HAL
open science

Variabilité fonctionnelle intraspécifique: quantification in situ et implications dans une vallée alpine

Cécile H. Albert

To cite this version:

Cécile H. Albert. Variabilité fonctionnelle intraspécifique: quantification in situ et implications dans une vallée alpine. Ecologie, Environnement. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT : . tel-00450219

HAL Id: tel-00450219

https://theses.hal.science/tel-00450219

Submitted on 25 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ JOSEPH FOURIER- GRENOBLE

ÉCOLE DOCTORALE DE CHIMIE ET SCIENCES DU VIVANT

THÈSE

présentée par

Cécile ALBERT

pour obtenir le grade de

DOCTEUR

Spécialité : BIODIVERSITÉ - ÉCOLOGIE - ENVIRONNEMENT

VARIABILITÉ FONCTIONNELLE INTRASPÉCIFIQUE: QUANTIFICATION *IN SITU* ET IMPLICATIONS DANS UNE VALLÉE ALPINE

Soutenue le 4 décembre 2009 devant le jury suivant

M. David Mouillot, Professeur, Université de Montpellier	Président du jury
M. Jean-François Soussana, Directeur de recherches, INRA	Rapporteur
M. Niklaus Zimmermann, Directeur de recherches, WSL Zürich	Rapporteur
M. Jérome Chave, Directeur de recherches, CNRS	Examinateur
Mme Sandra Lavorel, Directeur de recherches, CNRS	Directrice de thèse
M. Wilfried Thuiller, Chargé de recherches, CNRS	Directeur de thèse

Laboratoire d'Ecologie Alpine UMR 5553 – 2233 Rue de la Piscine 38041 Grenoble

UNIVERSITÉ JOSEPH FOURIER- GRENOBLE

ÉCOLE DOCTORALE DE CHIMIE ET SCIENCES DU VIVANT

THÈSE

présentée par

Cécile ALBERT

pour obtenir le grade de

DOCTEUR

Spécialité : BIODIVERSITÉ - ÉCOLOGIE - ENVIRONNEMENT

VARIABILITÉ FONCTIONNELLE INTRASPÉCIFIQUE: QUANTIFICATION *IN SITU* ET IMPLICATIONS DANS UNE VALLÉE ALPINE

Soutenue le 4 décembre 2009 devant le jury suivant

M. David Mouillot, Professeur, Université de Montpellier	Président du jury
M. Jean-François Soussana, Directeur de recherches, INRA	Rapporteur
M. Niklaus Zimmermann, Directeur de recherches, WSL Zürich	Rapporteur
M. Jérome Chave, Directeur de recherches, CNRS	Examinateur
Mme Sandra Lavorel, Directeur de recherches, CNRS	Directrice de thèse
M. Wilfried Thuiller, Chargé de recherches, CNRS	Directeur de thèse

Laboratoire d'Ecologie Alpine UMR 5553 – 2233 Rue de la Piscine 38041 Grenoble

'L'homme se trouve devant deux destins possibles : périr dans son berceau de sa propre main, de son propre génie, de sa propre stupidité, ou s'élancer, pour l'éternité du temps, vers l'infini de l'espace, et y répandre la vie délivrée de la nécessité de l'assassinat.
Le choix est pour demain.
Il est peut être déjà fait.'

R. Barjavel *La faim du tigre, 1966*

Avant-propos

Cette étude a été conduite au Laboratoire d'Ecologie Alpine de Grenoble (UMR 5553). Elle a été financée par une bourse de thèse ministérielle AMX et grâce à divers projets français et européens :

ANR Project QDiv: *Quantifying the effects of global environmental change on terrestrial plant diversity* 2006-2008. Granted by the Agence Nationale pour la Recherche (ANR-CNRS) and coordinated by Prof. Paul Leadley (Orsay University, France)

European project MACIS : *Minimisation of and Adaptation to Climate change: Impacts on biodiversity* 2007-2008 granted by European Commission and leaded by Dr. I. Kuhn (Centre for Environmental Research Leipzig-Halle, Germany)

ANR Project DIVERSITALP: *Forecasting the impacts of global changes on French Alpine Flora: distribution of specific, functional and phylogenetic diversities, simulations and conservation strategies* 2008-2011. Granted by the Agence Nationale pour la Recherche (ANR-CNRS) and coordinated by Dr. Wilfried Thuiller (Grenoble, France)

LECA Interface project "Vertical Ecology": *Past and present distribution, dynamic and functioning of high altitude biotic communities.* Granted by the Laboratoire d'ecologie alpine & le Centre National pour la Recherche Scientifique (LECA-CNRS) coordinated by Dr. Patrick Saccone, Dr. S. Ibanez and C. Albert.

Ce document est une synthèse en français des travaux de recherche effectués durant ma thèse. Il se compose d'une introduction, de quatre chapitres et d'une conclusion. Les publications issues ou en lien avec ce travail de thèse sont listées sur la page suivante, elles sont citées dans la suite du document sous la forme « Article X ». Celles qui sont en lien direct avec le manuscrit sont regroupées dans les Annexes.

Remerciements

Grenoble, le 9 Décembre 2009

Commencer à écrire ces lignes… C'est une page qui se tourne, la concrétisation d'un rêve : « quand je serai grande je serai chercheuse », l'aboutissement des études post-bac : 9 ans déjà !, la fin d'une aventure scientifique dont vous pourrez lire le récit au cours des pages qui suivent et c'est aussi réaliser qu'une thèse n'est pas le produit d'un individu, mais bien également le résultat d'une aventure humaine.

C'est pourquoi en premier lieu je souhaite remercier mes professeurs de majeur éco-sciences, ***Jacques Istas***, ***Denis Couvet*** *&* ***Pierre-Henri Gouyon*** *qui m'ont fait découvrir l'écologie et m'ont conseillée dans mes choix.*

Quant à mes débuts en écologie, je souhaite également remercier vivement ***Sandra Lavorel*** *de m'avoir acceptée en stage de Master 1 – malgré une lettre de motivation qui pourrait certainement servir d'exemple à ne pas suivre –, puis de m'avoir orientée sur un sujet de Master 2 passionnant et de m'avoir recommandée à Wilfried qui arrivait alors au LECA et enfin de m'avoir encadrée, conseillée et encouragée durant ma thèse.*

De fil en aiguille, j'ai donc été amenée à travailler en Master 2 puis en thèse avec ***Wilfried Thuiller****. Wilfried, je crois que je me souviendrai toujours de la première fois que je suis entrée dans ton bureau pour te parler d'un papier de Levine et al. sur les patrons d'invasibilité que je devais commenter, aujourd'hui ça me parait complètement incongru… Encadrer, conseiller, être disponible et à l'écoute, encourager tu as su – ou parfois appris à – le faire au fil des ans. Tu as aussi réussi à être très exigent, mais pas trop, tout en t'inquiétant parfois de voir que je travaillais peut-être plus que j'en avais l'air. Tour à tour compagnon de course, « élève » en escalade, collègue, modèle ou chef, tu t'es adapté aux situations et a réussi à tempérer les passages houleux d'une première année de terrain difficile, tout en me faisant confiance et en me considérant en égal. Merci ! Je suis heureuse que tu m'aies encadrée, d'avoir eu l'honneur –et les inconvénients ?☺- d'être ta première étudiante et je suis fière aussi d'une certaine manière d'avoir « amené sur le terrain » un grand modélisateur comme toi.*

Mes remerciements vont également aux membres du jury, ***Jérôme Chave***, ***David Mouillot***, ***Jean-François Soussana*** *&* ***Niklaus Zimmermann*** *qui ont accepté de consacrer du temps à la lecture et à l'évaluation de ce manuscrit et de sa présentation orale.*
Un grand merci à toutes les personnes qui m'ont aidée de prêt ou de loin dans mon travail de thèse ; en particulier ***Paul Leadley***, ***Antoine Guisan***, ***Sébastien Lavergne***,

Stefan Dullinger, pour leur participation à mes comités de thèse et leurs conseils; également Gilles Nigel Yoccoz pour son aide en statistique et pour son accueil à Tromsö, Ian davies, Georges Kunstler & Anne Dubuis pour les nombreuses interactions que j'ai pu avoir avec eux ; également la Station alpine – ou RDPL 55- et en particulier Serge Aubert, Rolland Douzet – aussi pour les sorties musique, théâtre, ciné-, Richard Hurstel, Pascal Salze, Karl Grigulis & Christophe Perrier pour leur aide botanique et la vie au Lautaret ; également le Parc National des Ecrins et en particulier Richard Bonnet, Hervé Cortot et Pierre Salomez ; également toute l'équipe de soutien du LECA, Olivier Lontin, Florence Sagnimorte, Kim Pla, Joëlle Claustre, Gwenaëlle Clément, Marie-Pascale Colace, Geneviève Girard, Bruno Courchaud. Je remercie aussi tout particulièrement les personnes m'ayant aidée sur le terrain, du grand beau temps à la catastrophe météo. Spécialement merci à Patrick Saccone sans qui mon travail de terrain aurait bien pu s'arrêter tout d'un coup après de trop nombreuses journées difficiles… Merci aussi à Alex Soudan & Julien Icard pour leur aide et de m'avoir fait expérimenter la dure réalité d'être chef. Merci à Amélie Saillard pour sa confiance et son travail acharné. Merci à Florian Boucher pour son amitié, son calme et de grands moments de rigolade.

Au fur et à mesure que ma thèse avançait, une nouvelle équipe s'est constituée au sein du LECA et je suis heureuse d'en faire partie encore pour quelques temps et de pouvoir collaborer avec vous toutes (et tous …). Je remercie en particulier Laure, Tamara, Isa, Cristina, Maria, Martha & Fra pour leur aide et soutien et pour une vie hyperactive en dehors du labo. Merci aussi au bureau 111 ancienne et nouvelle générations : Pierrot, Nico G. -et la TDR-, Flo, Flore, Fabrice – …dommage que tu aies fui le bureau, j'aimais bien les ménages de printemps et les discussions bougons –, Abde, Aurélie T., Hanna, Jonath – enfin de retour ! Et bientôt dans le ED –, Nico L., Pénélopi – heureusement qu'il était grand ce lit et que la voisine du dessus n'a pas tapé des pieds un matin de plus…. –. Merci à TDE, Irène, Jean-Christophe, Philippe, Fabien – ja ja- et aux tarés du Lautaret 2005-2006-2007-2008-2009 ; Yohann, Sandra B., Katlyn, Seb le trolle, David, Mason & Pierre-Raphaël –la « taupe » team-, Fanny la chiastoquète, Marion l'entomo, Rémi L. -Riri-, Florent -Fifi-, Loïc -Loulou-, Marika, Simon, Maxime, Sandrine, Manue, Eric, Coline, Harold – « câlin chalet » –, Louise B.. Merci aussi à l'ensemble du LECA et ses doctorants et intermittents, Alice, Angélique, Babar, Béné, Guillaume L., Guillaume T., Hamid, Lucie, Marco, Margot, Mickael, Osgur, Rodolphe, Said, Sophie, Tarafa, Tony, et tous ceux que j'oublie !

Merci les amis, Grenoblois, Parisiens, grimpeurs, alpinistes, boxeurs… Adil, Alex C., Alexandra, Bery, Caro, Christophe G., Claire C., Cyril, Florent, François A., François L., Hervé, Jérémie – qui me racontait sa thèse il y a des années lors des séances d'athlé à Charléty–, Ju, Jue, Julia, J.B., Kolia, Lolo, Ludo, Marie, Manu, Maryam, Mike, MTP, Patou, Pierre M., Pierre-Marie, Renaud, Sandra la Québécoise de Trömso, Sam T., Seb, Sim, Sönke, Vivi.

*Et puis bien sur un grand merci à mes **parents** -qui m'ont donné cette chance de choisir ma vie et ma voie et qui m'ont appris que l'on pouvait concilier un travail passionnant avec une vie de famille, du sport & de la culture-, à mes petits frères **Nicolas** et **Sylvain** - qu'ils trouvent eux aussi le chemin qui leur plait-, à **Mamie** –qui m'a fait le plaisir de venir à Grenoble pour ma thèse -, à **Xin** et aux **Z'albert** et **Leguets**.*

*Enfin un très grand merci à **Rémi** qui a supporté le poids de cette thèse depuis trois ans, qui a été un soutien infaillible et un relecteur patient de ce manuscrit ; et surtout qui m'a permis de m'évader, dans les montagnes, sur les falaises en France et ailleurs et avec qui se construisent à chaque instant les rêves et projets pour l'avenir.*

Finir d'écrire cette page c'est aussi le début d'une nouvelle aventure, scientifique et personnelle, ici et ailleurs avec toujours plus d'enthousiasme.

Juillet 2007 : Androsace helvétique au col du Galibier (2600 m). Photo de la Station Alpine Joseph Fourier

‘*À mes* professeurs de gymnastique, *à mes entraîneurs, à mes guides de haute montagne, qui m'ont appris à penser*’

Michel Serre, *philosophe & membre de l'académie française.*

Liste des publications publiées et en préparation

Article 1
Albert C.H., Thuiller W., Lavorel S., Davies I., Garbolino E., *Land-use change and subalpine tree dynamics: colonization of Larix decidua in French subalpine grasslands*. Journal of Applied Ecology, 2008, 45: 659-669.

Article 2
Thuiller W., **Albert C**., Araujo M.B. , Berry P.M., Cabeza M., Guisan A., Hickler T., Midgley G. , Paterson J., Shurr F.M., Sykes M.T., Zimermann N.E., *Predicting global change impacts on plant species' distributions: Future challenges.* Perspectives in Plant Ecology, Evolution and Systematic, 2008, 9 (3-4):137-152.

Article 3
Albert C.H. & Thuiller W., *Favourability functions versus probability of presence: advantages and misuses.* Ecography, 2008, 31: 417-422.

Article 4
Gross N., Robson, M.T. Lavorel, S., **Albert C**., Le Bagousse Pinguet Y., Guillemin R., *Plant response traits mediate the effects of subalpine grasslands on soil moisture.* New phytologist, 2008, 180: 652-662.

Article 5
Midgley G. F., Davies I. D., **Albert C. H.,** Altwegg R., Hannah L., Hughes G. O., Ries L. P. & Thuiller W. *BioMove – an integrated platform simulating the dynamic response of species to environmental change*. Sous presse à Ecography

Article 6
Thuiller W.*, **Albert C.H.*,** Dubuis A., Randin C., Guisan A., *Variation in habitat suitability models does not always relate to variation in species' plant functional traits.* Sous presse à Biology Letters: doi:10.1098/rsbl.2009.0669). *Contributions équivalentes des deux premiers auteurs.

Article 7
Albert C.H., Thuiller W., Yoccoz N.G., Soudan A., Boucher F., Saccone P., Lavorel S. *Intraspecific functional variability: extent, structure and sources of variation within an alpine valley.* (En revision mineure à Journal of Ecology)

Article 8
Albert C.H., Thuiller W., Dubuis A., Guisan A., Yoccoz N.G., Aubert S., Douzet R., Lavorel S., *Intraspecific functional variability: on its relative importance in respect to interspecific variability.* (Soumis à Ecology)

Article 9
Albert C.H., Yoccoz N.G, Graham C., Edwards T., Zimmermann N.E., Thuiller W. *Sampling in ecology and evolution - integrating questions and designs.* (Soumis à Ecography)

Article 10
Kunstler G., **Albert C.H**., Courbaud B., Lavergne S., Thuiller W., Vieilledent G., Zimmerman N.E., Coomes D.A., *Variation of the intensity and the importance of tree-tree interaction along climatic gradients across the French Alps and the Jura Mountains.* (En préparation)

Article 11
Gallien L., **Albert C.H.**, Boulangeat I., Muenkemueller T., Thuiller W., *Modelling invasions: where are we getting from there?* (Soumis à Diversity & distributions)

Table des matières

Introduction. ---------- **19**

Chapitre 1 – Echantillonnage : principes et applications

1. L'échantillonnage en écologie : vers une mise en œuvre des concepts ---------- **43**

1.1. Pourquoi échantillonner? Raisons et définitions ---------- *43*

1.2. Estimation de courbes de réponse : problèmes et solutions ---------- *45*

2. Base de données Guisane 2080 : réalités de l'échantillonnage ---------- **47**

2.1. Objectifs ---------- *47*

2.2. Sites d'étude et choix des gradients ---------- *49*

2.3. Seize espèces modèles ---------- *53*

2.4. Mise au point d'une stratégie d'échantillonnage ---------- *57*

2.5. Mesures de terrain: La base Guisane 2080 ---------- *65*

Chapitre 2 – Variabilité fonctionnelle intraspécifique: quantification, structure, sources et comparaison avec l'interspécifique

1. Contexte et objectifs ---------- **77**

2. Variabilité dans le temps et date d'échantillonnage ---------- **79**

3. Quantification et structure des variabilités fonctionnelles intra- et interspécifique ---------- **85**

3.1. Quelle est l'ampleur de la variabilité fonctionnelle intraspécifique ? ---------- *85*

3.2. Quelle est la structure spatiale de la variabilité fonctionnelle intraspécifique au sein des espèces ? ---------- *89*

3.3. La variabilité fonctionnelle intraspécifique est elle négligeable devant la variabilité interspécifique ? ---------- *91*

3.4. Qu'implique la variabilité fonctionnelle intraspécifique sur la définition des stratégies fonctionnelles ? *93*

4. Rôle des gradients ---------- **97**

4.1. Rôle des gradients environnementaux dans la variabilité fonctionnelle ---------- *97*

4.2. Modèle conceptuel ---------- *99*

5. Conclusion & perspectives ---------- **103**

Chapitre 3 – Signification des modèles d'habitat et lien avec les traits fonctionnels et la performance

1. Les modèles statistiques d'habitat ---------- **109**

1.1. Principe ---------- *109*

1.2. Construire un modèle d'habitat ---------- *109*

1.3. Hypothèses de construction ---------- *111*

1.4. Algorithmes utilisés et sélection de modèle ---------- *113*

1.5. Utilisations des modèles d'habitat ---------- *117*

1.6. Limites techniques et conceptuelles ---------- *117*

1.7. Trouver un sens ---------- *119*

2. Tests de la signification des modèles d'habitat ---------- 123

2.1. Les modèles d'habitat comme une description fonctionnelle de la niche ---------- *123*

2.2. Une relation distendue entre modèles d'habitat et démographie : manipulation de germination et survie ---------- *129*

2.3. Une relation distendue entre modèles d'habitat et démographie : démographie des forêts des Alpes françaises ---------- *135*

3. Conclusion ---------- 141

Chapitre 4 – Prise en compte de la variabilité fonctionnelle intraspécifique et implications

1. Pourquoi et comment prendre en compte la variabilité fonctionnelle intraspécifique? ---------- 151

1.1. Quoi faire changer ? ---------- *153*

1.2. Comment faire pour une espèce ? ---------- *153*

1.3. Comment faire pour plusieurs espèces ? ---------- *155*

1.4. Quel lien avec les modèles d'habitat ? ---------- *155*

2. Implication dans l'évaluation de la diversité fonctionnelle ---------- 157

2.1. Le concept de diversité fonctionnelle ---------- *157*

2.2. Effet de la variabilité fonctionnelle intraspécifique, de sa structure et de ses sources sur la stabilité des indices de diversité fonctionnelle ---------- *159*

3. Implication dans la modélisation de la dynamique d'un paysage ---------- 169

3.1. Vers des modèles « hybrides » ---------- *169*

3.2. L'exemple de Lamos/Biomove ---------- *171*

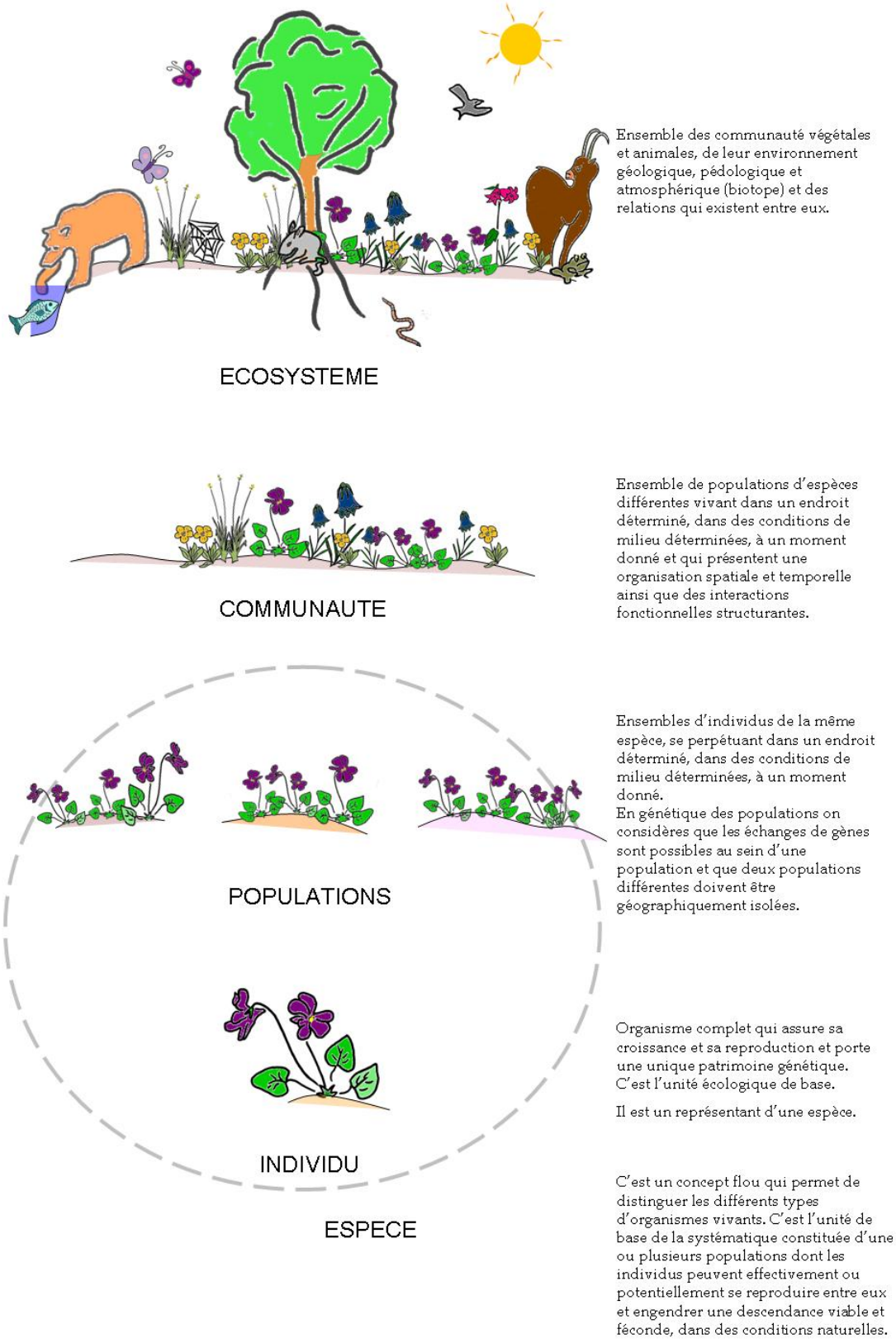
4. Conclusion ---------- 185

Conclusion générale ---------- 189

Bibliographie ---------- 197

Annexes ---------- 210

Introduction



Encadré 0.1 Définitions écologiques des différents niveaux d'organisation des organismes vivants.

La biosphère : un système complexe

Sur notre planète, la Terre, tout est lié et participe à une même dynamique, des cycles de l'eau, de l'air et du carbone au cycle de la vie qui permet la régénération permanente du vivant. Ce vivant qui persiste depuis plusieurs milliards d'année est un système d'une grande complexité. Il est constitué de plusieurs millions d'espèces dont une grande partie demeure inconnue, c'est la biodiversité à l'échelle spécifique. Plantes, champignons, microbes, animaux vertébrés et non vertébrés cohabitent et interagissent dans tous les milieux de la planète et sous des climats aussi bien désertiques que tropicaux, montagnards ou tempérés pour former des populations, des communautés et des écosystèmes (Encadré 0.1) très diversifiés, c'est la biodiversité écosystémique. Ces espèces ne sont pas des entités figées. Elles sont issues de l'Evolution et continuent d'évoluer du fait de pressions de sélection qui leur sont imposées par l'environnement dans lequel elles vivent. De ce fait tous les individus d'une même espèce ne portent pas nécessairement exactement le même génome et n'ont donc pas nécessairement les mêmes « comportements » face à l'environnement dans lequel ils vivent, c'est la biodiversité génétique. La diversité et la complexité du monde vivant se retrouvent ainsi à différentes échelles spatiales imbriquées.

L'Homme a pris part au cycle de la vie et aux différents cycles biogéochimiques depuis quelques 200 000 ans. Depuis le siècle dernier, l'impact de l'Homme sur son environnement et sur ces cycles est sans précédent. Du fait d'une augmentation incessante de la population humaine (Fig. 0.1a), créant un besoin croissant en ressources et en énergie, l'Homme émet des gaz à effet de serre dans l'atmosphère (Fig. 0.1b) -entrainant une hausse des températures (IPCC 2007)-, rejette des produits polluants dans les eaux et dans l'air (Zvereva *et al.* 2008), influence de manière négative la nature sauvage (ex. extinction des grands prédateurs : Treves & Karanth 2003, raréfaction des polinisateurs : Cox & Elmqvist 2000), participe de manière active à l'homogénéisation du vivant *via* des introductions d'espèces exogènes aux quatre coins du globe (Mack *et al.* 2000; Mooney & Hobbs 2000), modifie les régimes naturels de perturbations (ex. feux, de Groot *et al.* 2003) et influence les habitats naturels *via* la destruction (ex. déforestation massive : Wright & Muller-Landau 2006), la fragmentation et les changements d'utilisation des terres. Ces bouleversements rapides induits par l'Homme et ajoutés aux fluctuations naturelles des processus –dont l'ensemble constitue ce que l'on appelle les changements globaux- dégradent les cycles biogéochimiques et altèrent la diversité du monde vivant. Bien que la diversité du vivant ne puisse se résumer au monde végétal seul, la suite du texte se focalisera exclusivement sur les plantes.

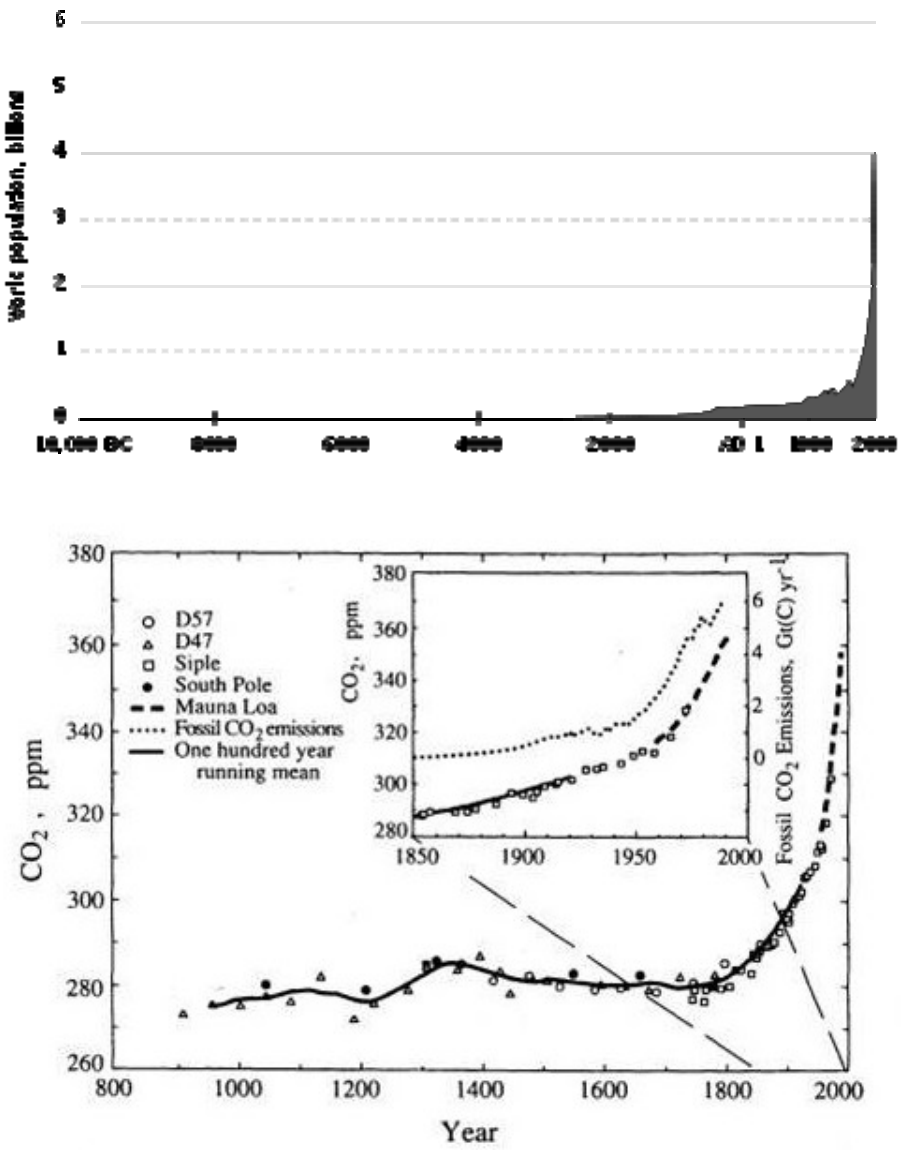


Fig. 0.1. Démographie et dioxyde de carbone dans l'atmosphère.
(En haut) Evolution de la démographie humaine au cours des millénaires (source : http://fr.wikipedia.org/wiki/Population_mondiale);(En bas) Evolution de la teneur atmosphérique en CO^2 au cours des siècles, avec un zoom sur les derniers 150 ans (source : http://acces.inrp.fr/acces/terre/CCCIC/activiteCycle/image/evolcar.jpg/image).

L'écologie fonctionnelle : simplifier pour comprendre et prédire

La multiplicité des organismes vivants et des interactions se produisant entre eux et avec leur milieu physico-chimique rend la tâche des écologues qui les étudient laborieuse, les amenant à se demander s'il existe réellement des lois universelles en écologie (Lawton 1999). Face à cette complexité l'écologie fonctionnelle propose de décrire les espèces par leurs caractéristiques biologiques et leurs fonctions au sein de leur environnement (abiotique et biotique) au lieu de les décrire par leur identité taxonomique (Calow 1987). Plusieurs espèces distinctes sur le plan taxonomique peuvent ainsi avoir les mêmes fonctions et caractéristiques biologiques. Ce domaine conceptuel s'est fortement développé au cours des vingt dernières années, se dotant progressivement de définitions et de protocoles standardisés (Weiher *et al.* 1999; Lavorel & Garnier 2002; Cornelissen *et al.* 2003; Violle *et al.* 2007). Elle repose en particulier sur la définition des traits fonctionnels végétaux qui sont toute grandeur morpho-phéno-physiologique mesurable au niveau individuel et ayant un impact indirect sur la valeur sélective de l'individu par un effet sur sa survie, sa croissance ou sa reproduction (Lavorel *et al.* 1997; Violle *et al.* 2007). Ces traits peuvent être des indicateurs de la réponse (traits de réponse) des individus aux variations du milieu, favorisant la dominance de certaines espèces dans les communautés végétales et/ou décrivent un effet (traits d'effet) des individus sur leur communauté ou le fonctionnement de l'écosystème auquel ils appartiennent (Lavorel & Garnier 2002 ; Garnier *et al.* 2004 ; Article 4). Mesurés sur des individus, ces traits peuvent être moyennés au niveau de la population ou de la communauté pour faire le lien entre plusieurs niveaux d'organisation (plante individuelle, population, communauté, écosystème, paysage, Encadré 0.1). La valeur (ou le niveau) prise par un trait fonctionnel (attribut de trait) donne une indication sur une fonction de la plante individuelle ou sur un aspect des composantes de sa valeur sélective : survie, croissance, fécondité (Violle *et al.* 2007). L'ensemble de ses valeurs (ou les niveaux) de traits (syndrome de traits) résulte de compromis (Díaz & Cabido 1997) entre les différentes fonctions de la plante et est nécessaire pour comprendre son rôle au sein de sa communauté. Jusqu'à présent les traits fonctionnels ont principalement été utilisés sous forme de moyenne à l'échelle des populations, des communautés et des espèces ; la mesure des valeurs prises par quelques individus dans quelques populations permettant d'attribuer à chaque espèce une valeur moyenne de trait. La sélection d'un petit jeu de traits fonctionnels non-redondants (Westoby 1998; Weiher *et al.* 1999 ; Lavorel & Garnier 2002) permet alors de décrire et de regrouper les espèces ayant des syndromes moyens de traits similaires en groupes fonctionnels (PFT pour « plant functional type », Lavorel *et al.* 1997) et permet aussi de rattacher chaque espèce étudiée à des stratégies fonctionnelles. La quête de stratégies fonctionnelles pour classifier les espèces et décrire des comportements généraux et contrastés (comme par exemple le triangle de Grime, Grime 1977; Westoby 1998; Reich *et al.* 2003) a été présentée tour à tour comme la

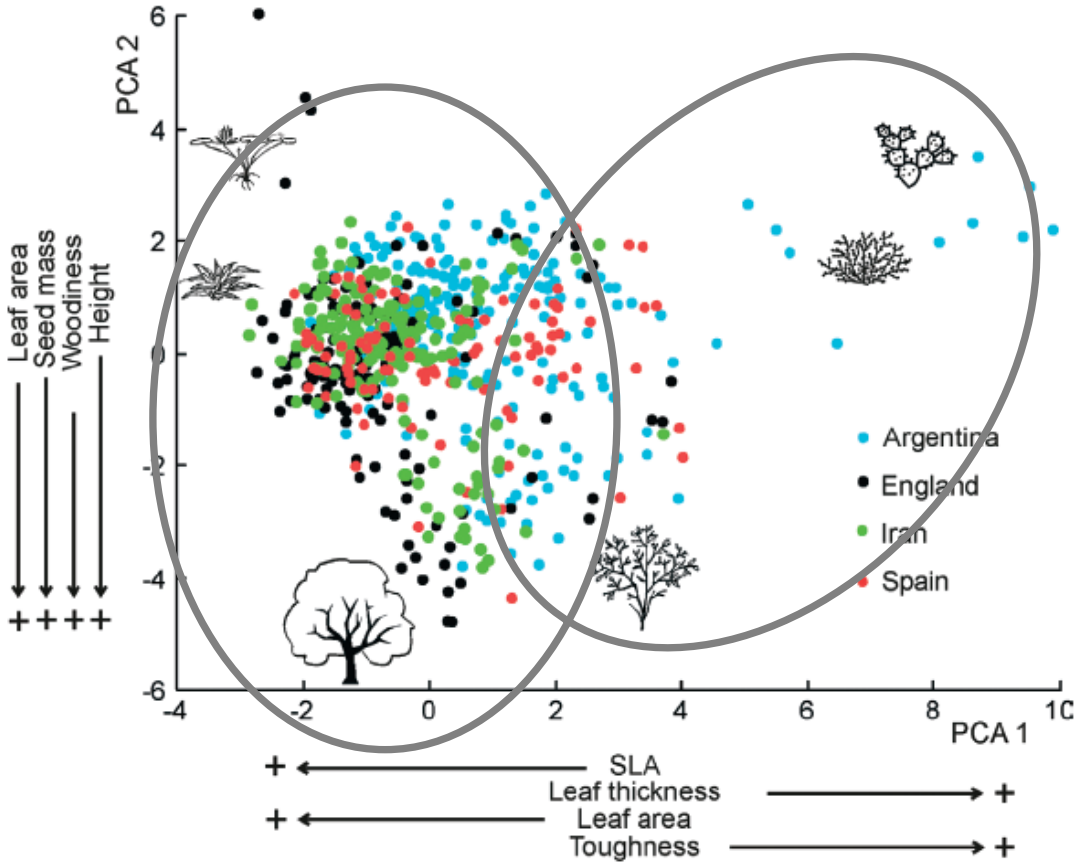


Fig. 0.2. Ordination par analyse en composante principale (ACP) sur 640 espèces provenant d'Argentine, Angleterre, Iran et Espagne, sur la base de 12 traits. *Sont présentés les deux premiers axes de l'ACP, les traits sont présentés d'autant plus près des axes qu'ils leur sont corrélés. Les schémas indiquent les formes de vie extrêmes telles que aquatiques ou a larges feuilles à gauche de l'axe 1 (forte SLA) et les cactacées à droite de cet axe 1 (faible SLA). Adaptée de Diaz et al. 2004 dans Journal of Vegetation Science.*

recherche d'un tableau périodique écologique (Steffen 1996) ou comme le Saint-Graal de l'écologie fonctionnelle (Lavorel & Garnier 2002). L'espèce reste cependant un concept flou et complexe (ex : hybridations inter espèce fréquentes chez les plantes). De plus les individus d'une même espèce ne sont chacun qu'une réalisation d'un potentiel génétique commun et peuvent être très différents les uns des autres malgré des ressemblances (Bonnier 1887) du fait des adaptations locales et/ou de la plasticité phénotypique (Byars *et al.* 2007). Par analogie avec l'espèce humaine, Westoby *et al.* (2002) a donc proposé de comparer les stratégies fonctionnelles avec les typologies psychologiques. Les deux quêtes en écologie et psychologie sont similaires et passent par la définition d'axes de variabilité permettant de décrire les grandes lignes des personnalités ou des fonctions des individus/espèces et de prédire leur comportement dans différentes situations. Les syndromes de traits attribués aux individus /espèces permettent alors d'exprimer la position de chacune le long des axes de variabilité (Westoby *et al.* 2002). Cela nécessite que les traits utilisés soient « robustes», c'est-à-dire qu'ils varient plus entre espèces qu'au sein de chacune, permettant une classification significative et pertinente des espèces (Garnier *et al.* 2001; Shipley 2003; Roche *et al.* 2004; Al Haj Khaled *et al.* 2005; McGill *et al.* 2006). En écologie fonctionnelle, un axe principal de variabilité et de spécialisation décrivant la gestion des ressources par les plantes a été mis en évidence à plusieurs reprises et à différentes échelles au niveau interspécifique (Grime *et al.* 1997; Diaz *et al.* 2004; Gross *et al.* 2007; Pierce *et al.* 2007). Cet axe résulte d'un compromis (ici, corrélation négative) entre la teneur des feuilles en matière sèche (LDMC, pour « leaf dry matter content ») et la surface spécifique foliaire (SLA, pour « specific leaf area »). Les espèces ayant une forte SLA (faible LDMC) possèdent des tissus foliaires peu denses, un taux de photosynthèse élevé et une forte vitesse de croissance, entrainant une forte capacité d'acquisition des ressources (Fig. 0.2). Ces espèces dites « exploitatrices » dominent dans les milieux fertiles (Grime *et al.* 1997; Reich *et al.* 1999). Au contraire les espèces ayant une faible SLA (forte LDMC) possèdent des tissus foliaires denses et ont une croissance lente, entrainant une forte capacité de conservation des ressources dans les tissus. Ces espèces dites « conservatrices » dominent dans les milieux peu fertiles. Jusqu'à présent la définition des stratégies fonctionnelles et la recherche de compromis de traits et d'axes de variabilité ont été menées essentiellement à l'échelle interspécifique, alors que la définition des traits est située à l'échelle individuelle.

Utilisations des traits fonctionnels

L'écologie fonctionnelle est donc un domaine conceptuel de l'écologie qui se veut intégrateur et qui permet le passage d'un niveau d'organisation à un autre favorisant la compréhension des propriétés de l'ensemble. Elle permet d'une part d'étudier la réponse du fonctionnement

d'un système végétal aux conditions environnementales et aux variations naturelles et anthropiques de celles-ci. D'autre part elle permet d'analyser la manière dont le fonctionnement de chaque niveau d'organisation affecte les services rendus par les écosystèmes, c.à.d. l'ensemble des bénéfices que l'Homme retire de leur fonctionnement, tels que l'alimentation, le recyclage de l'eau et de l'air, la pollinisation, le tourisme (Schröter *et al.* 2005; Tscharntke *et al.* 2005; Diaz *et al.* 2007). Pour ces raisons, l'approche fonctionnelle et les traits fonctionnels ont été beaucoup utilisés sous forme de moyennes de traits -au niveau population, espèce ou communauté- dans divers buts : comme des indicateurs de la dynamique des communautés (Choler *et al.* 2001) ; pour établir des classifications fonctionnelles des organismes (Lavorel *et al.* 1997) ; pour quantifier la diversité fonctionnelle des communautés (de Bello *et al.* 2009) ; pour décrire les relations à large échelle entre traits fonctionnels (Wright *et al.* 2004) ou entre traits et gradients environnementaux (Thuiller *et al.* 2004) ; pour examiner les relations entre le fonctionnement des écosystèmes et les biens et services associés (Diaz *et al.* 2007) ; et aussi pour paramétrer des groupes fonctionnels dans des modèles de dynamique du paysage (Woodward & Diament 1991, Article 1, Article 2) ou des individus dans des modèles de traits ou individu-centrés (Norberg *et al.* 2001; Lehsten & Kleyer 2007; Savage *et al.* 2007; Jeltsch *et al.* 2008).

Les traits fonctionnels dans l'environnement

L'importance des gradients environnementaux pour expliquer les variations dans les caractéristiques des plantes a été reconnue très tôt (Bonnier 1887). Cependant, si les évolutionnistes admettent l'importance de la variabilité intraspécifique sur la valeur sélective et la démographie (McGraw & Antonovics 1983), les « écologues savent que les traits sont variables au sein des espèces, mais feignent souvent de ne pas le savoir » (Cianciaruso *et al.* 2009). Etudier la variabilité des traits le long des gradients environnementaux de l'échelle régionale à l'échelle du globe est une méthode de choix pour comprendre comment les plantes répondent à des conditions environnementales contrastées. Il existe un grand nombre d'études sur ce sujet traitant de la réponse des traits fonctionnels ou écophysiologiques et des paramètres démographiques aux gradients (Cordell *et al.* 1998; Reich *et al.* 1999; Ryser & Aeschlimann 1999; Wright *et al.* 2004). Ces études sont le plus souvent menées au niveau interspécifique (Reich *et al.* 1999 ; Wright *et al.* 2004) et analysent la manière dont les moyennes de traits par espèces sont agencées par des compromis de traits généraux (Wright *et al.* 2005) ou des gradients environnementaux (Thuiller *et al.* 2004). Dans ces études, les espèces ou au mieux les populations sont considérées comme des entités homogènes et décrites par des valeurs de trait moyen obtenues à partir de mesures faites sur un nombre souvent restreint d'individus et de populations. Ces études ignorent

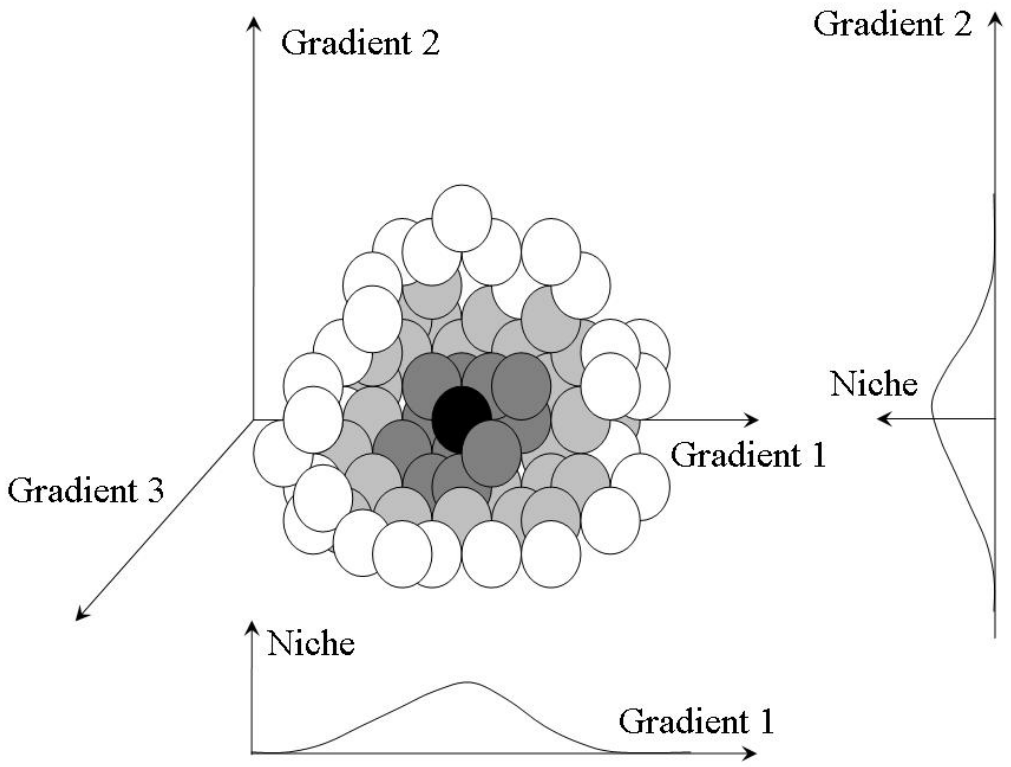


Fig. 0.3. Niche d'un organisme dans un espace environnemental à trois dimensions.
Le dégradé de gris va du cœur de la niche en noir à la périphérie en blanc.

fréquemment la variabilité intraspécifique des traits. D'autres études menées à l'échelle intraspécifique visent à étudier la variabilité des traits fonctionnels ou des paramètres démographiques au sein de chaque espèce sans chercher à généraliser les patrons obtenus. Ces études sont souvent menées sur une ou quelques espèces (Woodward 1986; Cordell *et al.* 1998; Ryser & Aeschlimann 1999; Dangasuk & Panetsos 2004; Milla *et al.* 2009), voire sur des genres ou des formes de vie (Körner *et al.* 1986; Körner 1989; Wegner *et al.* 2003). Enfin certaines études visent à comprendre et quantifier pour plusieurs espèces la réponse intraspécifique des traits à des conditions contrastées (Meziane & Shipley 1999; Shipley 2000; Dyer *et al.* 2001). Cependant elles restent le plus souvent menées en conditions contrôlées et les réponses des traits aux traitements y sont considérées par l'intermédiaire des moyennes de traits par traitement, sans chercher à prendre en compte la variabilité individuelle. Les traits fonctionnels restent donc relativement peu décrits à l'échelle intraspécifique. L'étude de la bibliographie ne permet actuellement pas de définir des courbes de réponse générales des traits fonctionnels aux gradients environnementaux à l'échelle intraspécifique, comme elle le permet à l'échelle interspécifique (Reich *et al.* 1999; Wright *et al.* 2004). S'interroger sur la réponse des traits de chaque espèce aux gradients environnementaux, nécessite de connaitre la distribution de l'espèce le long de ces gradients, afin d'étudier pour chacune les portions de gradients appropriées.

