the less is the weight) and the regression is re-estimated using these weights until the fit converges. The significance of the general trend was assessed with a Fisher test on the coefficient of determination R^2, after removing outlier individuals that had been identified by the bisquare regression. All analyses were carried out using R (R Development Core Team, 2008) and the package "adehabitat" (Calenge, 2006).

A.3 Results

A.3.1 Establishment of individual home ranges

The first two axes of the K-select accounted for 82% of the individuals' marginality (Fig. A.2a) and were retained in the analysis (Broken-stick method Jackson, 1993). Major anthropogenic structures (distances to settlements and public roads) contributed most to the first axis, whereas the second axis was explained primarily by an elevational gradient (Fig. A.2b; Table A.2). These results mirrored the structure of the study area itself (elevational gradient, isolated public roads and settlements). No obvious common pattern of habitat selection by individuals was apparent. Instead, individual home ranges presented roughly every characteristics of the study area (Fig. A.2c), some far from major anthropogenic structures (females W0624, W0716) and others close to them (females W0422, W0411, W0212, W9806, W0503, W0303), and at high (females W0303, W0217, W9403), medium (females W9806, W0503, W0716, W0624), and low elevations (females W0212, W0410, W0411,W0422).

The landscape composition regarding vegetation types was quite homogeneous, with coniferous forests and regenerating forests covering 67% and 19% of the study area, respectively. Therefore, there was almost no qualitative variation in vegetation composition among the individual home ranges. However, the proportion of each vegetation type was different. Indeed, the quantitative composition of vegetation of each home range was variable, ranging from 10% (W9403) to 39% (W0422) regenerating forest. Bogs (7% of the whole study area) were concentrated in a particular area, at high elevation, where two individuals established (W0303 and W0716) and included about 20% bogs in their home ranges. Isolated houses and private roads were found equally in each home range. At this scale, the bears did not show a strong avoidance of public roads and settlements, because public roads were found in most home ranges and some home ranges were located around settlements (W0212, W0411).

A.3.2 Use of habitats within home ranges

The first two axes accounted for 68% of the individuals' marginality (Fig. A.3a) and were retained in the analysis (Broken-stick method Jackson,