Le concept de niche écologique

Le concept de Niche écologique semble avoir été inventé par J. Grinnell en 1917 (Grinnell 1917). Il s'agissait alors d'une propriété de l'environnement abiotique, une niche environnementale donnée pouvant accueillir une espèce ayant des exigences physiologiques appropriées. En 1927, C. Elton propose une définition alternative en définissant la niche par rapport à l'environnement biotique (Elton 1927), comme une profession pouvant être attribuée à une espèce dans l'écosystème. En 1957, G.E. Hutchinson définit la niche de manière abstraite, comme un hyper volume à *n* dimensions, où chaque dimension est une ressource ou une condition limitante de l'environnement (ex. eau, lumière, nourriture). La niche devient alors une propriété de l'occupant, étant l'ensemble des conditions environnementales nécessaires pour permettre à l'organisme d'exister indéfiniment (Hutchinson 1957). Pour concilier les différentes approches historiques de la niche et le principe d'exclusion compétitive (Gause 1934) - qui veut que deux espèces en compétition pour une même ressource ne peuvent cohabiter durablement dans un environnement stable- G.E. Hutchinson distingue la niche fondamentale de la niche réalisée. La niche fondamentale est l'ensemble des conditions environnementales dans lesquelles l'espèce peut en théorie se maintenir en l'absence d'interactions biotiques, c.à.d. que les populations y ont des taux de

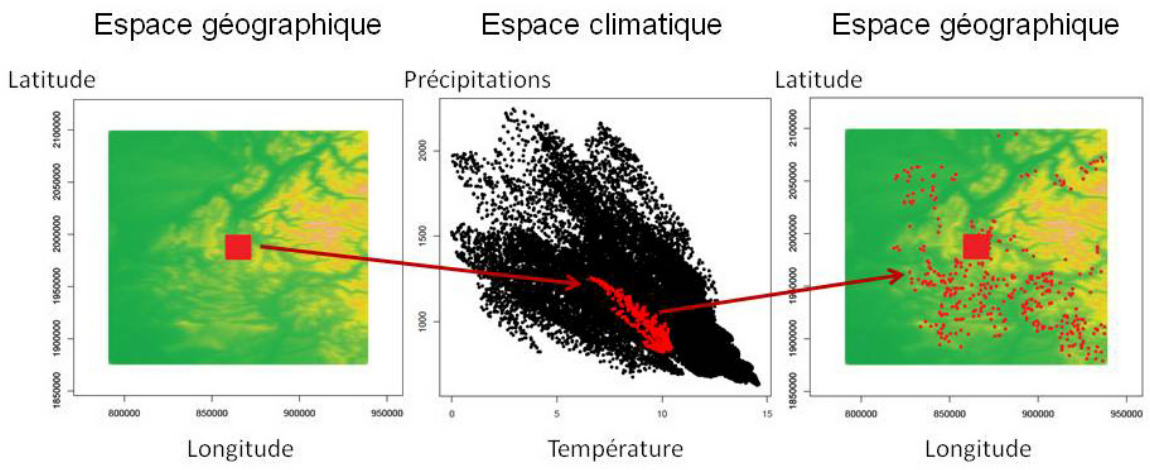


Fig. 0.4. Passage de l'espace géographique à l'espace climatique et réciproquement.
La projection d'un espace dans un autre n'est pas une simple bijection. (A gauche) la distribution fictive d'un organisme. Cette distribution géographique correspond à un certain type de conditions climatiques (Au milieu). Ces conditions climatiques (Au milieu) correspondent à plus de localisations géographiques que la distribution initiale (A droite). Les endroits climatiquement favorables où l'organisme n'est pas présent sont des endroits de distribution potentielle limités par la compétition et la dispersion.

croissance net positifs. La niche réalisée est la partie de la niche fondamentale qui est effectivement occupée par l'espèce une fois les interactions biotiques prises en compte, mais aussi due à la limitation par la dispersion (des zones parfaitement adéquates peuvent se trouver inaccessibles) et les relations sources-puits (une population pouvant être en dehors de la niche fondamentale mais se maintenir par l'apport constant d'individus d'une population voisine avec un taux de croissance positif, Pulliam 2000).

En accord avec les redéfinitions récentes du concept (Pulliam 2000; Chase & Leibold 2003), nous garderons finalement comme définition de la niche écologique: l'ensemble des conditions environnementales dans lesquelles une espèce peut avoir des populations à taux de croissance positifs. Cette niche écologique, est donc réalisée et peut être quantifiée d'une part par une description quantitative des variables environnementales et d'autre part par une quantification de la performance de l'organisme en termes de taux de croissance de ses populations (Fig. 0.3). On peut alors distinguer le cœur de la niche, qui correspond aux conditions optimales pour l'organisme et les zones périphériques qui correspondent à des conditions viables mais sous-optimales (performance plus faible).

La distribution (géographique) d'une espèce, reflète dans l'espace géographique (ex. longitude, latitude) sa niche écologique définie dans l'espace environnemental (ex. climat, sol, ressource), ainsi que sa capacité à disperser ses propagules et l'influence du milieu biotique (ex. prédateurs, compétiteurs). Ainsi distribution et niche se reflètent l'une l'autre sans pour autant coïncider (Fig. 0.4). Certaines aires géographiques peuvent en effet être adéquates en termes abiotiques sans pour autant que l'espèce considérée y soit présente du fait de la compétition ou d'une limitation par la dispersion. Au contraire celle-ci peut aussi être présente - sans pour autant former des populations à taux de croissance positifs - dans certaines aires géographiques qui sont en dehors de sa niche écologique, comme dans le cas des situations sources-puits.

Après une période d'oubli dans les années 80 au profit des concepts d'écologie des communautés et des modèles mécanistes de compétition (Mac Arthur 1972), la niche écologique est redevenue un concept et un outil à la mode dans le contexte des changements globaux. Elle connait un essor en particulier en biogéographie prédictive pour expliquer et prédire les relations entre végétation et environnement (Guisan & Thuiller 2005), mais aussi en évolution avec l'émergence du concept d'évolution de la niche (Wiens & Graham 2005; Ackerly *et al.* 2006; Pearman *et al.* 2008) et en écologie des communautés à propos des mécanismes de coexistence et d'assemblage (Keddy 1992; Chave 2004; Ackerly & Cornwell 2007). Ce nouvel essor du concept de niche s'est accompagné de la mise en place de modèles de distribution statistiques, dits « modèles d'habitat » (Guisan & Thuiller 2005), et également mécanistes (Chuine *et al.* 1999; Kearney & Porter 2009) visant à quantifier la niche des espèces.

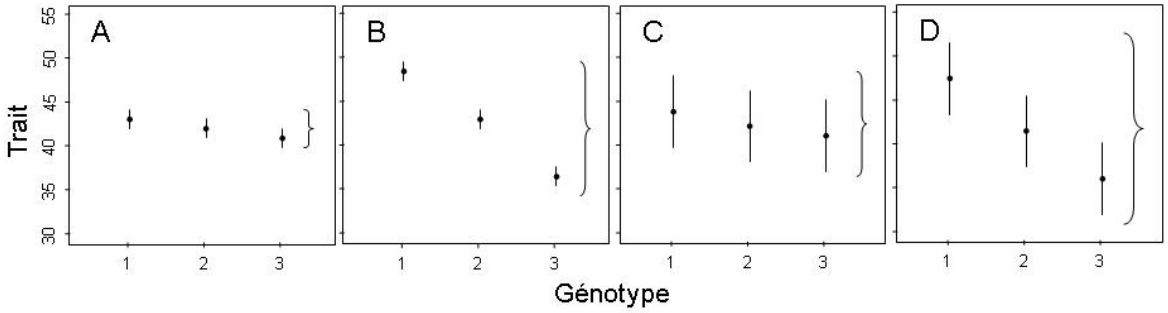


Fig. 0.5. Origines de la variabilité fonctionnelle intraspécifique.
A. Faible variabilité intraspécifique due à de faibles différences de traits moyens entre génotypes et à une faible plasticité des différents génotypes ; B. Forte variabilité intraspécifique due à de fortes différences de traits moyens entre génotypes mais à une faible plasticité des différents génotypes ; C. Forte variabilité intraspécifique due à une forte plasticité des différents génotypes mais à de faibles différences de traits moyens entre génotypes ; D. Forte variabilité intraspécifique due à de fortes différences de traits moyens entre génotypes et à une forte plasticité des différents génotypes.

Objectifs généraux de la thèse

La variabilité fonctionnelle intraspécifique

La variabilité fonctionnelle intraspécifique ou variabilité intraspécifique des traits exprime l'étendue des valeurs prises par les traits d'une espèce ainsi que la variabilité de ce qui est fonctionnellement vécu par les individus d'une espèce vivant dans différentes conditions environnementales. Cette variabilité résulte à la fois de la diversité génétique de l'espèce et de la plasticité phénotypique des différents génotypes (Fig. 0.5), les différences adaptatives et plastiques pouvant être complémentaires et non pas exclusives (Joshi *et al.* 2001 ; Byars *et al.* 2007). La diversité génétique d'une espèce résultant d'adaptations locales entraine une multiplicité des génotypes au sein de l'espèce, ces différents génotypes pouvant avoir des valeurs moyennes de trait différentes. La plasticité phénotypique est le potentiel de production d'une multitude de phénotypes à partir d'un unique génotype en fonction des conditions environnementales biotiques et abiotiques rencontrées (DeWitt *et al.* 1998). La variabilité intraspécifique résultant indique le potentiel de l'espèce à survivre, croître et se reproduire dans de nouvelles conditions environnementales (Joshi *et al.* 2001 ; Byars *et al.* 2007). Elle conditionne la réponse des populations et des espèces aux changements environnementaux et aux changements résultants dans les communautés et écosystèmes. C'est une sorte d'assurance-vie pour les espèces, une espèce peu diverse et/ou peu plastique ayant peu de chance de se maintenir lors d'importants changements environnementaux. La variabilité fonctionnelle intraspécifique peut être considérée à différents niveaux, pouvant provenir : (1) de différences entre moyenne de traits de populations vivant dans des conditions différentes ; (2) de différences entre valeurs de traits d'individus d'une même population ; (3) de différences entre feuilles d'un même individu (Shipley 1995).

Dans la suite de ce manuscrit je me suis intéressée à la variabilité fonctionnelle intraspécifique dans son ensemble, sans chercher à distinguer les effets adaptatifs des effets plastiques. Tout d'abord parce que cette variabilité résulte des deux phénomènes et donne donc une idée de la capacité des espèces à « s'adapter » -au sens large- à de nouvelles conditions (Milla *et al.* 2009). D'autre part parce que l'approche fonctionnelle s'intéresse essentiellement à l'expression phénotypique et à ses conséquences sur les niveaux d'organisation supérieurs. Il n'est pas nécessaire de savoir d'où vient la variabilité pour étudier ce qu'elle implique. Enfin le lecteur désirant en savoir plus sur cette partition de la variabilité fonctionnelle entre effets génétiques et plastiques pourra lire la thèse de Fabrice Grassein (Grassein 2009).

Des études récentes montrent que la variabilité fonctionnelle intraspécifique peut avoir des effets significatifs sur le fonctionnement des écosystèmes, incluant les cycles du carbone et de l'azote (Lecerf & Chauvet 2008), la réponse à l'herbivorie (Boege & Dirzo 2004), la disponibilité en lumière sous couvert forestier (Vieilledent 2009). Par exemple, Lecerf & Chauvet (2008) ont mis en évidence une forte variabilité intraspécifique des traits foliaires à l'échelle continentale chez une espèce d'arbre riparienne. La variabilité des traits entrainait

une forte variabilité des taux de décomposition, ceux-ci variant autant au sein des espèces qu'entre espèces. La variabilité intraspécifique commence aussi à être prise en compte dans les calculs de diversité fonctionnelle (Leps *et al.* 2006; Cianciaruso *et al.* 2009), même si il n'y a pour le moment pas de résultat général sur la nécessité de la prendre en compte ou non (ex. Lake & Ostling 2009).

Les attentes

On sait donc que les traits fonctionnels varient au sein des espèces et que leur variabilité a des conséquences dans différents types d'utilisation, cependant:

(1) Dans la pratique, les traits sont le plus souvent mesurés sur des individus (quelques) et moyennés à l'échelle des populations ou des espèces selon les cas, la variabilité intraspécifique étant négligée. On peut se demander si dans un environnement hétérogène et changeant, l'utilisation de traits moyens n'est pas un peu trop simpliste.

(2) A l'échelle interspécifique, des compromis de traits et des stratégies fonctionnelles ont été décelés à partir de traits moyens par espèces. En effet, si l'on retrouve plutôt des espèces à forte SLA (resp. faible) dans les milieux fertiles (resp. peu fertiles), on pourrait s'attendre également au sein d'une espèce à trouver des individus/populations ayant une forte SLA (resp. faible) dans les milieux fertiles (resp. peu fertiles). A ma connaissance, les stratégies fonctionnelles et les compromis de traits ont peu été testés à l'échelle individuelle au sein des espèces (Fonseca *et al.* 2000).

(3) L'utilisation des traits fonctionnels repose sur l'hypothèse que ceux-ci varient plus entre espèces qu'au sein de chacune («robustesse»). Cependant jusqu'à présent peu d'études ont proposé une comparaison entre variabilités inter- et intraspécifique (Garnier *et al.* 2001; Roche *et al.* 2004 ; Wilson *et al.* 1999), c'est donc un présupposé qui demande à être vérifié.

(4) Bien que les traits fonctionnels soient variables au sein des espèces, on sait très peu de choses de cette variabilité, sa structure ou son importance. Il est très difficile à partir des études existantes dans la littérature de se faire une idée de la structure de la variabilité entre espèces, et au sein des espèces : entre populations, entre individus d'une même population et au sein des individus d'une même population. De plus il est également difficile d'obtenir des comparaisons entre traits et entre espèces et de construire des courbes de réponse des traits aux gradients environnementaux.

(5) Les modèles de dynamique de la végétation sont souvent construits de telle sorte que les espèces ou PFT sont caractérisés par des paramètres fixes, les différentes espèces/PFT ayant des valeurs de traits différentes mais ces valeurs étant constantes au sein des espèces. Une mauvaise connaissance de la variabilité fonctionnelle

intraspécifique entraine aussi une mauvaise connaissance de ses effets potentiels sur les propriétés des communautés végétales et le fonctionnement des écosystèmes.

Inclure la variabilité intraspécifique dans les théories écologiques et en particulier en écologie fonctionnelle semble aujourd'hui essentiel pour affiner notre compréhension de la biodiversité et du fonctionnement des écosystèmes et prédire leurs réponses aux changements globaux. MgGill *et al.* (2006) ont proposé d'utiliser les traits fonctionnels pour faire de l'écologie une science plus quantitative, suggérant en particulier comme axe prioritaire de recherche l'étude de la réponse des traits fonctionnels aux gradients environnementaux.

Les objectifs

A l'aide d'approches de terrain et de modélisation, j'ai essayé dans cette thèse de combler un certain nombre d'attentes quant à notre connaissance de la variabilité fonctionnelle intraspécifique. Ce manuscrit présente mon travail comme une synthèse constituée de quatre chapitres complémentaires.

Chapitre 1. Echantillonnage : principes et applications

Ce chapitre comporte deux parties distinctes. Une première partie décrit un cadre théorique permettant de mener à bien un échantillonnage prenant en compte les particularités des questions et patrons écologiques impliqués. Ce cadre doit permettre de concilier les concepts théoriques de l'échantillonnage avec les spécificités de l'écologie de terrain. Une deuxième partie décrit l'échantillonnage concrètement mis en place durant ma thèse pour étudier la variabilité fonctionnelle intraspécifique et les données en résultant.

Chapitre 2. Variabilité fonctionnelle intraspécifique : quantification, structure et sources

Ce chapitre traite de deux questions essentielles quand à la variabilité fonctionnelle intraspécifique. (i) Comment caractériser la variabilité fonctionnelle intraspécifique en termes de quantification, structures et sources ? (ii) Quelle est l'importance relative de la variabilité fonctionnelle par rapport à la variabilité interspécifique?

Chapitre 3. Sens des modèles d'habitat et lien avec les traits fonctionnels et la performance

L'utilisation des sorties des modèles d'habitat (biogéographie prédictive) sous leur forme brute repose sur une hypothèse forte qui est : les fortes fréquences de l'espèce correspondent au cœur de la niche, c'est-à-dire aux conditions les plus adéquates pour cette espèce donnant lieu localement à de bonnes performances individuelles. Cette hypothèse a été jusqu'ici très peu testée. Ce chapitre propose de tester cette hypothèse en confrontant les valeurs des modèles d'habitat décrivant la qualité de l'habitat (i) avec les données de traits collectées durant ma thèse (Chapitre 1) ; (ii) avec des données démographiques de germination et de

survie ; (iii) avec des données de croissance d'arbres adultes récoltées durant ma thèse ; (iv) avec des données de l'Inventaire Forestier National de croissance d'arbres adultes sur l'ensemble des Alpes françaises.

Chapitre 4. Prise en compte de la variabilité fonctionnelle intraspécifique et implications

Ce chapitre utilise les résultats des autres chapitres pour résumer les différentes questions nécessaires à la prise en compte de la variabilité fonctionnelle intraspécifique et tester les implications de sa prise en compte dans deux cas d'étude. Une première étude teste l'effet de la variabilité intraspécifique et de sa structure sur le calcul d'indices de diversité fonctionnelle sur des communautés de la vallée de la Guisane (site d'étude principal de ma thèse). Une deuxième étude teste l'effet de la prise en compte de la variabilité fonctionnelle intraspécifique dans un modèle de paysage, par le couplage de celui-ci avec un modèle d'habitat.

Une dernière partie intitulée synthèse et perspectives conclu ce travail de thèse et apporte les perspectives vers de futures recherches.

Chapitre 1

Echantillonnage : principes et applications

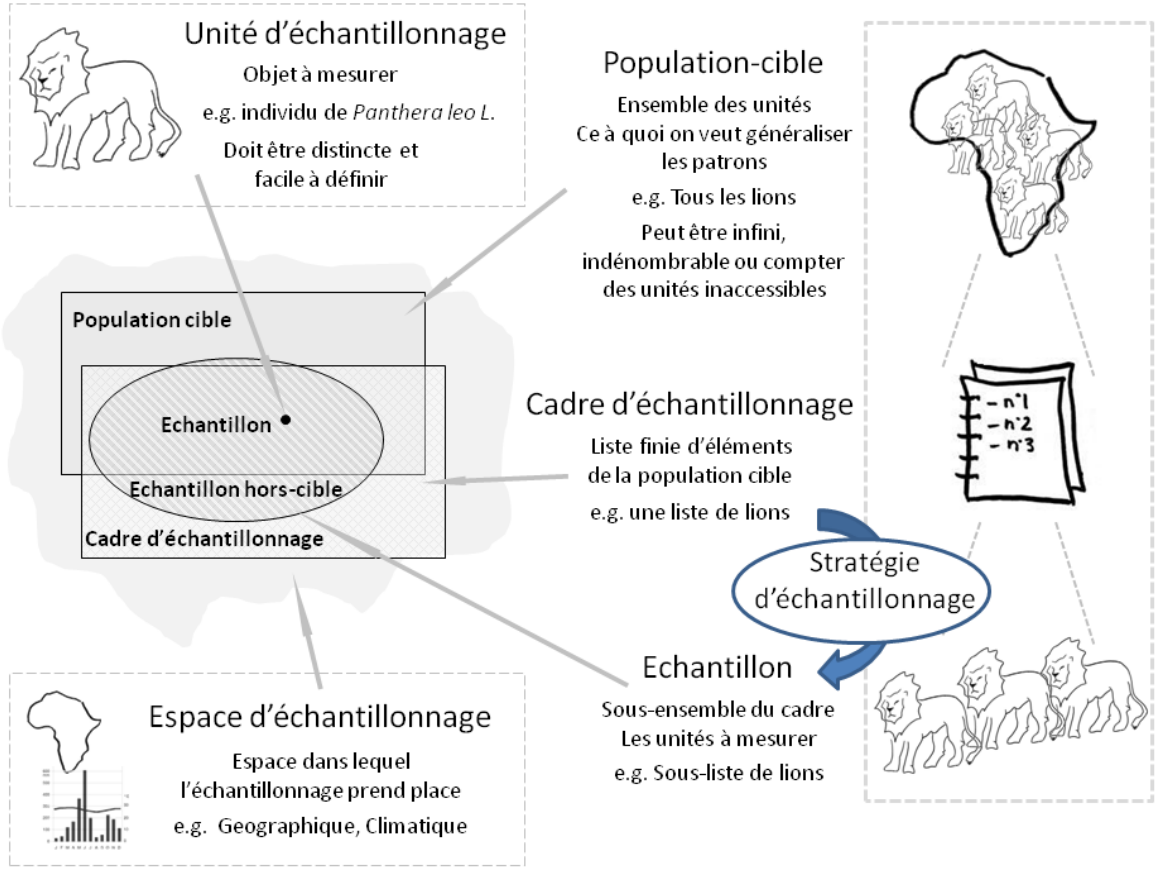


Fig. 1.1. Définitions des concepts d'échantillonnage

1. L'échantillonnage en écologie : vers une mise en œuvre des concepts

// Cette section correspond à l'Article 9

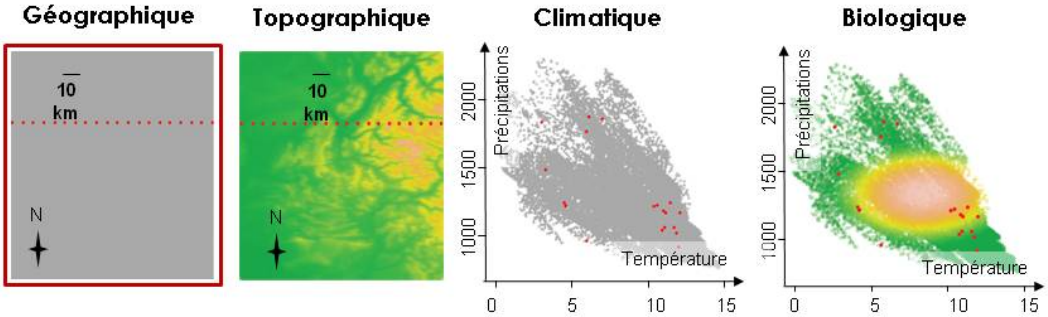
1.1. Pourquoi échantillonner? Raisons et définitions

Répondre à une question écologique nécessite généralement de mesurer une variable d'intérêt (ex. diversité, abondance) sur des objets d'étude (ex. communautés végétales, individus). Or il est le plus souvent impossible de mesurer cette variable pour tous les objets auxquels la question se réfère; il faut alors définir un échantillon représentatif de ces objets - appelés unités d'échantillonnage - dans l'ensemble des possibles - appelé cadre d'échantillonnage (Fig. 1.1). La théorie de l'échantillonnage offre aux écologues un cadre théorique permettant de sélectionner un échantillon représentatif puis d'estimer des paramètres (ex. moyenne, variance) sur la variable biologique étudiée. La méthode de sélection de l'échantillon est appelée stratégie (Fig. 1.1), il en existe deux grands types : probabilistes et non-probabilistes, qui diffèrent par leur fiabilité et par les méthodes d'estimation associées. Dans une stratégie probabiliste, toutes les unités du cadre ont une probabilité non nulle et connue d'être échantillonnées, ce qui conduit à des estimations non biaisées des paramètres recherchés (ex. moyenne, variance). Au contraire les stratégies non probabilistes portent des biais inconnus pouvant fausser les patrons mis en évidence. Le choix d'une méthode d'échantillonnage influence fortement les patrons pouvant être mis en évidence, la fiabilité des estimateurs et la validité des extrapolations (Lemeshow & Levy 1999; Olser *et al.* 1999). L'importance de choisir une stratégie appropriée à la question posée a été soulevée à plusieurs reprises dans la bibliographie écologique concernant des recensements d'espèces rares (ex. papillons, Haddad *et al.* 2008), des questions à retombées économiques (ex. ressources halieutiques, Courbois *et al.* 2008; Hughes & Peck 2008) ou encore des questions de santé publique (ex. radioactivité, Scott *et al.* 2008). Cependant ces études restent centrées sur un type de problème particulier et aucune d'entre elles ne semble aborder des questions plus générales comme le choix de l'espace d'échantillonnage (ex. géographique, climatique, Fig.1 dans Hutchinson 1957; Fig. 1.1 et Encadré 1.1).

Par ailleurs l'échantillonnage développé par les théoriciens et celui pratiqué sur le terrain par les écologues se ressemblent très peu. D'une part les écologues ne sont pas toujours très au fait des subtilités théoriques qui semblent de toute façon incompatibles avec les nombreuses contraintes logistiques rencontrées sur le terrain. De ceci résulte souvent des échantillonnages par transect ou le long des routes (ex. US Breeding Bird Survey, Droege 1990; Peterjohn 1994; Bart *et al.* 1995). D'autre part la théorie s'attache surtout à estimer des paramètres simples (ex. moyenne, variance) et ne donne pas les clés suffisantes pour estimer sans biais des courbes de réponse ou des modèles qui sont au cœur d'un grand nombre de questions écologiques; par exemple si l'on veut - comme dans cette thèse - estimer la réponse

Encadré 1.1. Espace d'échantillonnage

L'échantillonnage commence par la définition d'une liste d'objets à mesurer (Fig. 1.1). L'échantillonnage aléatoire ne nécessite pas d'information supplémentaire. Par contre dès que l'on veut par exemple stratifier, il faut rajouter des variables décrivant les objets à mesurer. Ces variables peuvent être de différents types: géographiques, climatiques, biologiques… L'espace d'échantillonnage est alors un espace multidimensionnel défini par plusieurs variables du même type: par exemple l'espace géographique défini par des gradients de latitude et longitude ou l'espace climatique défini par les gradients de température, radiation,… Ces espaces offrent différentes manières de regarder un même objet (Fig. ci-dessous: transfert d'un espace à l'autre). Une zone d'étude définie dans l'espace géographique (Fig. ci-dessous) pourra être projetée dans l'espace climatique; plusieurs points de l'espace géographique pouvant se superposer dans l'espace climatique (Fig. ci-dessous: points rouges= transect latitudinal). Au contraire un système d'étude défini dans une zone climatique pourra être projeté dans l'espace géographique; un même point de l'espace climatique pouvant représenter plusieurs points de l'espace géographique. Un transect latitudinal peut donc donner un échantillonnage très déséquilibré dans l'espace climatique.



d'un trait fonctionnel à un gradient environnemental. Faire face à ces lacunes, à une époque où il devient très facile d'engranger de grandes quantités de données très rapidement sur de larges échelles spatiales, doit être une priorité. Cela permettrait d'améliorer la fiabilité des estimations et la validité des extrapolations faites à partir des données de terrain; pour cela la qualité doit primer sur la quantité (Bjorklund & Bergek 2009).

1.2. Estimation de courbes de réponse : problèmes et solutions

La fiabilité des estimations de paramètres ou de courbes de réponse peut souffrir d'une mauvaise définition préalable de l'espace et du cadre d'échantillonnage. Dans la pratique il existe deux problèmes majeurs liés à l'estimation de courbes de réponse: (1) *Les gradients indirects* (ex. latitude, altitude) *sensu* Austin (1987) sont des combinaisons inconnues de multiples gradients environnementaux directs ayant un impact direct sur la physiologie des organismes (ex. température, photopériode). Leur utilisation est très fréquente en écologie (Cordell *et al.* 1998; Wegner *et al.* 2003) alors qu'ils ne sont pas transférables dans le temps et l'espace, par exemple une différence d'altitude de 1000 m n'a pas le même impact sur les conditions environnementales dans les Montagnes Rocheuses ou sur le Mont Kenya (Körner 2007; McCain 2009). Ils peuvent donc être utilisés de manière descriptive mais ne sont pas des substituts pertinents des gradients directs. (2) *Les gradients tronqués* – La sélection de l'ampleur et de l'intensité de l'étude sont cruciales et doivent se faire de manière à prendre en compte l'ensemble de la courbe de réponse variable étudiée (Mohler 1983; Austin 1987), car effectuer des estimations à partir de gradients tronqués peut conduire à des inférences erronées (Van Horn 2002; Thuiller *et al.* 2004a ; Fig. 2.13 A).

Ces différentes limitations nous ont amené à réfléchir dans l'Article 9 à un cadre conceptuel unifié pour l'échantillonnage dans les différents domaines de l'écologie. Nous avons en particulier mené une expérience semi-virtuelle illustrant sur un paysage réel les conséquences de l'échantillonnage sur les inférences résultantes. Nous avons construit une variable écologique (ex. taux de croissance, abondance) contrainte par deux gradients de l'espace climatique (températures et précipitations). Dans cette étude le but était (1) d'estimer la moyenne de la variable dans la zone d'étude et (2) de tester l'estimation de l'optimum de la variable par une mesure de distance entre les optima « vrai » et estimé (Encadré 1.2). Nous avons testé en particulier cinq stratégies d'échantillonnage différentes : aléatoire simple, suivant les routes principales, stratifiée par altitude, systématique dans l'espace climatique, basée sur modèle. Pour cette dernière une première étape permettait de faire une estimation préliminaire de l'optimum de la variable (Encadré 1.2) à l'aide d'un échantillonnage aléatoire simple; une deuxième étape suivait une stratégie dérivée de la méthode des surfaces de réponse (Encadré 1.3) centrée sur l'estimation préliminaire de l'optimum. Grâce à cette expérience semi-virtuelle, nous avons montré que l'efficacité de l'échantillonnage sur la

Encadré 1.2. Estimation d'une courbes de réponse ou de son optimum

En faisant l'hypothèse qu'une variable écologique répond de manière unimodale le long de deux gradients orthogonaux U et V, on peut l'estimer par le modèle linéaire (Fig. à droite):

$$z \sim c + a_1 \times U + a_2 \times U^2 + b_1 \times V + b_2 \times V^2 + d \times U \times V$$

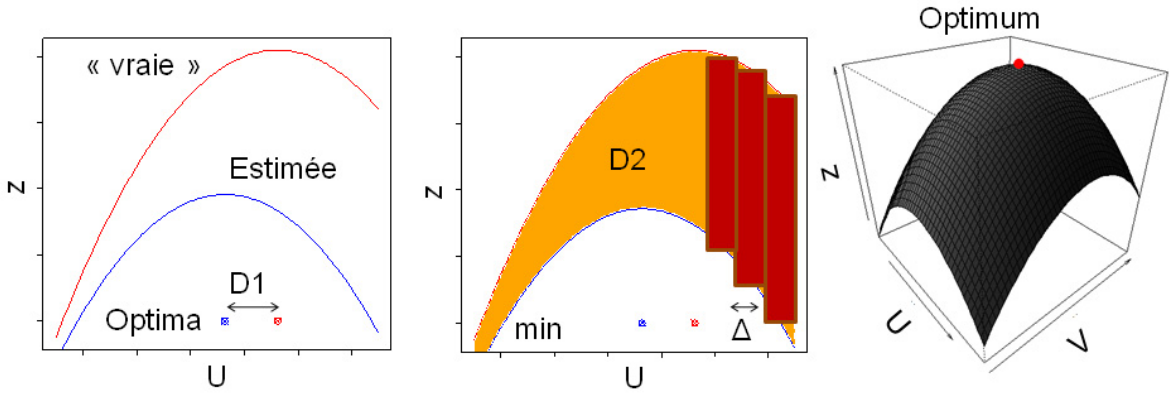
Une fois estimés les différents paramètres (c, a_1, a_2, b_1, b_2 et ab) du modèle, on peut estimer la position de l'optimum (conditions environnementale correspondant à la plus haute/basse valeur de la variable) à partir des dérivées partielles en U et V (Fig. à gauche et à droite):

Optimum.U=$(d*b_1-2*b_2*a_1)/(4*a_2*b_2-d^2)$
Optimum.V=$(d*a_1-2*a_2*b_1)/(4*a_2*b_2-d^2)$

Exemple de deux mesures de la qualité des estimations d'optimum et de courbe:

D1- La distance Euclidienne entre les optima "vrai" et estimé, illustrée ci-dessous à gauche sur un unique gradient, donne une mesure de la précision et du biais de l'estimation de l'optimum.

D2- La « distance » entre les courbes (resp. surface) de réponse « vraie » et estimée est la surface (resp. volume) entre ces deux courbes (resp. surfaces). On peut l'estimer à l'aide d'une somme de Riemann en approximant la surface (resp. volume) par une somme d'aires de rectangles (resp. pavés) définis sur k intervalles:
$\Delta * \Sigma$ [courbe.vrai(min+k*Δ)- courbe.estimée(min+k*Δ)]
, où min est le bas du gradient étudié et Δ la largeur des k intervalles (fig. au milieu).



précision des estimations résultantes dépendait de la question posée et des paramètres à estimer. Un échantillonnage aléatoire permettait d'estimer avec une bonne précision la moyenne de la variable; tandis que pour l'estimation de l'optimum de la variable, la stratégie basée sur modèle, dérivée de la méthode des surfaces de réponse était la plus appropriée (Fig. 1.2). Il n'existe donc pas d'échantillonnage type et celui-ci doit donc être construit en fonction de la question posée. Les éléments de l'échantillonnage (unités, population cible, cadre et espace d'échantillonnage) doivent en particulier être définis et explicités avec soin pour permettre la mise en œuvre de la stratégie la plus appropriée.
De plus nous avons vu que pour estimer des relations entre une variable écologique et des gradients environnementaux (ex. climatiques), il est nécessaire que le cadre d'échantillonnage englobe l'ensemble de la courbe de réponse pour éviter les erreurs liées aux *gradients tronqués* (1.2). Ceci requiert une connaissance *a priori* du système d'étude et des hypothèses fortes sur les résultats attendus pour mener à des inférences fiables et généralisables. Une telle connaissance *a priori* peut se faire grâce à des avis d'experts ou des données de bibliographie, grâce à un échantillonnage préliminaire (comme celui proposé dans notre expérience virtuelle) ou grâce à la construction d'expériences virtuelles permettant de simuler différentes stratégies, leur coût, et leur impact sur les estimations attendues. Pour le moment la mise en place de telles expériences virtuelles a été largement sous-utilisée en écologie, elle n'a été mise en œuvre que dans de rares cas d'utilisation de modèles d'habitat (Chapitre 3) pour affiner une stratégie de rééchantillonnage (Brito *et al.* 1999; Guisan *et al.* 2006).

2. Base de données Guisane 2080 : réalités de l'échantillonnage

Nous nous sommes servis de ces réflexions conceptuelles pour construire un échantillonnage orienté sur la question posée et nous avons pris les principales décisions à l'aide d'exemples virtuels, comme proposé ci-dessus (voir 1.2).
Ici je parlerai de base de données par extension de vocabulaire pour parler de l'ensemble complexe des données récoltées sur le terrain, mais ces données ne sont pas réellement structurées comme une base de données. Elles seront intégrées par la suite dans la base de traits de la végétation des Alpes (Androsace, qui regroupe les bases de données LEDA, Knevel *et al.* 2003 ; BioFlor, Kühn *et al.* 2004 ; VISTA, Garnier *et al.* 2007; Diversitalp, Thuiller Unpublished).

2.1. Objectifs

L'un des principaux objectifs de ma thèse était de quantifier et d'analyser la structure et les vecteurs de la variabilité fonctionnelle intraspécifique (Chapitre 2) et de faire le lien entre

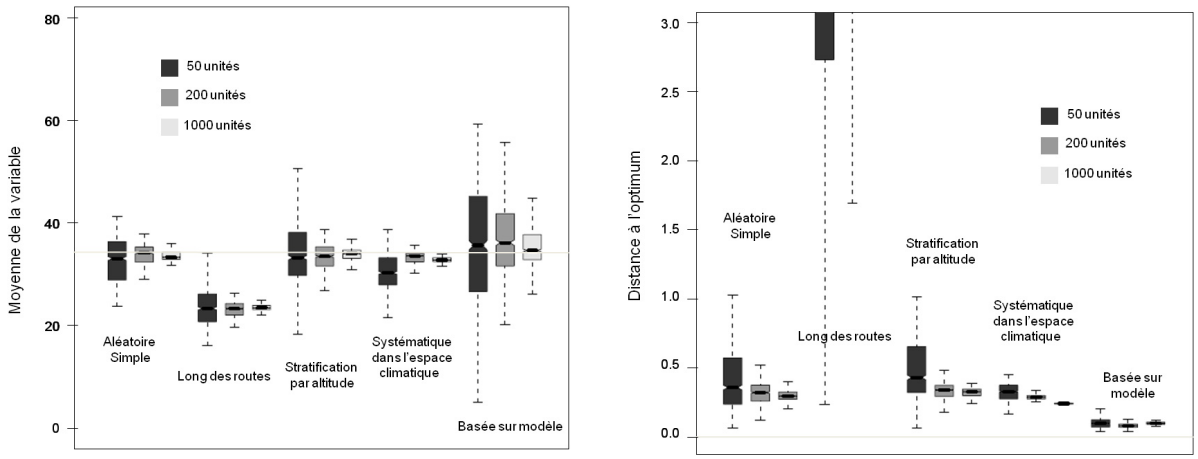


Fig. 1.2. Effets de la stratégie d'échantillonnage et de la taille de l'échantillon sur la précision de l'estimation des paramètres

Illustration de l'effet de la stratégie d'échantillonnage sur la précision et le biais des estimations de deux paramètres biologiques: (a) la moyenne de la variable et (b) la distance D1 entre les optima « vrai » et estimé (cf Encadré 1.2); l'optimum « vrai » correspondant aux conditions environnementales optimales pour la variables.
Les résultats sont donnés pour cinq stratégies différentes: (i) aléatoire simple, (ii) le long des routes, (iii) stratifié par altitude, (iv) systématique dans l'espace climatique et (v) basé sur modèle.
Les graphiques sont des boxplots des distributions des valeurs estimées pour 1000 simulations avec chaque taille d'échantillon (50, 200, 1000) ; ils donnent en particulier la médiane et les premier et troisièmes quartiles des distributions.

(à gauche)- Distribution des estimations de la moyenne de la variable sur la zone d'étude. La ligne grise correspond à la vraie valeur sur l'ensemble de la zone d'étude. Cette moyenne est contingente à l'aire étudiée.

(à droite)- Distribution des distances D1 estimées entre les optima « vrai » et estimé. La ligne grise correspond à 0, la meilleure valeur attendue.

cette variabilité, la performance des individus et la niche de l'espèce considérée (Chapitre 3). Cet objectif nécessitait donc d'estimer des moyennes et variances de traits fonctionnels pour différentes populations d'une même espèce, mais aussi des courbes de réponse des traits le long des gradients environnementaux et le long de la niche de chaque espèce. Les données de la bibliographie et des bases de données ne suffisant pas à mener de telles analyses (Chapitre 2), il a été nécessaire de mettre en place un échantillonnage pour collecter de nouvelles données.

L'idée était de mettre en place une stratégie unique permettant d'évaluer et de comparer les courbes de réponse des traits aux gradients pour un nombre substantiel d'espèces (voir § 2. 3) structurantes des communautés végétales. Nous avons choisis d'effectuer notre étude en milieu alpin. Les gradients alpins se prêtaient particulièrement bien à l'étude de la variabilité intraspécifique étant donné qu'ils offrent sur des distances courtes des conditions environnementales très contrastées contraignant l'expression phénotypique des espèces (Körner 1999). Nous avons fait l'hypothèse que les traits pourraient répondre de manière unimodale aux gradients environnementaux. Cette hypothèse repose sur le fait que de nombreuses études en écologie ont mis en évidence des optima dans les relations entre performance des espèces et gradients (Austin 1987; Gaston 2003). Tester cette hypothèse permet implicitement de tester l'hypothèse plus simple selon laquelle les traits pourraient répondre de manière linéaire aux gradients. Par contre cela ne permet pas de tester des hypothèses plus complexes (ex. relations bimodales ou unimodale asymétrique) qui nécessiteraient un échantillonnage plus complexe (Encadré 1.3). Un trait peut alors s'exprimer le long de deux gradients G1 et G2 selon (Encadré 1.2):

$$Trait \approx a1 \times G1 + a2 \times G1^2 + b1 \times G2 + b2 \times G2^2 + ab \times G1 \times G2 + c \quad (1)$$

où c est la constante du modèle, a_i les paramètres à estimer pour le gradient *G1* et b_i pour le gradient *G2*, *ab* étant le paramètre à estimer pour l'interaction entre les deux gradients.

2.2. Sites d'étude et choix des gradients

2.2.1. Deux sites d'étude

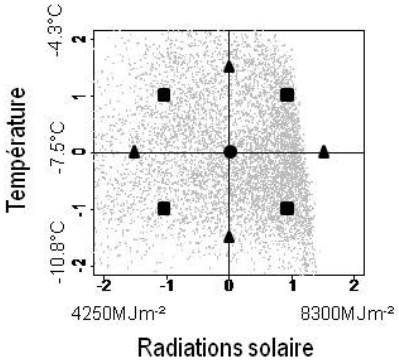
Le site principal de la Guisane (6°21' à 6°40' E; 44°5' à 45°4' N) est situé au cœur des Alpes françaises (05) et s'étend de Briançon (1200 m) à Combe Roche Noire (2600 m, Fig 1.3). Ce site a été choisi pour son hétérogénéité environnementale et sa proximité avec la Station Alpine Joseph Fourier (http://sajf.ujf-grenoble.fr/), nécessaire au travail en laboratoire et l'hébergement de longue durée. Ce site se caractérise par un climat très hétérogène avec des températures annuelles moyennes allant de 0°C à 6,3°C et des précipitations annuelles allant de 600 mm à 1200 mm. Les principaux substrats y sont le calcaire et le schiste.

Encadré 1.3. Méthode des surfaces de réponse (RSM)

La méthode des surfaces de réponse (RSM pour « Response Surface Methodology ») est une méthode d'échantillonnage utilisée pour évaluer la réponse d'une variable à la variation simultanée de plusieurs facteurs orthogonaux (Box & Draper 1959, Menke 1973, Clancy & King 1993). Partant du principe que l'effort d'échantillonnage est limité (temps, effort, budget), cette méthode recouvre un large potentiel de plans expérimentaux pour trouver des conditions optimales. Elle a été largement utilisée en physique (Menke 1973) et chimie (Adinarayana 2002) mais très peu en écologie (Inouye 2001).
Cette méthode requiert une connaissance *a priori* de la forme de la réponse (linéaire, quadratique, cubique) et fait le postulat que deux niveaux de chaque facteur impliqué suffisent à déterminer une réponse linéaire et trois pour déterminer une réponse quadratique. Le plan d'échantillonnage classique pour déterminer une relation quadratique le long de deux facteurs orthogonaux de type:

$$z \sim c + a_{1x}U + a_{2x}U^2 + b_{1x}V + b_{2x}V^2 + d_xU_xV$$

est le plan factoriel central composite, il comprend un point central permettant de déterminer c (rond), quatre points axiaux permettant de déterminer les effets propres à chaque facteur (a1, b1, a2 et b2; triangles) et quatre points sur les bissectrices des facteurs permettant de déterminer l'interaction entre les facteurs (d, carrés). L'exemple est donné avec U= Radiation solaire et V=Température.



Ce site est le site principal et une grande majorité des explications et des exemples s'y rapportent. Les mesures de terrain en Guisane devant aussi servir à la calibration d'un modèle de paysage sur la vallée et à la prédiction de l'évolution future de ce paysage en fonction de scénarii climatiques et d'utilisation des terres ébauchés jusqu'en 2080, le projet a par conséquent été présenté aux collectivités locales sous le nom de : « Guisane 2080 ».
Par ailleurs dans le cadre d'une collaboration avec Antoine Guisan et Anne Dubuis (Université de Lausanne) le travail de terrain a été répliqué sur un second site en Suisse nommé Anzeindaz (6°60′ à 7°10′ E; 46°10′ à 46°30′ N) pour permettre de conforter les résultats. Ce site se trouve dans la partie ouest des Alpes suisses au nord ouest de Lac Léman et s'étage entre 1070 m et 2310 m. Le climat de ce site est lui aussi très fortement hétérogène, il s'étale dans la même fourchette de températures annuelles moyennes que le site de la Guisane (-3,5°C à 8°C) mais est nettement plus humide (précipitations annuelles : 1300 mm à 2300 mm). Le substrat y est principalement calcaire.
Les deux sites englobent des formations végétales variées comprenant des forêts de conifères (ex. *Pinus sylvestris* et *P. uncinata*, *Larix decidua*), des landes (ex. *Rhododendron ferrugineum*, *Vaccinium myrtillus* et *V. uliginosum*), des prairies subalpines (ex. *Festuca paniculata*) et des pelouses alpines (ex. *Carex sempervirens*, *Nardus stricta*). Les zones fortement anthropisées (villes et villages) ont été retirées d'office des zones d'étude.

2.2.2. Caractériser les principaux gradients de la vallée de la Guisane

Afin de caractériser la vallée de la Guisane, nous avons collecté vingt-cinq variables environnementales de source et de résolution différentes (Tableau 1.1): seize variables topographiques et climatiques à une résolution de 50 x 50 m amicalement partagées par C. Piedallu de l'ENGREF Nancy ; deux variables topographiques dérivées des premières avec ArGIS® ; cinq variables topographiques et climatiques à une résolution de 200 x 200 m amicalement partagées par Niklaus Zimmermann du WSL Zurich ; une variable grossière de texture du sol provenant de la base de données des sols d'Eurasie (Soil European bureau - Joint Research Centre-Ispra) ; et une variable de couverture végétale dérivée de Corine Land Cover à 100 x 100 m (http://www.ifen.fr/donIndic/ Donnees/corine).

A partir de ces données, nous avons cherché à stratifier les conditions environnementales de la vallée. (1) Nous avons tenu à privilégier l'utilisation de gradients environnementaux directs afin que les courbes de réponse trouvées soient comparables à d'autres et projetables dans le temps et l'espace. (2) Nous avons privilégié les données climatiques et topographiques fines comme étant potentiellement celles influant le plus sur la végétation en milieu alpin. (3) Nous avons choisi de sélectionner uniquement deux gradients par souci de simplicité et de faisabilité. (4) Nous avons orienté notre sélection vers deux gradients orthogonaux.

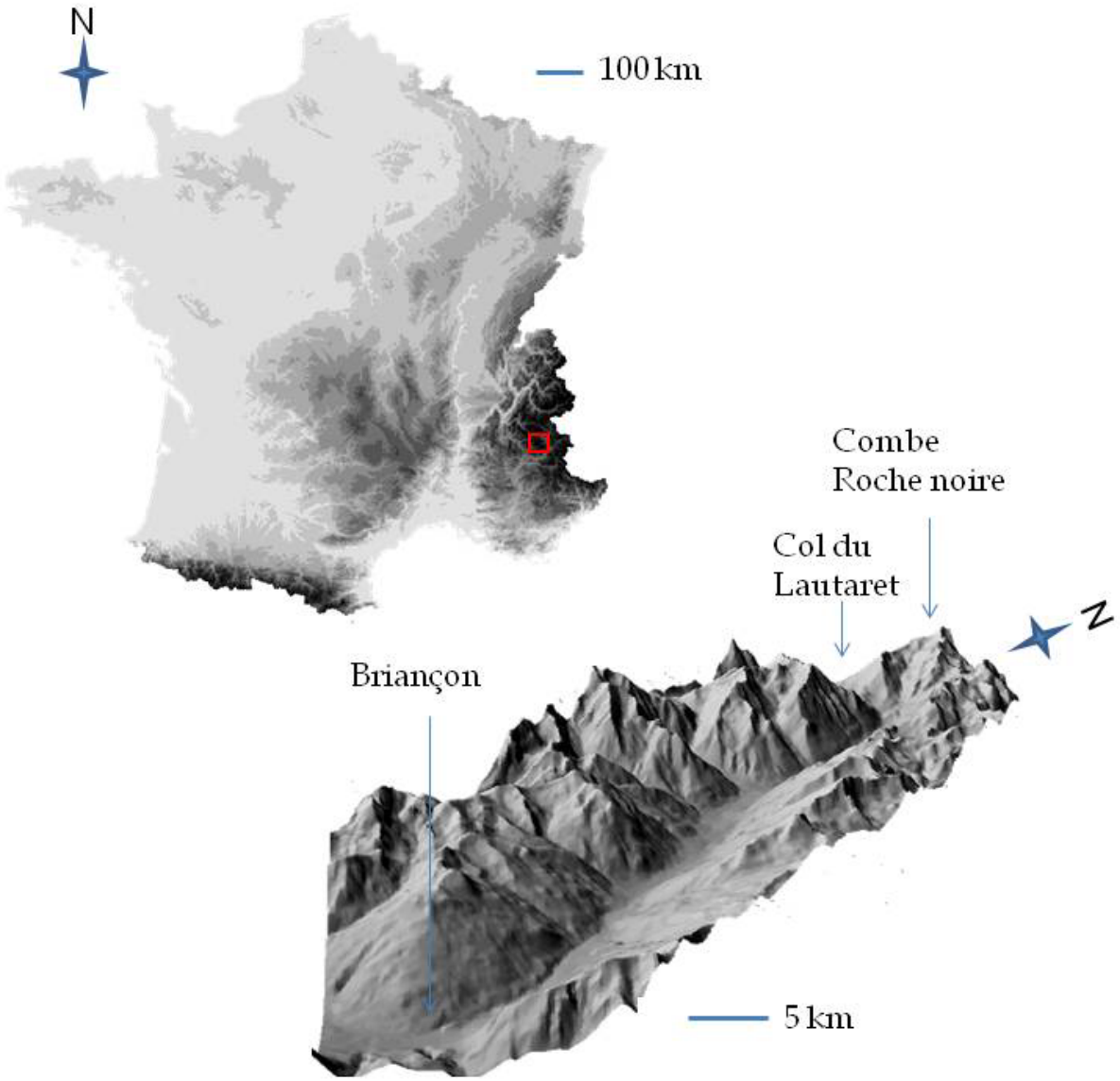


Fig. 1.3.Site d'étude de la vallée de la Guisane
En haut : Localisation de la vallée par rapport à la France entière (carré rouge), le dégradé de gris indique l'altitude au dessus du niveau de la mer : gris clair pour les basses altitudes et noir pour les hautes.
En bas : le zoom sur la vallée est une vue en 3D depuis l'Est. Le dégradé de gris indique les radiations solaires au mois d'aout : gris clair pour les forts niveaux de radiation et gris foncé pour les faibles niveaux de radiation.

L'orthogonalité des gradients étant en effet nécessaire pour rendre possible la discrimination des effets des deux gradients et de leur interaction (*Equation (1)*).
Suivant ces quatre critères, nous avons effectué une analyse en composante principale (ACP) sur un jeu condensé de huit variables extraites avec ArcGIS® sur la zone d'étude à 50 x 50 m (Fig. 1.4). Nous avons choisi les deux gradients les plus orthogonaux (pas nécessairement par rapport à leur corrélation avec les axes) dans le premier plan de l'ACP (43 et 35 % de la variabilité expliquée par les axes 1 et 2): la température minimale en hiver (Tminh) et les radiations solaires du mois d'août (Rad8). La zone d'étude géographique (Fig. 1.3) a ensuite été projetée dans l'espace climatique définit par Tminh et Rad8 (Fig. 1.7) ; les pixels de la zone d'étude à 50 x 50 m n'y sont pas répartis de manière uniforme, une grande majorité de ceux-ci se trouve dans des conditions de fortes radiations solaires en août et de températures intermédiaires.
Pour Anzeindaz, un cheminement similaire a conduit à la sélection de la température moyenne en janvier et des radiations en août à 25 x 25 m (Zimmermann & Kienast 1999) pour stratifier le climat de la vallée.

2.3. Seize espèces modèles

2.3.1. Choix des espèces

Le choix des espèces modèles pour notre étude a reposé sur différents critères biologiques et logistiques. (1) Nous avons orienté notre choix pour obtenir une quinzaine d'espèces pérennes, ce qui nous semblait un bon compromis entre le fait d'avoir un grand nombre d'espèces rendant les analyses plus robustes et la faisabilité sur le terrain. (2) Les espèces devaient être des espèces communes, c.à.d. que l'on puisse trouver fréquemment dans la vallée et dans des conditions contrastées. (3) Elles devaient aussi être des espèces structurantes, c.à.d. avoir un effet sur la structure de la communauté végétale et du paysage attenant, au moins sur une partie de leur distribution. Trois espèces supplémentaires ont été choisies en 2008 comme des espèces communes et non structurantes afin de voir si cela amenait à des résultats différents (*). (4) Afin de représenter les espèces structurantes des différents types de végétation de montagne (forêts, landes, prairies, pelouses) et d'obtenir des patrons généralisables (au final plus on représente de forme de vie, moins celles-ci sont représentées, notre dispositif n'est donc pas puissant pour comparer les formes de vie), les espèces choisies devaient appartenir à différentes formes de vie : graminoïdes, herbacées, légumineuse, ligneux hauts et bas, à aiguilles ou feuilles larges et décidus ou sempervirents. (5) Les espèces devaient différer par leurs exigences écologiques. Il est en effet possible de penser *a priori* que des espèces que l'on échantillonne par exemple dans des conditions plus « froides » que leurs conditions optimales, ou au contraire plus « chaudes » ou encore autour

Types	Variables	Résolution (m^2)	Catégoriel	Définition	Sélection
Topographique	X			Latitude (m)	
	Y			Longitude (m)	
	Mnt50	50x50		Modèle digital d'altitude (mètres au dessus du niveau de la mer)	
	Slope50	50x50		Pente (degrés entre 0 et 90)	
	Asp50	50x50		Orientation (degrés sens des aiguilles d'une montre depuis le Nord)	
	Flowdir	200x200	X	Direction potentielle des flux d'eau de chaque cellule vers ses voisines (degrés)	
	Flowacc	200x200	X	Accumulation potentielle de l'eau dans chaque cellule	
	Twi	200x200		Index topographique d'humidité	X
	Positopo	50x50	X	Positions topographiques, vues sous l'angle du ruissellement.	
	Toposcale	200x200		Exposition topographique (convexité, concavité)	X
Climatique	Long-season	50x50	X	Nombre de mois pendant lesquels l'évapotranspiration de Turc est positive	
	Rad1	50x50		Radiations solaires en Janvier ($kJ.m^{-2}.jour^{-1}$)	
	Rad8	50x50		Radiations solaires en Août ($kJ.m^{-2}.jour^{-1}$)	X
	Dd5	200x200		Cumul sur l'année des températures journalières au delà de 5.56°C	
	Stsup0	50x50		Cumul sur l'année des températures journalières au delà de 0°C	X
	Frost_e	200x200		Nombre de jours de gel	
	Tmoy8	50x50		Température moyenne en Août (°C)	
	Tmoy1	50x50		Température moyenne en Janvier (°C)	
	Tminh	200x200		Température minimale en hiver (°C x 100)	X
	PPwinter	50x50		Somme des précipitations de Janvier à Mars (mm)	
	PPsummer	50x50		Somme des précipitations de Juillet à Septembre (mm)	
	Dhctu7	50x50		Déficit hydrique en Juillet	X
	Etptu7	50x50		Evapotranspiration de Turc en Juillet	X
	Gams	50x50		Indice de continentalité hydrique	X
	Demartone	50x50		Indice d'aridité	
Sol & végétation	CorineLC	100x100	X	Types de couverture végétale	
	Soil	1000x1000	X	Texture du sol	

Tableau 1.1. Variables environnementales pour caractériser les principaux gradients de la vallée de la Guisane

de leurs conditions optimales ne vont pas répondre aux gradients de manière similaire. Nous avons donc choisi des espèces subalpines, alpines et d'autres plutôt de plaine. De plus les espèces choisies devaient aussi si possible (6) être des espèces « connues » sur lesquelles il existe déjà une connaissance dans la bibliographie ou au laboratoire; (7) être des espèces facilement reconnaissables pour éviter les ambiguïtés taxonomiques ; (8) être des espèces le plus possible non apétantes pour limiter les effets du pâturage ; (9) avoir des feuilles «mesurables», c.à.d. pas des feuilles très petites ou très découpées afin de faciliter les mesures de traits sur des dizaines d'individus. Nous avons aussi décidé de nous affranchir du contexte phylogénétique en ayant 16 espèces réparties dans 11 familles différentes.

Pour représenter les herbacées nous avons choisi, avec l'aide de Pierre Salomez (responsable flore du Parc National des Ecrins) et Rolland Douzet (botaniste à la Station Alpine Joseph Fourier), quatre graminoïdes : *Festuca paniculata Schinz & Thell.* (FP), *Dactylis glomerata L.* (DG)*, *Sesleria caerulea (L.) Ard* (SC)*, *Carex sempervirens Vill.* (CS)*; quatre dicotylédones non légumineuses: *Polygonum viviparum L.* (PV), *Geum montanum L.* (GM), *Silene nutans L.* (SN) et *Leucanthemum vulgare Lam.* (LV); et une légumineuse : *Trifolium alpinum L.* (TA). Pour les ligneux bas nous avons choisi deux arbustes nains alpins: *Dryas octopetala L.* (DO)* et *Salix herbacea L.* (SH); deux ligneux bas subalpins sempervirents, un à larges feuilles: *Rhododendron ferrugineum L.* (RF) et un conifère : *Juniperus communis L.* & *sibirica Lodd.* (J) et un ligneux bas décidu à larges feuilles: *Vaccinium myrtillus L.* (VM)*. Pour les ligneux hauts, nous avons sélectionné deux conifères, un sempervirent : *Pinus sylvestris L.* & *uncinata Ramond* (P) et un décidu : *Larix decidua Mill.* (LD). Toutes ces espèces ont été échantillonnées en Guisane et seulement celles marquées d'un astérisque à Anzeindaz.

Nous avons utilisé deux agrégats taxonomiques (*Juniperus* et *Pinus*) à cause de l'apparente hybridation des espèces dans la zone d'étude (comm. pers. Pierre Salomez). Nous avons vérifié après coup que les distributions des traits de ces agrégats étaient bien unimodales, validant l'idée de continuums fonctionnels.

Bien évidemment cette liste finale (Tableau 1.2) est loin d'être parfaite et certaines espèces comme *D. glomerata* ou *V. myrtillus* semblent présenter des écotypes. Cependant cette liste combine au mieux les différents critères ci-dessus pour pouvoir permettre de remplir nos objectifs. De plus les treize espèces non-arborées font partie des cent quinze espèces qui représentent la majorité (plus de 80% en couverture) de la biomasse des prairies et landes en Guisane.

Pour chacune des espèces sélectionnées, l'ensemble des populations existant en vallée de la Guisane constitue la population-cible de l'échantillonnage. Dans la suite nous appellerons « population » (sens biologique) l'ensemble des individus d'une même espèce présents dans les stations étudiées (Fig. 0.1), même si ce n'est pas *sensu stricto* une population au sens de la génétique des populations.

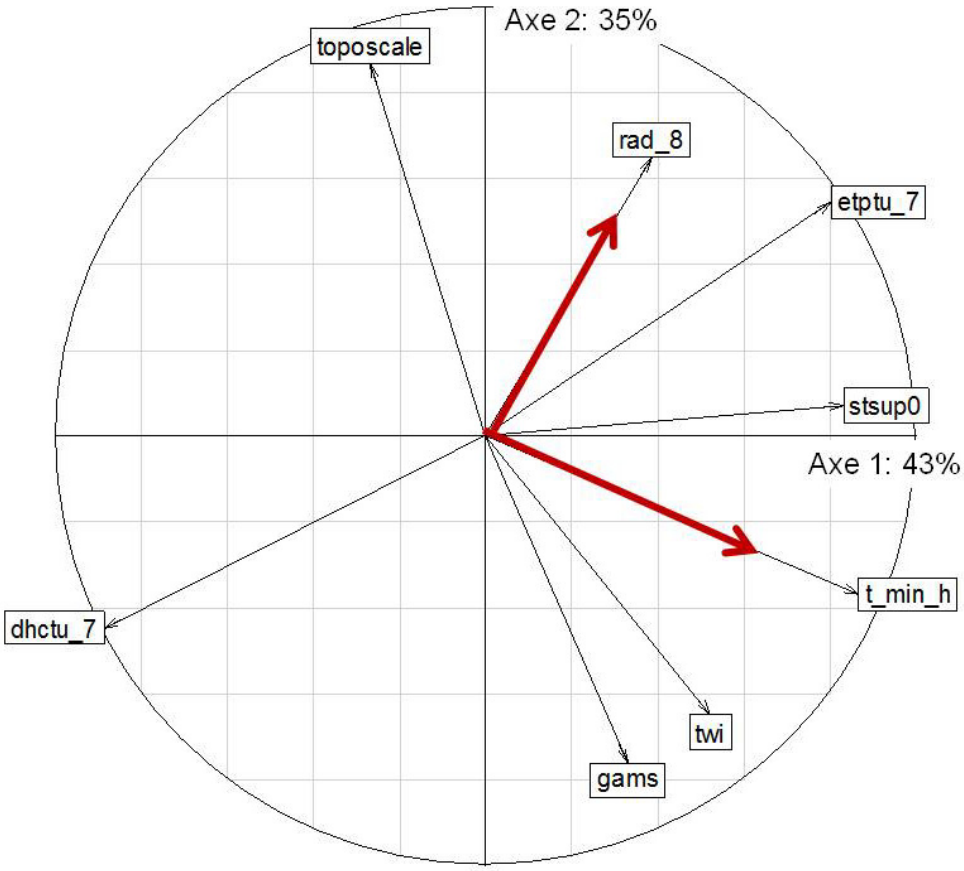


Fig. 1.4. Cercle de corrélation de l'ACP sur 8 données environnementales.
Les deux gradients directs les plus orthogonaux sont Tminh et Rad8.

2.3.2. Données de présence et cadre d'échantillonnage

Afin de ne pas mener notre échantillonnage au hasard, nous nous sommes appuyés sur des données phyto-sociologiques pour répertorier pour chaque espèce sélectionnée les stations de présences connues. En Guisane, nous avons utilisé les 500 relevés exhaustifs de végétation (10 x 10 m) géo-référencées (à 100 m près) du Conservatoire Botanique National Alpin (CBNA). La nomenclature de ces relevés suit l'index synonymique de la flore de France proposé par Kergélen (1993). A Anzeindaz, nous avons utilisé une base de 550 relevés de végétation (8 x 8 m) choisies selon une stratégie aléatoire-stratifiée hors forêts. Ces listes de présences connues pour chaque espèce définissent le « cadre » de notre échantillonnage.

2.4. Mise au point d'une stratégie d'échantillonnage

2.4.1. Effort d'échantillonnage: étude virtuelle préliminaire

Principe

Cette première étude avait pour but de tester l'influence de l'effort d'échantillonnage sur la précision des patrons estimés. Le système d'étude virtuel a été construit pour simuler la réponse quadratique d'un trait (ex. la hauteur végétative H) à un gradient de température (T) s'échelonnant de 3 à 15°C. Construire un tel système défini par une fonction H(T) nécessite de faire une hypothèse sur la forme de la réponse (ici quadratique avec un optimum à 11,25°C) et son ampleur (hauteur entre 30 et 50 cm), mais aussi sur la variabilité du trait au sein des populations (ici distribution normale de moyenne H(T) et d'écart-type sd= 0,5, 2 ou 4). Simuler l'échantillonnage de ce système nécessite ensuite de choisir un nombre de populations (p de 5 à 50) à échantillonner, qui se décompose en nombre de niveaux de température (t) et nombre de répétitions pour chaque niveau (r ici 1) et un nombre d'individus (i de 3 à 30) à mesurer dans chaque populations. Les paramètres p et i constituent l'effort d'échantillonnage, c.à.d. le nombre de mesures qui vont être effectuées (p x i). Il faut ensuite décider de la manière de répartir cet effort, c.à.d. choisir une stratégie d'échantillonnage (ici un transect de température). A chaque échantillonnage virtuel, nous avons estimé avec un modèle linéaire ($H \sim T+T^2$) la réponse de la hauteur à la température. Afin d'évaluer l'efficacité des différents efforts d'échantillonnage sur la précision des patrons estimés, nous avons choisi deux mesures: la distance entre les optima estimé et vrai, c.à.d. l'écart de température entre les deux (D1, Fig. 1.6) et la distance entre les courbes de réponse estimée et vraie, c.à.d. la surface entre ces deux courbes (D2) (Encadré 1.2). Chaque type d'échantillonnage a été répliqué cent fois.

Résultats & discussion

Augmenter la précision des patrons estimés nécessite d'augmenter le nombre de populations échantillonnées (p) ou le nombre d'individus mesuré par population (i) (Fig. 1.5). Pour avoir

Species	Label	Family	Life Form	Sites	Individuals	Leaves	LNC	Niche 1	Niche 2	Sampling	Spe	Environmental requirements and Sampling
Carex sempervirens Vill.	CS	*cyperaceae*	G	12	108	1	SP	897	1990	899	0.54	
Dactylis glomerata L.	DG	*poacea*	G	10	90	1	I	1448	1073	624	0.62	
Dryas octopetala L.	DO	*rosaceae*	S	15	135	10	I	771	2040	878	0.52	
Festuca paniculata (L.) Schinz & Thell	FP	*poacea*	G	16	144	1	I	605	2040	831	0.48	
Geum montanum L.	GM	*rosaceae*	H	17	153	1	I	781	2308	760	0.46	
Juniperus communis L. / sibirica Lodd. Ex Burgsd	J	*cupressaceae*	S	14	126	10	SP	1597	1504	1156	0.60	
Larix decidua Mill.	LD	*pineaceae*	T	18	162	10	I	870	1883	747	0.62	
*Leucanthemum vulgare Lam. **	LV	*asteraceae*	H	12	108	1	SP	1385	1278	920	0.53	
Pinus sylvestris L. / uncinata Ramond ex.DC	P	*pineaceae*	T	9	81	10	SP	1329	1221	720	0.59	
*Polygonum viviparum L.**	PV	*polygonaceae*	H	12	108	1	SP	770	2247	525	0.57	
Rhododendron ferrugineum L.	RF	*ericaceae*	S	7	63	10	I	737	1981	626	0.55	
Salix herbacea L.	SH	*salicaceae*	S	7	63	10	SP	492	2471	329	0.45	
Sesleria caerulea (L.) Ard	SC	*poacea*	G	11	99	1	SP	1456	1748	916	0.61	
*Silene nutans L. **	SN	*caryophyllaceae*	H	12	108	1	SP	1386	1621	895	0.58	
Trifolium alpinum L.	TA	*fabaceae*	H	13	117	1	SP	585	2279	871	0.43	
Vaccicium myrtillus L.	VM	*ericaceae*	S	14	126	10	I	1004	1683	719	0.60	
Total	**16**	**11**	**4**	**199**	**1791**	**8595**	**1179**	-	-	-	-	0 m.a.s.l. 2000 4000

Tableau 1. 2. Caractéristiques des seize espèces étudiées

Pour chacune des espèces : leur nom (les espèces ont été échantillonnées en 2007, excepté celles avec un astérisque qui ont été échantillonnées en 2008), leur code, leur famille, leur forme de vie (G : graminoïdes, H : autres herbacées, S : ligneux bas, T= arbre), le nombre de populations échantillonnées, le nombre d'individus mesurés, le nombre de feuilles mesurées par individu (1 ou 10), la manière dont la teneur des feuilles en azote (LNC) a été mesurée (SP : regroupement par sous-plot, I : individuelle) ;

Niche1 : amplitude altitudinale de la distribution de l'espèce (intervalle inter-quantiles 10-90%), d'après les données de 2.3.2. ; Niche 2 : l'altitude optimale de l'espèce (mode de la distribution altitudinale des populations connues, données de 2.3.2.) ; Sampling : l'amplitude altitudinale échantillonnée, Spe : l'indice de spécialisation de Fridley.

Les "préférences environnementales" décrivent la distribution de l'espèce entre 0 et 4000 m :

(1) les rectangles blancs indiquent la distribution altitudinale totale de l'espèce d'après les données de 2.3.2. (2) les rectangles hachurés indiquent la limite inter-quantiles 10-90 % de la distribution altitudinale ; (3) les rectangles noirs indiquent la zone échantillonnée.

une précision de +/- 0,2°C sur l'estimation de l'optimum de température (D1), la solution la moins coûteuse en termes d'effort d'échantillonnage (sd=4) est d'échantillonner une dizaine de populations représentées par une dizaine d'individus (Fig. 1.5 et correspond à D2<5). Plus le trait est variable au sein des populations, moins la précision est grande ; ce même effort d'échantillonnage donnera en moyenne une précision de +/- 0,1°C sur D1 pour sd=2. Il est toutefois important de noter que cette évaluation dépend certainement fortement de la forme de la réponse et surtout de la position de l'optimum le long du gradient (Fig. 1.6), comme suggéré par Mohler (1983). Nous avons orienté notre échantillonnage afin de mesurer une dizaine de populations par espèce et une dizaine d'individus par population, ce qui semblait être un compromis acceptable entre faisabilité et précision.

2.4.2. Choix d'une stratégie: test sur une surface semi-virtuelle

Principe

Cette étude virtuelle avait pour but de tester l'influence de la stratégie d'échantillonnage sur la précision des patrons estimés. Le système d'étude virtuel a été construit pour simuler la réponse quadratique d'un trait (ex. la hauteur végétative H) aux deux gradients d'échantillonnage normalisés (voir 2.2.2): Tminh s'étale de -3 à +3 écarts types autour de la moyenne et Rad8 de -4,6 à 1,6. Nous avons choisi une réponse du trait (de 20 à 50) quadratique le long des deux gradients (optimum au point (+0,5 ; +0,5)). Au sein de chaque population le trait suit une distribution normale de moyenne H (Rad8, Tminh) et d'écart type sd=4. Comme nous avons montré que la méthode des surfaces de réponse était appropriée à l'étude de courbes de réponse et d'optimum (voir 1.2), nous avons testé douze stratégies dérivant de cette méthode et différant par le nombre de populations échantillonnées et la répartition de celles-ci dans l'espace climatique (Rad8, Tminh, Fig. 1.7). Chaque stratégie a été simulée cent fois avec dix individus mesurés par population. A chaque échantillonnage virtuel, nous avons estimé avec un modèle linéaire ($H \sim Tminh + Tminh^2 + Rad8 + Rad8^2$) la réponse de la hauteur aux gradients. Afin d'évaluer l'efficacité des différentes stratégies sur la précision des patrons estimés, nous avons utilisé les mêmes mesures que précédemment : la distance D1 entre les optima estimé et vrai et le volume D2 séparant les surfaces de réponse estimée et vraie (Encadré 1.2).

Résultats & discussion

Les différentes stratégies que nous avons testées sont centrées sur le point (0 ; 0), c.à.d. les conditions moyennes (ex. E1->E8) ou sur (0,5 ; 0,5) qui correspond à l'optimum (ex. E9->E12); elles s'étalent autour du centre jusqu'à +/- 1 écart-type autour de la moyenne (ex. E1), ou plus largement (ex. E12); elles se contentent des neuf points nécessaires à la méthode des surfaces de réponse (Encadré 1.3 ; ex. E1) ou tentent de couvrir d'avantage d'hétérogénéité environnementale en échantillonnant plus de populations (ex. E4, E8 ; Fig. 1.7).

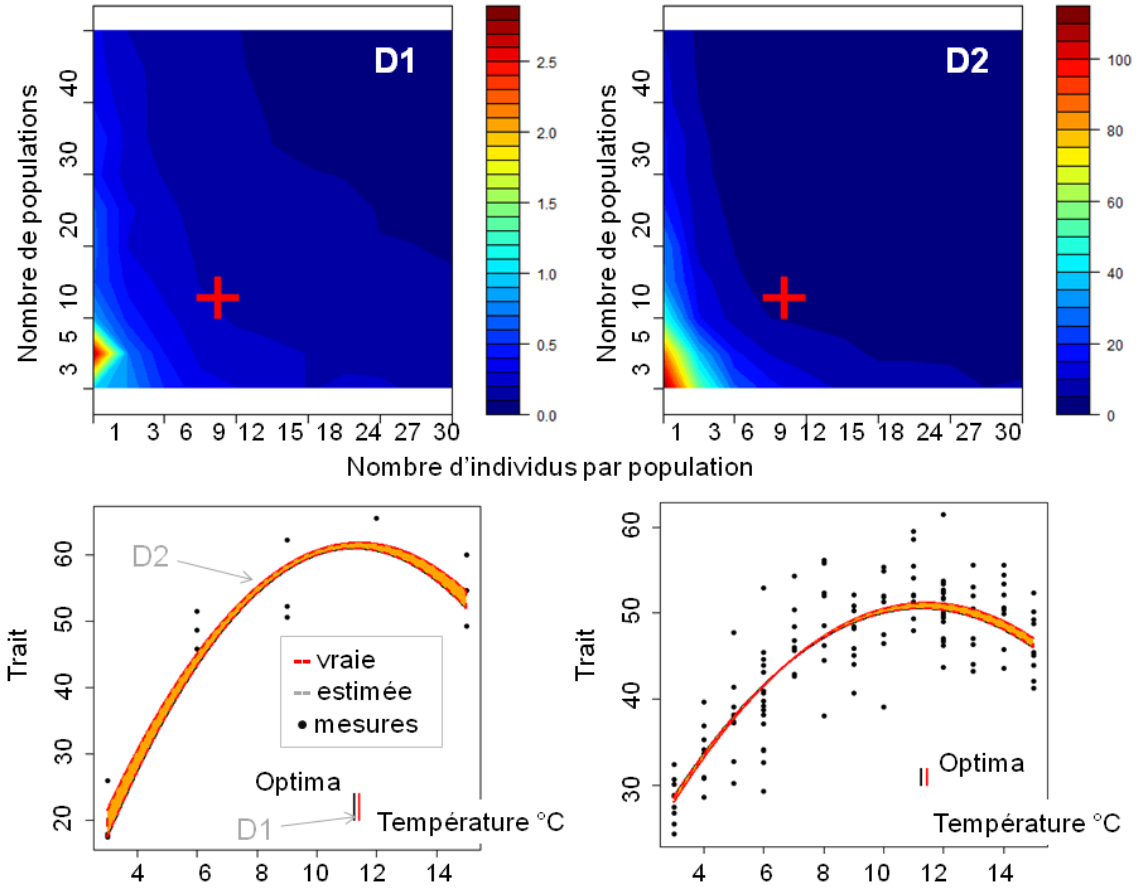


Fig. 1.5. Influence de l'effort d'échantillonnage sur la précision des estimations de courbes de réponse et d'optimum.

Cette figure présente les résultats d'une expérience virtuelle le long d'un gradient unique de température. Le trait à échantillonner suit une réponse unimodale au gradient de température et les résultats donnés ici sont pour un écart-type intra-population de sd= 4.

En haut sont représentées les estimations moyennes de D1 (distance entre les optima « vrai » et estimé, voir Encadré 1.2) et D2 (Surface entre les deux courbes de réponse « vraie » et estimée) pour 100 simulations de chaque niveau d'effort d'échantillonnage, c.à.d. un nombre de populations (entre 3 et 50) et un nombre d'individus (entre 1 et 30). Les croix rouges correspondent à un échantillonnage de 10 populations et 10 individus par population.

En bas sont représentés des cas de simulations d'échantillonnage obtenues avec (à gauche) 5 populations et 3 individus par population et (à droite) 13 populations et 10 individus par population. Dans chaque cas sont représentées les distances D1 et D2 entre les courbes de réponse vraies et estimées.

Toutes ces stratégies amènent à des distances entre optima (D1) équivalentes, ou en tous cas D1 ne permet pas de mettre en évidence une stratégie beaucoup plus efficace que les autres (Fig. 1.8) certainement parce que l'optimum « vrai » tel que nous l'avons choisi se situe au cœur de l'espace climatique assez proche des conditions climatiques moyennes (ex. Fig. 1.6 & 1.7). Par contre toutes les stratégies ne sont pas équivalentes pour la distance entre les surfaces de réponse « vraie » et estimée (D2). Les stratégies les plus efficaces sont E8>E4>E7>E12>E6>E3; ce classement de précision est assez logique puisque E8 est la stratégie où le plus gros effort d'échantillonnage est mis en place avec une large couverture de l'hétérogénéité. La stratégie E4 donne un bon compromis entre l'effort d'échantillonnage (p=12 pour E4 contre p=20 pour E8) et précision de la surface estimée (Fig. 1.8). Cette étude virtuelle n'est bien entendu pas source de vérité, elle ne fait que donner des indications sur les choix à faire, tout en supposant que les hypothèses posées sont les bonnes. Nous avons ici choisi de supposer que la variabilité au sein des populations était de sd=4, c.à.d. relativement forte (Fig. 1.5) et aussi que l'optimum de hauteur se situait au cœur de la zone d'étude, ce qui risque de ne pas être le cas pour toutes les espèces, puisque nous avons choisi des espèces ayant des exigences écologiques différentes (voir 2.3). Cela pourrait entrainer entre les différentes espèces des différences d'efficacité de l'échantillonnage pour mettre en évidence des courbes de réponse des traits aux gradients. En connaissance de cause, nous avons donc choisi de suivre la stratégie E4 et d'échantillonner une dizaine d'individus par population.

2.4.3. Choix des populations: intersection entre théorie et données de présence

Pour chaque espèce nous disposions de cent à cinq cents populations référencées (voir 2.3.2). Pour chacune des espèces nous avons donc adapté l'échantillonnage théorique E4 à ces présences connues. Les points théoriques trop loin (dans l'espace climatique) de toute population connue ont été supprimés et certains points théoriques ont été ajoutés dans le cas de populations présentes dans des conditions climatiques particulières. Nous avons ensuite sélectionné les populations référencées les plus proches (+/- 0,2 écart-type en général) des points théoriques dans l'espace climatique. Un point théorique était le plus souvent proche de plusieurs populations référencées, il a donc fallu sélectionner parmi ces populations, une unique représentante de chaque point théorique. Nous avons d'abord pris soin d'éliminer les points trop difficile d'accès et nous avons ensuite sélectionné au hasard parmi les possibilités. Il y a eu, toutes espèces confondues, cent quatre vingt dix neuf populations échantillonnées sur le terrain (Tableau 1.2 et Fig. 1.11). Ces populations échantillonnées ne correspondent finalement pas exactement à celles sélectionnées sur le papier; en effet, certaines populations n'ont pas été retrouvées, certaines se sont révélées inaccessibles, des groupements ont parfois été effectués afin de mesurer plusieurs espèces sur une même placette et des populations ont été ajoutées selon les opportunités. Notre échantillonnage réalisé s'est donc appuyé sur une stratégie théorique complexe, mais reste un

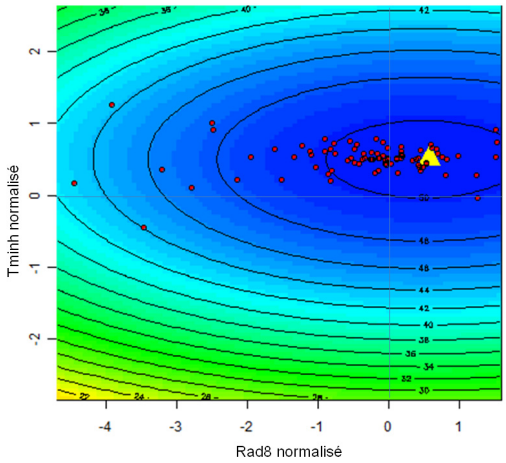


Fig. 1.6. Réponse virtuelle d'un trait le long de deux gradients environnementaux : radiation et température normalisées. *. Ce trait répond de manière unimodale le long de chacun des gradients (grande valeurs en bleu et petites en vert). Le triangle jaune est le vrai optimum de la surface. Les points rouges sont les optima estimés lors des 100 échantillonnages avec la stratégie E4 (Fig. 1.7)*

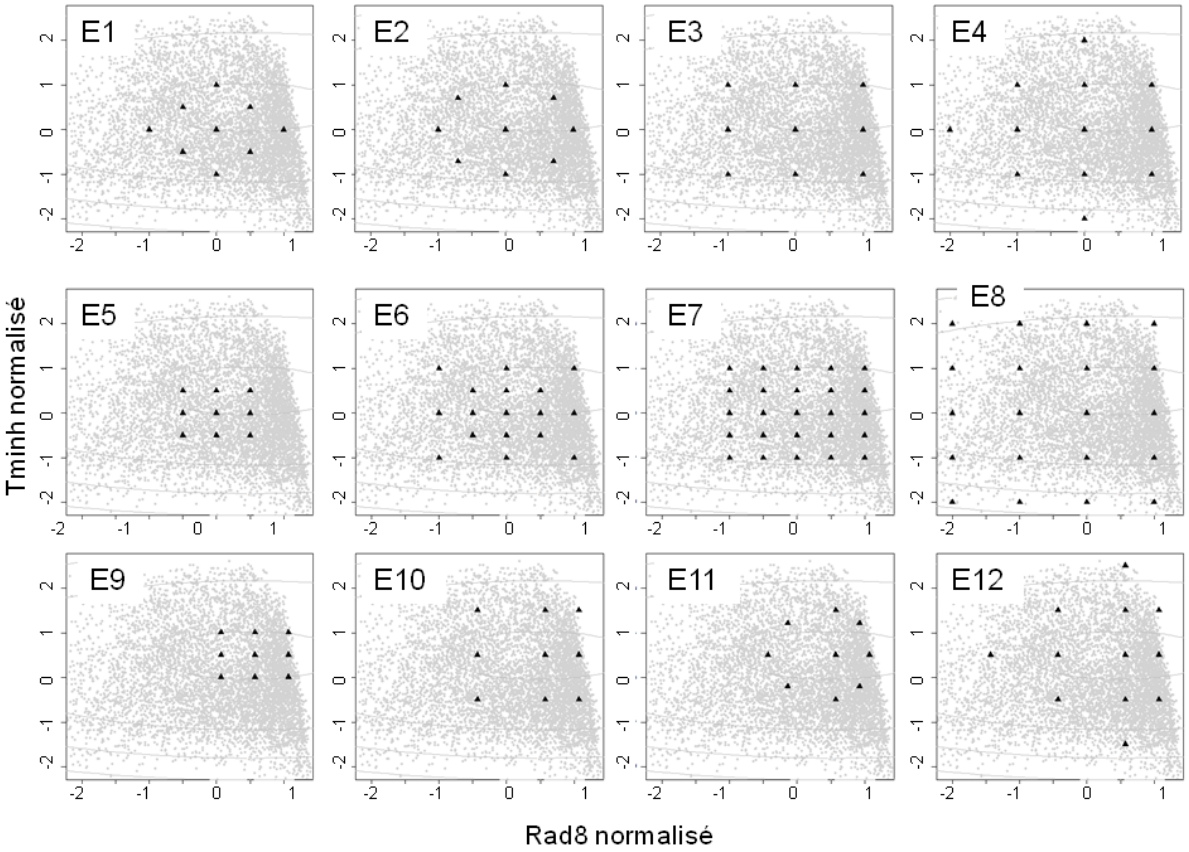


Fig. 1.7. Les douze stratégies testées virtuellement

Ces stratégies sont dérivées de la méthode des surfaces de réponse (Encadré 1.3) ; elles sont centrées sur (0,0), c.à.d. les conditions moyennes (ex. E1->E8) ou sur (0.5, 0.5) qui correspond à l'optimum (ex. E9->E12); elles s'étalent autour du centre jusqu'à +/- 1 écart-type autour de la moyenne (ex. E1), ou plus largement (ex. E12); elles se contentent des neuf points nécessaires à la méthode des surfaces de réponse (ex. E1) ou tentent de couvrir d'avantage d'hétérogénéité environnementale en échantillonnant plus de populations (ex. E4, E8)

échantillonnage de terrain, avec toute la complexité et les imprévus que cela implique. Il est difficile de quantifier l'écart des stratégies réalisées par rapport à la stratégie théorique, nous avons donc utilisé plusieurs critères pour déterminer l'éloignement du réalisé au théorique (Fig. 1.10). En moyenne sur les seize espèces nous avons échantillonné 12,4 populations par espèce, 67% de la l'amplitude théorique sur l'axe des radiations et 72% sur l'axe des températures. La distance moyenne entre les populations dans l'espace climatique était en moyenne 67% de la distance entre les populations théoriques. On voit donc qu'en moyenne on a échantillonné plus de populations que prévu (moins dans certains cas comme *S. herbacea*, *R. ferrugineum*) et que ces populations sont un peu moins étalées que prévu sur les axes de radiations et températures. Même si notre échantillonnage empirique s'est éloigné de l'échantillonnage théorique prévu, les études théoriques préliminaires ont permis : (1) d'ajuster notre échantillonnage en termes d'effort d'échantillonnage (nombre de populations et d'individus et répartition dans l'espace climatique) ; (2) d'échantillonner l'hétérogénéité environnementale au mieux en fonction des présences des espèces.

2.4.4. Une stratégie hiérarchisée

Afin de déterminer la structure spatiale de la variabilité fonctionnelle intraspécifique, nous avons affiné notre stratégie pour la rendre hiérarchique (Fig. 3.4). Comme nous venons de le voir, nous avons sélectionné entre 6 et 18 populations (Tableau 1.2) de chaque espèce le long de deux gradients orthogonaux. Chaque population a été délimitée dans une zone de 50 x 50 m autour du point géo-référencé afin d'être cohérent avec les données climatiques à 50 x 50 m. Au sein de chacune de ces populations, nous avons sélectionné -si possible au hasard, par tirage aléatoire de coordonnées- trois sous-populations (50 x 50 cm ou 1 x 1 m pour les herbacées et les ligneux bas et 10 x 10 m pour les ligneux hauts). La taille de ces sous-populations dépendait de la taille des individus de l'espèce étudiée, car le but était d'avoir des sous-populations rencontrant des environnements homogènes pour que les différences entre sous-populations intègrent les micro-hétérogénéités. Dans chaque sous-population, nous avons sélectionné - si possible au hasard - trois individus matures définis selon les espèces comme des touffes ou comme des ramets. Cette définition des individus était parfois problématique, nous avons donc pris le parti d'avoir différentes définitions des individus tout en restant cohérent au sein d'une même espèce. Pour des raisons de disponibilité, nous avons prélevé une feuille par individu chez les espèces herbacées et dix feuilles par individu chez les espèces ligneuses; ces feuilles étaient bien développées, non sénescentes et non pâturées.