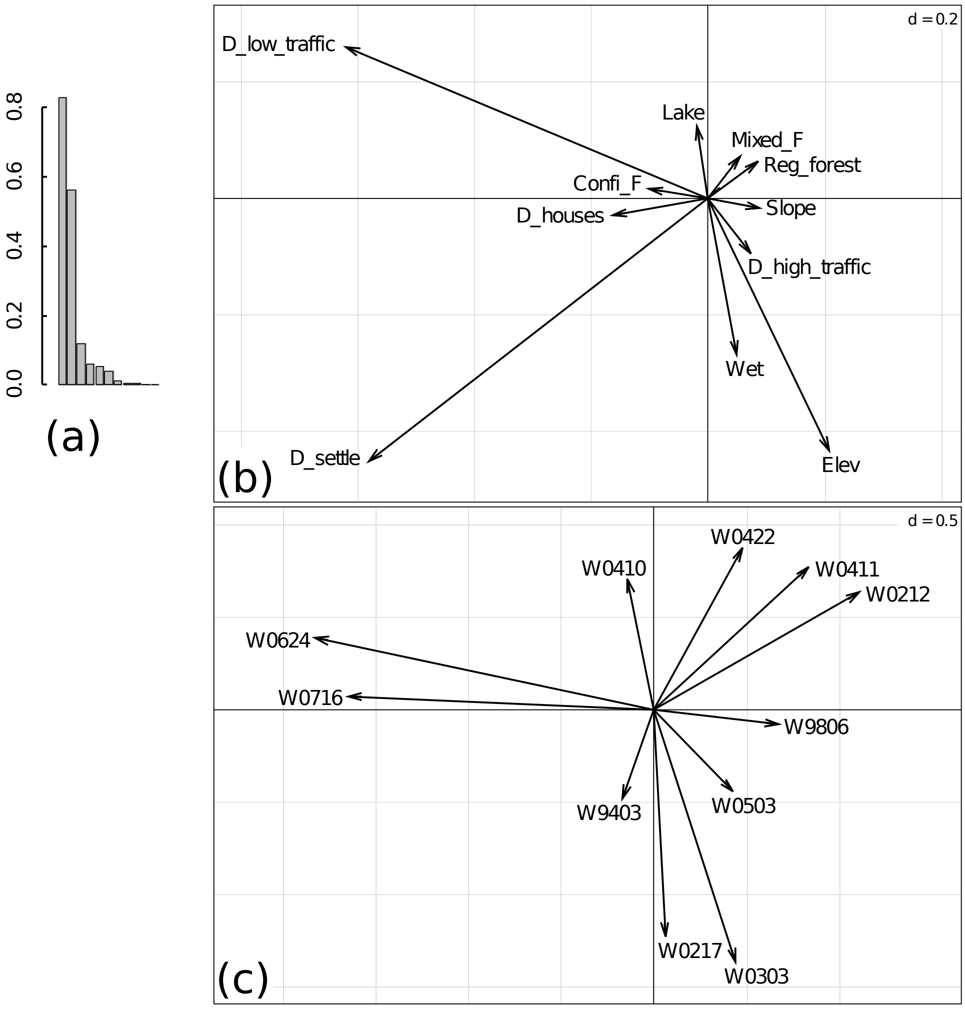


Figure A.2: Habitat selection of female brown bears in south-central Sweden at the 2^{nd} order of study. (a) Bar chart of the eigenvalues of the K-select, measuring the mean marginality explained by each factorial axis. (b) Variable loadings on the first two factorial axes. The projection of the variables on the factorial axes gives a representation of their contribution to these axes. (c) The individuals' marginality vectors on the first factorial plane. The projection of individuals on the factorial plan allows the interpretation of the habitat selection by these individuals; the longer the arrow, the stronger the marginality. As availability is the same for all animals, the arrows are centered on the origin of the factorial plane.

Habitat variables	2^{nd} order		3^{rd} order	
	Axis 1	Axis 2	Axis 1	Axis 2
Elevation	0.207	-0.432	-0.118	0.085
Slope	0.088	-0.017	-0.362	-0.004
Urban	0.019	0.019	0.017	-0.016
Pastures	-0.001	0.001	0.000	0.000
Agriculture	-0.002	0.015	-0.069	0.117
Deciduous forest	-0.000	0.000	0.000	0.000
Coniferous forest	-0.102	0.015	0.097	0.041
Mixed forest	0.054	0.072	0.000	0.015
Regen. forest	0.085	0.063	-0.356	-0.035
Wetlands	0.049	-0.268	0.221	0.098
Water	-0.019	0.122	0.143	-0.175
Dist. high traffic roads	-0.621	0.259	-0.022	0.016
Dist. low traffic roads	0.072	-0.095	-0.044	-0.111
Dist. houses	-0.163	-0.028	0.003	-0.187
Dist. settlements	-0.581	-0.450	-0.003	0.031

Table A.2: Scores of variables on the two axes of the K-select analyses regarding habitat selection of female brown bears in south-central Sweden. The higher the absolute value, the higher the contribution of the variable on the axis.

1993). In contrast to the level of home range establishment, there was a common pattern of habitat selection within home ranges. All females selected slopes and young forests, and avoided bogs (Fig. A.3c, d). Although there was a strong common pattern of selection, variability between individuals was well illustrated by the K-select analysis. W0411 and W0624 selected areas at high elevations but close to houses and private roads, whereas W9403 selected areas at low elevations but far from houses and private roads. The other individuals did not particularly select for these features.

It is important to note that there was no correlation between slope and regenerating forest in the overall study area (r = 0.09), meaning that the choices for slope and regenerating forest were independent. Additionally, there was no clear difference in slope according to vegetation type (regenerating forest compared to forested areas, Fig. A.4). Thus, female bears did not seem to select steeper slopes when in regenerating forest than in older forest.