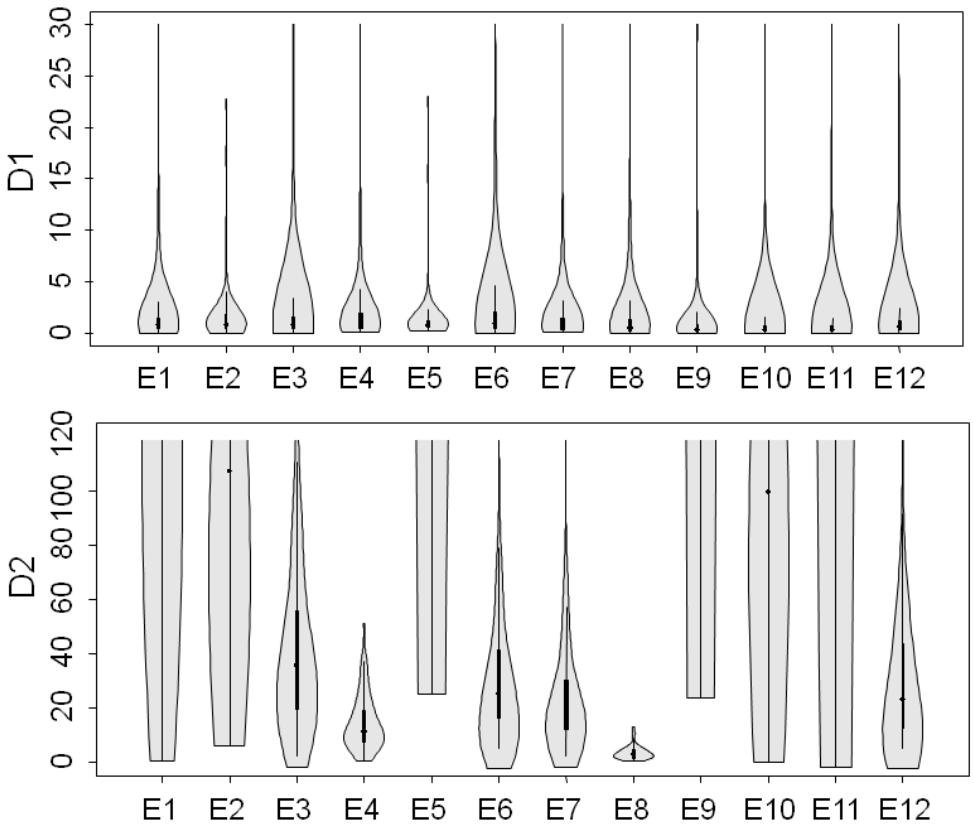


Fig. 1.8. Distributions des estimations des distances D1 et D2 en fonction des douze stratégies testées (Fig. 1.7 et Encadré 1.2).

Cette figure est représentée sous la forme de graphique en violons qui sont une combinaison d'une boîte à moustache et d'un kernel de densité. Ils présentent (i) la densité des données estimée par une méthode de kernel (en gris) (ii) la médiane des valeurs (tirets noirs) (iii) l'amplitude interquartiles entre le 1° et le 3° (segments verticaux noirs).

(En haut)- estimations de la distance D1, distance entre les optima « vrai » et estimé

(En bas)- estimations de la distance D2, distance (volume) entre les surfaces « vraie» et estimée.

2.5. Mesures de terrain: La base Guisane 2080

2.5.1. Mesures individuelles

Sur chaque individu sélectionné, nous avons effectué différentes mesures de traits fonctionnels (Chapitre 2) décrivant différents aspects de la fonction de l'individu dans sa communauté et de sa réponse à l'environnement et différentes étapes du cycle de vie.

2.5.1.1. Traits morphologiques

(1) La hauteur maximale végétative (Hmax) est la distance entre le sommet des tissus photosynthétiques et le sol en centimètres; elle est associée à la capacité compétitive de la plante et tend à être corrélée allométriquement à la biomasse aérienne (Cornelissen *et al.* 2003).

(2) L'étalement latéral a été estimé ici comme la surface projetée de la plante au sol ; elle a été calculée pour les herbacées en simplifiant selon les cas la surface par un rectangle défini par le plus grand diamètre et le diamètre orthogonal correspondant ou par un disque défini par le diamètre de l'individu. Cette mesure donne une autre estimation de la capacité compétitive de la plante. Ce trait n'ayant été mesuré que sur quelques espèces et à des niveaux différents (individu, talle) il n'a pas été inclus dans les analyses.

(3) Le diamètre des arbres (DBH pour « diameter at breast height ») est le diamètre de l'arbre par-dessus l'écorce à environ 1,37 m au dessus du sol du côté amont de l'arbre. Ce diamètre dépend évidement de l'âge de l'arbre, mais il dépend aussi de sa croissance qui résulte d'effets environnementaux et d'effets individuels propres.

(4) La biomasse aérienne est la masse de tissus aériens, nous avons mesuré cette biomasse comme la masse des individus entiers ou comme la masse d'une talle ou d'un ramet dans le cas de certaines espèces clonales pour lesquels l'individu était particulièrement difficile à définir (ex. *C. sempervirens*, *T. alpinum*) ; pour les espèces ligneuses basses, la biomasse a été mesurée par quadrat de 10 x 10 cm pour les espèces les plus basses (ex. *S. herbacea*, *D. octopetala*) ou de 50 x 50 cm pour les espèces plus hautes (ex. *V. myrtillus*, *R. ferrugineum*). Ce trait n'ayant été mesuré que sur quelques espèces et selon des protocoles différant entre espèces il n'a pas été inclus dans les analyses.

2.5.1.2. Traits foliaires

(1) La teneur en matière sèche (LDMC pour « leaf dry matter content ») est la masse de la feuille séchée en étuve divisée par la masse fraîche saturée en eau (Cornelissen *et al.* 2003), ici exprimée en $mg.g^{-1}$. C'est un trait considéré comme peu variable et robuste (Roche *et al.* 2004) qui est souvent négativement corrélé au taux de croissance relatif (RGR pour « relative growth rate », Weiher *et al.* 1999). La LDMC a été ici mesurée en suivant la méthode de réhydratation partielle de Vaieretti *et al.* (2007).

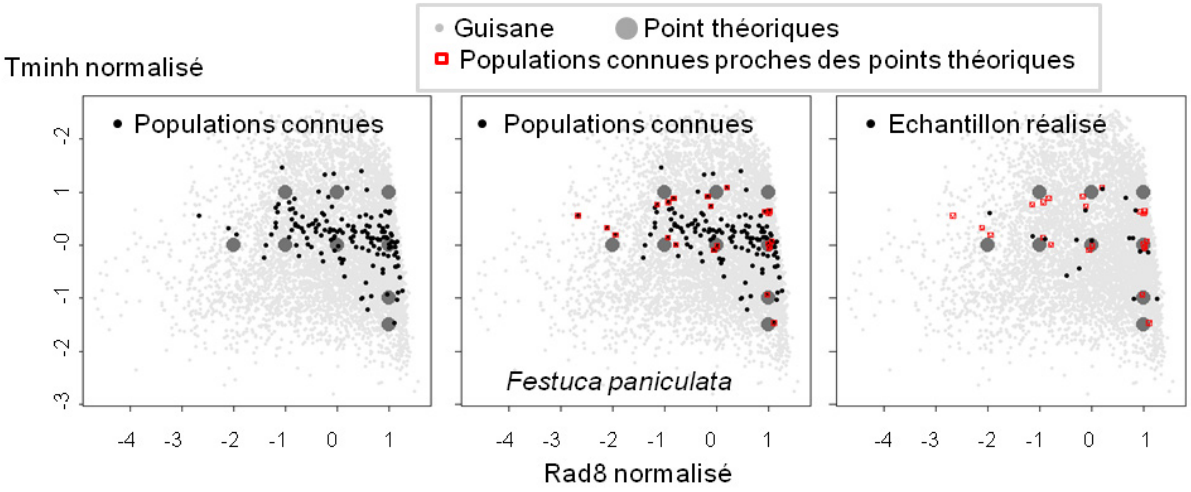


Fig. 1.9. Ecart entre populations échantillonnées et points définis théoriquement : cas de Festuca paniculata.

L'ensemble des points de la zone géographique étudiée (Guisane) est représenté en gris clair. Seuls les points noirs (à gauche) sont des populations connues de Festuca paniculata. Nous avons donc adapté les points théoriques de la stratégie E4 à ces populations connues (Gros points gris). Nous avons ensuite sélectionné les populations connues les plus proches des points théoriques (carrés rouges). Les populations finalement échantillonnées (points noirs à droite) sont différentes de celles choisies en théorie mais recouvrent quand même l'ensemble de l'hétérogénéité climatique où l'espèce est présente.

(2) La surface spécifique foliaire (SLA pour « specific leaf area ») est la surface d'un côté d'une feuille fraîche divisée par la masse de cette même feuille séchée en étude, en $m^2.kg^{-1}$; la SLA est généralement corrélée positivement au RGR et négativement à la LDMC (Cornelissen *et al.* 2003).

(2) La teneur en azote (LNC pour «leaf nitrogen content») est la quantité totale d'azote par unité de masse sèche, en $mg.g^{-1}$ (resp. par unité de surface, en $mg.mm^{-2}$) ; cette grandeur est reliée à la capacité photosynthétique par unité de masse (resp. de surface) (Cornelissen *et al.* 2003).

(3) La teneur en carbone (LCC pour «leaf carbon content») est la quantité totale de carbone par unité de masse sèche, en $mg.g^{-1}$, elle est reliée à l'assimilation de la plante en carbone.

LNC et LCC ont été mesurées sur des échantillons de feuilles séchées et broyées (3-5 mg) avec un analyseur FlashEA 1112 elemental analyzer (Thermo Fisher Scientific Inc. made in Milan, Italia). Ces mesures ont été effectuées au niveau individuel ou en regroupant les trois individus d'un même sous-plot dans le cas d'espèces à petites feuilles (*C. sempervirens*, *L. vulgare*, *P. viviparum*, *S. herbacea*, *S. caerulea*, *S. nutans*, *T. alpinum).*

2.5.1.3. Traits reproducteurs

Nous avions comme première idée de mesurer la masse des graines qui est liée à l'allocation des individus à la reproduction et au succès de recrutement des plantules après germination (Weiher *et al.* 1999). Ce trait nous aurait permis d'avoir les trois traits recommandés par Westoby (1998) pour sa méthode LHS (pour « leaf, height, seed ») permettant de déterminer les grandes stratégies fonctionnelles végétales. Cependant cette mesure aurait demandé de retourner plusieurs fois sur chaque site et nous avons donc abandonné. A défaut, nous avons mesuré des indicateurs de fertilité comme le nombre d'inflorescences (ex. *F. paniculata*), le nombre de fleurs par inflorescences (ex. *T. alpinum*), le nombre de cônes jeunes et matures, la taille des inflorescences (ex. pour *L. vulgare* qui est une composée), et la hauteur des inflorescences. Mais ces traits ne sont pas comparables entre espèces et leurs valeurs suivent pour la plupart des distributions de poisson très fortement asymétriques vers zéro ; le nombre de données que nous avons n'a donc pas permis d'en tirer des résultats intéressants, il aurait certainement fallu mesurer plus d'individus sur moins de populations pour pouvoir traiter ces données. De plus il faudrait aussi réaliser des allométries avec la fécondité réelle.

2.5.1.4. Age des arbres

Tous les arbres échantillonnés durant la campagne Guisane 2080 ont été carottés à cœur à environ 1,30m à l'aide d'une tarière de Pressler de 5 mm de diamètre intérieur. Les carottes ont été poncées et scannées. Le décompte des cernes nous a donné l'âge approximatif des arbres.

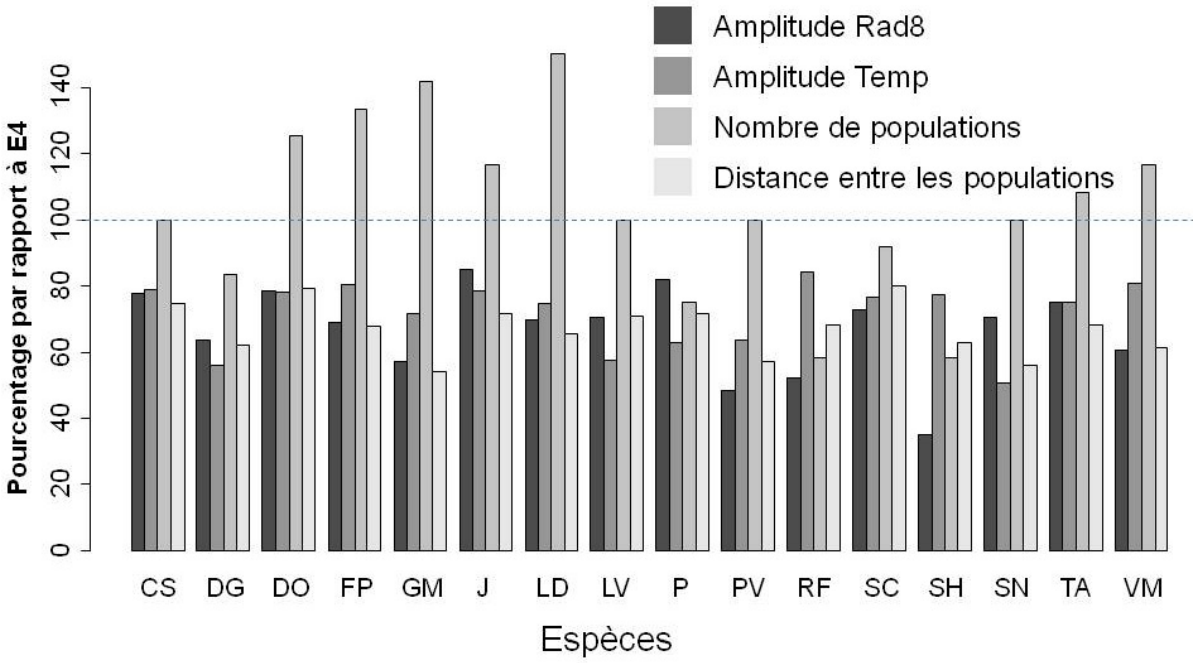


Fig. 1.10. Ecart des échantillonnages réalisés à l'échantillonnage théorique. *Pour chacune des seize espèces étudiées, de gauche à droite (du plus foncé au plus clair) et en pourcentage de ce qui était prévu avec la stratégie E4 : l'amplitude échantillonnée sur l'axe Rad8, l'amplitude échantillonnée sur l'axe Temp, le nombre de populations mesurées, la distance moyenne entre les populations échantillonnées dans le plan (Rad8,Temp).*

2.5.2. Mesures de l'environnement au niveau population ou sous-population

2.5.2.1. Caractérisation du sol

Au niveau de chaque sous-population, nous avons collecté des échantillons de sol dans les premiers horizons (jusqu'à 10 cm). Sur ces échantillons nous avons mesuré le pH par dilution dans l'eau (protocole standard : Robertson *et al.* 1999). Nous avons également mesuré la teneur en matière organique par perte au feu à 550°C sur les échantillons groupés par population (protocole standard : Robertson *et al.* 1999). La texture du sol a été déterminée (fraction de sable, argile et limons) par le Laboratoire d'Analyses des Sols d'Arras (http://www.arras.inra.fr/index.htm).

Des analyses en composantes principales (ACP, Rao 1964) inter et intra-groupes (Encadré 2.4) sur l'ensemble de ces données de sol ont montré que la structure entre espèces n'était pas négligeable (ratio=0,16), c.à.d. que les seize espèces sélectionnées n'avaient pas les mêmes exigences édaphiques. Nous avons donc effectué des ACP pour les caractéristiques de sol de chaque espèce et nous avons gardé les deux premiers axes de chaque ACP pour caractériser les sols.

2.5.2.2. Perturbations anthropiques

Comme la vallée de la Guisane présente une forte diversité de perturbations anthropiques (ex. urbanisation, pâturage, fauche, gestion des forêts), nous avons essayé de décrire de manière synthétique les perturbations anthropiques subies par les populations échantillonnées. Un jeu de données du Parc National des Ecrins (PNE, Delphine) décrit les perturbations en prairies de type fauche et pâturage. La fauche est décrite selon trois niveaux : jamais, biennale, annuelle. Le pâturage est décrit selon quatre niveaux: absence, léger, fort, surpâturage. Nous avons agrégé ces informations en (0) pas de perturbation et (1) fauche ou pâturage. Ces perturbations n'affectent pas les arbres adultes (nous n'avons pas d'information sur les perturbations de type exploitation/gestion forestière).

2.5.2.3. Compétition

Nous avons choisi des espèces communes et structurantes (sauf pour *S. nutans*, *L. vulgare* et *P. viviparum*), c.à.d. des espèces qui dominent dans les milieux correspondant pour elles à des conditions environnementales favorables. Elles représentaient ainsi une grande partie de la biomasse et la hauteur moyenne des communautés était donc largement influencée par la hauteur des espèces étudiées et n'était donc pas un bon indicateur de la compétition subie par ces espèces. Nous avons donc estimé visuellement des taux de fermeture du milieu à trois niveaux : fermeture basse au niveau de la strate herbacée, fermeture intermédiaire au niveau de la strate des ligneux bas et fermeture haute pour la strate arborée. Pour une espèce donnée, nous avons donc considéré que la compétition résultait de la fermeture des strates

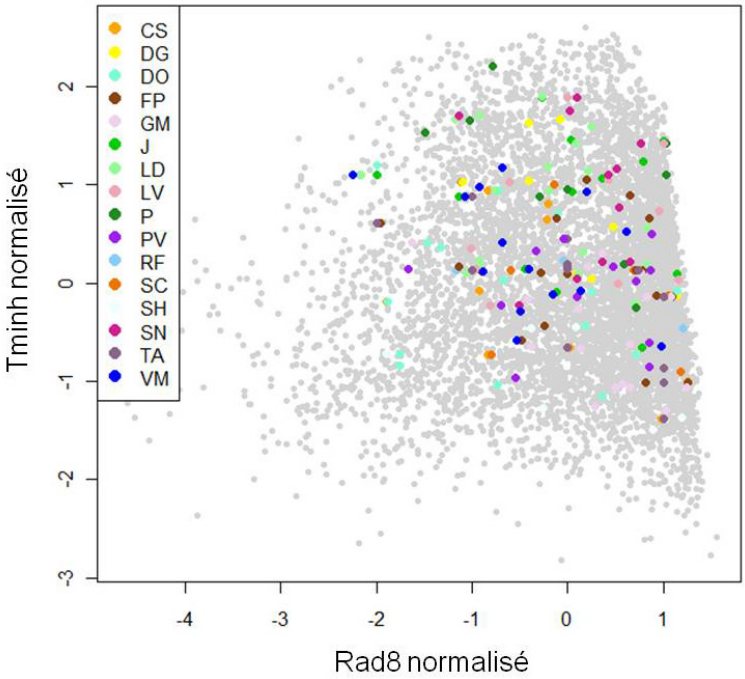
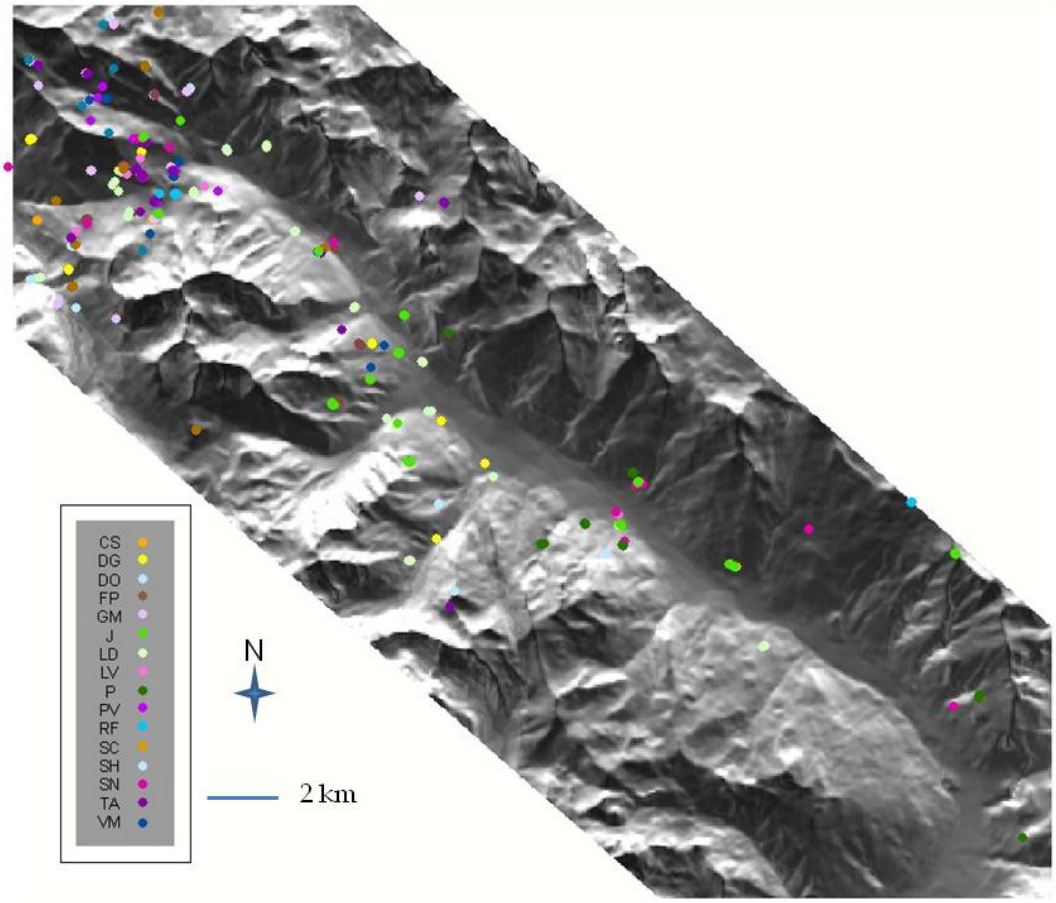


Fig. 1.11. Les cent quatre vingt dix neuf populations échantillonnées sur le terrain

A. Représentation dans l'espace climatique des différentes populations – les populations des différentes espèces sont représentées par des points de différentes couleurs. Les points gris clair représentent l'ensemble des points de la zone géographique étudiée (Guisane).

B. Représentation dans l'espace géographique des différentes populations – les populations des différentes espèces sont représentées par des points de différentes couleurs (même couleurs que pour A). Le fond de carte représente Rad8, selon un dégradé de gris : les fortes radiations sont représentées en noir et les faibles en blanc.



supérieures (ex. pour *D. glomerata* nous avons considéré les strates intermédiaire et haute). La somme des fermetures des niveaux supérieurs a ensuite été transformée en (0) absence de compétition et (1) compétition si la somme est supérieure à 25%.

Pour les trois espèces non structurantes d'après les relevés de végétation (voir 2.5.2.4) : *S. nutans*, *L. vulgare* et *P. viviparum,* la compétition a été prise en compte comme la hauteur moyenne de la végétation environnante (dans chaque sous-population).

Pour les arbres (*L. decidua* et *Pinus sp.)* nous avons considéré deux indices de compétition. Comme indicateur de la compétition racinaire nous avons considéré le nombre d'arbres, plus grands ou plus petits que l'arbre mesuré, présents dans un cercle de 2,5 m autour de l'arbre mesuré. Comme indicateur de la compétition aérienne, nous avons considéré le nombre de « côtés » de l'arbre mesuré ayant accès à la pleine lumière. Nous avons admis pour cela que l'arbre avait quatre côté et que chacun pouvait être occulté par un ou plusieurs arbres dominants. Cet indice était donc à quatre niveaux (0 ; 0,50 ; 0,75 ; 1), 1 étant caractéristique d'un arbre isolé.

2.5.2.4. Biodiversité

Des relevés de végétation (10 x 10 m) ont été effectués en Guisane par Gilles Pellet (en 2007) sur cent soixante quatorze placettes correspondant à des populations échantillonnées (442 espèces différentes relevées). Ces relevés utilisent l'indice de couverture de Braun-Blanquet *et al.* (1952) à six niveaux: 5 = >75%, 4= 50-75%, 3 = 25-50%, 2 = 10-25%, 1= 1-10%, +=<1%. Nous avons agrégé les niveaux 1 et + ensemble. Cette couverture représente la surface occupée (à une hauteur donnée) par l'espèce dans la communauté. La somme de toutes les couvertures peut donc être supérieure à 100% étant donné que toutes les espèces ne vivent pas à la même hauteur.

Pour une étude du Chapitre 4, nous avons sélectionné les relevés correspondant à des prairies ou landes (sans arbres) et pour lesquels les données de la base de traits Androsace permettaient de décrire plus de 80% de l'abondance pour Hmax et LDMC. Or l'hypothèse du ratio de biomasse (Grime 1998) propose en effet que le fonctionnement des écosystèmes est essentiellement déterminé par les espèces dominantes et leurs caractéristiques. Nous avons donc obtenu au final un sous jeu de quarante et un relevés utilisés dans le Chapitre 4 pour tester les implications de la variabilité fonctionnelle sur le calcul d'indices de diversité.

2.5.2.5. Modèles d'habitat

Pour chacune des seize espèces étudiées, nous avons calibré des modèles d'habitat (Guisan & Thuiller 2005; Chapitre 3) à partir des 500 relevés de végétation issue de la base du Conservatoire Botanique National Alpin dans la vallée de la Guisane (cf 2.3.2). Il y avait pour chacune des espèces entre 37 et 180 présences. Nous avons utilisé la méthode de moyennage des modèles (Encadré 2.2) sur des GAM construits sur trois variables explicatives : les précipitations annuelles, les degrés jour à 5°C et les radiations solaires du mois d'aout

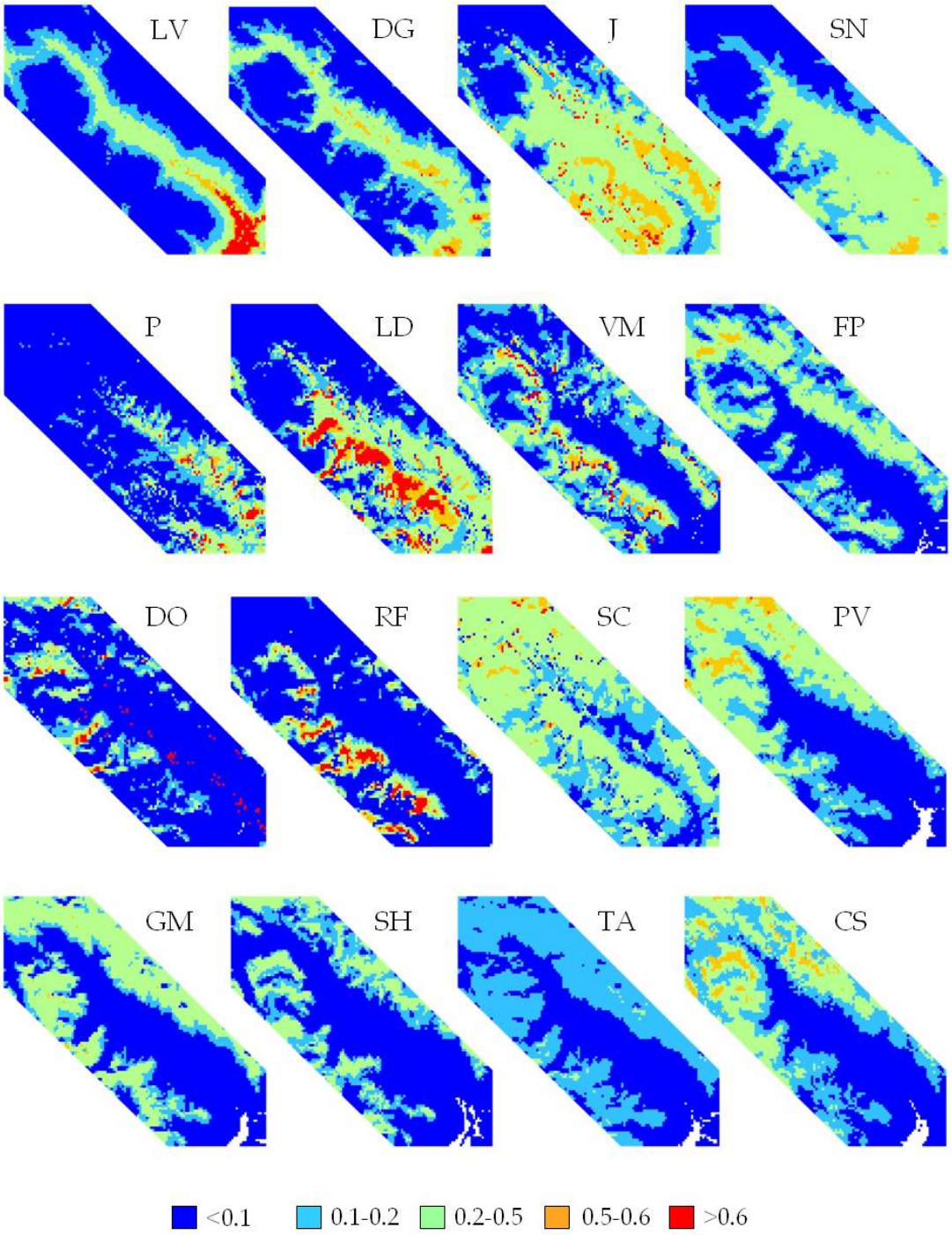


Fig. 1.12. Projection des modèles d'habitat des seize espèces étudiées en vallée de la Guisane.
Les lignes du haut donnent les projections pour les espèces de basse altitude (espèces de plaine et subalpines), celles du bas pour les espèces plus fréquentes en haute altitude (espèces subalpines et alpines). Les projections sont données pour les sorties brutes des modèles d'habitat.

(Tableau 1.1). La projection de ces modèles dans la vallée (Fig. 1.12) montre que toutes les espèces n'ont pas les mêmes exigences écologiques, certaines étant des espèces de basse altitude (*D. glomerata*, *L. vulgare*), d'autres des espèces subalpines (*F.paniculata*, *L.decidua*) et d'autres des espèces alpines (*C. sempervirens*, *S. herbacea*).

Chapitre 2

Variabilité fonctionnelle intraspécifique : quantification, structure, sources et comparaison avec l'interspécifique

	Hmax		*LDMC*		*LNC*	
Espèces	*Inter-pop*	*Intra-pop*	*Inter-pop*	*Intra-pop*	*Inter-pop*	*Intra-pop*
Carex sempervirens	*17/0.63*	*0.4*	*38/0.08*	*0.04*	*14/0.12*	
Dactylis glomerata		*0.6*		*0.06*		
Dactylis (Ryser 1999)			*25/0.11*			
Dryas octopetala	*6/0.43*	*0.46*	*32/0.22*	*0.05*		
Festuca paniculata	*54/0.35*	*0.24*				
Larix decidua			*29/0.05*	*0.06*		
Leucanthemum vulgare	*26/0.67*		*13/0.15*			
Pinus			*30/0.39*			
Polygonum viviparum	*10/0.47*	*0.25*	*23/0.07*	*0.09*		
Sesleria caerulea	*18/0.56*	*0.36*	*39/0.02*	*0.05*		
Salix herbacea	*4.2/0.48*					
Silène nutans	*32/0.58*					
Vaccinium myrtillus	*40/0.76*	*0.6*			*1.88/0.29*	
Vaccinium (Gerdol 2005)					*1.92/0.09*	

Tableau 2.1. Moyennes et coefficients de variation issus de l'analyse des bases de données et de la bibliographie (issu de l'Article 7). *Les moyennes et coefficients de variation ont été calculés pour trois traits:Hmax, LDMC et LNC. Entre populations (Inter-pop): les moyennes et variances ont été calculées à partir des moyennes des populations/relevés disponibles dans la bibliographie et les bases de données. Les coefficients de variation entre populations en découlent : CV= écart-type/moyenne. Les résultats sont donnés sous la forme moyenne/CV. Au sein des populations (Intra-pop): on peut calculer un coefficient de variation pour chacune des populations/relevé pour lesquelles on dispose de données de variabilité (variance, écart-type) avec CV=écart-type/moyenne. Le CV intra-pop donné ici est la moyenne des CV obtenus pour chaque population.*

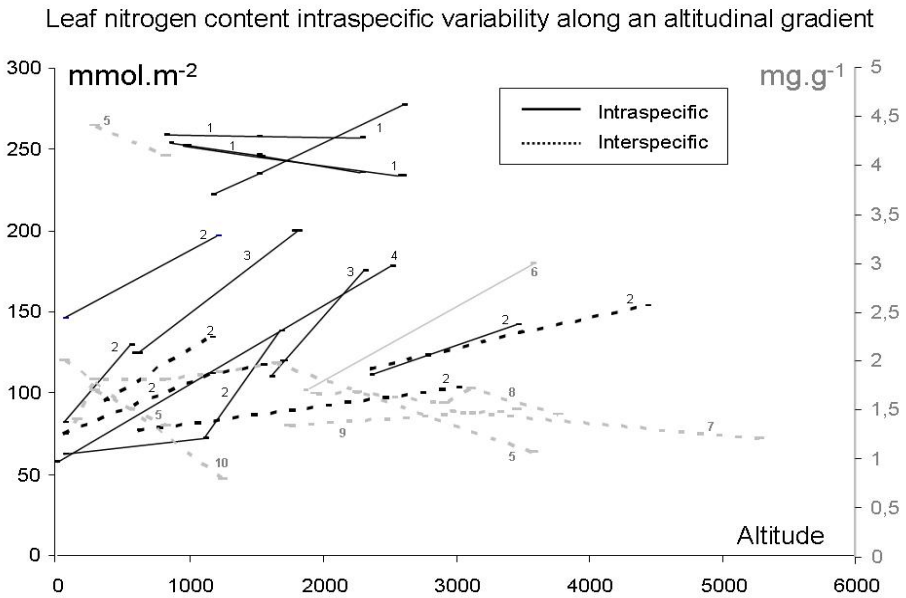


Fig. 2.1. Variation de LNC en réponse à l'altitude: méta-analyse (issue de l'Article 7). *Données de la bibliographie sur la variabilité de la teneur des feuilles en azote le long d'un gradient d'altitude. Selon la source les données sont en mmol.m^-2 ou en mg.g-1 et sont au niveau intra ou interspécifique: 1(Hultine & Marshall, 2000), 2(Körner, 1989), 3(Sparks & Ehleringer, 1997), 4(Cordell et al., 1998), 5(Reich et al., 1999), 6(Bowman , 1995), 7(He et al., 2006), 8(Luo et al., 2005), 9(Wegner et al., 2003), 10(Craine, 2003). Les différentes courbes données pour une même référence correspondent à différentes espèces ou groupes d'espèces.*

1. Contexte et objectifs

Comme nous l'avons vu précédemment les traits fonctionnels végétaux - comme la surface spécifique foliaire (SLA) ou la teneur des feuilles en matière sèche (LDMC) - sont de plus en plus utilisés à différentes fins. Ces traits sont variables entre les différentes espèces et au sein de celles-ci du fait d'adaptations locales et de la plasticité (Byars *et al.* 2007). Les différences de valeurs de traits entre espèces ont été largement étudiées notamment en termes de réponse aux gradients, de compromis de traits et de stratégies fonctionnelles (Reich *et al.* 2003 ; Diaz *et al.* 2004; Thuiller *et al.* 2004b ; Wright *et al.* 2005). Toutefois, la variabilité intraspécifique demeure mal connue alors que plusieurs études récentes montrent qu'elle peut avoir une grande influence sur le fonctionnement des écosystèmes (Boege & Dirzo 2004; Lecerf & Chauvet 2008). Prendre en compte une telle variabilité intraspécifique - aussi bien dans les études théoriques qu'empiriques - semble aujourd'hui indispensable pour affiner notre connaissance de la biodiversité, notre compréhension du fonctionnement des écosystèmes et pour être à même de prédire leur réponse aux changements environnementaux à venir (McGill *et al.* 2006). Il est donc indispensable de bien connaitre cette variabilité : son ampleur, sa structure et la manière dont les traits varient dans l'environnement.

Un regroupement des données de Hmax, LDMC et LNC provenant de la base Androsace, et de la bibliographie (Welker *et al.* 1997; Ryser & Aeschlimann 1999; Reich *et al.* 2003; Gerdol 2005; Peratoner *et al.* 2007; Thomas *et al.* 2007) pour les seize espèces étudiées permet de voir que tous les traits ne sont pas autant représentés, les traits les plus complexes et les plus onéreux à mesurer (comme LNC) étant les moins décrits. Ces données montrent aussi que peu d'information est disponible sur la variabilité des traits en elle-même. Le calcul des coefficients de variation entre populations et de la moyenne des coefficients de variation au sein des populations indique là encore que les informations disponibles sur la variabilité sont fortement incomplètes (Tableau 2.1). Il est souvent possible d'estimer des coefficients de variation entre populations (dès que plusieurs moyennes de traits sont enregistrées pour une même espèce), mais il est plus rarement possible d'avoir une idée de la variabilité des traits fonctionnels au sein de populations collectées dans des conditions différentes.

Les données actuelles des bases de connaissance et de la bibliographie ne suffisent donc pas pour quantifier la variabilité fonctionnelle intraspécifique en quantité et structure. Elles ne suffisent pas non plus à dégager des courbes de réponse génériques entre traits et gradients au niveau intraspécifique, comme c'est le cas pour la réponse de la LNC à l'altitude (Fig. 2.1). Cela pourrait être du à une absence de patron, à l'idiosyncrasie des courbes de réponse ou au fait que les études existantes ne peuvent pas véritablement être comparées. En effet les études qui traitent de la variabilité fonctionnelle intraspécifique sont souvent menées le long de gradients indirects tels que l'altitude (Chapitre 1), qui ne sont pas comparables d'une

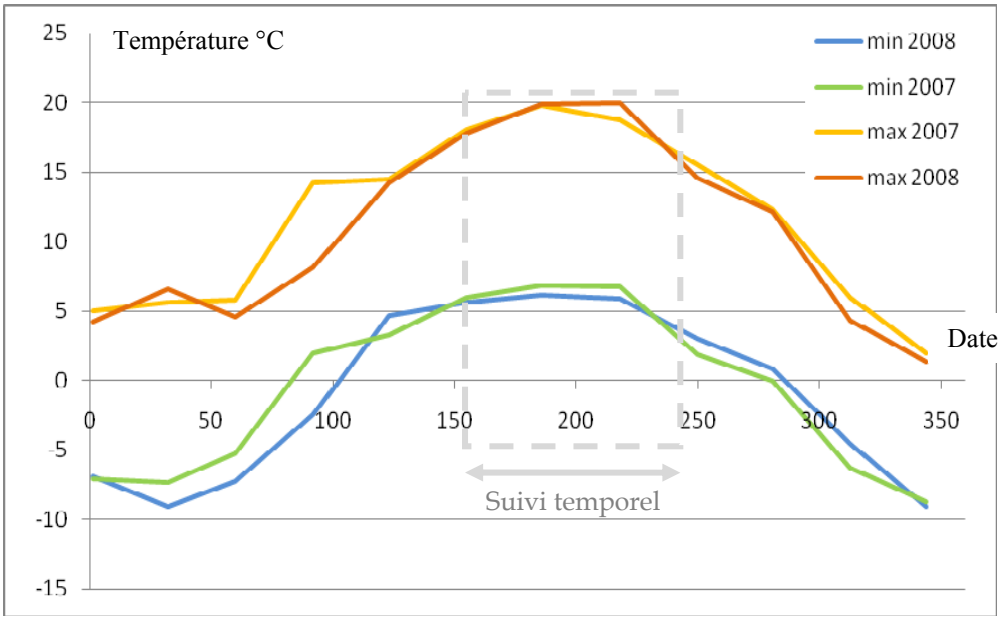


Fig. 2.2. Données de température de la Station météorologique de Villar d'Arène *(Météo France 05181002 : 45°01'54"N 6°21'42"E, 1665 mètres altitude) pour les années 2007 et 2008. Pour chaque année sont données les courbes des moyennes de températures maximales et minimales pour chaque mois. Les dates sont en jours Juliens (1 Janvier =1).*

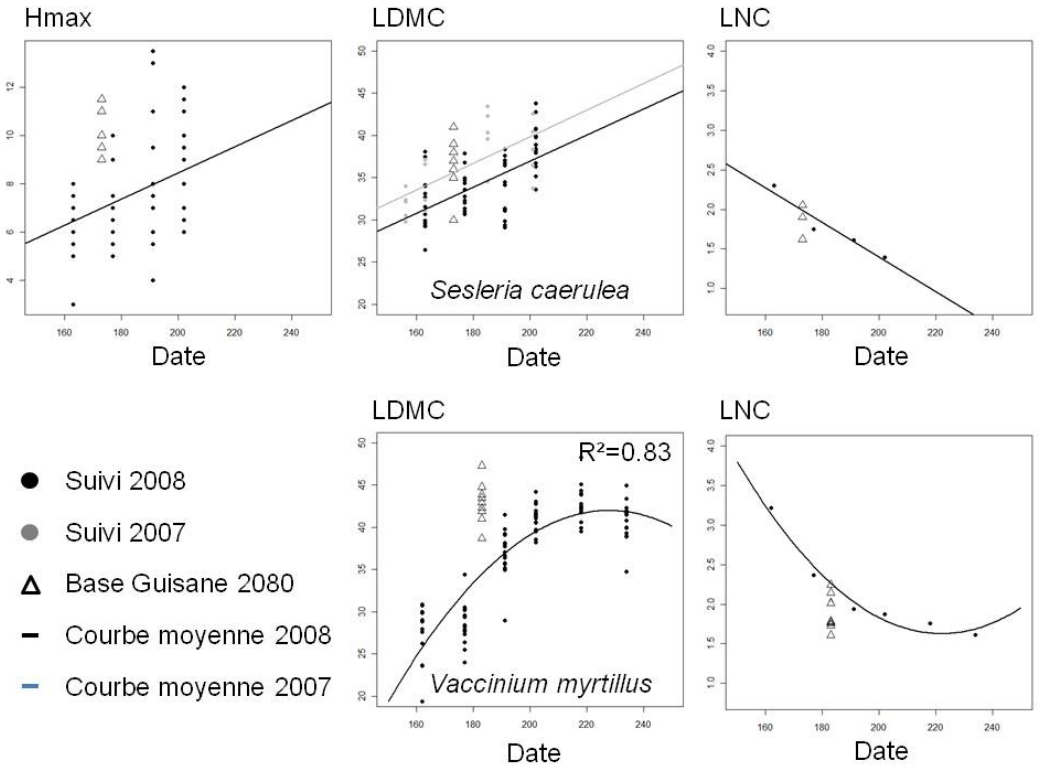


Fig. 2.3. Variation des traits fonctionnels au cours de la saison de végétation. *La variation des traits dans le temps est donnée par rapport aux dates en jours juliens (1=premier janvier). Les résultats sont donnés ici pour deux espèces représentatives : une herbacée Sesleria caerulea et une espèce de ligneux bas Vaccinium myrtillus.*

zone d'étude à une autre. D'autre part ces études n'utilisent pas toutes les mêmes méthodes d'échantillonnage.

Deux grandes questions permettent de résumer la quête de quantification de la variabilité intraspécifique : (i) Comment caractériser la variabilité fonctionnelle intraspécifique en termes de quantification, structures et sources ? (ii) Quelle est l'importance relative de la variabilité fonctionnelle intraspécifique par rapport à celle interspécifique? Afin de répondre à ces questions et d'améliorer notre compréhension de la variabilité intraspécifique, ce chapitre propose une analyse détaillée de la variabilité fonctionnelle. Les données récoltées sur seize espèces d'une vallée alpine selon une unique stratégie hiérarchique stratifiée (Chapitre 1) devraient permettre de mettre en évidence des patrons de variabilité s'il y en a et de comparer les espèces et formes de vie. L'analyse de la variabilité fonctionnelle qui suit se présente en trois parties : (1) tout d'abord je présente une analyse préliminaire sur la variabilité des traits fonctionnels dans le temps. Cette analyse doit permettre de quantifier la variabilité des traits au cours de la saison de végétation et de repérer une éventuelle date d'échantillonnage optimale ; (2) puis suit une analyse précise de la structure de la variabilité intraspécifique et une comparaison des variabilités inter- et intraspécifique. Cette analyse doit apporter des réponses aux deux questions posées ; (3) enfin la variabilité observée est remise dans un contexte d'environnements hétérogènes par l'étude des courbes de réponse des traits aux gradients.