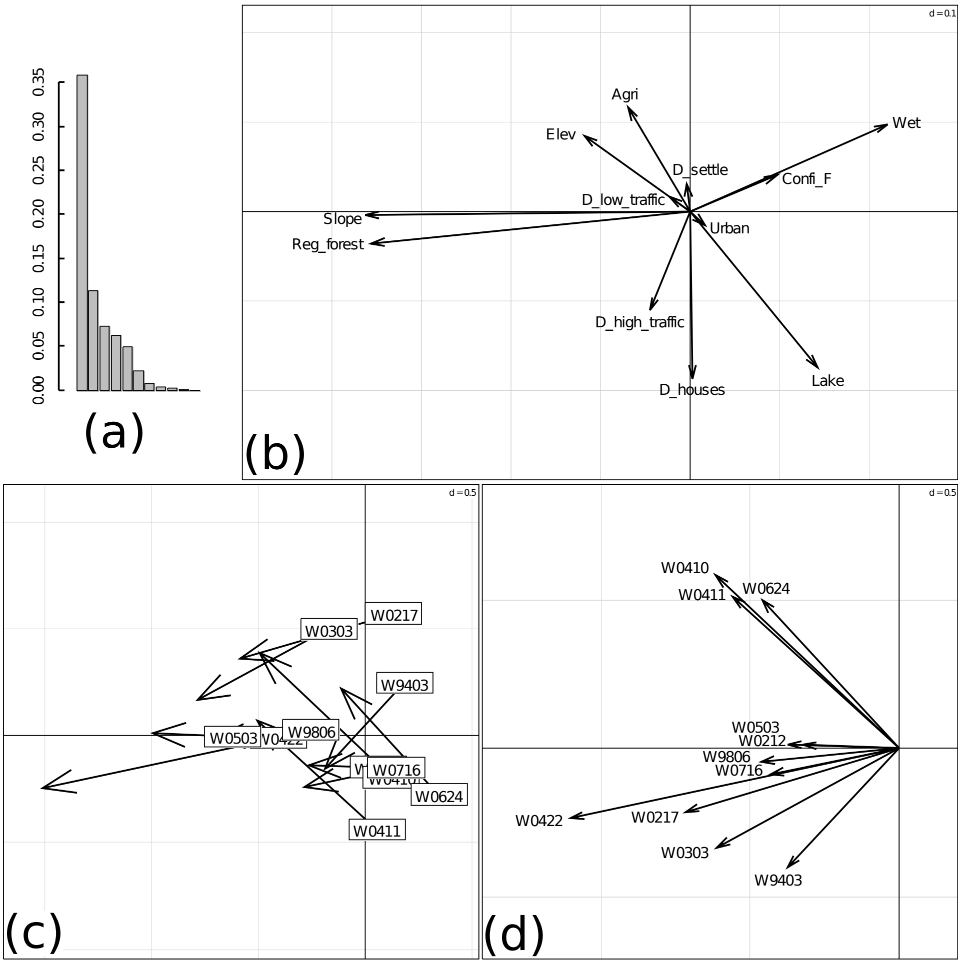


Figure A.3: Habitat selection of female brown bears in south-central Sweden at the 3^{rd} order of study. (a) bar chart of the eigenvalues of the K-select, measuring the mean marginality explained by each factorial axis; (b) variable loadings on the first two factorial axes; (c) the individuals' marginality vectors on the first factorial plane, labels corresponding to the average habitat availability; (d) the individuals' marginality vectors after re-centering each individual habitat availability.

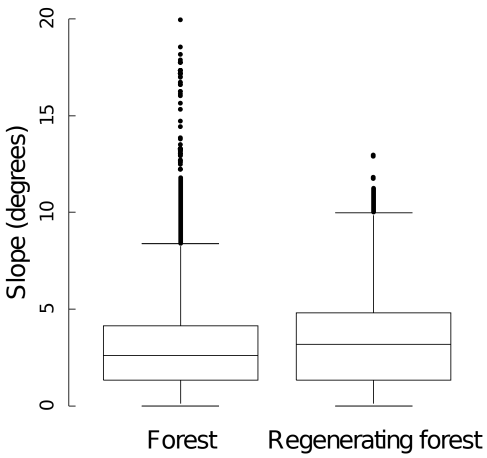


Figure A.4: Distribution of relocations of female brown bears on slopes according to the vegetation type.

A.3.3 Influence of human disturbance

There was a positive relationship between the mean human disturbance index within the home range and the strength of selection for slopes greater than 7% (Fig. A.5), although one individual (W0422) exhibited a very strong selection of slopes and had a moderate disturbance index. The weight attributed by the bi-square regression to this individual was almost 0, whereas the other individuals had weights around 1. Therefore, we removed this outlier individual before performing the regression test and found that the higher the human-caused disturbance in a home range, the more individuals selected slopes ($R^2 = 0.8$, $P < 0.001$).

On the other hand, the relationship between the human disturbance index within a home range and the strength of selection of regenerating forest was negative (Fig. A.6). Again, one individual (W0217) showed a strong selection for regenerating forest with a moderate mean disturbance in its home range. For this individual, the weight estimated by the bi-square regression was around 0, contrary to other individuals (weights of 1). We therefore removed this individual before carrying the regression test and found that the higher the human-caused disturbance in the home range, the more individuals selected regenerating forests ($R^2 = 0.8$, $P < 0.001$).

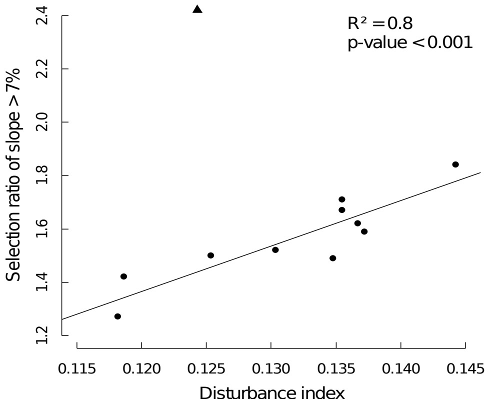


Figure A.5: Relationship between the mean human disturbance index of individual female brown bear home ranges in south-central Sweden and selection ratios of slopes $> 7\%$ for all individuals (one point per individual). As slope was primary a continuous variable, we converted it into a categorical variable with 2 classes: slopes $\leq 7\%$ and slopes $> 7\%$. The black line corresponds to the regression fit without outlier individuals (triangle).

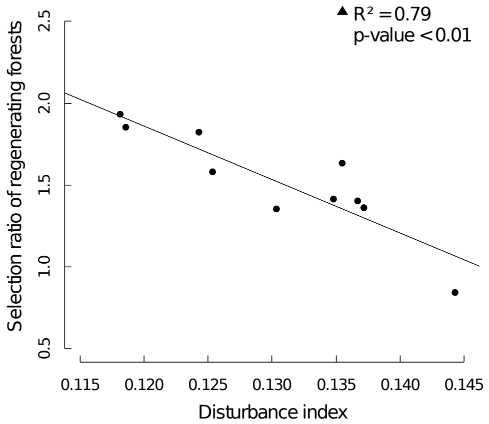


Figure A.6: Relationship between the mean human disturbance index of individual female brown bear home ranges in south-central Sweden and selection ratios of regenerating forests for all individuals (one point per individual). The black line corresponds to the regression fit without outlier individuals (triangle).

A.4 Discussion

Our hypotheses were that bears should (i) select undisturbed areas while establishing their home range and (ii) select habitat providing food items inside their home range. Our first hypothesis clearly was rejected, whereas the second one was supported. Rettie and Messier (2000) suggested that limiting factors that potentially can reduce individual fitness should drive selection at coarser scales. They proposed a direct relationship between these limiting factors and the selection level of individuals, suggesting a continuum of scales at which these factors should impact individual fitness differentially. Our results do not agree completely with this hypothesis. Indeed, female brown bears exhibit an avoidance of human structures at the landscape scale (Katajisto, 2006) and at a fine scale (within home ranges), but no particular pattern of avoidance of these structures at an intermediate scale, that of home range establishment.

At the level of home range establishment, the bears showed no general pattern of habitat selection of the variables we considered and did not particularly avoid anthropogenic structures, such as public roads or human settlements. Instead, individual home ranges seemed to be distributed evenly throughout the study area and were composed of various proportions of habitat types. This apparent lack of habitat choice can be explained partly by the social behaviour of female brown bears. The home ranges of unrelated females show little overlap (Støen et al., 2005), suggesting a form of territorial behaviour. Beckmann and Berger (2003) also highlighted an ideal-despotic distribution of black bears *Ursus americanus*, with larger males precluding females and smaller males from areas with abundant food resources. The same mechanism may occur among female brown bears, subordinate females being forced to establish in less suitable habitats by dominant females, i.e. in more disturbed areas. This phenomenon has been demonstrated in social species like wild chimpanzees Pan troglodytes, where dominant females compete with subordinates, forcing them to settle in less suitable areas (Murray et al., 2007). In addition, female brown bears are often philopatric (Støen et al., 2005), settling close to or within their mother's home range, although subdominant siblings are sometimes forced to disperse due to competition for philopatry (Zedrosser et al., 2007), but do not move far from their natal area (27 km in average Støen et al., 2006). Philopatry could provide benefits for females establishing matrilineal assemblages that may partly supplant habitat choice behaviour.