2. Variabilité dans le temps et date d'échantillonnage

Objectifs

Les traits fonctionnels des plantes et en particulier les traits foliaires varient au cours de la saison de végétation (Garnier *et al.* 2001, Al Haj Khaled *et al.* 2005) et au cours de la durée de vie des feuilles (ex. *R. ferrugineum* : Pornon & Lamaze 2007). Garnier *et al.* (2001) conseillent d'échantillonner les espèces herbacées en mai et les ligneux en juillet en milieu méditerranéen. Or nous voulons échantillonner pour chaque espèce étudiée une dizaine de populations dans des conditions contrastées (Chapitre 1). Pour être certain que les différences de traits entre les populations sont dues aux conditions environnementales et non à la date d'échantillonnage, il faut déjà avoir une idée de la variabilité des traits dans le temps. Deux facteurs interagissent quant à notre échantillonnage sur le terrain : (1) il n'est pas possible de tout échantillonner en même temps ; aller sur les sites de populations et faire les mesures en laboratoire prend du temps et nécessite d'étaler la campagne de trait dans la saison ; (2) les différentes populations que l'on veut échantillonner pour une même espèce ne sont pas toutes autant avancées de la même manière dans la saison de végétation. Sur les sites de basse altitude, les individus pourrons avoir déjà fleuri et fané mi-juillet alors que sur les sites de haute altitude le déneigement (et donc le début de la saison de végétation) viendra seulement d'avoir lieu. Cette étude préliminaire vise donc à déterminer quelle est la

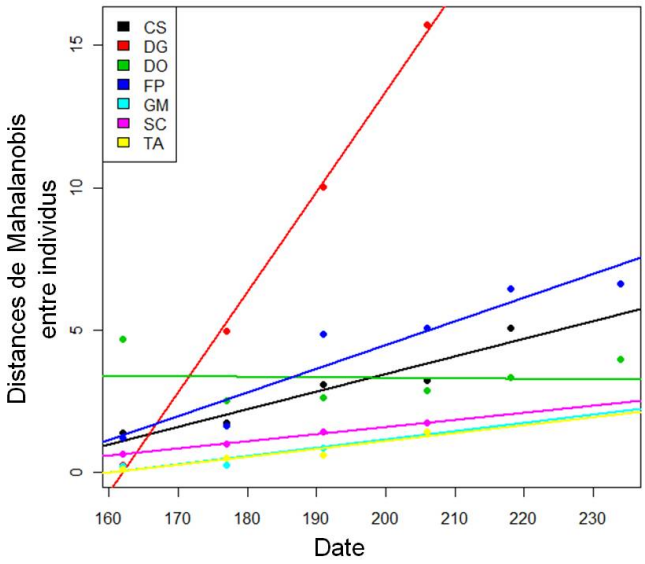


Fig. 2.4. Distances fonctionnelles moyennes entre les moyennes de populations mesurées à la date i et celles mesurées au début de saison. *Le temps est donné en jours juliens (1=premier janvier). Les distances sont des distances de Mahalanobis (Encadré 2.3).*

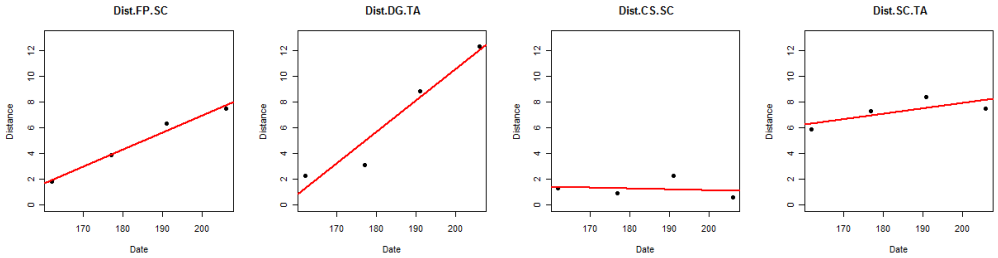


Fig. 2.5. Distances fonctionnelles entre les populations de couples d'espèces à différentes dates. *FP.SC= Festuca paniculata. Sesleria caerulea (pente=+0.13, p<0.05, R^2=0.96), DG.TA=Dactylis glomerata. Trifolium alpinum (pente=+0.24, p<0.05, R^2=0.89), CS.SC= Carex sempervirens. Sesleria caerulea (pente =0), SC.TA = Sesleria caerulea. Trifolium alpinum (pente=0). Le temps est donné en jours juliens (1=premier janvier). Les distances sont des distances de Mahalanobis (Encadré 2.3).*

bonne date d'échantillonnage, si les traits varient de la même manière pour toutes les espèces et si les stratégies fonctionnelles des différentes espèces changent au cours de la saison de végétation.

Principe

Nous avons effectué un suivi temporel de trois traits fonctionnels : Hmax (uniquement herbacées, les ligneux n'ayant pas une croissance significative d'une semaine à l'autre), LDMC et LNC (moyenné par population) pour quatre espèces de la base Guisane 2080 en 2007 (*F. paniculata*, *D. glomerata*, *S. caerulea*, *D. octopetala*) et dix espèces en 2008 (*F. paniculata*, *D. glomerata*, *S. caerulea*, *T. alpinum*, *G. montanum*, *V. myrtillus*, *L. decidua*, *R. ferrugineum*, *D. octopetala*, *J. sempervirens*). Ces suivis ont été menés sur une population de chaque espèce, ces populations étaient toutes situées vers 2000 m dans les environs de la Station Alpine Joseph Fourier. Pour chacune des populations, nous avons mesuré les traits toutes les deux semaines, soit quatre à six fois durant l'été avant le passage des moutons (Fig. 2.2) ; les mesures étaient effectuées sur quinze individus différents à chaque fois. Nous avons analysé la variabilité des traits moyens des différentes populations à l'aide de régressions linéaires sur les moyennes par date. Nous avons également mesuré les distances fonctionnelles (Encadré 2.3) entre les individus des espèces herbacées à la fois au sein d'une même population entre dates et aussi pour chaque date entre espèces.

Résultats

Comme attendu, une forte variabilité des traits a été observée au cours de la saison. De plus nous avons observé une forte hétérogénéité de cette variabilité entre les espèces. (1) Nous avons observé une augmentation de Hmax durant la saison (Fig. 2.3) : rapidement pour les espèces à stratégie « exploitatrice » comme *D. glomerata* (+0,93 cm/jour, $p<10^{-3}$) et moins rapidement pour les espèces à stratégie « conservatrice » comme *F. paniculata* (+0,4 cm/j, $p<10^{-3}$). (2) Nous avons également observé une augmentation significative de la LDMC (Fig. 2.3) au cours de la saison (+0,13 à +0,27 unité/j) et la formation d'un plateau correspondant probablement à la maturité des feuilles (nous n'avons pas cherché à quantifier les critères officiels de maturité tels que la fin d'extension du limbe ou l'apparition d'une ligule chez les graminées). (3) Nous avons finalement observé une diminution significative de la LNC au cours du temps (Fig. 2.3). Excepté pour *D. glomerata* (-0,07 u/j, $p<10^{-3}$), cette diminution était assez faible (-0,01 u/j) voire nulle pour les espèces sempervirentes *Juniperus sp.* et *R. ferrugineum*. (4) La comparaison entre les suivis temporels des populations de 2007 et 2008 révèle un certain décalage temporel entre les deux saisons, 2007 étant plus précoce que 2008 (Fig. 2.3). Ce résultat concorde avec les données météo des années correspondantes, l'été 2007 ayant été précocement chaud (Fig. 2.2). (5) L'analyse des distances fonctionnelles entre individus montre que les individus d'une même espèce s'éloignent significativement au cours du temps des individus mesurés à la première date, la vitesse d'éloignement est très

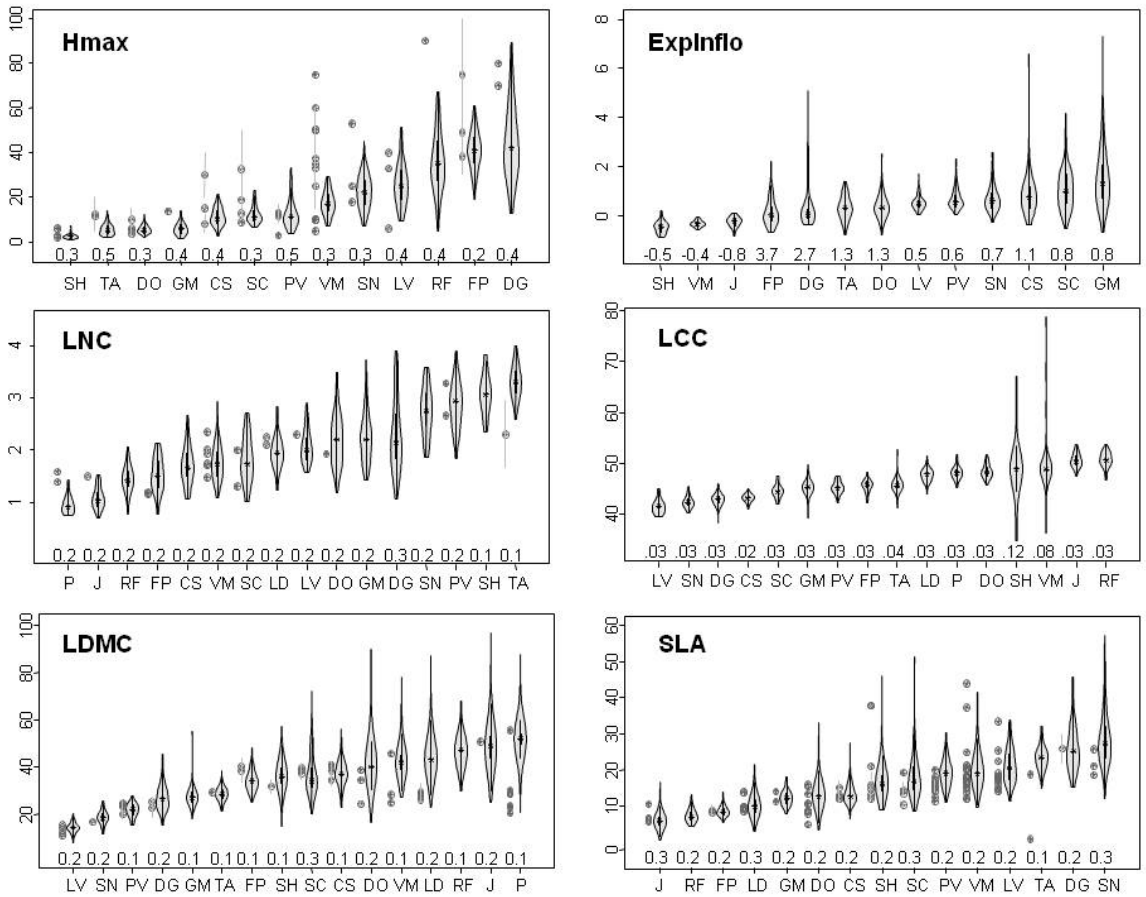


Fig. 2.6. Ampleur de la variabilité fonctionnelle intraspécifique

Les graphiques en violons sont une combinaison d'une boîte à moustache et d'un kernel de densité. Ils présentent (i) la densité des données estimée par une méthode de kernel (en gris) (ii) la médiane des valeurs (tirets noirs) (iii) l'amplitude inter-quartiles entre le 1° et le 3° (segments verticaux noirs). Les figures représentent l'ampleur de la variabilité intraspécifique pour (de haut en bas et de gauche à droite) la hauteur végétative (cm), l'exposition de l'inflorescence en hauteur (HauteurInflo-Hmax)/Hmax, la teneur des feuilles en azote (%), la teneur des feuilles en carbone (%), la teneur des feuilles en matière sèche (%) et la surface spécifique foliaire (mm^2/mg^{-1}) pour chaque espèce.

Pour chaque trait, les résultats sont ordonnés par les valeurs moyennes de chaque espèce. Pour chaque trait et chaque espèce les coefficients de variation sont donnés en dessous de chaque violon (ex. 0.2 pour la LDMC de J).

Les astérisques représentent les valeurs de traits provenant de la bibliographie et des bases de données existantes (LEDA, Androsace).

différente entre les espèces (+0,024 à +0,35 u/j) (Fig. 2.4), sauf pour *D. octopetala* qui est sempervirent et dont les traits moyens de population changent peu au cours de la saison. D'autre part les distances entre espèces augmentent significativement ($p<0,05$; Fig. 2.5) comme entre *F. paniculata* et *S. caerulea* ou restent identiques pour des espèces fonctionnellement très différentes (*S. caerulea* et *T. alpinum*) ou pour des espèces fonctionnellement similaires (*S. caerulea* et *C. sempervirens*).

Discussion

Nous avons observé une forte variabilité des traits au cours de la saison de végétation mais peu entre années pour les quatre espèces qui ont été suivies en 2007 et 2008, ce qui est en accord avec la bibliographie (Garnier *et al.* 2001, Al Haj Khaled *et al.* 2005). Les traits foliaires des herbacées étaient plus variables que ceux des ligneux caducifoliés et plus variables que ceux des sempervirents, ce qui peut s'expliquer par un turnover plus rapide des ressources dans les feuilles ayant une durée de vie plus courte. La variabilité des traits pour les espèces ayant des feuilles à longue durée de vie étant visible sur de plus longues périodes d'observation (Pornon & Lamaze 2007).

En milieu alpin, l'évolution des traits au cours de la saison est probablement gouvernée par la date de déneigement et la somme des degrés-jours, ce sont ces variables qui influencent notamment la phénologie des espèces (Rome 2009). Les populations d'une même espèce devraient donc voir leurs traits évoluer de la même manière au cours de la saison, mais avec des départs différés par les dates de déneigement. Comme nous n'avons suivi l'évolution des traits que pour une population de chaque espèce nous n'avons pas pu faire le lien entre la variabilité des traits et les mesures de degrés-jours. Finalement nous avons échantillonné en suivant au mieux l'avancée de la saison de végétation en prenant la floraison comme repère et avec un échantillonnage des espèces herbacées en juillet et des espèces ligneuses en août. Les résultats présentés ici n'ont pas été pris en compte quantitativement pour apporter une éventuelle correction aux traits mesurés ou pour quantifier la variabilité due à la durée de l'échantillonnage pour chaque espèce. Il aurait fallu pour cela : (1) effectuer pour chaque espèce des suivis temporels de traits sur plusieurs populations situées dans des environnements contrastés. (2) faire des mesures locales de température au sol (par exemple avec des iButtons) pour connaitre la date de déneigement réel subie par chaque population. Le déneigement local n'est en effet pas uniquement lié aux gradients régionaux et dépend très fortement de la microtopographie et de la micro-orientation des sites.

Il est intéressant de noter que dans les données présentées ici (Fig. 2.3), il existe une forte variabilité individuelle des traits. Nous avons utilisé ici les moyennes de traits au niveau population afin de voir l'évolution générale des populations. Cependant à une date donnée, les mesures individuelles sont largement dispersées autour de la moyenne de la population (la quantification de cette variabilité individuelle est étudiée dans la section suivante). Cela est probablement du à divers facteurs : adaptation locale, plasticité, ontogénie, phénologie.

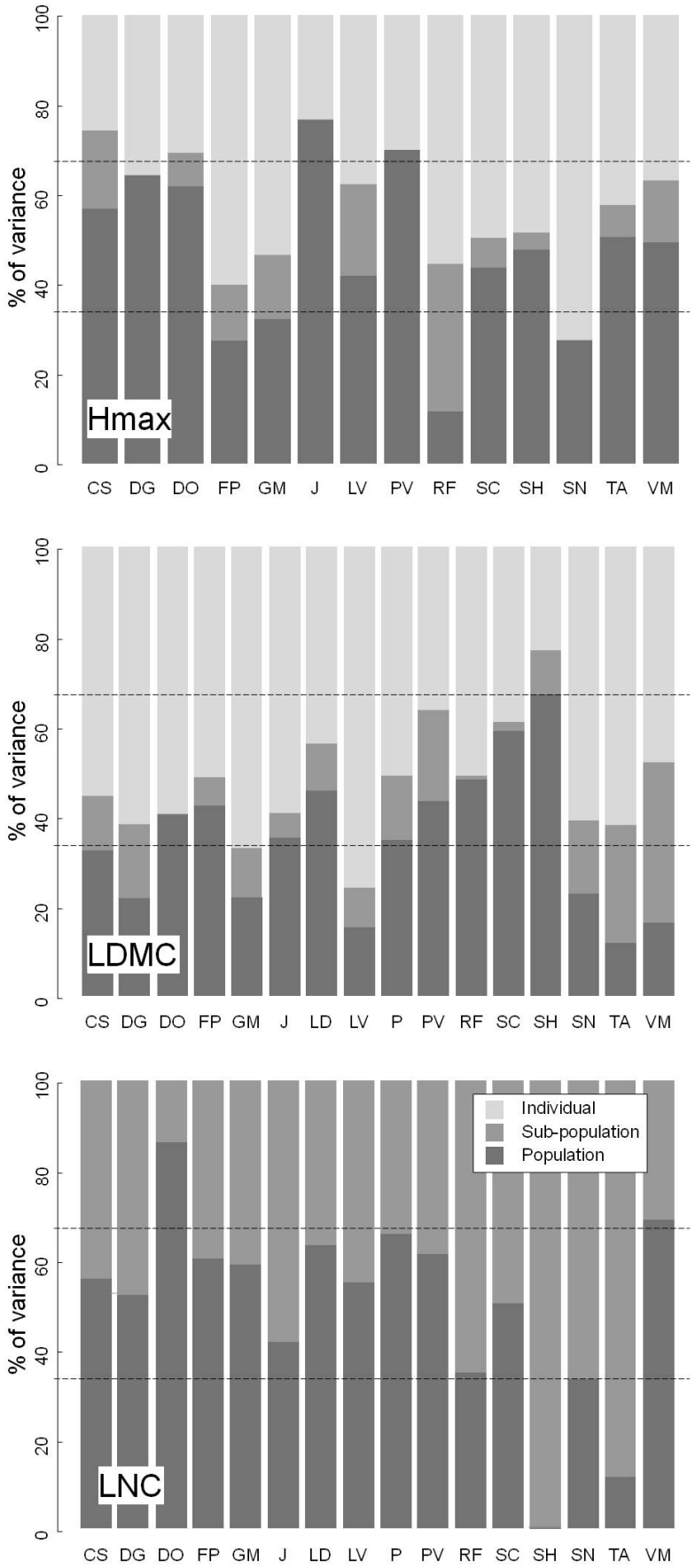


Fig. 2.7. Décomposition de la variance (issue de l'Article 7). *Décomposition de la variance relative des traits aux niveaux des individus, des sous-populations et des populations pour (a) la hauteur végétative (14 espèces) ; (b) la teneur des feuilles en matière sèche (16 espèces) et (c) la teneur des feuilles en azote (16 espèces). Les espèces sont classées par ordre alphabétique. Les seuils à 33% et 66 % sont donnés par les lignes pointillées.*

Cela ajoute de la complexité à la définition d'une date optimale d'échantillonnage pour une population d'une espèce donnée.

L'évolution temporelle simultanée des différents traits fonctionnels (augmentation de Hmax et LDMC et diminution de LNC) entraîne un « glissement » des espèces dans l'espace des traits (Fig. 2.4 & 2.5) et une modification des stratégies fonctionnelles moyennes des populations (éloignement d'elle-même et des autres dans l'espace des traits). La distance entre les populations de deux espèces dans cet espace des traits reste constante ou augmente avec l'avancement de la saison ce qui entraîne une conservation de la classification fonctionnelle des espèces au cours de la saison en accord avec la bibliographie (Garnier *et al.* 2001, Al Haj Khaled *et al.* 2005 ; Fig. 2.5).

// Les sections 3 & 4 correspondent aux Articles 7 & 8, pour cette raison je n'y détaillerai pas les méthodes utilisées.

3. Quantification et structure des variabilités fonctionnelles intra- et interspécifique

J'ai orienté l'étude de la variabilité fonctionnelle intraspécifique en suivant quatre questions/objectifs principaux :

1. Quelle est l'ampleur de cette variabilité ?
2. Quelle est la structure spatiale de cette variabilité au sein des espèces ?
3. Cette variabilité est elle négligeable devant la variabilité interspécifique ?
4. Qu'implique-t-elle sur la définition des stratégies fonctionnelles ?

3.1. Quelle est l'ampleur de la variabilité fonctionnelle intraspécifique ?

Pour répondre à cette première question nous avons calculé les coefficients de variation pour la hauteur végétative (Hmax), la teneur en matière sèche des feuilles (LDMC), la teneur en azote (LNC) et en carbone (LCC) des feuilles, la surface spécifique foliaire (SLA) et l'exposition de l'inflorescence (ExplInflo) pour les seize espèces échantillonnées en vallée de la Guisane. Pour chacune de ces espèces entre six et dix huit populations ont été échantillonnées en maximisant l'hétérogénéité des conditions environnementales prises en compte (Chapitre 1). Grâce à notre stratégie d'échantillonnage, nous avons réussi à échantillonner pour chaque espèce une part relativement large de sa distribution altitudinale (Tableau 1.2). Cette quantification a aussi permis de comparer la variabilité des différentes formes de vie (graminoïdes, herbacées, ligneux bas et haut), des espèces ayant des exigences écologiques différentes (amplitude de leur distribution altitudinale dans les Alpes françaises, altitude moyenne de cette distribution, amplitude altitudinale échantillonnée, degré de spécialisation Tableau 1.2).

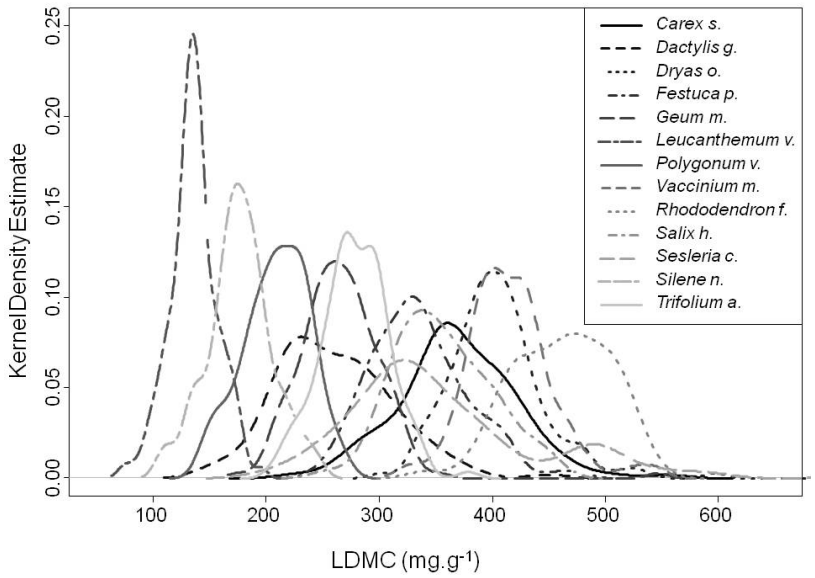


Fig. 2.8. Distributions de la teneur en matière sèche des feuilles entre et au sein des espèces (issue de l'Article 8) *L'estimation par densité de kernel est une méthode non paramétrique pour estimer la densité de probabilité d'une variable aléatoire. Ici elle donne l'estimation de la distribution de la teneur des feuilles en matière sèche pour les 13 espèces étudiées.*

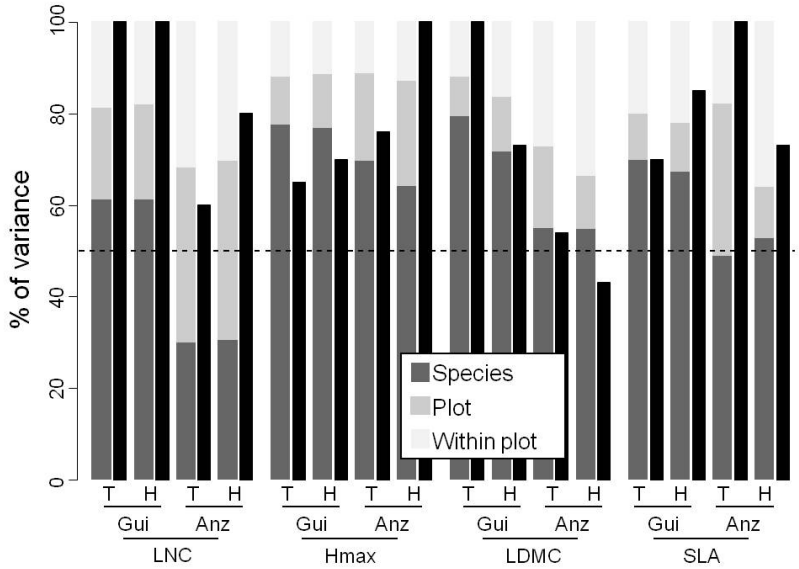


Fig. 2.9. Décomposition de la variance en contributions inter- et intraspécifique pour chaque trait (issue de l'Article 8). *Pour quatre traits (LNC= teneur des feuilles en azote, Hmax= Hauteur végétative, LDMC= teneur des feuilles en matière sèche et SLA= Surface spécifique foliaire) la figure donne la décomposition de la variance en pourcentages dus aux différents niveaux: espèce, entre populations et au sein des populations (niveaux de gris). La décomposition est donnée pour chaque site d'étude (Gui pour Guisane et Anz pour Anzeindaz) et pour toutes les espèces (T) ou seulement les espèces herbacées (H). Les barres noires indiquent la variance observée relativement à la variance la plus élevée observée dans les données pour chaque trait : par exemple pour SLA, la variance de toutes les espèces observées dans Anzeindaz est la plus forte et se voit donc affecter 100%, les autres variances étant calculées comme un pourcentage de cette variance.*

Les six traits mesurés se sont révélés fortement variables, aussi bien entre espèces (LDMC moyen entre 142 $mg.g^{-1}$ à 518 $mg.g^{-1}$), qu'au sein de chaque espèce. Hmax était particulièrement variable au sein des espèces (CV de 0,19 à 0,49); SLA était également assez fortement variable (CV de 0,13 à 0,32), plus que LNC (CV de 0,09 à 0,29) et que LDMC (CV de 0,07 à 0,24). Ces différences de variabilité entre traits corroborent les résultats de Garnier *et al.* (2001), Roche *et al.* (2004) et Wilson *et al.* (1999) qui prônent l'utilisation de la LDMC comme trait le plus robuste pour localiser les espèces sur un axe d'acquisition/conservation des ressources (Fig. 2.6).

Toutes les espèces ne se sont pas montrées autant variables et tous les traits mesurés pour une même espèce n'étaient pas toujours autant variables. Par exemple *T. alpinum* était l'espèce la plus variable pour Hmax (CV=0,46) et la moins variable pour LNC (CV=0,09). La question de la variabilité doit donc se poser non pas à l'échelle de l'espèce mais plutôt à l'échelle du trait. D'ailleurs nous n'avons trouvé aucun patron de variabilité par rapport aux formes de vie, aux exigences écologiques ou à l'ampleur de l'échantillonnage effectué pour chaque espèce (analyses non présentées). Cette absence de patron pourrait être due au caractère idiosyncratique de la variabilité intraspécifique, cependant notre dispositif n'est pas très puissant pour montrer des différences entre formes de vie par exemple (une seule légumineuse).

La forte variabilité fonctionnelle intraspécifique que nous avons observée et dont l'ampleur dépendait à la fois des traits et des espèces suggère que l'utilisation de moyennes de traits pour décrire une espèce est une approche très restrictive de la réalité. Il est évident que la variabilité que nous avons mesurée sur le terrain résulte d'un très grand nombre de facteurs et processus environnementaux, biologiques et écologiques comme : l'âge des individus (impossible à connaitre pour les espèces herbacées) et des feuilles (normalement standardisé par le protocole de mesure : c'est la dernière feuille mature qui doit être mesurée) ; les conditions abiotiques, les conditions biotiques, les perturbations et les différences d'avancement phénologique entre individus. Comme nous l'avons déjà évoqué dans la section précédente, les possibles différences d'avancement phénologiques entre individus et entre populations sont une limite évidente de cette étude. En effet ces différences ne sont pas facilement mesurables et sont donc inclues implicitement dans nos mesures de terrain, du fait de la nécessité pour notre étude de couvrir beaucoup de populations en une saison. Cependant nous avons procédé comme cela est fait usuellement dans les études reposant sur des traits fonctionnels en suivant les protocoles établis (Garnier *et al.* 2001 ; Cornelissen *et al.* 2003 ; Garnier *et al.* 2001), nous avons donc des erreurs de mesures qui doivent être du même ordre de grandeur que celles existant dans la littérature et les bases de données.

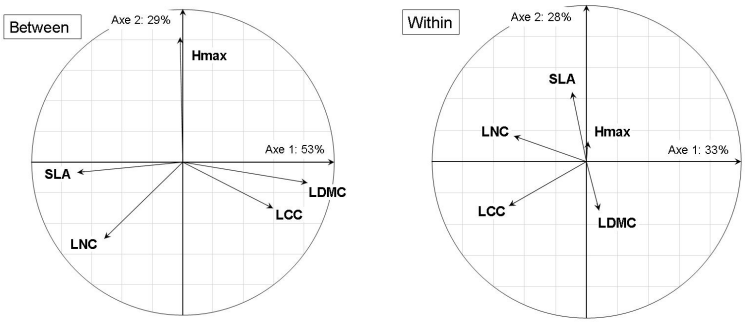


Fig. 2.10. Structure fonctionnelle multidimensionnelle et compromise de traits à l'échelle interspécifique (issue de l'Article 8). *Cercles de corrélation entre les 5 traits mesurés (LNC= teneur des feuilles en azote, Hmax= Hauteur végétative, LDMC= teneur des feuilles en matière sèche ; SLA= Surface spécifique foliaire et LCC= teneur des feuilles en carbone), obtenus à partir d'ACP inter-groupes.*

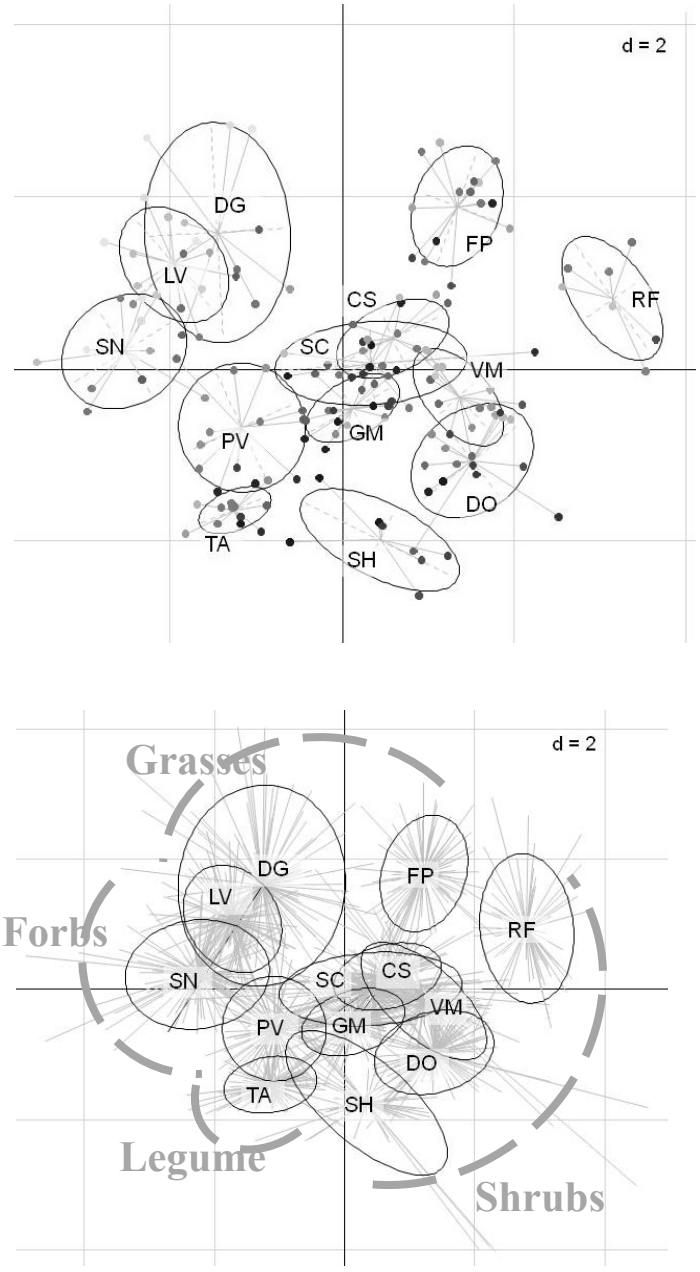


Fig. 2.11. Dispersion des différents niveaux d'organisation étudiés : espèces, populations et individus dans l'espace des traits (issue de l'Article 8).
Les ellipsoïdes d'inertie englobent 65% des données individuelles de chaque espèce, pour une question de lisibilité. On retrouve la projection des centres de gravité de chaque espèce et celle des individus (étoiles grises) définie à partie d'ACP inter-groupe. (A) Projection des centres de gravité (moyennes) des espèces et des populations dans le plan de l'espace des traits défini par les axes 1 et 2 de l'ACP inter-groupe. Les traits moyens de chaque population étudiée sont représentés par des points dont la couleur correspond à la température (Tminh) de la population. Les points les plus foncés (resp. clairs) correspondent aux populations les plus froides (resp. chaudes). (B) Projection des centres de gravité des espèces et de tous les individus dans l'espace des traits défini par les deux premiers axes de l'ACP inter-groupe. Les demi-cercles gris montrent l'agencement des différentes formes de vie dans l'espace des traits

3.2. Quelle est la structure spatiale de la variabilité fonctionnelle intraspécifique au sein des espèces ?

Etant donnée l'ampleur de la variabilité observée au paragraphe précédent, nous avons cherché à comprendre comment celle-ci était structurée spatialement (individu, sous-population, population et gradient). Nous avons décomposé la variabilité en différents niveaux grâce à des modèles mixtes, adaptés à l'analyse de données hiérarchisées (Encadré 2.1).

En moyenne, une grande part de la variabilité était due aux différences entre populations (Hmax: 47%, LDMC: 35% et LNC: 50%) résultant potentiellement de différences génétiques et des conditions environnementales contrastées (Fig. 2.7). Etonnement une grande partie de la variabilité restante était due aux différences entre individus au sein d'une sous-population (en moyenne 53% pour LDMC) et une très faible part de la variabilité était due aux sous-populations en elles-mêmes (environ 10%). En sélectionnant des individus dans de petites sous-populations contrastées on s'attendait au contraire (1) à ce que les sous-populations intègrent une large partie de la micro-hétérogénéité de l'environnement et (2) à ce que les individus d'une même sous-population soient génétiquement liés et phénotypiquement plus similaires. D'autant plus que pour certaines espèces chez lesquelles il est difficile d'identifier les individus (*Dryas octopetala*, *Salix herbacea*, *Vaccinium myrtillus* and *Trifolium alpinum*) nous avons mesuré des ramets, plusieurs ramets d'une même sous-population provenant donc potentiellement d'un même individu génétique. Cela signifie donc qu'il existe une très forte variabilité phénotypique au sein d'un même individu, comme cela a déjà été montré chez les arbres (Maguire & Bennett 1996).

Au final il y a donc peu de structuration spatiale de la variabilité fonctionnelle, ce qui pourrait résulter (i) soit d'un fort flux de gènes qui ferait que des individus proches spatialement ne sont pas nécessairement apparentés ; (ii) soit d'une forte plasticité des génotypes en réponse à des conditions micro-environnementales très hétérogènes (iii) soit d'un manque de puissance de notre dispositif (seulement trois individus mesurés dans chaque sous-population) pour montrer cette structuration au sein des populations. De plus comme nous l'avons dit en introduction, notre étude ne permet pas de trancher entre adaptations locales et plasticité car la variabilité fonctionnelle est une mesure de la variabilité phénotypique et résulte des deux mécanismes (Byars *et al.* 2007). Déterminer l'importance relative des contributions génétiques et environnementales dans la réponse des espèces aux gradients environnementaux reste un problème non résolu (Byars *et al.* 2007), d'autant plus que tous les traits fonctionnels ne sont probablement pas déterminés par les mêmes mécanismes (Woodward 1986; Cordell *et al.* 1998), et que le degré de plasticité et d'adaptation pourrait dépendre des espèces (Hovenden & Vander Schoor 2003). Nos résultats vont dans ce sens puisque la décomposition de la variabilité intraspécifique s'est révélée très différente entre traits et entre espèces (Fig. 2.7). Pour la plupart des espèces, la

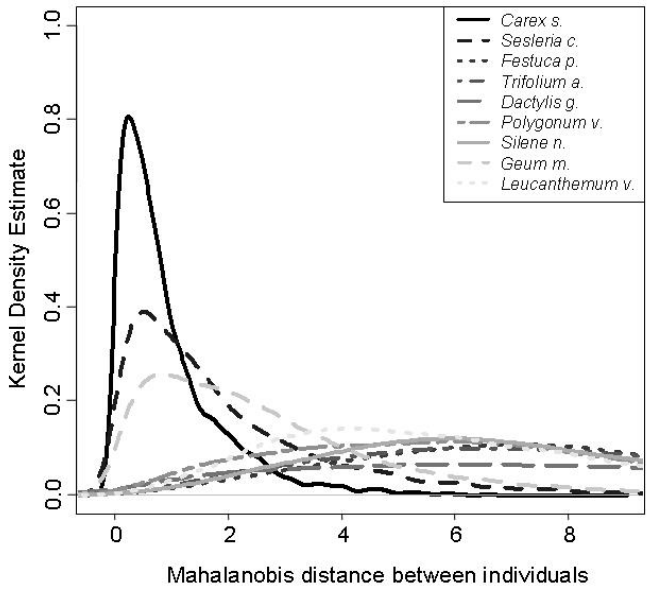


Fig. 2.12. Distances fonctionnelles entre individus (issue de l'Article 8). *L'estimation par densité de kernel est une méthode non paramétrique pour estimer la densité de probabilité d'une variable aléatoire. Ici l'estimation est faite pour la distance de Mahalanobis entre les individus dans l'espace des traits. On a représenté uniquement la distance des individus de chaque espèce aux individus de Carex sempervivens.*

décomposition de la variabilité des traits était notamment très différente, par exemple pour *T. alpinum* 50% de la variabilité de Hmax était due aux différences entre populations, contre 12% pour LNC et LDMC.
Encore une fois nous n'avons trouvé aucun patron de décomposition de la variabilité par rapport aux formes de vie, aux exigences écologiques ou à l'ampleur de l'échantillonnage effectué pour chaque espèce.

3.3. La variabilité fonctionnelle intraspécifique est elle négligeable devant la variabilité interspécifique ?

Afin de répondre à cette question, nous avons comparé les variabilités inter et intraspécifiques des traits fonctionnels collectés dans la base Guisane 2080 pour les deux vallées d'étude : Guisane et Anzeindaz (Chapitre 1). D'une part l'analyse a été menée sur chaque trait séparément en décomposant leur variance en composantes inter et intraspécifique à l'aide de modèles mixtes (Encadré 2.1). D'autre part, afin de savoir si plusieurs traits considérés simultanément permettaient de mieux représenter les différences entre espèces, la décomposition a été effectuée dans un espace multidimensionnel de traits (Hmax, LDMC, SLA, LNC et LCC) à l'aide d'ACP inter et intra-groupes (Encadré 2.4).

La variabilité intraspécifique provoque un recouvrement des attributs de traits entre espèces, résultant en un continuum fonctionnel à l'échelle interspécifique (Fig. 2.8). Malgré ce fort recouvrement une grande partie de la variabilité (en moyenne 70%) reste due aux différences entre espèces (Fig. 2.9). Des partitions similaires ont été obtenues pour les données de Guisane réduites aux espèces herbacées (Fig. 2.9). Par contre nous avons observé moins de variabilité interspécifique pour les données de Anzeindaz du fait que moins d'espèces (et des espèces moins contrastées) y ont été échantillonnées. L'approche multidimensionnelle a donné la même décomposition: 70% d'inter et 30% d'intra. Ces décompositions dépendent naturellement beaucoup du choix des espèces, cependant nous avons tenté de maximiser la variabilité interspécifique (en Guisane) en choisissant au sein des différentes formes de vie des espèces connues pour leurs stratégies fonctionnelles différentes (ex. *D. glomerata*, *F. paniculata* et *S. caerulea* voir Gross *et al.* 2007).

L'hypothèse clé sur laquelle repose l'utilisation des traits fonctionnels est que ces traits sont plus variables entre qu'au sein des espèces, voire que la variabilité intraspécifique est négligeable devant l'interspécifique (Garnier *et al.* 2001; Shipley 2003; Roche *et al.* 2004; Al Haj Khaled *et al.* 2005; McGill *et al.* 2006). Nos résultats ne sont pas complètement en accord avec cette hypothèse. La variabilité intraspécifique demeure généralement plus faible que la variabilité inter-, mais les analyses uni- et multi-variées que nous avons menées permettent toutes deux d'identifier une composante intraspécifique forte dans la variabilité (autour de

Species / Trait	*Hmax*	*LDMC*	*LNC*
CS	*10 (10)*	*79 (86)*	*28 (82)*
DG	*78 (100)*	*11 (17)*	*4 (55)*
DO	*72 (76)*	*11 (11)*	*44 (53)*
FP	*0 (0)*	*0 (0)*	*25 (0)*
GM	*65 (100)*	*0 (26)*	*0 (4)*
J	*79 (98)*	*0 (0)*	*23 (56)*
LD		*3 (31)*	*36 (50)*
LV	*13 (13)*	*36 (36)*	*0 (0)*
P		*50 (100)*	*15 (58)*
PV	*0 (0)*	*0 (0)*	*22 (22)*
RF	*0 (0)*	*48 (48)*	*0 (0)*
SC	*49 (95)*	*0 (58)*	*0 (17)*
SH	*20 (20)*	*96 (96)*	*0 (0)*
SN	*40 (40)*	*8 (8)*	*0 (0)*
TA	*0 (0)*	*41 (83)*	*100 (100)*
VM	*49 (81)*	*29 (100)*	*35 (35)*

Tableau 2.2. R² des modèles mixtes (issu de l'Article 7). *Partie de la variance entre population expliquée par les gradients de radiation et température. Entre parenthèse, les R² des meilleurs modèles calibrés parmi les différents modèles incluant les gradients, le sol, la perturbation et la compétition.*

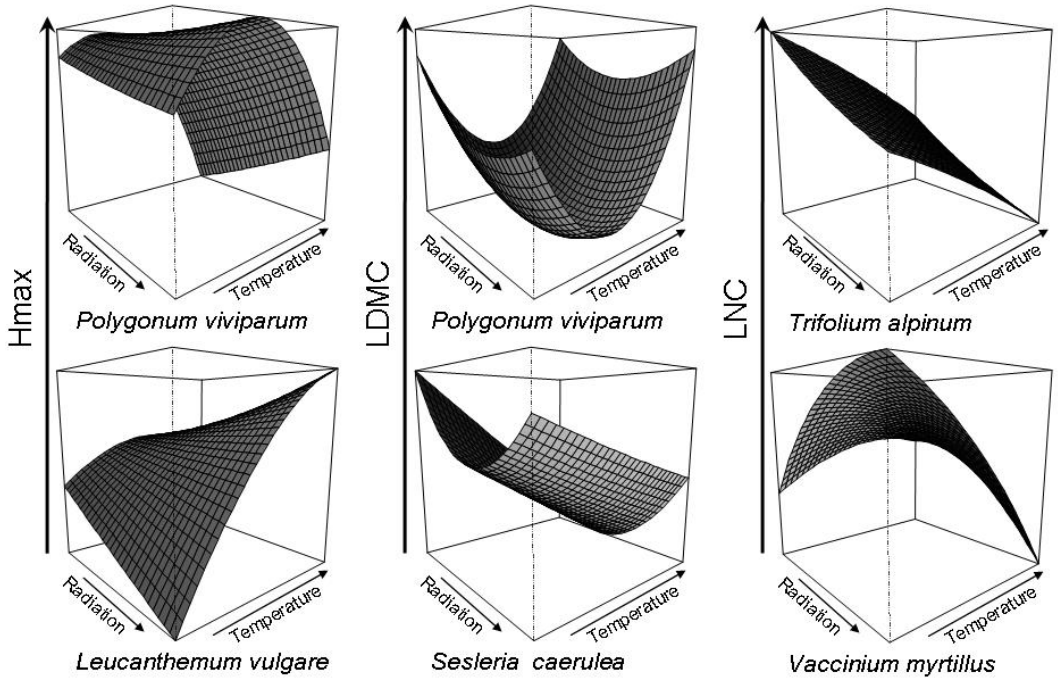


Fig. 2.13. Surfaces de réponse des traits le long des gradients (issue de l'Article 7).
Deux exemples de surfaces de réponse sont donnés pour chaque trait mesuré. Ces surfaces résultent des modèles linéaires avec les traits comme variable de réponse et des termes linéaires et quadratiques et interaction des gradients comme variables explicatives. Polygonum viviparum et Trifolium alpinum ont été mesurées au milieu de leur distribution altitudinale. Leucanthemum vulgare, Sesleria caerulea et Vaccinium myrtillus ont été mesurées dans la partie supérieure de leur distribution altitudinale.