At the scale of habitat selection within the home ranges, a strong common pattern arose towards a selection of slopes and regenerating forests. The use of regenerating forests, including clear-cuts, has previously been reported in studies of grizzly bears in Alberta (Nielsen et al., 2004a). Several important food items in the Scandinavian brown bears' diet occur in these forest stands. In particular, several species of ants prefer regenerating forest stands (Rolstad et al., 1998, 2000), and ants represent a significant part of the food intake (Swenson et al., 1999) by bears. Clear-cuts and regenerating forests also offer a more diverse and abundant herbaceous material (Apps et al., 2004; Nielsen et al., 2004b) and clear-cutting favours grasses and herbs (Rolstad et al., 2000), which constitute 12-18% of diet of Scandinavian brown bears (Dahle et al., 1998). Moreover, crowberries *Empetrum nigrum*, a major species of berries eaten by bears, grow in open areas after clear-cutting (Mallik, 2003). Finally, moose (*Alces alces*) forage preferably in regenerating forest stands and clear- cuts (Edenius et al., 2002; Nikula et al., 2004; Cassing et al., 2006), as they provide the greater available biomass for browsing (Kalen and Bergquist, 2004). Moose calves represent an important food item of the Scandinavian brown bears diet (14–30% in spring Dahle et al., 1998) and bears eat 26% of calves born on the study area each year (Swenson et al., 2007).

Several studies at different scales have documented the selection of steep slopes or rugged terrain by bears (Apps et al., 2004; Nellemann et al., 2007). Slopes can provide some security to the bears, by guaranteeing a greater visibility and the wind would more easily deliver olfactory information (S. Brunberg, Pers. Com.). Steeper slopes can also provide higher food availability, especially within clear-cuts (Nielsen et al., 2004a,b), anthills are more abundant on dry warm slopes (Nielsen et al., 2004b), and terrain ruggedness is also known to positively influence the amount of herbaceous food resources (Nellemann and Thomsen, 1994). Therefore, we would expect that female brown bears would prefer regenerating forest in steeper areas. However, they used slopes in forests and regenerating forests equally (Fig. A.4). Hence, bears seemed to use slopes, regardless of vegetation type, and then regenerating forests were used for their food resources.

Selection ratios	Individuals					
	W0212	W0411	W0217	W0313	W0410	W0422
Reg_forest	0.84	1.36	2.56	1.63	1.58	1.82
Slopes > 7%	1.84	1.59	1.67	1.71	1.50	2.42

Selection ratios	Individuals				
	W0503	W0624	W0716	W9403	W9806
Reg_forest	1.35	1.85	1.93	1.41	1.40
Slopes > 7%	1.52	1.42	1.27	1.49	1.62

Table A.3: Selection ratios of each female brown bear in south-central Sweden regarding regenerating forest and slopes > 7%.

Although all bears seemed to select slopes and regenerating forests, there was a great variability in the strength of this selection among individuals (Table A.3). The degree of human disturbance in individual home ranges appeared to be a potential predictor of habitat use behaviour (Fig. A.5 & A.6). Bears used slopes even more when their home range was located in an area with higher human disturbance, probably due to the need for increased security (Nielsen et al., 2004b). On the contrary, the strength of selection for regenerating forests was lower when human disturbance in the home range was higher. This is may be due to good visibility in artificially regenerating forests and clear-cuts that hence provide less security for bears. Nielsen et al. (2004a) observed that grizzly bears used clear-cuts during crepuscular and night periods, which also suggests that they provided little security during the day. Thus, females that may not have the choice to establish their home range in less disturbed areas may compensate for this by adjusting their habitat selection at a finer spatial scale, showing less selection for young forests and greater selection for more slopes than females located in less disturbed areas. At this scale, they seem to trade food intake in favour of more secure, forested areas on steeper slopes, before less secure areas with higher food resources.