30%). L'importance relative des parts inter- et intraspécifique de la variabilité dépend évidemment des espèces considérées, cependant dans cette étude nous avons pris soin de choisir des espèces particulièrement contrastées au sein de la flore alpine dans le but de maximiser la variabilité interspécifique. Admettre ou non que la variabilité intraspécifique est négligeable va donc fortement dépendre du système étudié et des traits et espèces sélectionnés. Il faudrait maintenant tester les implications d'une telle variabilité intraspécifique sur les utilisations (ex. indices de diversité fonctionnelle) des traits afin de décider au cas par cas si l'hypothèse de robustesse est applicable ou non, c'est ce que nous ferons dans le Chapitre 4.

De plus il est essentiel de noter que les mesures de traits individuelles (par opposition avec les traits moyennés par population) ont un rôle considérable dans nos résultats. En effet les partitions de variance recalculées à partir de données moyennées par population donnent en moyenne 85% de variance inter et 15% de variance intraspécifique. Soit 15% d'intraspécifique de moins qu'avec les données individuelles. La variabilité intraspécifique est donc largement sous-estimée lorsque les traits sont moyennés au niveau de la population. Or ces mesures moyennées par populations sont ce qu'il y a de plus courant dans les études fonctionnelles (ex. Shipley 2000; Garnier *et al.* 2001; Roche *et al.* 2004). Nos résultats préconisent donc l'utilisation de mesures individuelles de traits et l'analyse des syndromes de traits à l'échelle individuelle (Fonseca *et al.* 2000; Vile *et al.* 2006; Kattge *et al.* in prep), plutôt qu'à l'échelle de la population ou de l'espèce (Wright *et al.* 2004, Diaz *et al.* 2004). Du moins il serait intéressant de disposer de données sur les syndromes de traits des individus dans les bases de données (Kattge *et al.* in prep, Madin *et al.* 2007) pour permettre de faire des analyses à différents niveaux, les études au niveau individuel et au niveau des populations ou des espèces étant différentes et complémentaires. Les données stockées au niveau individuel peuvent toujours être moyennées par la suite.

3.4. Qu'implique la variabilité fonctionnelle intraspécifique sur la définition des stratégies fonctionnelles ?

La variabilité fonctionnelle intraspécifique ne semble pas négligeable *a priori*. Nous avons donc cherché à savoir comment se répartissent, dans l'espace des traits (Hmax, LDMC, SLA, LNC et LCC), les individus autour des moyennes par espèce et ce qu'implique cette répartition des individus sur la définition des stratégies fonctionnelles. Afin de déterminer les principaux axes de variabilité structurant le jeu de données, nous avons utilisé des ACP inter et intra-groupes (Encadré 2.4). Nous avons également exécuté des ACP simples sur les données individuelles, les données moyennées par plot et les données moyennées par espèce afin de tester la robustesse des axes de variabilité aux différents niveaux de résolution. D'autre part nous avons calculé les distances fonctionnelles (Encadré 2.3) entre tous les individus mesurés.

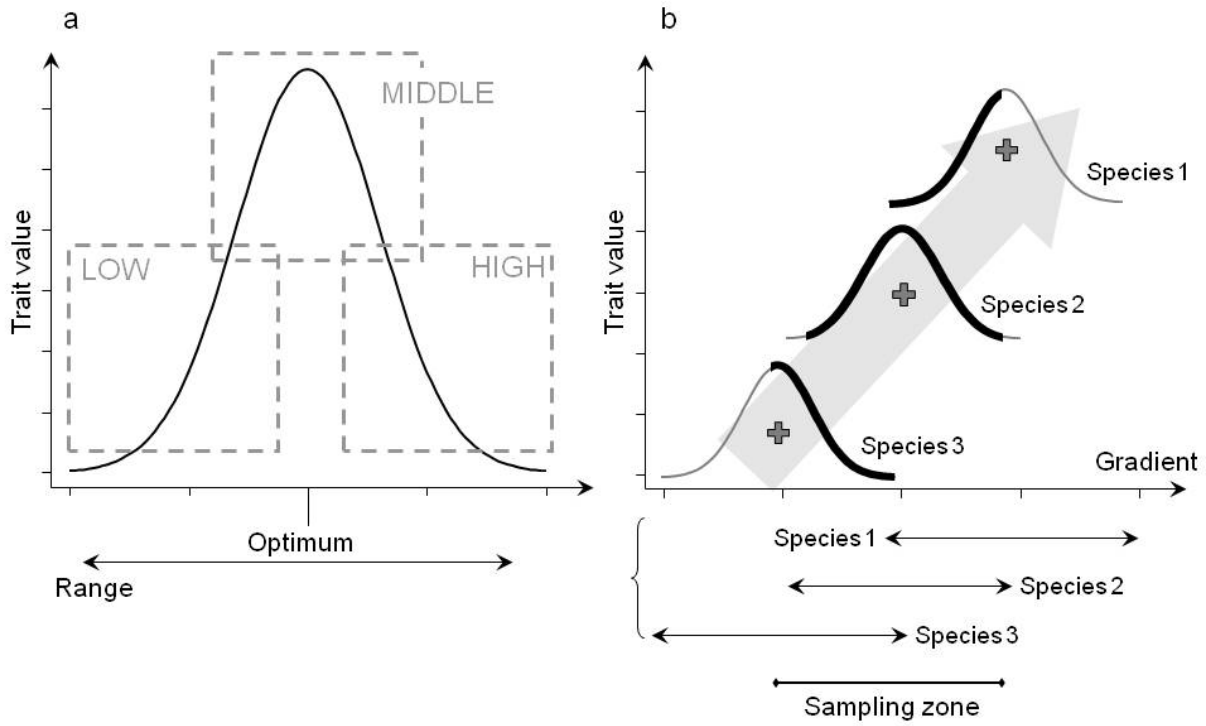


Fig. 2.14. Réponse des traits le long des gradients aux échelles inter- et intraspécifique : modèle conceptuel (issue de l'Article 7).

(A) Forme générique supposée de la courbe de réponse d'un trait le d'un gradient pour l'ensemble de la distribution d'une espèce le long de celui-ci. Le trait suit une courbe unimodale, atteignant une valeur maximale de trait pour les conditions environnementales optimales pour l'espèce considérée et décroissant de part et d'autre vers les extrémités de la distribution. Le trait peut-être mesuré sur le terrain dans la partie « basse » (low), « intermédiaire» (middle) ou supérieure (high) de sa distribution le long du gradient, conduisant à observer des courbes de réponse entre trait et gradient différentes.

(B) Analyse théorique multi-espèces. Trois espèces hypothétiques sont échantillonnées le long du gradient. L'espèce 1 (Species 1) est échantillonnée sur la partie basse de sa distribution, l'espèce 2 (Species 2) est échantillonnée sur la partie intermédiaire de sa distribution et l'espèce 3 (Species 3) sur la partie haute de sa distribution. A l'échelle intraspécifique, les relations trait-gradient semblent donc idiosyncratiques. De plus la relation trait-gradient à l'échelle interspécifique est très différente de la relation générique intraspécifique (A).

Le jeu de données individuelles était structuré par un premier axe expliquant 53% de la variance et fortement corrélé à la LDMC (positivement) et à la SLA (négativement) et un deuxième axe expliquant 29% de la variance et principalement corrélé à Hmax (Fig. 9). Les structures obtenues avec des données individuelles, moyennées par population et par espèces étaient similaires, de même que les structures obtenues avec les données d'Anzeindaz ou avec des jeux de données réduits aux espèces herbacées ou aux ligneux bas. Ces résultats corroborent les résultats classiques sur les compromis de traits liés à la dynamique des ressources et exprimés par la corrélation négative entre SLA et LDMC (Grime *et al.* 1997; Diaz *et al.* 2004; Gross *et al.* 2007; Pierce *et al.* 2007). Ces résultats montrent que la définition des compromis de traits et des stratégies fonctionnelles selon l'axe de dynamique des ressources est robuste aux mesures individuelles et donc à la variabilité intraspécifique. On retrouve dans l'espace des traits (Fig. 2.11) un regroupement des centres de gravité des espèces par formes de vie: graminoïdes, autres herbacées et ligneux bas ; et une séparation des espèces à stratégie « exploitatrice » (forte SLA et faible LDMC comme *D. glomerata.* et *L. vulgare*) et des espèces à stratégie « conservatrice » (faible SLA et forte LDMC comme *F. paniculata* et *C. sempervirens*). Par contre au niveau intraspécifique, les compromis de traits semblent moins forts avec seulement 30% de la variance expliquée par l'axe de dynamique des ressources ; ils sont de plus variables entre espèces.

L'inclusion des mesures individuelles ne modifie pas les axes de variations, et donc les grandes stratégies, mais ajoute un bruit aux patrons observés et conduit au recouvrement des nuages d'individus de chaque espèce dans l'espace des traits (ellipse=65% des individus inclus). De la même manière qu'avec un seul trait (Fig. 2.8 & 2.11) on retrouve donc un continuum fonctionnel plutôt qu'une distribution discrète d'attributs de traits avec des espèces différenciées. L'analyse des distances fonctionnelles entre individus confirme l'existence de ce continuum. Les individus d'une espèce ayant des attributs de traits particulièrement marginaux peuvent être fonctionnellement plus proches d'individus d'autres espèces que des individus de sa propre espèce (Fig. 2.12). Il ne semble donc pas évident d'attribuer une stratégie donnée à une espèce. Il peut par conséquent sembler grossier de décrire les espèces par des moyennes de trait. Il serait peut-être plus réaliste de considérer les traits de chaque espèce non plus comme des valeurs moyennes, mais comme des distributions de traits autour des valeurs moyennes, comme cela a été suggéré par Lavorel *et al.* (2007) et Kattge *et al.* (2009). Cependant trancher entre les études empiriques et théoriques pour lesquelles il suffit d'utiliser des moyennes et celles pour lesquelles il faudrait plutôt utiliser des distributions est certainement une question d'échelle d'observation. Nous reviendrons sur ce point.

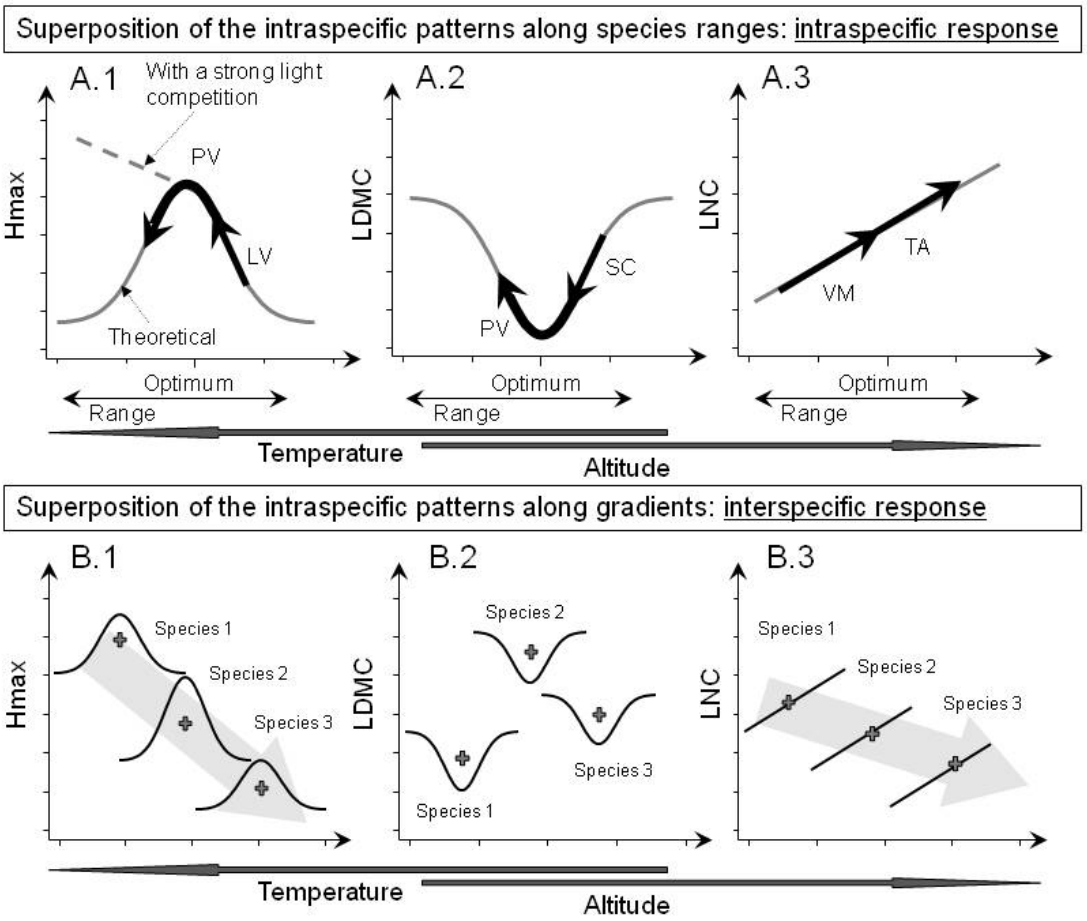


Fig. 2.15. Réponse des traits le long des gradients aux échelles inter- et intraspécifique : applications aux observations des trois traits étudiés (issue de l'Article 7).

(A) Variation intraspécifique générique supposée de Hmax, LDMC et LNC (ligne grise), et superposition avec les observations pour les deux espèces de la Fig. 2.12 (flèches noires) le long de leur distribution.

(B) Variation interspécifique de Hmax, LDMC et LNC (flèches grises) exprimée comme la superposition des réponses intraspécifiques de trois espèces hypothétiques le long d'un gradient de température (décroissant). Agréger les réponses intraspécifiques le long de la distribution de chaque espèce peut conduire à des patrons interspécifiques très différents des patrons génériques intraspécifiques.

4. Rôle des gradients

Nous avons vu à la section précédente qu'il existait une forte variabilité intraspécifique et qu'une partie variable de celle-ci était due aux différences entre populations. Cette nouvelle section vise à expliquer cette variabilité entre populations par les gradients environnementaux régionaux et les conditions locales.

4.1. Rôle des gradients environnementaux dans la variabilité fonctionnelle

J'ai cherché à répondre à trois questions:
1. Pour chaque espèce et chaque trait, quelle est la part de la variance entre populations expliquée par les gradients régionaux ?
2. Des variables locales de sol, compétition et perturbations améliorent-elles la part de variance expliquée ?
3. Comment les différents traits répondent-ils aux gradients en termes de courbes de réponse ?

Nous avons fondé notre échantillonnage sur deux gradients directs orthogonaux, un gradient de température et un gradient de radiation (Chapitre 1), en faisant l'hypothèse que les traits suivraient une réponse linéaire (dans le cas d'un gradient fortement limitant) ou une courbe en cloche (exprimant un optimum de trait le long du gradient). L'équation suivante pourrait donc décrire la variation des traits le long des gradients :
Trait ~ Température + Température^2 + Radiation + Radiation^2 + Température* Radiation
Nous avons utilisé des modèles mixtes - avec l'équation ci-dessus comme partie fixe et les niveaux de variabilité (population, sous-population, individu) comme partie aléatoire - pour tester cette hypothèse et déterminer la part de variance entre populations due aux gradients (Encadré 2.1). Nous avons également testé l'inclusion de facteurs locaux dans les modèles: caractéristiques du sol, perturbation anthropique, compétition (Chapitre 1).

Les différents modèles calibrés convergeaient et avaient des effets significatifs. (1) La part de variance entre populations expliquée par les gradients (R^2) était en moyenne plus importante pour Hmax (30%) et LDMC (26%) que pour LNC (21%). Il y avait encore une fois de larges différences entre espèces, différences non expliquées par les formes de vie ou les exigences écologiques des espèces (Tableau 2.2 et voir 3.1). (2) Inclure des variables locales comme le sol ou la perturbation anthropique dans les modèles a fortement accru la part de variance expliquée pour certains traits et certaines espèces (43% pour Hmax et LDMC; Tableau 2.2). Des patrons attendus sont ressortis comme par exemple une sensibilité de *F. paniculata* à la perturbation (Jouglet 1999). (3) La majorité des surfaces de réponse obtenues étaient à tendance quadratique avec de forts effets interactifs entre gradients (Fig. 2.13). Aucun patron général n'est pour autant ressorti. Ces surfaces semblaient fortement idiosyncratiques et

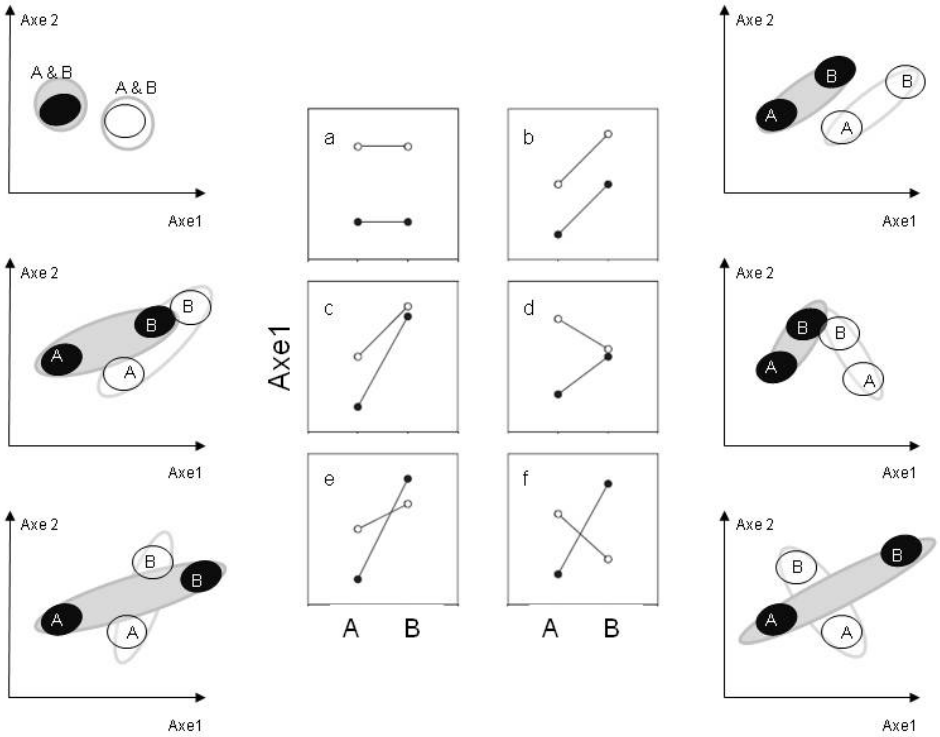


Fig. 2.16. Variation des traits et classement des espèces.
Panneau central (adapté de Garnier 2001): Variation hypothétique des valeurs de traits de deux espèces (une noire une blanche) poussant dans différentes conditions environnementales (A et B). Les variations représentées dans les panneaux a à d mènent à des classements des espèces similaires dans les deux types de conditions. Les cas présentés dans les panneaux e et f montrent des situations pour lesquelles le classement n'est pas conservé.

Autour : pour chaque cas du panneau central, la même situation est montrée en deux dimensions. Les distributions des valeurs de traits prises par une espèce sont représentées par des ellipses (grise pour l'espèce noir et blanche pour l'espèce blanche). Les valeurs de traits des populations vivant dans les conditions A et B sont données par les disques noirs et blancs, marqués d'un A ou d'un B selon les conditions. Cas a : pas de variation des traits entre les environnements A et B ; Cas b : les deux espèces ont la même réponse (même sens et même force) au changement de conditions ; Cas c : les deux espèces répondent dans le même sens mais pas avec la même force, la différence des forces de réponse n'est pas assez forte pour amener à une modification du classement des deux espèces ; Cas d : les deux espèces ne répondent pas dans le même sens, les réponses ne sont pas suffisamment fortes pour amener à une modification du classement des deux espèce; Cas e : les deux espèces répondent dans le même sens mais pas avec la même force, la différence des forces de réponse est pas assez forte pour amener à une modification du classement des deux espèce ; Cas f : les deux espèces ne répondent pas dans le même sens, les réponses sont suffisamment fortes pour amener à une modification du classement des deux espèce;

aucun rapprochement n'a pu être effectué entre espèces d'une même forme de vie ou ayant des exigences écologiques similaires par exemple (Tableau 1.2).
Nous avons vu au paragraphe 3.2 que les différences de traits entre populations d'une même espèce pouvaient résulter à la fois d'adaptations génétiques locales et de plasticité phénotypique. Le manque de patron général dans les courbes de réponse pourrait être du à la présence de différents génotypes ayant des réponses différentes à l'environnement. Cependant F. Grassein (2009) a montré que différents génotypes au sein d'une espèce (ex : *F. paniculata* et *D. glomerata*) pouvaient avoir des réponses similaires (force et direction) aux changements environnementaux. Cette unicité de réponse serait accompagnée d'une large part de variabilité intraspécifique due aux conditions environnementales avec 70% d'effets environnementaux contre 30% d'effets génétiques. Dans ce cas, connaître les courbes de réponses des espèces dans leur globalité (c.à.d. incluant les différents génotypes) serait pertinent en soi et le caractère idiosyncratique des courbes de réponse aurait donc une autre origine.

La section suivante est destinée à la description d'un modèle conceptuel visant à expliquer le patchwork complexe et partiellement inexpliqué de nos résultats.

4.2. Modèle conceptuel

Malgré les efforts particuliers que nous avons faits pour avoir un dispositif d'échantillonnage comparable entre les différentes espèces étudiées, nous n'avons pas obtenu plus de patron général qu'avec l'étude des données des bases de traits et de la littérature: l'apparente complexité et incohérence des résultats n'est donc pas due à des problèmes de comparabilité des données. De plus des patrons ont déjà été mis en évidence entre traits et gradients à l'échelle interspécifique (Reich *et al.* 1999; Wright *et al.* 2005).
Nous proposons donc un modèle conceptuel (Fig. 2.14 et 2.15) pour expliquer nos résultats. Supposons qu'au sein d'une espèce, les traits aient –en absence d'interactions biotiques- une réponse générique aux gradients, suivant par exemple une courbe en cloche, avec des paramètres différant entre espèces, avec par exemple un extremum de trait dans les conditions optimales pour l'espèce (Fig. 2.14-A). Ainsi à l'échelle intraspécifique, si l'on échantillonne uniquement la partie basse du gradient on obtiendra une relation positive entre trait et gradient ; tandis que si l'on échantillonne la partie haute on obtiendra une relation négative et si l'on échantillonne la partie intermédiaire on obtiendra une relation quadratique (Fig. 2.14-A). En fonction de la partie du gradient qui est échantillonnée et des exigences écologiques des espèces étudiées (largeur et optimum de niche, Chapitre 3), les relations observées entre trait et gradient pourront donc être très différentes et apparaître incohérentes et idiosyncratiques (Fig. 2.14-B), alors que le patron de réponse générique est unique (Fig. 2.14-A). D'un côté, appréhender la relation trait-gradient dans l'espace écologique (c.à.d. le long la niche) en recombinant les différents « morceaux » de courbes

Encadré 2.1. Modèles linéaires mixtes

librarie R: nlme - fonction: lme()

Principe (Pinheiro & Bates, 2000)

Ce sont des modèles linéaires contenant des effets fixes (valeurs fixées) et des effets aléatoires (valeurs issues de la réalisation d'une loi gaussienne de variance inconnue). Ils permettent en particulier d'analyser une structure hiérarchique dans les données qui induit une non-indépendance des unités observées au même niveau, rendant l'hypothèse d'indépendance simple (faite dans un modèle linéaire classique) inappropriée. Un modèle mixte avec un effet fixe et un effet aléatoire peut s'écrire:

$$Y_{ij} \sim \mu + F_j + A_i + e_{ij}$$

, où μ est la moyenne des observations, F un effet fixe , A un effet aléatoire dont les niveaux (A_1,.. A_i) suivent une loi normale $N(0,\sigma_a^2)$ avec σ_a^2 inconnue, les résidus e_{ij} suivent une loi normale $N(0, \sigma^2)$ avec σ^2 inconnue. Yj suit donc une loi normale $N(\mu+F_j, \sigma_a^2+\sigma^2)$.

Décomposition de la variance (Xu, 2003)

A partir des estimations des paramètres σ_a^2 et σ^2 (maximisation de la vraisemblance restreinte REML) on peut donc décomposer la variance en ses composantes aux différents niveaux hiérarchiques (ex. entre et au sein des populations). Dans notre exemple la part de variance due aux différences entre niveaux est donnée par $\sigma_a^2/(\sigma_a^2+\sigma^2)$ et celle au sein des niveaux par $\sigma^2/(\sigma_a^2+\sigma^2)$.

Part de variabilité expliquée par un facteur fixe

On peut déterminer cette part en calculant:

$$R^2 = 1 - \frac{\sigma a^2}{\sigma a_0{}^2}$$

, où σ_a^2 est la variance du facteur aléatoire dans l'exemple ci-dessus et σ_{a0}^2 la variance du facteur aléatoire du même modèle sans facteur fixe.

observés pour chaque espèce dans les parties basse, intermédiaire ou haute du gradient (par rapport à l'optimum) conduira à la réponse générique comme sur la Figure 2.14. D'un autre côté, appréhender la relation trait-gradient dans l'espace environnemental (ex. climatique) en replaçant chaque espèce le long du gradient en fonction de ses exigences écologiques conduira à un patron interspécifique pouvant être très différent de la réponse générique (Fig. 2.14, Thuiller *et al.* 2004b).

Recombiner les courbes de réponse obtenues pour chaque trait en suivant ce modèle conceptuel nous a permis d'avoir une nouvelle interprétation de nos résultats. (1) Nous avons obtenu une courbe uni-modale pour Hmax le long du gradient de température avec une hauteur maximale obtenue pour des températures intermédiaires plus favorables (Fig. 2.15). Pour certaines espèces (*D. glomerata*, *S. caerulea*) Hmax augmentait avec la température du fait de conditions plus favorables et probablement d'une forte compétition pour la lumière (Woodward 1986). A l'échelle interspécifique, Hmax augmentait avec la température (R^2=17%) et nous n'avons pas trouvé de relation avec la radiation (Fig. 2.15). (2) La reconstruction conceptuelle pour LDMC a conduit à une cloche inversée, donnant des valeurs de LDMC plus faibles – correspondant à un taux de croissance plus élevé chez les herbacées (Weiher *et al.* 1999)-dans les conditions environnementales les plus adéquates au cœur de la niche (Fig. 2.15). LDMC augmentait ou suivait une courbe en cloche avec la radiation comme dans Shipley (2003). A l'échelle interspécifique, la bibliographie fait état de relations positives (Roche *et al.* 2004) ou négatives (Reich *et al.* 1999) entre LDMC et altitude, nous n'avons trouvé aucune réponse de LDMC aux gradients. (3) A l'échelle intraspécifique, LNC était décroissant le long du gradient de température (Fig. 2.15) avec un patron plus clair pour la teneur en azote surfacique ($mmol.m^{-2}$) que pour la teneur massique ($mg.g^{-1}$). Ce résultat est en accord avec la bibliographie éco-physiologique à l'échelle de l'espèce ou de la forme de vie (Woodward 1986; Körner 1989; Cordell *et al.* 1998). La réponse de la teneur en azote aux gradients a été largement étudiée dans le passé puisque l'azote est un élément clé dans de nombreuses fonctions métaboliques liées à la capacité photosynthétique des individus (Reich *et al.* 1997). A l'échelle interspécifique, LNC ($mg.g^{-1}$) était décroissant le long du gradient de température (R^2=10%), contrairement à ce qui a été trouvé à plus large échelle (Craine 2003).

Ainsi les patrons interspécifiques peuvent cacher des patrons intraspécifiques très différents (Fig. 2.15), comme cela a été conceptualisé par Shea & Chesson (2002) pour la relation entre nombre d'invasions réussies et richesse spécifique (voir aussi Austin 2007).

Redéfinir les relations trait-gradient par l'intermédiaire des exigences écologiques des espèces nous a permis d'expliquer le caractère idiosyncratique des courbes et surfaces de réponse trait-gradient et de concilier des relations trait-gradients différant aux échelles inter et intraspécifiques. Ce modèle demande à être testé avec une quantification des « exigences écologiques » des espèces, c'est ce que nous ferons au Chapitre 3.

Encadré 2.2. Méthode du moyennage de modèles librarie R: Biomod

GLM et GAM (McCullagh & Nelder 1989, Hastie & Tibshirani, 1990)
Les modèles linéaires généralisés (GLM) permettent de modéliser des relations paramétriques complexes entre une variable de réponse (ex. trait fonctionnel, présence d'espèce) et des variables explicatives (ex. variables de milieu) à l'aide de la distribution de la variable de réponse (ex. binomial, poisson) et d'une fonction de lien (ex. identité, logit). Les modèles additifs généralisés (GAM) permettent de modéliser des relations non paramétriques entre variables de réponse et explicatives et d'établir des courbes de réponse précises. Les fonctions des variables explicatives y sont des fonctions de lissage, par exemple de type «spline» cubique qui sont des enchaînements continus et dérivables de polynômes de dégré 3 au maximum définis sur de petits intervalles.

AIC et sélection de modèle
Le critère d'information d'Akaike (AIC, Akaike 1974) permet de sélectionner le modèle le plus efficace et de ne garder que les variables ayant un effet explicatif significatif.

La méthode du moyennage de modèles (Burnham & Anderson 2002; Link & Barker 2006)
Plus robsute que la sélection de modèles, cette méthode propose de moyenner les 2^n modèles possibles à partir d'un jeu de n variables explicatives, chaque modèle étant pondéré par son poids d'évidence:

$$\overline{P} = \sum_{i=1}^{R} w_i P_i \qquad w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_{r=1}^{R}\exp(-\frac{1}{2}\Delta_r)} \qquad \Delta_i = AIC_i - \min AIC$$

, où i correspond au iè modèle, Pi est la probabilité d'occurence en résultant, w_i son poids d'évidence d'Aikaike, AIC_i son AIC et minAIC la plus petite valeur d'AIC obtenue. Plus AIC_i est grand moins le modèle est plausible.

5. Conclusion & perspectives

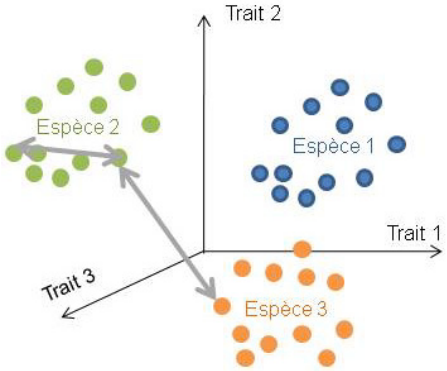
Le but de ce chapitre était de quantifier l'ampleur et la structure de la variabilité fonctionnelle intraspécifique et de tester l'hypothèse principale sur laquelle repose l'utilisation des traits fonctionnels, c.à.d. que ceux-ci varient plus entre espèces qu'au sein de chacune. Malgré le fait que nous n'avons trouvé aucun facteur expliquant les différences de variabilité et de structure de variabilité entre les différentes espèces, notre étude donne un cadre quantitatif nouveau pour pouvoir tester plus avant les effets de la variabilité fonctionnelle dans diverses applications des traits fonctionnels. Il est en effet essentiel pour faire ces tests d'avoir des idées claires des ordres de grandeur impliqués (Chapitre 4). Il serait par la suite utile d'avoir ce genre de quantifications pour d'autres systèmes que les systèmes alpins. Cette première quantification nous a déjà permis de mettre en évidence (1) que l'hypothèse de robustesse des traits –variabilité intraspécifique négligeable devant la variabilité interspécifique- ne devait pas être supposée *a priori* ; (2) que la détermination des stratégies fonctionnelles et l'axe majeur de dynamique des ressource demeuraient valides à l'échelle individuelle ; (3) mais que la classification des individus des différentes espèces au sein de ces stratégies ne l'était vraisemblablement pas.

Dans notre étude, les populations se dispersent dans l'espace des traits autour des moyennes d'espèce (Fig. 2.11-A) et les individus se dispersent autour des moyennes de populations, conduisant à un recouvrement des nuages de points définis par les individus de chaque espèce. Comme nous l'avons déjà évoqué, cette étude ne permet pas de répondre à la question de l'origine de la dispersion des individus autour des moyennes de populations, cela aurait nécessité de mettre en œuvre des analyses génétiques pour départager les effets des adaptations locales et de la plasticité. Nous avons par contre montré que la dispersion des populations autour des moyennes d'espèce n'était pas directement structurée par les gradients environnementaux (Fig. 2.11-A), mais certainement plus par les exigences environnementales des espèces (modèle conceptuel et Chapitre 3). Par conséquent une étude menée dans des conditions environnementales données (ex. basses températures) ne conduira probablement pas exactement à la même classification des espèces dans l'espace des traits qu'une étude menée dans des conditions environnementales différentes (ex. hautes températures, Fig. 2.11 & Fig 2.16). Ces modifications potentielles de classification des espèces dans l'espace des traits pourraient conduire à des modifications de classement des espèces le long d'un unique gradient de trait. On pourrait alors se retrouver dans les cas e ou f suggérés par Garnier *et al.* (2001) (Fig. 2.16). Ces observations vont à l'encontre d'études sur le classement des espèces selon un unique trait (Meziane & Shipley 1999; Garnier *et al.* 2001; Roche *et al.* 2004). Elles suggèrent que les classifications d'espèces devraient être établis pour chaque type de conditions environnementales et non pas en utilisant des moyennes de traits provenant de mesures sur des populations vivant dans des conditions complètement différentes. Par contre nous avons vu -lors du suivi temporel des traits- que dans des

Encadré 2.3. Espace des traits et distances fonctionnelles

L'étude des syndromes de traits peut se faire par l'intermédiaire d'un espace des traits, défini comme un espace multidimensionnel (Fig.) organisé le long de gradients de traits fonctionnels (ex. SLA, LDMC, LNC, Hmax…). Dans cet espace, on peut placer les moyennes d'espèce correspondant aux points définit par les moyennes de l'espèce pour chacun des traits, des moyennes de populations d'une espèce dans des conditions particulière, ou encore des individus.

Entre les individus ainsi positionnés (Fig.), on peut calculer une distance fonctionnelle donnant la distance entre les individus (ou espèces, ou populations) par rapport à leurs syndromes de traits. Afin de d'intégrer dans cette mesure de distance les corrélations entre traits, nous avons choisi de calculer des distances de Mahalanobis pour évaluer cette distance fonctionnelle (Mahalanobis, 1936). On peut calculer cette distance entre individus d'une même espèce ou entre des individus de différentes espèces.



conditions environnementales fixées, les espèces s'éloignaient ou restaient à même distance fonctionnelle au cours de la saison de végétation. Cela suggère que sous des conditions environnementales fixées, la classification des espèces (et donc le classement le long d'un unique gradient de trait) pourrait être conservée dans le temps (Fig. 2.16 cas *a* à *d*). Ceci est en accord avec la bibliographie (Garnier *et al.* 2001; Al Haj Khaled *et al.* 2005).
L'ensemble de nos résultats montre donc que trancher entre les études pour lesquelles il est satisfaisant d'utiliser des moyennes traits – ignorant alors la variabilité intraspécifique - et celles lesquelles il faudrait plutôt utiliser des distributions de traits – prenant en compte la variabilité intraspécifique - est une question d'échelle d'observation. En effet, pour décrire des contrastes forts entre stratégies fonctionnelles comme c'est le cas dans les grandes études interspécifiques à l'échelle du globe (Reich *et al.* 1999; Wright *et al.* 2004; Wright *et al.* 2005), il n'est pas nécessaire de prendre en compte la variabilité intraspécifique. Par contre l'étude de la position fine des espèces ou populations le long d'axes de variation fonctionnelle à l'échelle locale peut conduire à des recouvrements (les individus n'étant pas toujours fonctionnellement différents entre espèces proches) et nécessite de prendre en compte la variabilité intraspécifique, aussi bien celle induite par l'environnement (différences entre populations) que la variabilité individuelle.

Ces considérations théoriques ont des implications directes sur l'utilisation des traits fonctionnels. (1) La forte variabilité des traits que nous avons observée et ce qu'elle implique sur la classification fonctionnelle des espèces pourrait soulever des inquiétudes quant à l'utilisation de bases de données de traits ne donnant que des valeurs uniques (souvent une moyenne) par espèce. Cependant je ne pense pas que nos résultats aillent à l'encontre d'une utilisation massive de grandes bases de traits (Weiher *et al.* 1999; Garnier *et al.* 2001) mais plutôt qu'ils soulignent la nécessité de donner une grande importance à la qualité et à la nature des données présentes dans ces bases. Il est en effet indispensable que les mesures de traits en tant que telles soient accompagnées d'informations sur la localisation, l'environnement, le protocole de mesure et si possible la variabilité des mesures (Garnier *et al.* 2007; Madin *et al.* 2008; Lavorel *et al.* 2009; Kattge *et al.* in prep). Parallèlement il est aussi essentiel d'être vigilant lors de l'utilisation de telles bases et de ne comparer que des grandeurs comparables. (2) Comme nous l'avons évoqué au paragraphe 3.3 il semblerait plus adéquat d'utiliser dorénavant des mesures de traits et des syndromes de traits au niveau des individus et non plus des populations ou espèces (ex. Fonseca *et al.* 2000; Westoby *et al.* 2002; Vile *et al.* 2006), c'est de toute façon à ce niveau qu'elles sont mesurées sur le terrain, il faut ensuite conserver l'information par individu, en particulier dans les bases de données. (3) Enfin ces considérations peuvent avoir des implications importantes sur la manière d'inclure les espèces et mesures de traits dans les modèles de dynamique de la végétation, que ce soit à l'échelle des grands biomes ou à celle des individus. Je développerais ce point au Chapitre 4.

Encadré 2.4. Analyses en composantes principales (ACP) inter et intra-groupes

library: ade4

(Benzécri, 1983; Dodélec & Chessel, 1991)
Ce sont des types particuliers d'ACP incluant l'effet d'un vecteur catégoriel unique sur une matrice de variables continues.

1. L'ACP inter-groupes sépare au mieux deux ou plusieurs groupes par rapport aux niveaux du vecteur catégoriel; elle revient à faire une ACP sur les moyennes par groupe. Elle est basée sur la maximisation de la «variance entre groupes». Les éléments des groupes sont ensuite traités comme des éléments supplémentaires pour être projetés dans l'espace des variables.

2. L'ACP intra-groupes décrit la structure commune des corrélations au sein des groupes; elle revient à faire une ACP sur les données centrées sur les moyennes par groupe.

Ces deux analyses permettent de décomposer la variance totale dans un espace multidimensionnel. L'inertie calculée dans chacune de ces analyses représente la part de variance due respectivement aux niveaux inter et intra groupes.

Chapitre 3

Signification des modèles d'habitat et lien avec les traits fonctionnels et la performance

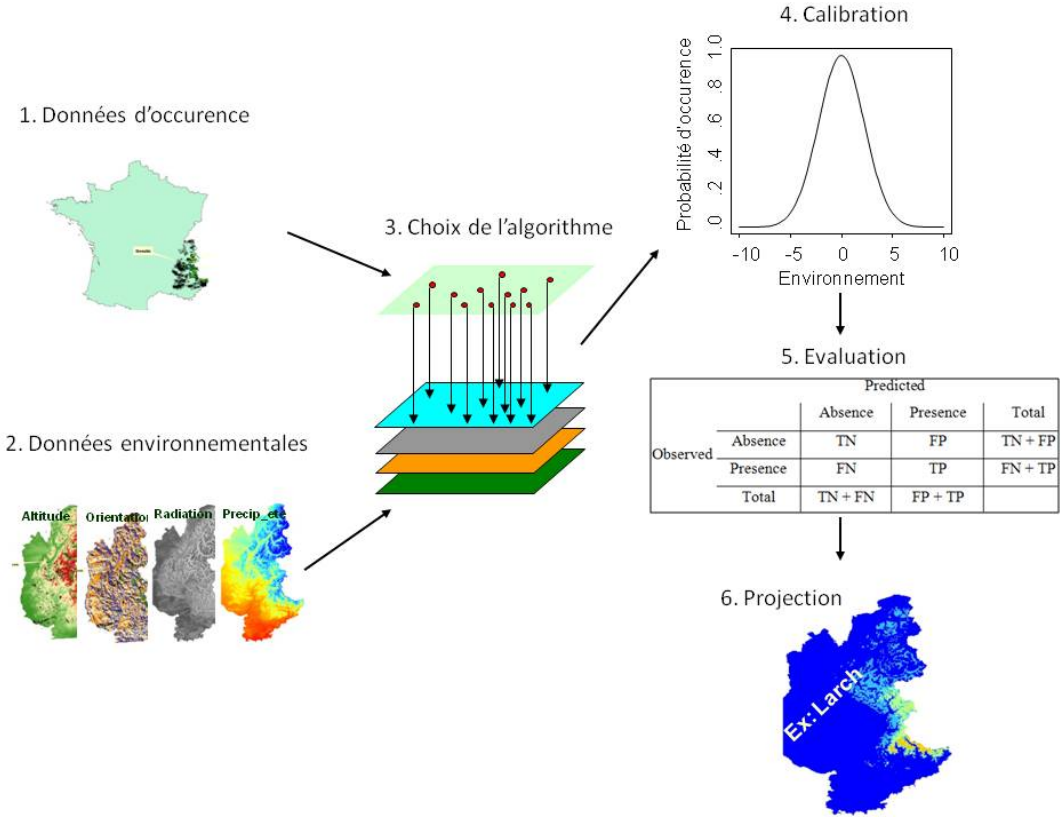


Fig. 3.1. Construire un modèle d'habitat : les grandes étapes

1. Les modèles statistiques d'habitat

1.1. Principe

Un des principaux objectifs de la biogéographie est de décrire et d'expliquer la distribution des êtres vivants à la surface du Globe. Aujourd'hui cet objectif est accompagné d'une volonté de prédire les modifications potentielles de ces distributions sous l'influence de variables de forçage (ex : changement climatique). Cette biogéographie prédictive, admet que la distribution des espèces est déterminée -au moins en partie- par les variables environnementales et que des approximations raisonnables de ces variables peuvent être estimées (Austin 2007). Elle propose par conséquent de caractériser les conditions environnementales adéquates à l'occurrence d'une espèce à l'aide de modèles de distribution statistiques et corrélatifs (Guisan & Thuiller 2005) ou mécanistes (Chuine *et al.* 1999; Kearney & Porter 2009). La disponibilité de données d'occurrence spatialement explicites pour un grand nombre d'espèces a conduit à une utilisation massive des modèles corrélatifs sous les noms d'enveloppe bioclimatique, modèle d'habitat, modèle de niche ou basé sur la niche, modèle de distribution des espèces, modèle d'adéquation de l'habitat (Austin *et al.* 1990; Westman 1991; Guisan & Thuiller 2005). Ces modèles visent à formuler le lien entre distribution actuelle d'un organisme (espèce, PFT, biome) et conditions environnementales, et permettent de prédire sa distribution géographique potentielle et également de comprendre et quantifier les facteurs impliqués (Rushton *et al.* 2004). Kearney (2006) a donné une définition précise de ces modèles que j'appellerai modèles d'habitat. Par la suite j'appellerai aussi les sorties brutes de ces modèles, qui sont statistiquement parlant des probabilités de présence, « qualité de l'habitat » comme une traduction de « habitat suitability ». Il s'agit de tout modèle qui rapproche de manière purement descriptive la nature physique d'une région à la distribution d'une espèce, sans lien mécaniste entre les variables environnementales et l'espèce. Au cours des dix dernières années, une large variété de techniques de modélisation ont émergé et de nombreuses études ont décrit ces modèles : leurs principes, leurs avantages, leurs inconvénients et les dangers dus à une mauvaise utilisation (Guisan & Zimmermann 2000; Guisan & Thuiller 2005; Austin 2007; Article 2, Article 3).

1.2. Construire un modèle d'habitat

Mettre en œuvre un modèle d'habitat passe par plusieurs étapes indispensables (Fig. 3.1). Il faut tout d'abord disposer d'informations représentatives (Chapitre 1) sur la distribution actuelle d'un organisme (1) sous la forme de relevés géo-référencés d'abondance, de présence et / ou d'absence. Il faut également disposer de mesures ou d'estimations des variables environnementales appropriées en ces points géographiques (2). Les variables

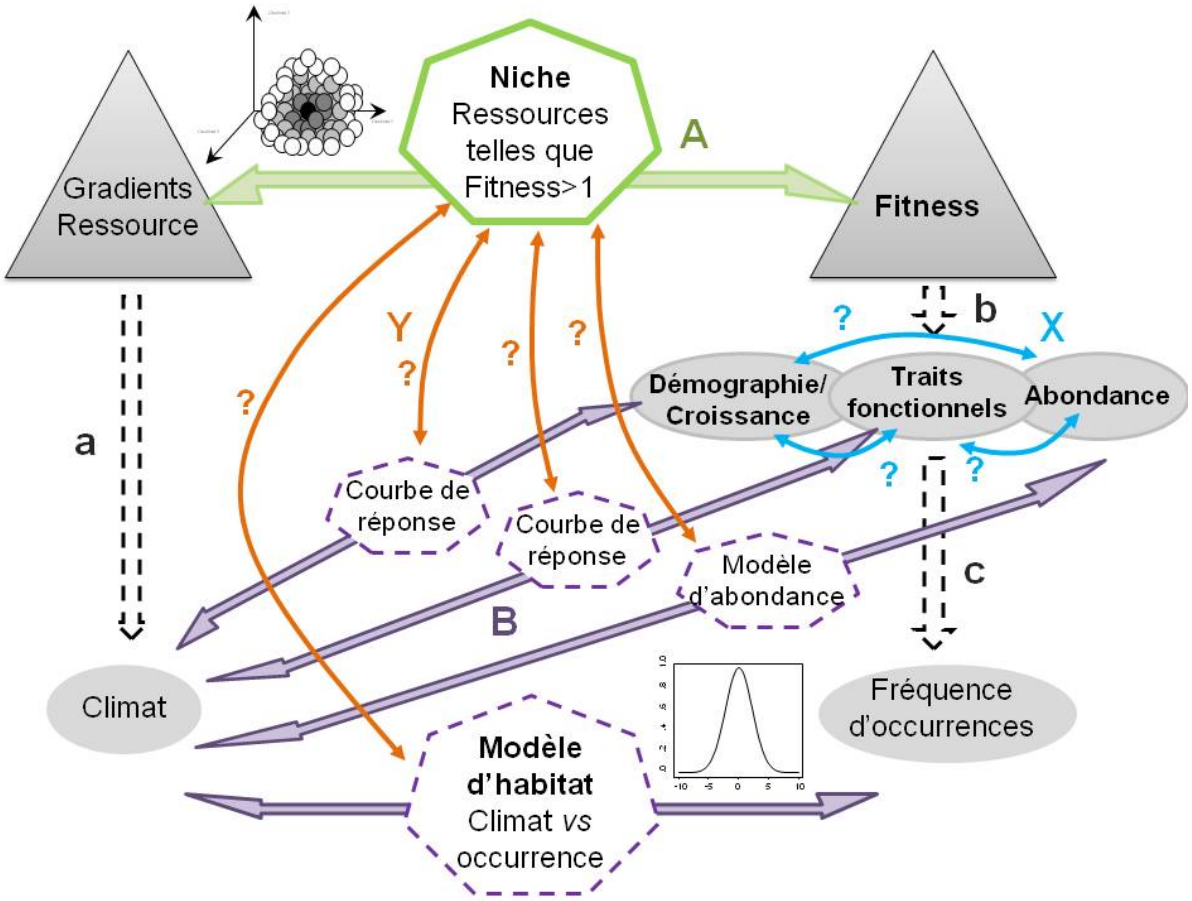


Fig. 3.2. Niche et modèles d'habitat: que modélisons nous?

La Niche écologique (vert) est le concept théorique de base, elle établi un lien entre fitness des individus/populations et gradients de ressource (A). Souvent on considère que les variables climatiques sont un estimateur des gradients de ressource (a) et que les paramètres démographiques, les traits fonctionnels et l'abondance sont des estimateurs de la fitness (b). On considère de plus par extension que la fréquence d'occurrence est une estimation elle aussi de cette fitness (c).

Les modèles d'habitat (violet) établissent un lien entre gradients climatique et fréquence d'occurrence (B). Mettant en lien un estimateur des gradients de ressources et un estimateur de la fitness, ils sont censés représenter la niche écologique (Y, orange). De la même manière (B) on peut construire des modèles d'abondance liant climat et abondance ou des courbes de réponse liant climat et traits fonctionnels ou paramètres démographiques. Tous ces modèles (B) peuvent donc être vus comme des estimateurs de la niche (Y, orange). Les liens existants entre paramètres démographiques, traits fonctionnels et abondance sont complexes et encore mal connus (X, bleu).

Sous le terme « climat » peuvent parfois être utilisés d'autres types de variables tels que des variables de couverture végétale ou d'utilisation des terres.

choisies doivent avoir une influence sur la physiologie de l'organisme étudié. Elles sont souvent des descripteurs de l'environnement climatique (température, précipitations) et physique (type de sol) et parfois biologique (ex. interactions biotiques : Araujo & Luoto 2007). Puis il faut choisir un algorithme (3) ou formuler le modèle permettant de lier ces deux types d'information en fonction de leur nature (présence/absence ou présences seules, variables environnementales continues ou catégorielles). Cela passe par exemple par le choix des formes de relations autorisées (ex. linéaire, quadratique, cubique). Austin (2007) prône l'utilisation de courbes unimodales asymétriques. L'algorithme est ensuite calibré dans l'espace écologique (4). Ceci correspond écologiquement à identifier les conditions environnementales associées à la présence de l'espèce et statistiquement à estimer les paramètres – du moins dans le cas d'algorithmes paramétriques- décrivant au mieux le lien entre distribution et environnement. La calibration d'un modèle passe par exemple par une mesure de réduction de la variance (ou de la déviance) ou de maximisation de la vraisemblance. Le modèle doit ensuite être évalué (5) pour déterminer ses pouvoirs explicatif (ex. R^2) et prédictif (ex. Kappa, AUC, Encadré 3.2, Fielding & Bell 1997; Pearce & Ferrier 2000; Allouche *et al.* 2006). L'évaluation du pouvoir prédictif passe par l'utilisation d'un jeu de données d'occurrences indépendant de celui utilisé pour la calibration. Comme il est rare de disposer de jeux de données indépendants on utilise souvent une partie aléatoire du jeu de données d'origine (ex. 70%) pour calibrer le modèle et la partie restante (ex. 30%) pour l'évaluer. Ceci pouvant être répété plusieurs fois pour mesurer la sensibilité du modèle aux conditions initiales (validation croisée, Thuiller *et al.* 2009). D'autre part, comme la plupart des algorithmes utilisés fournissent des résultats sous forme d'une distribution continue de valeurs de qualité de l'habitat (entre 0 et 1), il faut convertir ces sorties de modèle sous forme binaire pour pouvoir les comparer aux données observées. Cette conversion se fait à l'aide d'un seuil –on considère l'organisme comme présent (resp. absent) au dessus (resp. en dessous) de ce seuil- dont le choix est délicat (Liu *et al.* 2005). Si son pouvoir prédictif est satisfaisant (Encadré 3.2), le modèle peut finalement être projeté dans l'espace géographique sur l'ensemble de la zone d'étude (distribution potentielle) ou par extrapolation sur un nouveau jeu de données correspondant à des conditions environnementales futures ou d'aires inconnues.

1.3. Hypothèses de construction

La construction d'un modèle d'habitat repose sur deux hypothèses fondamentales. D'une part, l'utilisation de la distribution actuelle d'une espèce (ou organisme) pour décrire sa niche écologique nécessite de supposer que celle-ci est en quasi-équilibre avec son environnement. C'est-à-dire que l'espèce occupe l'ensemble des conditions écologiques adéquates pour elle et qu'elle n'occupe pas des conditions non adéquates. En réalité, les espèces sont peu probablement à l'équilibre et on considère donc qu'il suffit qu'elles soient

en pseudo-équilibre (Guisan & Thuiller 2005) avec l'environnement, c.à.d. que l'ensemble du jeu de données d'occurrences reflète la distribution comme si elle était à l'équilibre. Le degré d'équilibre d'une espèce avec l'environnement dépend des interactions biotiques (exclusion compétitive, facilitation, Araujo & Luoto 2007) et de sa capacité de dispersion (Pulliam 2000; Soberon 2007). Les espèces ayant une plus grande facilité à disperser sont donc supposées être d'avantage à l'équilibre (Araújo & Pearson 2005). D'autre part les données d'occurrence utilisées pour l'espèce étudiée doivent représenter l'ensemble de sa niche (cœur, périphérie et extérieur), sans omettre des parties de gradients (Chapitre 1, Thuiller *et al.* 2004a; Randin *et al.* 2006; Strauss & Biedermann 2007). Cette bonne représentation du volume occupé par l'espèce dans l'espace environnemental (ex. climatique) est essentielle pour pouvoir prédire correctement la distribution potentielle de l'espèce dans les zones de l'espace géographique où elle n'a pas été étudiée. Extrapoler, c.à.d. utiliser des modèles pour faire des prédictions hors de l'étendue des données utilisées pour la calibration peut conduire à des prédictions très incertaines, voire fausses (Thuiller *et al.* 2004a ; Elith & Graham 2009). Il faut en particulier être vigilant à l'effort et à la stratégie d'échantillonnage mis en œuvre pour récolter ou rassembler les données d'occurrence (Chapitre 1, Graham *et al.* 2004) et à la qualité des données, surtout pour des espèces à faible probabilité de détection (Mackenzie *et al.* 2004). Enfin l'utilisation des modèles d'habitat repose sur un certain nombre de simplifications implicites et de limites (Hampe 2004). Premièrement, les interactions biotiques (compétition et facilitation) n'y sont pas prises en comptes explicitement, mais elles sont incluses de manière implicite par l'intermédiaire des données de distribution actuelle résultant entre autre de leur impact. Deuxièmement, la projection des modèles d'habitat dans le temps (passé et futur) nécessite d'admettre que la relation entre distribution et environnement est figée, c'est-à-dire que ces relations sont les bonnes et ne seront pas modifiées, et qu'il n'y a pas d'adaptation génétique rapide des espèces.

1.4. Algorithmes utilisés et sélection de modèle

Il existe aujourd'hui une grande diversité d'algorithmes plus ou moins complexes qui peuvent être utilisés pour construire des modèles d'habitat et donner les probabilités de présence de l'espèce comme une fonction d'un jeu de variables environnementales. Ces algorithmes permettent d'identifier des relations potentiellement complexes et non-linéaires dans un espace multidimensionnel. Il existe entre autres des méthodes statistiques de type régression (modèle linéaire/additif généralisé: GLM ; GAM ; lissage de régression multivariée : MARS, Friedman 1991), classification (arbre de classification : Breiman *et al.* 1984 ; arbre de classification renforcés par les méthodes de « boosting » : Ridgeway 1999 ; ou de « bagging », random forest : Breiman 2001), analyses multivariées (Analyses des facteurs de niche écologique : ENFA, Hirzel *et al.* 2002), techniques d'apprentissage (entropie maximale : Maxent, Phillips *et al.* 2006 ; réseaux de neurones : Pearson *et al.* 2002 ; algorithme

génétique : Garp, Stockwell & Peters 1999). De nombreuses différences existent entre ces divers algorithmes. L'une des différences clé est probablement le type de données qu'ils requièrent. La plupart nécessitent en effet des données d'absence pour établir des contrastes entre les conditions dans lesquelles l'espèce est présente et celle dans lesquelles elle est absente. D'autres comme ENFA n'ont besoin que de données de présence pour modéliser les conditions où est l'espèce est présente. Afin de pallier au manque de données d'absence fiables, certaines méthodes proposent aussi de sélectionner des pseudo-absences permettant d'établir des contrastes entre les conditions dans lesquelles l'espèce est présente et celles dans lesquelles on suppose qu'elle ne l'est pas. La sélection des pseudo-absences : quantité, localisation, distance aux présences est l'objet de nombreux travaux (Engler *et al.* 2004; VanDerWal *et al.* 2009).

De nombreuses études détaillent les avantages et inconvénients de ces différents algorithmes et comparent leur capacité afin de déterminer le meilleur modèle (Segurado & Araújo 2004; Elith *et al.* 2006; Lawler *et al.* 2006). Il en ressort des différences substantielles mais aucune réponse-type, car le choix du modèle dépend entre autre des données disponibles et de la question posée. De plus une fois l'algorithme choisi, il faut encore choisir la forme du modèle, comme par exemple les relations linéaires ou quadratiques dans un GLM, le degré de lissage dans un GAM ou encore le nombre d'arbre dans un *random forest*. La sélection du meilleur algorithme et du meilleur modèle se fait souvent par des procédures automatisées comme par exemple la régression pas à pas basée sur des critères d'information comme le critère d'information d'Akaike (AIC, Akaike 1974). Une autre approche, plus récente, consiste à moyenner les modèles en les pondérant par leur pertinence (Encadré 2.2, Burnham & Anderson 2004; Rushton *et al.* 2004). Cette méthode permet également de connaitre l'importance d'une variable environnementale dans la distribution d'une espèce en considérant la pertinence (poids d'évidence) des modèles dans laquelle elle apparait. Quoi qu'il arrive le choix d'un algorithme donné ne fait pas tout, aussi complexe soit il, car il ne faut pas oublier que la démarche de modélisation est une recherche du modèle le plus parcimonieux possible. De plus il n'est pas judicieux de choisir un algorithme/modèle sur des critères purement statistiques –tels que le pouvoir prédictif- qui ne suffisent pas à estimer sa qualité, il faudrait pouvoir également tester son réalisme écologique (Austin 2007). En particulier certains algorithme complexes (Maxent, random forest) sont des méthodes purement prédictives et non descriptives et ne permettent pas de comprendre les pourquoi et comment des distributions observées. La plateforme de modélisation multi-méthodes Biomod (https://r-forge.r-project.org/projects/biomod/) permet à un utilisateur avisé de choisir ou de moyenner différents algorithmes/modèles selon différentes méthodes et critères (Thuiller *et al.* 2009), son utilisation permet de pallier aux inconvénients de n'utiliser qu'une unique méthode.

1.5. Utilisations des modèles d'habitat

Les modèles d'habitat se sont peu à peu révélés performants pour générer de l'information biogéographique pouvant être utilisée dans de nombreux domaines théoriques et appliqués de l'écologie. On les retrouve en particulier utilisés comme aide à la décision en biologie de la conservation: pour identifier et sélectionner de nouvelles aires de protection (Araújo & Williams 2000), guider la gestion des espèces sauvages (Milsom *et al.* 2000), trouver de nouvelles stations d'espèces rares ou menacées (Guisan *et al.* 2006) ou encore identifier des sites vulnérables à l'extinction locale (Gates & Donald 2000). Ils sont aussi utilisés en biologie de la restauration pour identifier des aires de réintroduction d'espèces en danger (Debeljak *et al.* 2001) ou encore en biologie des invasions pour identifier des sites vulnérables aux invasions biologiques (Higgins *et al.* 1999; Lopez-Darias *et al.* 2008). Ils sont largement utilisés pour explorer les conséquences potentielles des changements climatiques sur la distribution des espèces (Heikkinen *et al.* 2006; Thuiller *et al.* 2006), le risque d'émergence des maladies (Peterson *et al.* 2006) ou encore les impacts économiques (Perrings 2003). Les modèles d'habitat sont également employés en paléo-écologie pour étudier les extinctions passées et faire le lien entre distributions actuelles et passées des espèces (Martinez-Meyer *et al.* 2004) ; ou encore pour explorer le conservatisme de la niche écologique (Wiens & Graham 2005; Pearman *et al.* 2008) et les mécanismes de la spéciation (Kozak & Wiens 2006). Enfin ils peuvent être une aide à l'échantillonnage (Article 9; Chapitre 1) pour trouver des nouvelles populations d'une espèce connue (Engler *et al.* 2004) ou aider à la découverte de nouvelles espèces (Raxworthy *et al.* 2004).

1.6. Limites techniques et conceptuelles

Malgré un engouement croissant pour ces modèles d'habitat, il existe encore un certain nombre de problèmes techniques et conceptuels liés à leur utilisation et pour lesquels aucun consensus n'a jusqu'à présent été trouvé (Araujo & Guisan 2006; Austin 2007 ; **Article 2**). Les problèmes techniques qui mettent en cause la fiabilité des modèles d'habitat sont en grande partie liés à l'échantillonnage (Chapitre 1). Se posent entre autres des questions telles que : faut-il prendre en compte l'auto-corrélation spatiale dans les modèles d'habitat, sachant qu'elle peut modifier voire inverser les patrons trouvés (Dormann *et al.* 2007) ? Comment prendre en compte une mauvaise détectabilité des présences (Mackenzie *et al.* 2004)? Quelle étendue des données prendre et à quelle échelle (Stockwell & Peterson 2002)? Et en particulier dans le cas de la sélection des pseudo-absences (VanDerWal *et al.* 2009)? La plupart de ces questions tournent autour d'une question centrale : qu'est-ce qu'un bon jeu de données pour modéliser la qualité de l'habitat et quels sont les biais qui peuvent exister (Araujo & Guisan 2006)? Cette question est cruciale étant donné que ce qui sort des modèles reflète ce qui y rentre, c.à.d. si les données initiales comportent des biais et ne

représentent pas correctement la distribution, les sorties du modèle seront erronées, voire fausses (Phillips *et al.* 2009). Le problème est que l'on ne connait ni la distribution vraie de l'espèce, ni sa niche fondamentale, on ne connait que ce que l'on en a échantillonné. Dans l'Article 3 nous nous sommes intéressés à la question de la prévalence du jeu de données, c.à.d. le nombre de présences sur le nombre de données (Encadré 3.1). Dans le cas d'une modélisation par modèle linéaire généralisé binomial, la prévalence (nombre total de présence sur nombre total de site) du jeu de données influe directement sur la valeur moyenne des sorties du modèle, rendant les comparaisons entre modèles compliquées et peu pertinentes. Ainsi, les espèces modélisées avec peu de données de présence tendent à avoir des valeurs de probabilités de présence plus faibles que les autres, ce qui peut être problématique pour comparer les sorties des modèles entre espèces (qualité de l'habitat). Cela est d'autant plus problématique que les probabilités de présence obtenues pour une même espèce après calibration d'un GLM sur deux jeux de données différant par leur prévalence peuvent être très différentes (Article 3). Des palliatifs mathématiques ont émergé pour faire face à ce problème, comme le rééchantillonnage arbitraire ou la pondération des jeux de données visant à obtenir une prévalence de 0,5 (Teixeira *et al.* 2001) ou encore la formule proposée par Real *et al.* (2006) et maintenant (Real *et al.* 2009), mais il est difficile de trouver à ces palliatifs un sens biologique (Austin 2007).

En termes conceptuels, les modèles d'habitat soulèvent aussi des débats puisque personne n'est véritablement d'accord sur ce qu'ils représentent (Fig. 3.2) : habitat *vs* niche (Kearney 2006), niche fondamentale *vs* réalisée (Soberon 2007), distributions réalisées *vs* distributions potentielles (Soberon 2007). Les modèles d'habitat sont des modèles simples, corrélatifs, statiques, d'un côté ils ne prennent pas en compte les processus écologiques directement (dispersion, interactions biotiques, perturbation) et d'un autre côté ils les prennent en compte indirectement en se basant sur la distribution actuelle des espèces qui résulte de tous ces processus (Article 2). Ces modèles tentent en quelque sorte d'expliquer la niche réalisée par des variables caractérisant la niche fondamentale, ils ne représentent ni vraiment la niche fondamentale ni la niche réalisée, mais plutôt quelque chose d'intermédiaire qui dépend de la portion de la niche qui est représentée par les données. De plus les modèles d'habitat reposent sur la distribution en présence d'une espèce et sur la quantité de présences observées dans des conditions environnementales données, alors que le concept de niche comme nous l'avons vu est défini à partir du taux de croissance des populations. Il n'est donc pas simple de rattacher ces modèles d'habitat aux concepts écologiques et le débat sur leur signification est loin d'être clos.

1.7. Trouver un sens

En pratique, les résultats des modèles d'habitat (donnés entre 0 et 1) peuvent être interprétés de trois manières différentes: (1) soit les résultats sont transformés en sorties binaires à l'aide

d'un seuil (Liu et al 2005), un bon modèle étant alors un modèle permettant de bien discriminer entre présence et absence et donnant par exemple une bonne mesure d'AUC (Encadré 3.2). Dans ce cas une qualité de l'habitat de 0,4 est adéquate si 0,4>=seuil et non adéquate dans le cas contraire (2) soit les résultats sont utilisés de manière ordinale, comme une mesure de la qualité relative de l'habitat. Dans ce cas une qualité de l'habitat de 0,4 est meilleure qu'une qualité de 0,3 et moins bonne qu'une de 0,5. On peut dans ce cas rééchelonner les sorties entre 0 et 1, en fonction de la qualité la plus haute obtenue ; (3) soit les résultats sont interprétés comme une mesure absolue de la qualité de l'habitat et peuvent être comparées entre espèces. Dans ce cas une qualité de 0,4 est une qualité légèrement inférieure à la qualité moyenne. Dans ce dernier cas les modèles doivent être validés à l'aide d'un graphique de calibration (Pearce & Ferrier 2000, Encadré 3.2), ce sont les valeurs de sortie en elles mêmes qui comptent et non plus le pouvoir discriminant des modèles. Le premier cas (1) est le plus logique, il vise à utiliser les modèles en rapport avec la manière dont ils ont été construits, c'est-à-dire à partir de présence et de fréquence d'occurrence. Les deux cas suivants (2) et surtout (3) sont plus problématiques par rapport à ce que nous avons vu précédemment, ils supposent que les valeurs de sorties sont fiables (en particulier pour les valeurs intermédiaires voir l'Article 3 Fig. 3) et reposent sur une hypothèse forte qui est : les fortes fréquences d'occurrence de l'espèce –donnant de fortes valeurs de qualité de l'habitat- correspondent au cœur de la niche, c.à.d. aux conditions les plus adéquates pour cette espèce donnant lieu localement à de bonnes performances (fort taux de croissance des populations, forte valeur sélective individuelle). Cette simplification est demeurée jusqu'à présent une hypothèse de travail peu testée. Il faudrait pouvoir confronter les sorties des modèles d'habitat avec des mesures de performance des individus ou populations. Cependant il n'est pas facile de mesurer in situ la performance des espèces sur de larges gradients (Pulliam 2000, Violle & Jiang 2009) et il est donc encore moins facile de quantifier la niche d'une espèce (Fig. 3.2). La quantifier par l'intermédiaire des fréquences d'occurrence (ou probabilité de présence) est une solution de substitution qui s'est imposée du fait de sa facilité d'exécution (Article 2). Pulliam (2000) avait déjà suggéré de mesurer les paramètres démographiques de populations vivant en conditions contrastées pour quantifier la niche, Violle *et al.* (2009) ont suggéré récemment que cela pouvait être fait directement à partir des traits fonctionnels. En effet, chez les plantes, la performance est souvent évaluée par l'intermédiaire de mesures individuelle des composantes de la valeur sélective: survie, biomasse ou croissance et capacité reproductive (Violle *et al.* 2007 ; Violle & Jiang 2009). Or les traits fonctionnels sont des substituts à ces mesures des composantes de la valeur sélective. Des liens ont par exemple été établis entre SLA et croissance (Weiher *et al.* 1999 ; Wright & Westoby 2001; Poorter & Bongers 2006) ou entre Hmax et biomasse (dans nos données R^2 autour de 20-30% pour 8 espèces pour lesquelles on a la mesure de biomasse fraiche individuelle) ou encore entre le syndrome de trait dans son ensemble et la valeur sélective (Milla *et al.* 2009). Des études sont également menées sur les liens entre traits et

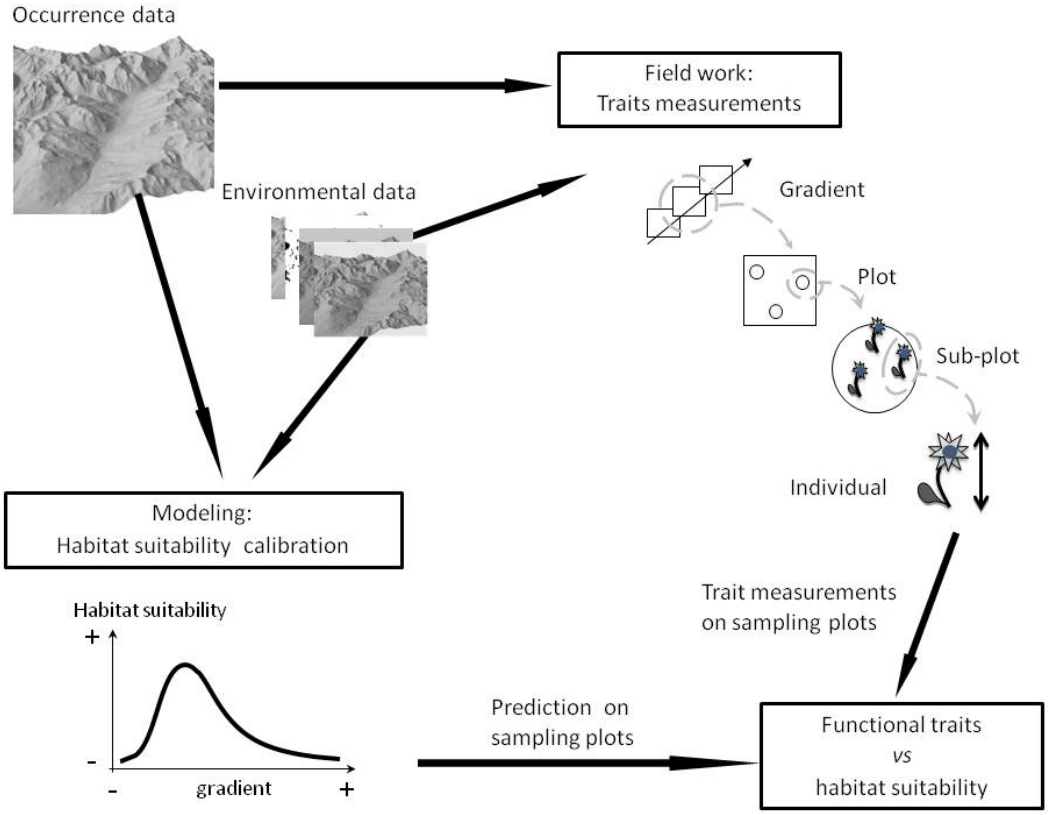


Fig.3.3. Principe de l'analyse : *construction d'un modèle d'habitat et échantillonnage hiérarchique conduisant à la confrontation de deux mesures de la niche. Issue de l'Article 6.*

paramètres démographiques (Poorter *et al.* 2008) ou sur les relations avec l'abondance (McGill *et al.* 2006). Dans ce chapitre, l'objectif est donc de confronter deux mesures différentes de la niche, l'une donnée par les modèles d'habitat dérivés des fréquences d'occurrence et l'autre donnée par des traits fonctionnels ou des paramètres démographiques ou de croissance sensés caractériser la valeur sélective des individus. Cette confrontation doit mener à une éventuelle validation écologique de l'hypothèse selon laquelle les plus fortes fréquences d'occurrence sont obtenues au cœur de la niche. Nous allons mener trois confrontations différentes et complémentaires entre sorties des modèles d'habitat et : (1) trois traits fonctionnels (Hmax, LDMC et LNC) mesurés in situ lors de la campagne Guisane 2080; (2) deux paramètres démographiques (taux de germination et taux de survie des plantules) supposés limiter la régénération et mesurés sur lors d'une expérimentation en vallée de la Guisane ; (3) l'accroissement radial d'arbres adultes dans l'ensemble des Alpes françaises.

2. Tests de la signification des modèles d'habitat

2.1. Les modèles d'habitat comme une description fonctionnelle de la niche

// Cette section correspond à l'Article 6

2.1.1. Objectifs

Comme décrit ci-dessus, les traits fonctionnels des plantes étant supposés être une première approximation de la performance des individus (Violle *et al.* 2007, Poorter *et al.* 2008) et devraient donc permettre de quantifier la niche de l'espèce (Violle & Jiang 2009). Nous avons donc confronté deux mesures différentes de la niche afin de voir s'il existait une signification fonctionnelle à la mesure de la niche (ou qualité de l'habitat) donnée par les modèles d'habitat. Dans ce cas la variabilité intraspécifique des traits fonctionnels devrait être expliquée par des différences de qualité de l'habitat.

2.1..2. Principe

L'étude a été menée sur les 16 espèces pérennes de la base Guisane 2080 et sur les deux sites d'étude (Chapitre 1). Nous nous somme intéressés à trois traits fonctionnels connus pour être en lien avec la stratégie fonctionnelle des espèces (Westoby *et al.* 2002) : la hauteur végétative (Hmax) supposée être liée à la biomasse et à la croissance, la teneur des feuilles en matière sèche (LDMC) supposée être liée à l'exploitation des ressources et indirectement à la croissance et la teneur des feuilles en azote (LNC) supposée être liée à la croissance et au taux photosynthétique. Ces traits ont été collectés selon la stratégie hiérarchique décrite au Chapitre 1 (Fig. 3.3).

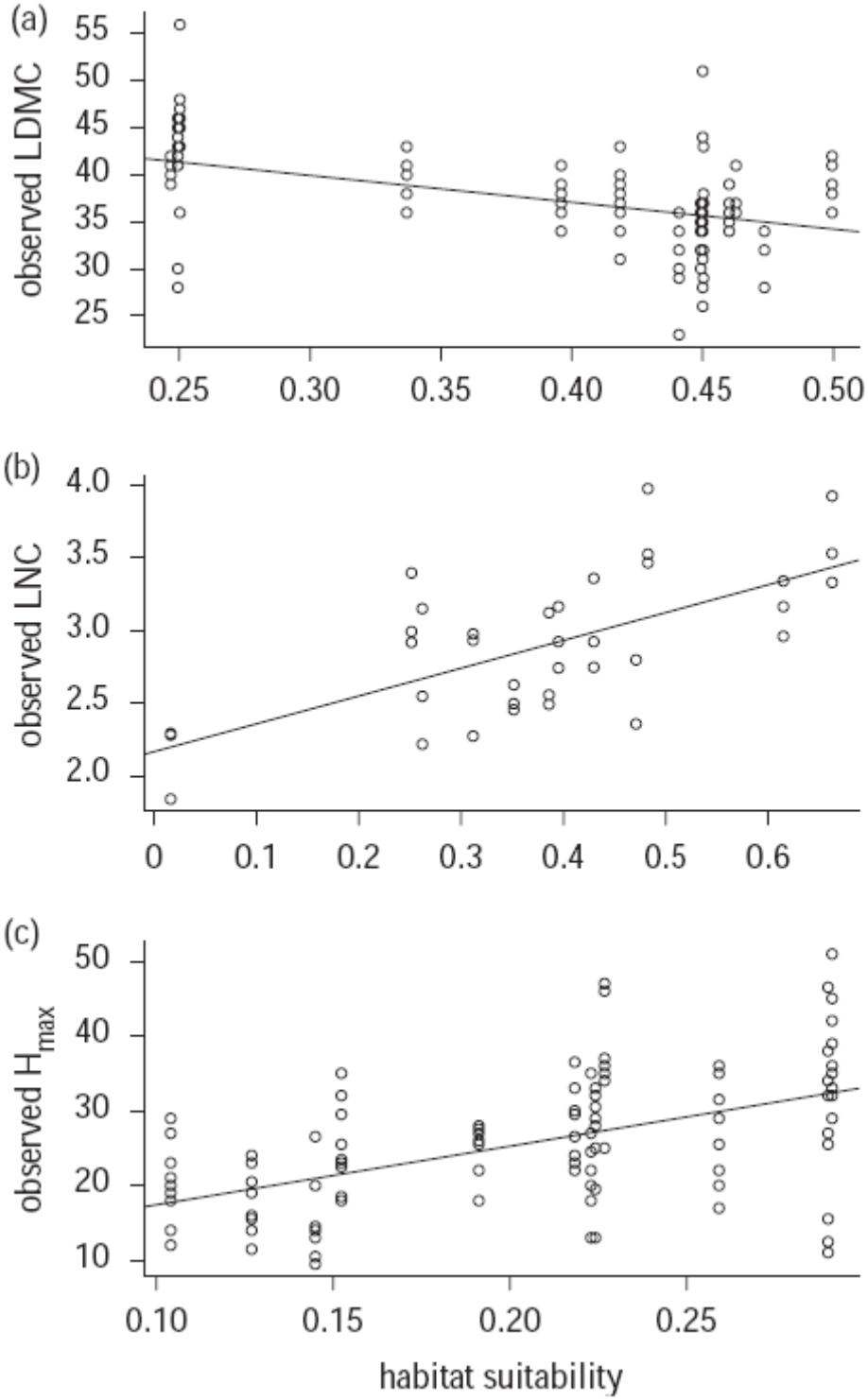


Fig. 3.4. Lien entre traits fonctionnels et qualité de l'habitat. *(a) LDMC (en %) de C. sempervirens; (b) LNC de P. viviparum ; (c) Hmax de L. vulgare. Issue de l'Article 6*

Pour chaque espèce et dans chaque site, nous avons également calibré des modèles d'habitat de type modèles additifs généralisés avec la méthode du moyennage de modèles (Encadré 2.2). En Guisane les modèles ont été paramétrés avec les variables à 50 x 50m de : pente, position topographique, degrés-jours, indice d'humidité, température minimale en hiver et radiations solaires provenant du modèle météorologique Aurelhy (Benichou & Le Breton 1987). A Anzeindaz ils ont été paramétrés avec les variables à 25 x 25 m de : pente, position topographique, température annuelle moyenne, radiations solaires annuelles et indice d'humidité (Zimmermann & Kienast 1999). Ces modèles ont ensuite été projetés sur les points de mesure des traits fonctionnels (Fig. 3.3). Pour chacun des modèles nous avons calculé son pouvoir prédictif à l'aide de l'AUC (Encadré 3.2).

La relation entre traits et qualité de l'habitat a été établie à l'aide de modèles mixtes qui permettent de prendre en compte la structure hiérarchique de l'échantillonnage (Encadré 2.1, Fig. 3.1). Ces modèles suivaient la formule :

$$Trait = fixe\ (Qualité.de.l.habitat),\ aléatoire\ (Population / Sous.population).$$

Nous avons mesuré la part de variabilité des traits expliquée (R^2) par la qualité de l'habitat (Encadré 2.1).

2.1.3. Résultats

Les modèles calibrés pour chaque espèce dans chaque site avaient un bon pouvoir prédictif (AUC entre 0,79 et 0,96) et avaient de bons graphiques de calibration (résultats non présentés). D'autre part comme nous l'avons vu au Chapitre 2, les traits fonctionnels mesurés *in situ* étaient très variables du fait de différences de moyennes de traits entre stations et d'une forte variabilité individuelle. Hmax était particulièrement variable avec des coefficients de variation (CV) entre 0,19 et 0,49 et LDMC et LNC l'étaient moins avec respectivement des CV entre 0,08-0,25 et 0,09-0,29.

Dans l'ensemble, le lien entre qualité de l'habitat et traits fonctionnels dépendait de l'espèce et du trait considéré en termes de force (R^2 dans 0-100%) et de sens (relation positive ou négative, Tableau 3.1). Le lien entre Hmax et la qualité de l'habitat était le plus souvent positif (Fig. 3.4 et Tableau 3.1) et assez forte (9 R^2 >20% parmi les 21 relations testées), le lien entre LDMC (4 R^2 >20%) ou LNC (11 R^2>20%) et qualité de l'habitat étaient très variables (Tableau 3.1). Les liens observés n'étaient pas similaires entre les deux sites d'étude, avec par exemple une corrélation positive entre Hmax et la qualité de l'habitat pour *V. myrtillus* en Guisane et négative à Anzeindaz. Pour certaines espèces aucun trait ne présentait de lien avec la qualité de l'habitat.

La force des liens obtenus ne dépend pas de la qualité prédictive des modèles ($p>0,05$).

	Guisane/Anzeindaz		
species	LDMC	LNC	H_{max}
C. sempervirens	69.6 [−]/—	34.1 [+]/—	—/18 [+]
D. glomerata	—/—	33.7 [−]/ 48.8 [−]	—/—
Dryas octopetala	—/—	42.8 [+]/—	—/45.3 [+]
F. paniculata	—	—	—
Geum montanum	21.5 [+]	—	23.3 [+]
Juniperus sp.	13 [+]	—	20.9 [+]
Larix deciduas	—	45 [−]	33.8 [+]
L. vulgare	—	64 [+]	58.5 [+]
Pinus sp.	37.4 [−]	—	31.0 [+]
P. viviparum	—	62.1 [+]	—
R. ferrugineum	—	—	—
Sesleria caerulea	—/—	41.7 [+]/—	—/33.7 [−]
Salix herbacea	5.1 [+]	100 [+]	18.8 [−]
Silene nutans	—	8.5 [−]	9.7 [+]
Trifolium alpinum	—	49.5 [−]	—
V. myrtillus	—/60.2 [+]	—/25.4 [−]	63.1 [+] /51.3 [−]

Tableau 3.1. *R^2 et signes des relations (modèles linéaires mixtes) entre qualité de l'habitat et traits fonctionnels (Seulement les R^2> 5% sont données). Issue de l'Article 6.*

2.1.4. Discussion

Malgré de fortes inégalités entre espèces et traits, dans certains cas les modèles d'habitat décrivent relativement bien la variabilité des traits (24 R^2>20% sur les 63 testés). Cela laisse penser que, modèle d'habitat et traits fonctionnelles peuvent donner une même quantification de la niche pour l'espèce considérée. Dans la plupart des cas, nous avons trouvé des relations positives entre Hmax et la qualité de l'habitat, indiquant qu'une meilleure qualité de l'habitat correspond à une plus forte croissance et une plus grande capacité compétitrice des individus. Cela est en accord avec le modèle conceptuel présenté au Chapitre 2. Les relations obtenues étaient plus contrastées pour les deux autres traits. En particulier, la variabilité de LDMC n'était pas souvent bien prédite par la qualité de l'habitat. La seule forte relation obtenue entre LDMC et qualité de l'habitat pour les espèces herbacées (*C. sempervirens*) était négative conformément au compromis de traits sur l'axe d'acquisition *vs* conservation des ressources et au modèle conceptuel du Chapitre 2. Dans un site caractérisé par une forte qualité d'habitat, les individus tendraient à exploiter les ressources (faible LDMC) au lieu de les conserver. Mais cela n'est pas confirmé par les résultats obtenus pour les autres espèces. Pour LNC on se serait attendu à obtenir comme pour Hmax une relation positive avec la qualité de l'habitat, puisque LNC est lié à la capacité photosynthétique et à la croissance des individus. Cependant cette relation positive n'a été obtenue que dans la moitié des cas (ex. *P. viviparum*).

Bien que certaines fortes relations aient été trouvées et soient en accord avec ce que l'on attendait, les résultats restent mitigés et ne permettent pas de valider complètement la confrontation entre modèles d'habitat et traits fonctionnels, ni le modèle conceptuel décrit au Chapitre 2.

Il pourrait y avoir plusieurs explications possibles à ces résultats :

Du point de vue des modèles d'habitat, (i) il se pourrait que les modèles ne soient pas assez représentatifs de la niche. En effet ils sont calibrés sur une vallée dans laquelle les espèces ne sont peut-être (surement ?) pas à l'équilibre, en particulier les espèces ligneuses dont la distribution actuelle est en deçà de la limite altitudinale supposée des arbres (vers 2600 m dans les Alpes). (ii) Les modèles d'habitat que nous avons construits sont établis à partir de données régionales et non locales. Or la performance des individus pourrait être d'avantage influencée par des facteurs environnementaux locaux. L'ajout de variables de sol (pH, teneur en matière organique) ne nous a pas permis d'améliorer la part de variance expliquée, mais nous n'avons pas testé d'autres variables telles que la disponibilité locale réelle (et non estimée) en eau et en lumière. (iii) Les modèles d'habitat ne prennent pas en compte les interactions biotiques qui jouent un rôle clé dans la performance locale des individus (Harper 1964).

D'un point de vue écologique : (i) LNC et Hmax qui sont le mieux décrits par la qualité de l'habitat en termes de force du lien sont peut-être plus directement liés à la performance que LDMC qui y est lié de manière indirecte par des corrélations avec LNC et SLA. (ii) La

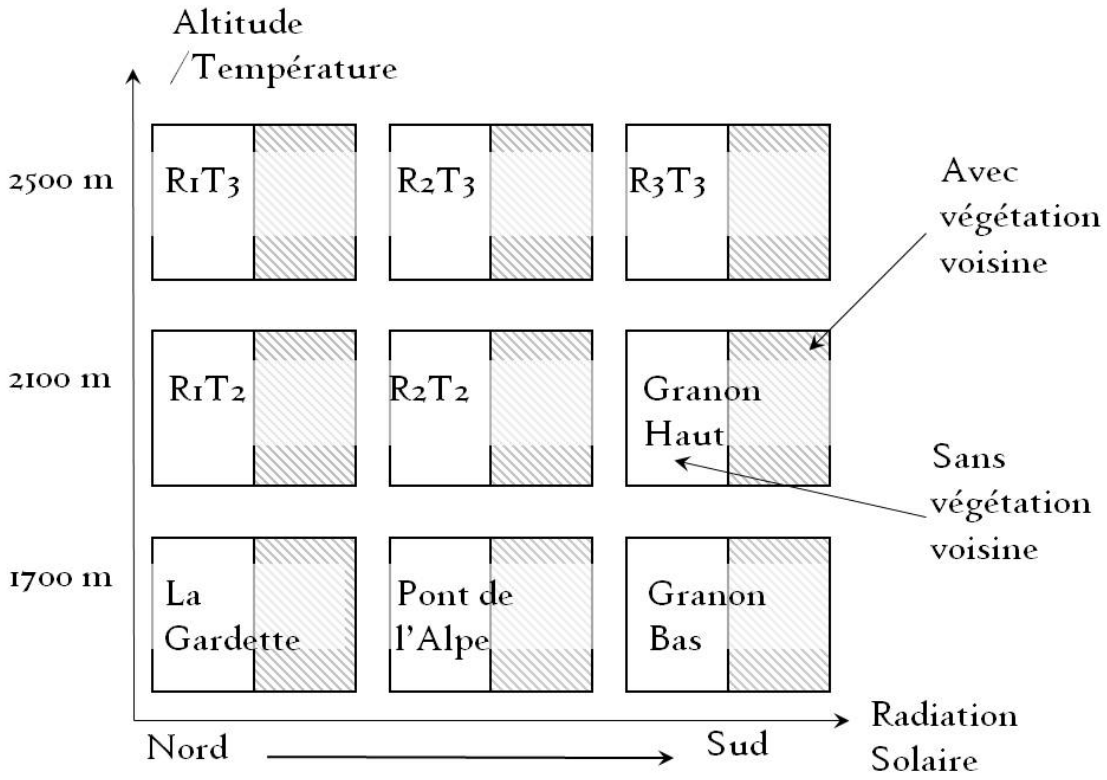


Fig. 3.5. Dispositif expérimental. *Organisation des sites le long des deux axes de stratification : trois niveaux de température hivernale (moyenne de la vallée et moyenne +/-1sd) et trois niveaux de radiations solaires estivales (moyenne de la vallée et moyenne +/-1sd).*
Dans chaque site deux placettes sont installées : l'un sans végétation (blanc) et l'autre avec végétation (gris).

Fig. 3.6. Dispositif expérimental. *Photo du site Granon Bas en Novembre 2007 (Photo P. Saccone). Gros plan sur les plantules émergeant après la germination en Juillet 2008.*

variation dans la valeur sélective des individus dépend des variations correspondantes et simultanées pour de multiples traits (Milla *et al.* 2009). Pour cette raison, la variation d'un seul trait ne permet peut-être pas une bonne description de la niche et c'est le syndrome dans son ensemble qu'il faudrait considérer.

2.2. Une relation distendue entre modèles d'habitat et démographie : manipulation de germination et survie

2.2.1. Objectifs

Pour aller plus loin que l'étude précédente, nous avons tenté, comme suggéré par Pulliam (2000) de quantifier la niche à l'aide de taux démographiques. Nous avons choisi de suivre deux paramètres démographiques clé : la germination et la survie des plantules à l'échelle de la vallée de la Guisane. Ce sont deux étapes cruciales et limitantes du cycle de vie des plantes et du processus de régénération (Baskin & Baskin 1998), car les graines et plantules sont des stades vulnérables aux facteurs biotiques et abiotiques (Kitijama & Fenner 2000; Harper 1977). De plus ces étapes influencent fortement la dynamique des populations et la composition des communautés (Harper 1977). L'objectif était de confronter les taux de germination et de survie dans des sites environnementalement contrastées avec la qualité de l'habitat de ces sites.