In this context, it is important to note that, at the individual scale, the reason older forests seemed to be used in proportion of their availability was because of their high availability on the study area and in home ranges. However, these forests are probably important for security and thermoregulation during resting periods. Old forests also provide some important food items like blueberries Vaccinium myrtillus, which are negatively affected by clear-cutting (Kind, 2001). Moe et al. (2007) showed that habitat selection must take into account diel behaviour of individuals. Based on few GPS-collared female bears, they documented a trend for females to select tall conifers for resting and young forests for foraging. A detailed analysis of diel activity pattern might indeed help us to better understand how bears use their habitat in relation to their activity, using more detailed environmental characteristics and associated with movement data. Sequential autocorrelation of relocations can indeed provide precise information on movement and activity of individuals that can be incorporated into habitat analyses (Martin et al., 2008).

Forest management affects habitat characteristics regarding both food resources and canopy cover (Edenius et al., 2002). In several North American studies, logged forests and clear-cuts were clearly avoided by bears, but natural open areas were used for foraging (Apps et al., 2004). In our study area, the landscape is mainly composed of managed forests and clearcuts are usually more common than natural openings, except for bogs, which the bears we studied seemed to avoid. However, our results suggest that present Swedish forestry methods are a less important factor regarding the habitat selection of bears than human disturbance. Actually, modern logging creates new habitats that can substitute the loss of meadow and pasture (Rolstad et al., 2000), and we suggest that maintaining clear-cuts and regenerating forests can help provide a high occurrence of several important bear food resources. However, this potential benefit is at the cost of an increased human disturbance through the development of a large road network.

The efficient management or protection of large carnivores requires a thorough understanding of how they use their environment at different scales. Conclusions at one particular scale may not always be valid at other scales (Wiens, 1989). This is the case for Scandinavian brown bears, not only regarding vegetation type, but also human-caused disturbance. Anthropogenic structures such as roads, houses and small human settlements are found within home ranges, and impact habitat selection by bears negatively at a fine spatial scale, resulting in trade-offs between food and security. Our results also suggested that present forestry management in Sweden can affect brown bear habitat positively.

Acknowledgments

The Scandinavian Brown Bear Research Project was funded by the Swedish Environmental Protection Agency, the Norwegian Directorate for Nature Management, the Swedish Association for Hunting and Wildlife Management, WWF Sweden and the Research Council of Norway. We are grateful to Clément Calenge (Université Lyon 1) and Ivar Herfindal for their valuable help on the methodological part.

Appendix B

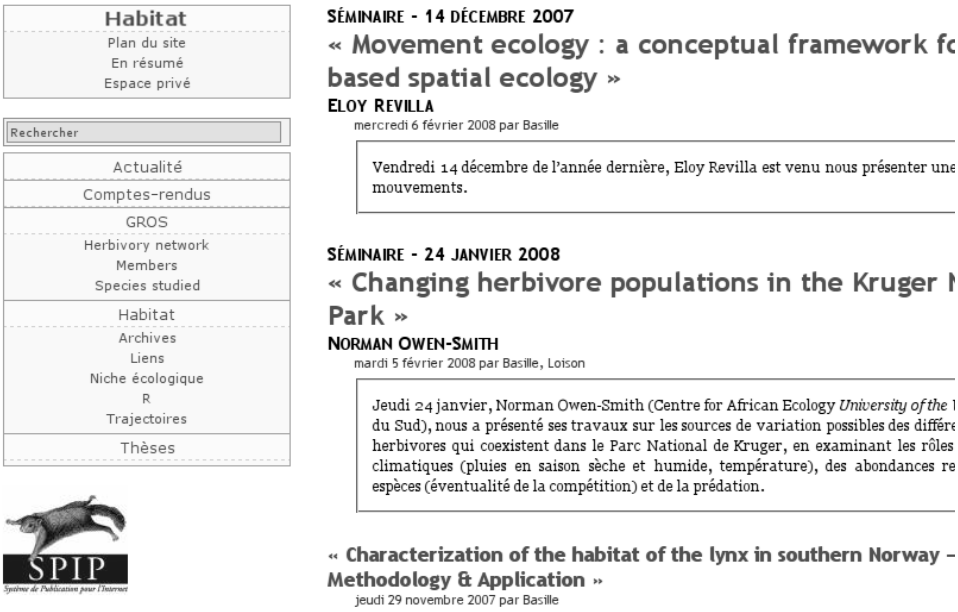
The Habitat network

> "I don't want to achieve immortality through my work. I want to achieve it through not dying."
>
> Woody Allen.

In parallel to my regular scientific work, I've been animating and administrating an informal group, focused on spatial ecology and in particular what may be called *habitat ecology* (see **Chapter 1**). Essentially, every person interested in habitat selection, trajectories, or any other spatial analyses of animals locations may be concerned by this group.

This structure was known under different names, GRHA (*Groupe de Recherche HAbitat*), GRhAS (*Groupe de Recherche on Habitat Selection*), after the example of the GROS (Groupe de Recherche sur les Ongulés Sauvages), another group born from the team's work in ungulate research. Anyone familiar with french will certainly recognize the reference in these acronyms ("gros" and "gras" are two synonymous for the English word *fat*, when it's relative to people) to the British humor, in the LARG (*Large Animal Research Group*) of Tim Clutton-Brock. Anyway, as the GRHA or GRhAS was an informal structure, I've been used to refer to it as the *Habitat network*.

The initial aim of this group, as formulated about three years ago, was to build a link between biologists and statisticians, based on common resources (data) and shared knowledge and problematics. With that objective in mind, the group was open for anybody (however, the language used in communications limited this largely to French-speaking people). The group quickly grew out to 36 members throughout France. However, with time, the group contracted again to an active hardcore, composed



SÉMINAIRE - 14 DÉCEMBRE 2007

« Movement ecology : a conceptual framework f
based spatial ecology »

ELOY REVILLA

mercredi 6 février 2008 par Basille

Vendredi 14 décembre de l'année dernière, Eloy Revilla est venu nous présenter une
mouvements.

SÉMINAIRE - 24 JANVIER 2008

« Changing herbivore populations in the Kruger
Park »

NORMAN OWEN-SMITH

mardi 5 février 2008 par Basille, Loison

Jeudi 24 janvier, Norman Owen-Smith (Centre for African Ecology *University of the*
du Sud), nous a présenté ses travaux sur les sources de variation possibles des différe
herbivores qui coexistent dans le Parc National de Kruger, en examinant les rôles
climatiques (pluies en saison sèche et humide, température), des abondances re
espèces (éventualité de la compétition) et de la prédation.

« Characterization of the habitat of the lynx in southern Norway –
Methodology & Application »

jeudi 29 novembre 2007 par Basille

Figure B.1: A screenshot of the home page of the Habitat website.

mainly of researchers based in the Laboratoire de Biométrie et Biologie Évolutive (LBBE).

The visible aspect of this work took the form of a website [1], currently hosted by the LBBE (see Fig. B.1). The website functions as a centralizing unit for reports, information, etc. related to the group. To date, the website hosts around 50 articles, more than half of them are meetings or seminar reports. Our regular meetings formed the occasion to discuss papers, or to present new projects or more or less advanced research results. A more specialized subunit was dedicated to the study of trajectories. Occasionally, our team also hosted external researchers, and their presentations formed also a valuable contribution to the website's content. Finally, the website functions as a portal to pass on information and announcements for national and international meetings related to spatial ecology, students theses, and new functions and tutorials for the `adehabitat` package. On top of this a mailing-list was created to easily reach all members.

Even if the initial goals got reduced over time, the group created a great opportunity to bring up new ideas and share thoughts, formalize new hypotheses, discuss papers, etc. In this sense, the group reached

---

[1] http://biomserv.univ-lyon1.fr/spip_habitat/, access controlled with a password.

one of its major objectives, as to build the foundations for an increased exchange between different researchers. From a personal perspective, it greatly helped me to put my own approach to science into practice, based on the sharing of ideas during regular meetings (remember the catalyst). This allowed me to broaden my perspectives and was certainly a great contribution to the thesis process as a whole.

“Calvin and Hobbes”, by Bill Watterson, *It’s A Magical World*, 1995.

Picture credit: Manuela Panzacchi.