2.2.2. Principe

Nous avons choisi pour cette manipulation les deux espèces d'arbre du projet Guisane 2080 : *L. decidua* et *P. uncinata*. Nous avons sélectionné dans la vallée de la Guisane neuf sites expérimentaux avec des conditions environnementales contrastées censées représenter différents niveaux de qualité de l'habitat pour chacune des deux espèces (Fig. 3.5 & 3.6). Ils ont été choisis suivant une double stratification: trois niveaux de température hivernale (moyenne de la vallée et moyenne +/-1sd) et trois niveaux de radiations solaires estivales (*idem*, voir Chapitre 1 et Fig. 3.5 & 3.7). Dans chaque site nous avons choisi la placette de manière à avoir la pente la plus faible possible et nous y avons aménagé des exclos de 60 cm de côté et 50 cm de hauteur afin d'éviter au maximum l'herbivorie et la prédation des graines (Fig. 3.6). Afin de tester l'influence des conditions biotiques sur la description de la niche par les paramètres démographiques, nous avons mis en place dans chacun des neufs sites une placette sans modification de la végétation locale et une placette sans végétation. Dans chaque placette nous avons planté pour chaque espèce 50 graines et 20 plantules. Les graines (Office National des Forêts, faculté germinative évaluée à 77%) et les plantules âgées de 2 ans (pépinière d'altitude Puthod, 74) provenaient des Alpes internes Nord. Durant l'été 2008, nous avons compté le nombre de succès de germination et de survie. Nous avons aussi mesuré des variables locales : pH, taux de matière organique et texture du sol, teneur volumétrique en eau du sol (Time Domain Reflectometry, miniTRASE system 1, Soil

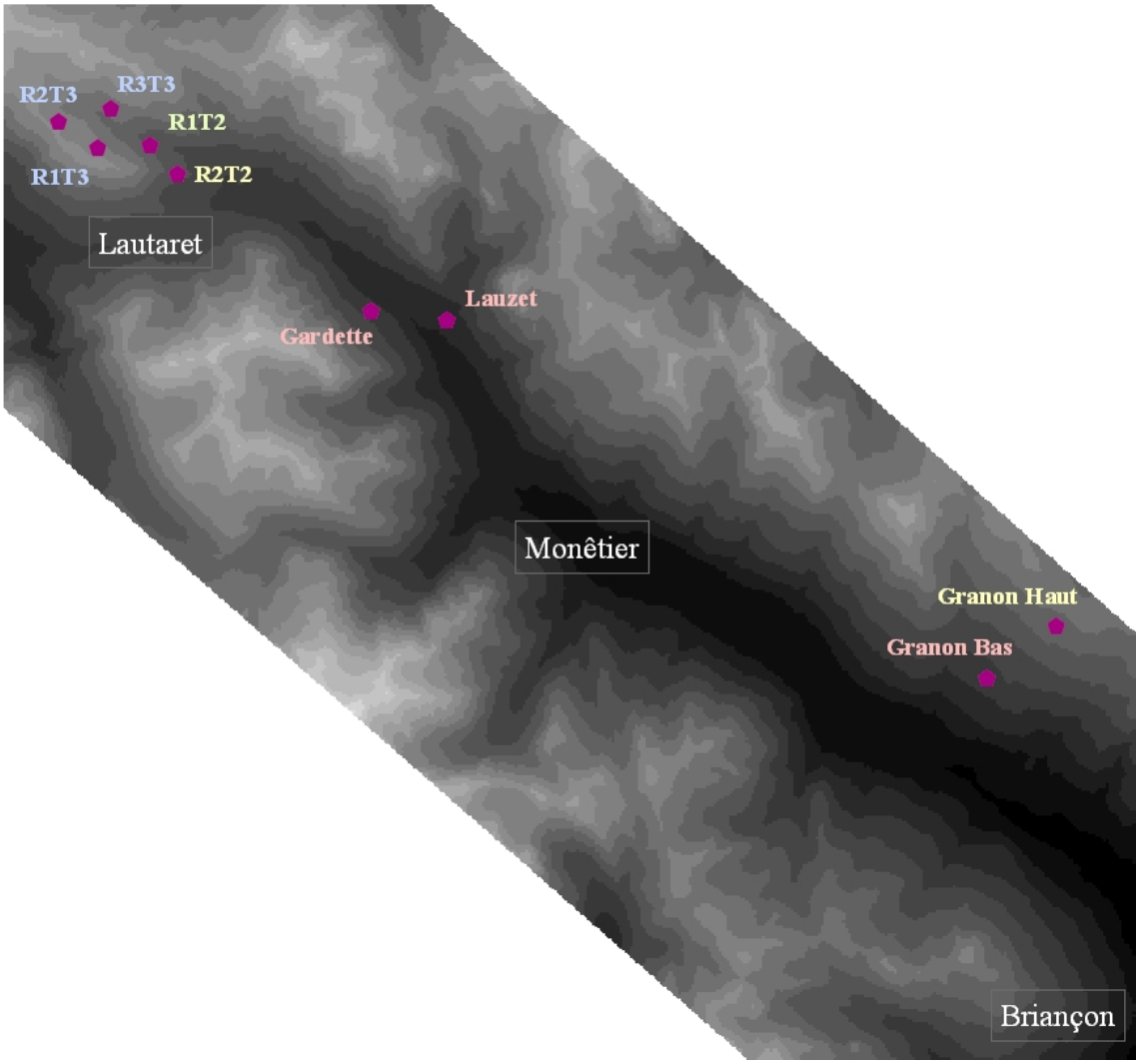


Fig. 3.7. Dispositif expérimental. *Répartition des sites expérimentaux dans la vallée de la Guisane. Le fond de carte est une stratification par altitude (niveaux de gris) : les basses altitudes sont en noir et les hautes altitudes en blanc. La localisation des sites est indiquée par un hexagone. Les trois sites correspondant aux plus fortes températures sont notés en rose ; les trois sites intermédiaires en jaunes et les trois sites correspondant aux plus faibles températures en bleu. L'orientation des sites peut-être lue sur la carte, celle-ci étant orientée vers le Nord.*

moisture equipment corporation, Santa Barbara, CA, USA), température du sol (mesures toutes les heures sur une année ; iButton Temperature loggers, Maxim Integrated Products, Sunnyvale, CA, USA), hauteur de végétation, relevés botaniques et nombre d'espèces présentes, quantité de sol nu, durée d'ensoleillement en août (horizontoscope, Tonne 1954). Les données ont été traitées à l'aide de modèles linéaires généralisés sur famille binomiale. Les modèles d'habitat pour les deux espèces étaient les mêmes que dans l'analyse précédente (1.2).

2.2.3. Résultats

Pour les deux espèces nous avons observé (Fig. 3.8) des taux de germination (entre 0 et 70%) et des taux de survie (entre 0 et 100%) très variables entre les placettes (test du X^2 tous largement significatifs). Les deux espèces ont des réponses similaires aussi bien en termes de germination qu'en termes de survie des plantules (comparaison des colonnes de la Fig. 3.8). Les deux paramètres démographiques ont des réponses différentes (comparaison des lignes de la Fig. 3.8). Pour les deux espèces la survie a été plus importante dans les sites caractérisés par une saison de végétation courte, c.à.d. un enneigement plus long caractérisé par une température au sol nulle (R^2=0,37 pour le mélèze et 0,38 pour le pin). Par contre aucun des facteurs mesurés (température, humidité du sol, pH du sol, matière organique, texture du sol, température) n'a permis d'expliquer la variabilité des taux de germination entre sites ($p>0,5$). Dans le site central R2T2, un névé persistant pourrait avoir provoqué un glissement des graines hors de la placette, ce qui expliquerait les très faibles taux de germination et peut-être l'absence de facteur explicatif.

Pour ce qui est de l'effet de l'environnement biotique (avec *vs* sans végétation), on a observé en moyenne moins de germination et de survie sous couvert (effets négatifs significatifs $p<10^{-4}$), mais les effets étaient contrastés entre les sites.

Excepté pour la survie de *L. decidua* sans végétation (R^2 =36%) nous n'avons trouvé que des liens assez faibles entre les variables démographiques et la qualité de l'habitat donnée par les modèles d'habitat des deux espèces (Tableau 3.2). La tendance était à des effets négatifs de la qualité de l'habitat sur la germination et positifs sur la survie.

Il est intéressant de noter qu'une caractérisation post-échantillonnage des sites d'études à l'aide d'une ACP sur les données mesurées *in situ* modifie fortement la typologie initiale des sites. Certains sites choisis pour différer uniquement par leur température (Tminh) comme R3T3 et Granon-Haut se sont révélés avoir des sols (humidité et teneur en sable) et des végétations très différents (Fig. 3.9). Or ces facteurs locaux sont susceptibles d'influencer plus les taux de germination et de survie des plantules que les facteurs régionaux. Cela souligne encore une fois l'importance de l'échantillonnage par rapport à la question posée et met en évidence l'importance des données indisponibles avant étude comme les données de sol (Chapitre 1).

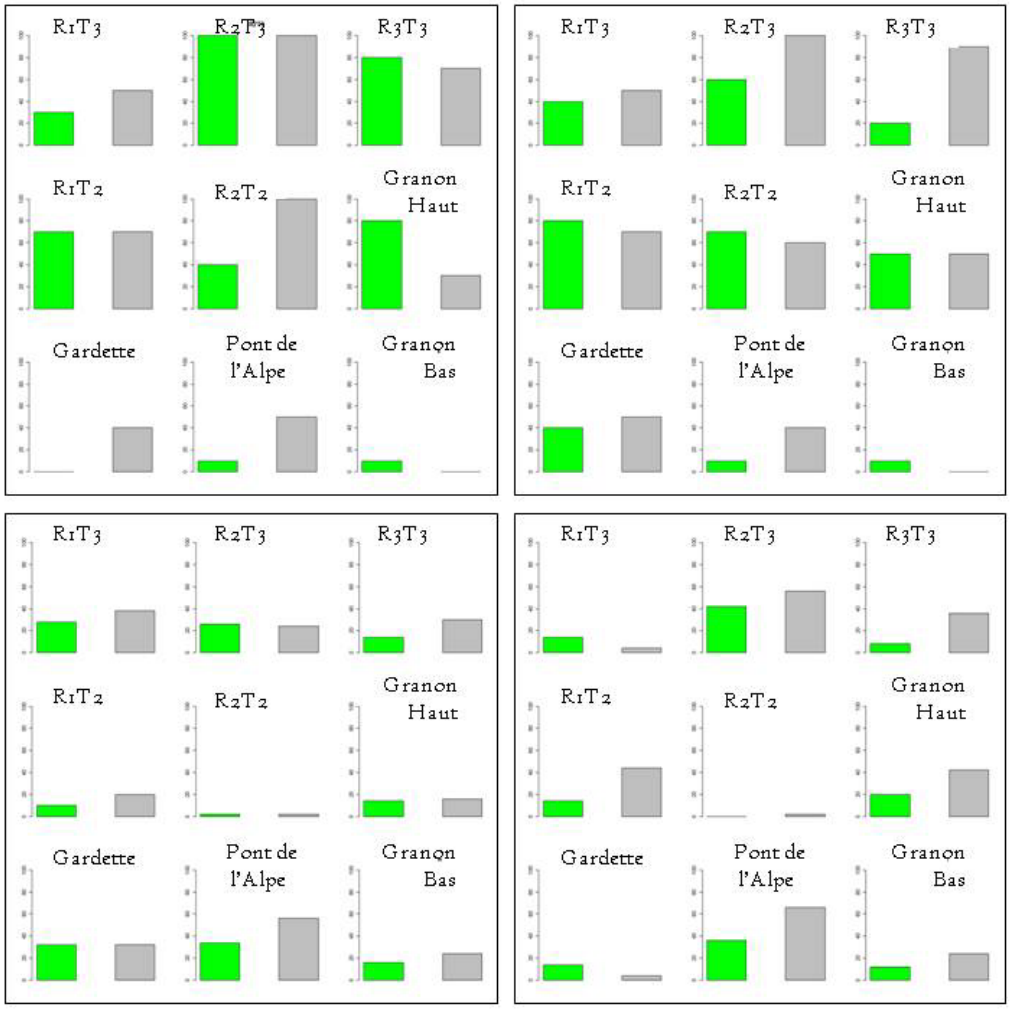


Fig. 3.8. Taux de Survie et taux de germination obtenus en 2008. *La première ligne donne les taux de survie en % ; la deuxième ligne les taux de germination en %. La colonne de gauche donne les résultats de L. decidua et celle de droite ceux de P. uncinata. Dans chaque panneau, les 9 sites sont présentés selon les axes de radiation/ orientation (Nord à gauche et Sud à droite) en abscisse et température/altitude en ordonnée (Basses températures en haut et hautes températures en bas), comme dans la Fig. 3.5. Pour chaque site sont donnés les résultats avec (vert, à gauche) et sans (gris, à droite) végétation.*

2.2.4. Discussion

Les résultats obtenus ici sont encore une fois très complexes et contrastés. Pour la germination il est possible que les résultats soient corrompus par la présence du névé persistant comme je l'ai dit. Cependant nous avons mis en évidence une relation forte (R^2=36%) positive entre qualité de l'habitat et survie des plantules de *L. decidua*. Ce résultat va dans le sens de travaux précédents ayant testé le lien entre qualité de l'habitat et taux de recrutement (Wright *et al.* 2006) et qualité de l'habitat et fécondité (Elmendorf & Moore 2008). Du point de vue des modèles d'habitat, l'absence de résultats plus clairs pourrait ici être du au fait (i) que l'échantillonnage a été mené sur deux gradients régionaux. Ces gradients régionaux sont relativement bien représentatifs des gros contrastes environnementaux. Nous avons trouvé des corrélations fortes entre Tminh (extrapolation régionale) et la température moyenne du mois de juillet 2008 mesurée (R^2=76%) et entre Rad8 (extrapolation régionale) et l'ensoleillement estimé localement (R^2=71%). Cependant ces gradients régionaux ne résument pas les différences de sol, de végétation et autres facteurs locaux et sont des moyennes sur plusieurs dizaines d'années qui ne représentent pas les variations saisonnières (ex. les différences entre les années 2007 et 2008, Fig. 2.2). (ii) Le choix des sites est plus adapté à obtenir de bons résultats pour *L. decidua* que pour *P. uncinata*, car les sites représentent des habitats de bonne, moyenne et mauvaise qualité pour *L. decidua*, mais surtout des sites de moyenne et mauvaise qualité pour *P. uncinata*. Nous avons volontairement choisi deux espèces n'ayant pas les mêmes exigences climatiques, ou du moins pas la même distribution actuelle. *L. decidua* est présent à des altitudes plus élevées que *P. uncinata*. Il aurait donc fallu mettre en place plus de sites ou des sites différents pour les deux espèces, pour obtenir des habitats de bonne qualité pour *P. uncinata*. Cela montre que ce genre de manipulation est difficile à mettre en place pour de nombreuses espèces ayant des exigences environnementales contrastées.

D'un point de vue écologique : (i) *Ontogénie* - germination et survie des plantules ne sont pas influencées de la même manière par les facteurs environnementaux et sont probablement influencés encore différemment de la croissance et du succès reproducteur des arbres adultes. Cela laisse penser qu'étudier les paramètres démographiques de manière indépendante n'est pas un bon estimateur de la performance (Violle *et al.* 2007 ; Milla *et al.* 2009), mais aussi que les conditions environnementales requises par l'espèce peuvent changer au cours de la vie des individus (ontogénie, Eriksson 2002; Cornelissen *et al.* 2003a). Il n'est donc peut-être pas approprié d'essayer de lier les modèles d'habitat établis à partir de la distribution d'arbres adultes et les processus démographiques des arbres juvéniles. Il faudrait donc peut-être réfléchir ici à une autre manière de quantifier la niche, qu'à partir des données de distribution des adultes. (ii) *Interactions biotiques* - Il est intéressant de noter que la plus forte relation (R^2 de 36% entre survie de *L. decidua* et qualité de l'habitat) a été obtenue

Taux	*Espèce*	*Sans Végétation*	*Avec Végétation*
Germination	*L. decidua*	*15 (-)*	*14 (-)*
	P. uncinata	*2 (-)*	*8 (-)*
Survie	*L. decidua*	*36 (+)*	*6 (+)*
	P. uncinata	*1 (+)*	*5 (+)*

Tableau 3.2. R^2 et signe des relations entre qualité de l'habitat et paramètres démographiques.
Les résultats sont données pour les deux paramètres démographiques mesurés : taux de germination et taux de survie des plantules ; pour les deux espèces étudiées : L. decidua et P. uncinata et pour les deux traitements : avec et sans végétation. Toutes les pentes étaient significativement non nulles (p<0,05).

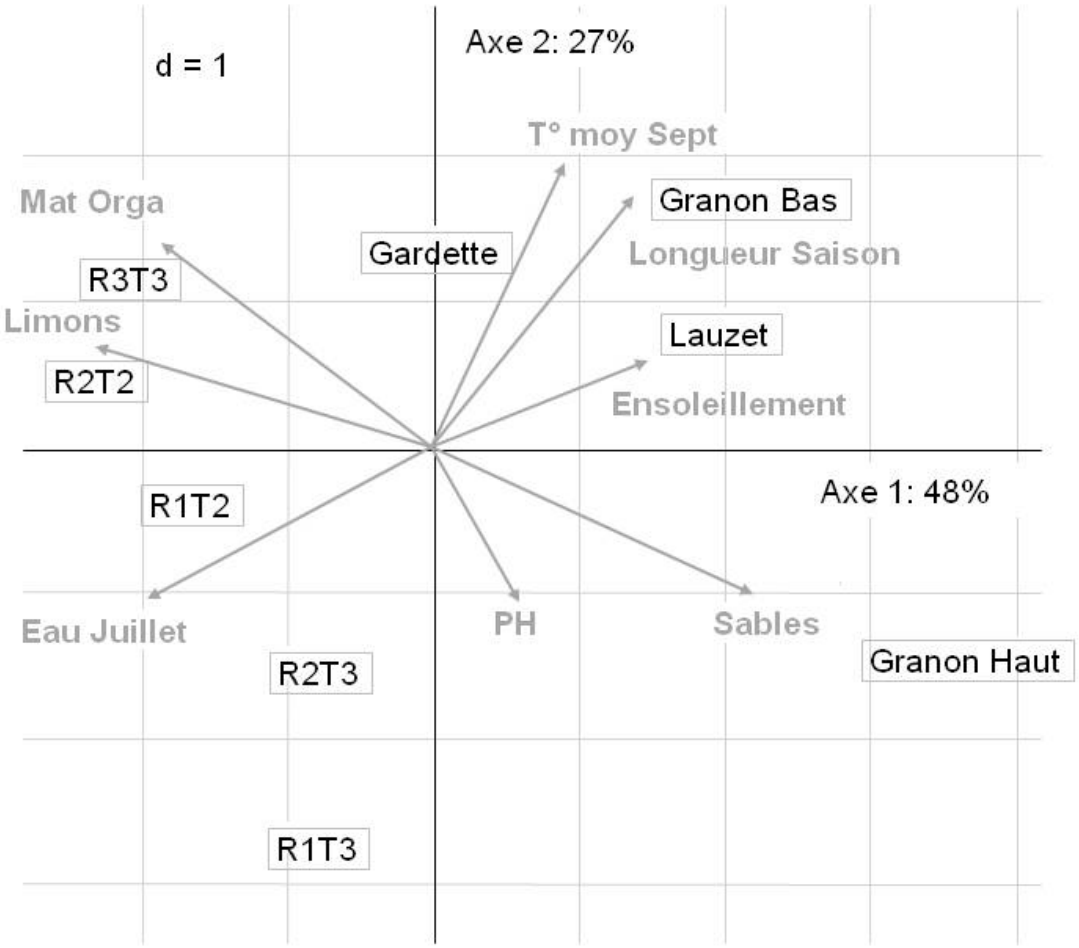


Fig. 3.9. Caractérisation des neufs sites expérimentaux.
La figure représente les deux premiers axes de l'ACP effectuée sur huit variables environnementales mesurées in situ : la température moyenne du mois de Septembre, la longueur de la saison de végétation (nombre de jours où la température du sol est supérieure à 0°C), la durée de l'ensoleillement évalué à l'horizontoscope et les taux de sable, de limons, de matière organique, d'eau du sol ainsi que son pH (en gris). Les deux premiers axes représentent respectivement 48 et 27% de la variance. Les stations sont localisées par les boites grises.

sans végétation, la relation avec végétation étant quasi inexistante (R^2=6%). Cela est cohérent puisque les modèles d'habitat ne prennent pas directement en compte ces interactions biotiques. A l'inverse Elmendorf & Moore (2008) ont montré que la prise en compte de ces interactions dans les modèles (modèles de co-occurrence) permettait également d'améliorer le lien entre fécondité et qualité de l'habitat. Pour ce qui est des interactions biotiques en tant que telles, nous attendions des effets plus contrastés avec éventuellement un inversement des effets depuis des effets négatifs (compétition) dans les conditions les moins stressantes (ex : hautes températures) vers des effets positifs (facilitation) dans les conditions les plus stressantes (ex : basses températures, review Brooker *et al.* 2008). Nos résultats ne montrent pas de tels effets par rapports aux gradients Tminh et Rad8. (iii) *Etape limitante* - Il est surprenant de voir les forts taux de survie qui ont eu lieu dans les sites froids, en bordure de la limite naturelle des forêts (~2300 m dans les Alpes). Ces forts taux de survie se retrouvent encore à la fin de la deuxième année (été 2009, résultats non présentés). Ces sites climatiquement extrêmes sont certainement assez favorables à une survie des plantules car la saison y est courte : durant les périodes critiques de gelées -au printemps et à l'automne- elles sont protégées par le manteau neigeux et durant l'été la chaleur et la sècheresse y sont moins marqués (en juillet en moyenne 30% d'humidité du sol et 30°C de température maximale au sol sous végétation dans les sites de haute altitude contre 20% et 40°C dans les sites de basse altitude). Ces résultats obtenus aussi bien avec et sans végétation vont dans le sens de l'hypothèse de C. Körner qui propose que la limite altitudinale des forêts est due à la stratégie « grande taille » des arbres, qui nécessite une capacité à produire de la structure incompatible avec les basses températures (Körner 1998, 1999). Ceux-ci ne pourraient survivre que tant qu'ils ne dépassent pas la hauteur des herbacées (stratégie basse adéquate). Cela signifierait que l'étape limitant l'établissement des arbres dans les conditions étudiées ne serait pas les taux de germination ou de survie des plantules, mais plutôt l'émergence hors du couvert herbacé. Il faudrait donc plutôt étudier cette étape pour quantifier la niche des espèces étudiées par des paramètres démographiques, puisque l'étape limitante influence le taux de croissance des populations.

2.3. Une relation distendue entre modèles d'habitat et démographie : démographie des forêts des Alpes françaises

2.3.1. Objectifs

L'une des limitations de l'étude précédente étant les éventuelles modifications des exigences climatiques des individus au cours de leur vie, nous nous sommes ici intéressés à des arbres adultes en accord avec les données d'occurrence d'arbres adultes utilisées dans les modèles d'habitat. La croissance est le phénomène biologique d'augmentation de la taille dans le temps qui implique la formation de nouveaux tissus. L'accroissement (ex. accroissement radial) est l'ampleur de l'augmentation de la taille (ex. diamètre) sur une période donnée du

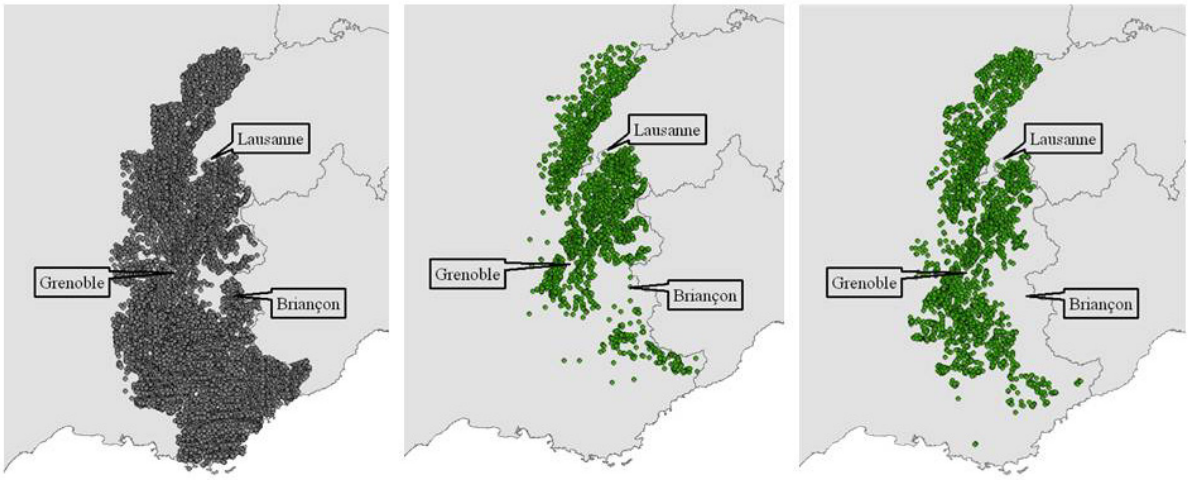


Fig. 3.10. Carte des points de données de l'IFN (à gauche) et distribution (présences) de Picea abies (au centre) et de Fagus silvatica (à droite).

fait de la croissance. Chez les arbres, il est possible d'évaluer un tel accroissement en mesurant la largeur des cernes annuels de croissance sur des carottes de bois prélevées à 1,30 m à l'aide d'une tarière de Pressler. Ces largeurs de cernes sont des intégrateurs complexes de multiples signaux (espèce, génotype, âge, taille, climat régional, conditions stationnelles, perturbations régionales, variabilité individuelle, hétérogénéité des ressources) correspondant aux caractéristiques intrinsèques de l'arbre et aux conditions environnementales dans lesquelles il croît. Il est connu que la croissance des arbres en forêts naturelles est extrêmement variable (Canham *et al.* 2004). Comprendre les variations de cette croissance est essentiel en écologie forestière puisque c'est un facteur clé structurant les forêts. De plus la croissance des individus -responsable de l'accroissement de biomasse- est l'une des composantes de leur fitness (Violle *et al.* 2007). On pourrait s'attendre d'une part à ce que la croissance et donc les accroissements radiaux des individus décroissent avec l'altitude (basses températures, courte saison de végétation, exposition croissante aux vents) et d'autre part à ce que les meilleurs taux de croissance soient observés au cœur de la niche, dans les habitats prédits de bonne qualité par les modèles d'habitat.

Etant données les limites des études locales (peu de données, non équilibre probable des espèces avec leur environnement) nous avons ici changé d'échelle d'étude pour passer à l'échelle des Alpes françaises. L'objectif était de confronter les accroissements radiaux d'arbres individuels avec la qualité de l'habitat.

2.3.2. Principe

Les données de l'inventaire forestier national (IFN) dans les Alpes comptent quelques 14 741 relevés (Fig. 3.10). Ces relevés contiennent les données individuelles d'accroissement radial (en m) sur les cinq dernières années (année en cours exclue, mesures effectuées entre 1992 et 2002) des arbres présents dans une placette circulaire de 15 m de rayon. Ne sont présentés ici que les résultats concernant *Picea abies* et *Fagus silvatica* qui sont les espèces les mieux représentées avec plus de 25 000 individus dans plus de 3 000 placettes de la zone d'étude (Fig. 3.10). Pour les deux espèces nous avons calibré des modèles d'habitat de type modèles additifs généralisés avec la méthode du moyennage de modèles (Encadré 2.2) sur les données en présence/absence des placettes IFN et les données environnementales à 100 x 100 m (Zimmermann & Kienast 1999): degré-jours à 5°C, radiation solaire annuelle, humidité relative et précipitations annuelles. Afin d'établir la part de variabilité des croissances individuelles expliquée par les sorties de modèles d'habitat, nous avons utilisé des modèles mixtes (Encadré 2.1) qui permettent de prendre en compte la structure hiérarchique de l'échantillonnage: $Croissance = fixe\ (Qualité.de.l.habitat),\ aléatoire\ (Placette)$.

Ces modèles nous ont permis de déterminer les parts de variabilité dues aux différences entre placettes et aux différences entre individus au sein des placettes et aussi par une comparaison avec le modèle sans effet fixe de déterminer la part de la variabilité (R^2) entre placettes due à la qualité de l'habitat (Encadré 2.1).

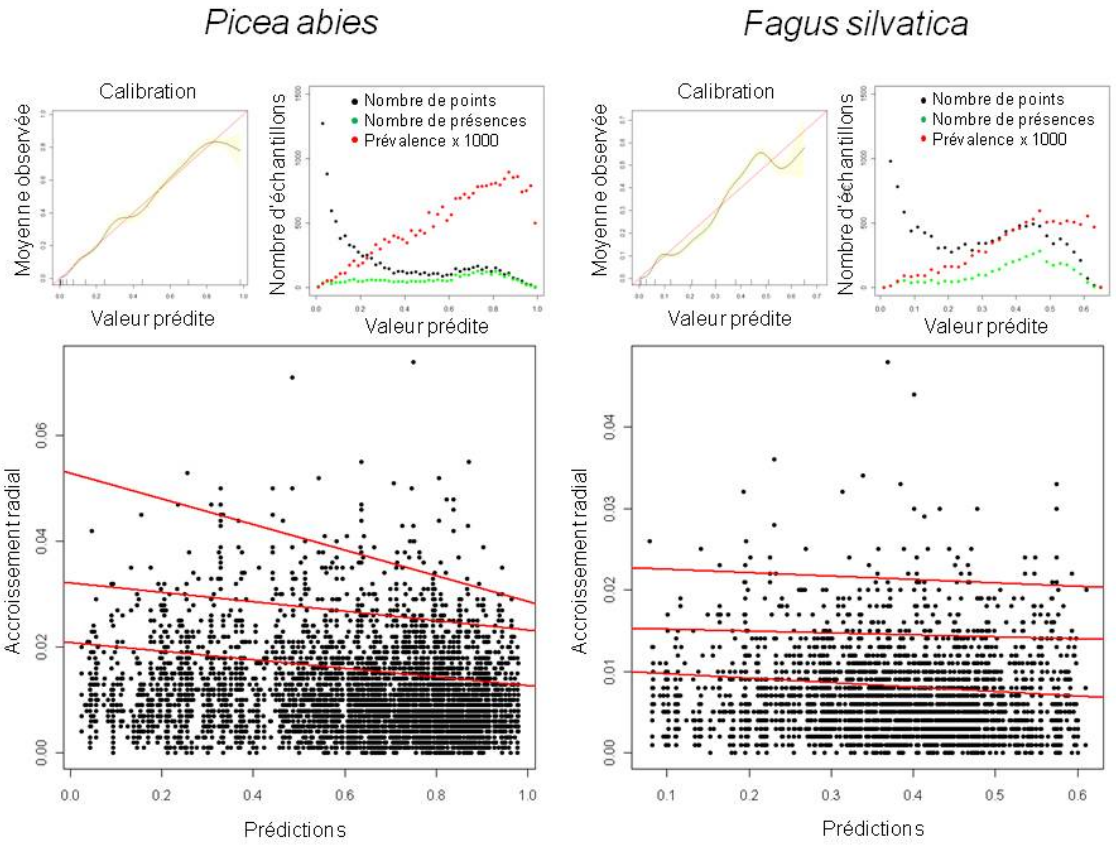


Fig. 3.11. Modèles d'habitat et confrontation avec les accroissements radiaux pour P. abies et F. silvatica.

Prédictions= Qualité de l'habitat prédite par les modèles d'habitat. Pour chaque espèce, (en haut à gauche) graphique de calibration (voir Encadré 3.2) ; (en haut à droite) nombre de points échantillonnés nombre de présence et prévalence en fonction des valeurs prédites ; (en bas) confrontation de l'accroissement radial des individus en fonction de la qualité de l'habitat prédite dans le site ou ils poussent. Les lignes rouges sont les régressions quantiles à 0.8, 0.9 et 0.95%. Seule la régression à 95% pour P. abies est significativement (p<0,05) négative.

2.3.3. Résultats

Les modèles d'habitat calibrés avaient des pouvoirs prédictifs corrects (AUC hêtre : 0,82 et épicéa: 0,91) ils étaient aussi correctement calibrés sauf pour les valeurs les plus hautes (Fig. 3.11). Pour le hêtre (resp. l'épicéa) il y avait 65% (resp. 52%) de la variabilité des croissances individuelles due aux différences entre placettes, le reste étant des différences entre les individus d'une même placette.

Cependant ces différences entre placettes n'étaient sont pas expliquées par les gradients environnementaux à l'échelle des Alpes (degré-jours, radiation solaire, humidité relative, R^2=0), ni par la qualité de l'habitat donnée par les modèles (R^2=0). La même analyse sur les données de chaque année (celles récoltées en 1999 p. ex.) donne la même absence de relation.

2.3.4. Discussion

L'absence de résultats pour cette étude n'est peut-être pas si surprenante. Paulsen *et al.* (2000) ont montré que, contrairement à la hauteur qui diminue lorsque l'on se rapproche de la limite altitudinale des forêts, la largeur des cernes de croissances était d'avantage influencée par la variabilité climatique interannuelle. De plus les 45 000 arbres étudiés ont des croissances très différentes, mais également très faibles. Pour *F. silvatica*, l'accroissement moyen sur 5 ans est de 5,8 mm, avec un maximum de 5,1 cm et un quantile à 75% de 8 mm. Pour *P. abies*, l'accroissement moyen sur 5 ans est de 1,1 cm, avec un maximum de 7,4 cm et un quantile à 75% de 1,5 cm. Seuls quelques individus, peut-être plus jeunes (pas de données d'âge) ou vivant dans des conditions vraiment particulières, présentent de forts accroissements. Il existe donc peu de contrastes dans les données. Par ailleurs diverses limites techniques n'ont pas pu être résolues pour le moment. (i) En utilisant un large jeu de données, nous pensions pouvoir mettre en évidence (en considérant les meilleures croissances dans chaque type de conditions) une éventuelle courbe potentielle de croissance. Différentes tentatives avec des régressions quantiles (Fig. 3.11) ou des calculs de moyenne, de quantile, de maximum sur des classes d'environnements n'ont rien donné (aucunes relations significatives). Cela pourrait être du à la répartition largement biaisée des données du fait d'un échantillonnage inapproprié à notre étude (Chapitre 1), car les données IFN sont des données de stations forestières et ne prennent donc que très peu en compte les arbres poussant en bordure haute de leur distribution. De ce fait en termes de degré-jours, 60% des données représentent le tiers central du gradient, les deux tiers correspondant aux sites à faibles et à fort degré-jours sont largement sous représentés (Fig. 3.12). (ii) D'autre part il n'est pas évident *a priori* de choisir les informations à prendre en compte dans cette analyse, car l'accroissement radial serait susceptible de changer avec les conditions environnementales, mais on pourrait également s'attendre à ce que l'âge, la hauteur, la circonférence (dBh) et les relations entre ces différentes grandeurs interdépendantes changent dans l'environnement. (iii) Enfin comme précédemment, les modèles d'habitat ne

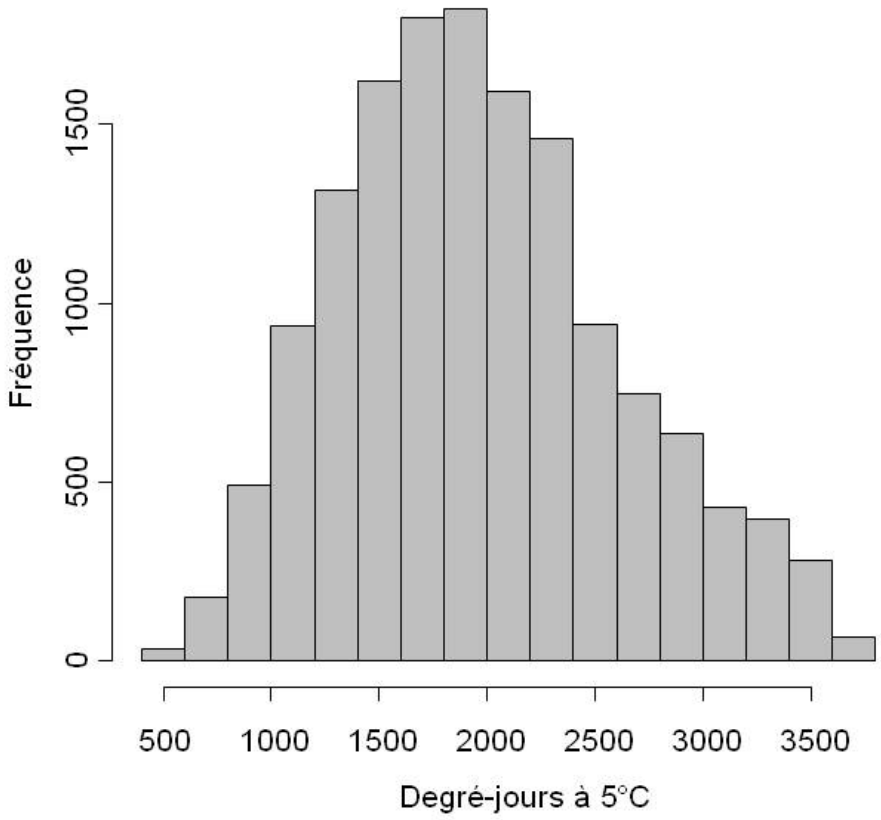


Fig. 3.12. Distribution des degré-jours à 5°C sur l'ensemble des placettes IFN étudiées

prennent pas en compte les interactions biotiques qui sont connues pour avoir une influence sur la croissance des arbres, voire pour modifier les relations entre croissance et taille (Coomes & Allen 2007).

Le point positif de cette étude concerne les modèles d'habitat. Ceux-ci sont calibrés sur des jeux de données très complets (plus de 3000 présences), décrivant vraisemblablement assez correctement la distribution actuelle des espèces. Les deux modèles ont de bons pouvoirs prédictifs et de bonnes courbes de calibration (Encadré 3.2). De plus sur l'ensemble des placettes de présence, on retrouve aussi bien des sites caractérisés par une mauvaise qualité de l'habitat, que par une moyenne, bonne ou très bonne qualité (Fig. 3.11), contrairement à ce que l'on avait obtenu pour le modèle de *P. uncinata* dans l'étude précédente. Cette représentation complète du gradient de qualité de l'habitat est un point clé à rechercher dans ce genre d'étude pour être à même de mettre en évidence des relations.

Je pense que pour pouvoir poursuivre cette étude il serait nécessaire de procéder à un ré-échantillonnage des données par rapport aux conditions environnementales -qui sont mal représentées par le jeu de données initiales- et également par rapport aux caractéristiques intrinsèques des individus (dBh) et aux interactions biotiques (présence ou abondance d'autres espèces dans la placette).

3. Conclusion

Le but de ce chapitre était de tester la principale question existant derrière une utilisation étendue des modèles d'habitat : « la qualité de l'habitat donnée par les modèles d'habitat basés sur la fréquence d'occurrence des espèces explique-t-elle les variations de performance de cette espèce ? ». Les trois tests que nous avons menés pour répondre à cette question visaient à confronter la qualité de l'habitat -donnée par les modèles d'habitat- avec des estimateurs supposés de la performance (traits fonctionnels, paramètres démographiques, croissance). Ils ont donné des résultats très inégaux mais encourageants, et en accord avec ceux des deux seules études que nous avons trouvées dans la bibliographie sur le test de cette hypothèse (Wright *et al.* 2006; Elmendorf & Moore 2008). Elmendorf & Moore (2008) ont montré dans une zone d'étude très hétérogène mais très petite (0.33 km², Guisane ~200 km²) un lien positif plus ou moins marqué entre qualité de l'habitat et fitness pour quatre espèces annuelles (R^2=30%, 25%, 14% et 3%). Ces études qui traitent d'espèces annuelles restent cependant très différentes des nôtres qui traitent d'espèces pérennes. Pour les espèces pérennes (plus de 90% de la végétation alpine) potentiellement sujettes à des modifications ontogéniques de la niche la question de lien qualité de l'habitat–performance peut-être comme nous l'avons vu plus difficile à tester. L'ensemble de ces résultats reste insuffisant mais prometteur quand à la validation de ce lien.

Dans un sens il est évident que les modèles d'habitat par construction sont faits pour donner la probabilité de présence d'une espèce dans des conditions environnementales données.

Encadré 3.1. Prévalence et sorties des régressions logistiques

Albert C.H. & Thuiller W. , Favourability functions versus probability of presence: advantages and misuses, Ecography 2008, 31:417-422

Contexte et problématique

Les GLM (logit) sont très utilisés en biogéographie prédictive sous forme de résultats binaires (transformation avec un seuil) ou de probabilités (valeurs plus fortes dans les sites les plus favorables). Mais ces probabilités dépendent de la prévalence (nombre de présences / nombre de données) du jeu de données de calibration (une grande prévalence implique de fortes probabilités), ce qui les rend incomparables entre espèces et entre modèles pour une même espèce. Pour corriger ce biais d'échantillonnage, nous avons testé deux méthodes de ré-échelonnage des probabilités, l'une considérant un jeu comme biaisé s'il a une prévalence différente de 50%, l'autre s'il a une prévalence différente de la vraie.

Méthodes et résultats

A partir d'un gradient de température virtuel, nous avons construit 3 espèces virtuelles différant par leurs « vraies » prévalences (20, 50, 80%). Pour chacune nous avons échantillonné aléatoirement de sorte à avoir des jeux de données de prévalences 20, 50 et 80%, puis nous avons calibré des GLM sur chacun des jeux et sur l'ensemble des données. Nous avons ensuite appliqué les formules de ré-échelonnage correspondant aux deux types de biais (R50 et RV). RV donne les sorties les plus cohérentes, avec des valeurs équivalentes pour les modèles calibrés pour une même espèces sur des jeux de prévalences différentes; c'est aussi RV qui améliore le mieux les prédictions quelle que soit la prévalence d'échantillonnage.

Conclusion

Une prévalence basse dans un jeu de données peu représenter une réalité ou un biais d'échantillonnage. Il n'y a pas de manière simple de corriger les sorties dans des cas réels, c'est pourquoi il faut avoir de bons échantillons au départ (Chapitre 1). L'utilisation des probabilités brutes doit être faite avec prudence plutôt de manière relative qu'absolue.

Cependant comme nous l'avons vu dans la première partie de ce chapitre, ces modèles sont souvent utilisés d'une manière plus large pour décrire une qualité de l'habitat ou la niche d'une espèce. Comme nous l'avons également vu, il n'existe pour le moment pas de consensus sur la signification « étendue » de ces modèles et c'est d'ailleurs la raison d'être de ce chapitre. Il n'est donc pas utile de rediscuter plus avant des limitations des modèles d'habitat en eux-mêmes. D'autres modèles plus complexes pourraient d'ailleurs être plus appropriés que les modèles d'habitat pour décrire la performance des espèces; Elmendorf & Moore (2008) ont montré que les modèles de co-occurrences d'espèces qui intègrent facteurs abiotiques et biotiques (Leathwick & Austin 2001) expliquaient mieux la fitness et l'abondance des espèces que les modèles d'habitat. On pourrait aussi penser à des modèles de types modèles statistiques d'abondance (Pearce & Boyce 2006; Potts & Elith 2006) ou encore modèles mécanistes physiologiques (Chuine *et al.* 1999; Kearney & Porter 2009). Mais ces modèles moins simplistes sont aussi beaucoup plus difficiles à calibrer. N'oublions pas que l'avantage premier des modèles d'habitat est qu'ils sont simples à utiliser (au moins en apparence) et qu'ils requièrent des données (occurrence) qui sont les plus disponibles (ex. atlas, relevés botaniques) et les plus faciles à collecter.

Il existe cependant une limitation qui est en lien direct avec notre question. Cette limitation est plus une limitation de l'usage que l'on fait des modèles que des modèles en eux-mêmes. Il s'agit du choix des variables environnementales. Le plus souvent ces variables sont des variables climatiques moyennes (ex. sur la période 1970-1999) extrapolées à partir de modèles météorologiques et de relevés satellites et réduites à des résolutions de quelques kilomètres ou centaines de mètres. Ces données sont donc grossières et comportent de fortes incertitudes (Seo *et al.* 2009). Trois problèmes peuvent survenir du choix de ce type de variables. (1) Il peut y avoir une incompatibilité des échelles temporelles. En effet les mesures *in situ* de traits fonctionnels ou de taux de germination reflètent les conditions climatiques de l'année de la mesure et non la moyenne des dix voire trente années passées. (2) Il peut y avoir une incompatibilité des échelles spatiales. En effet les données climatiques régionales dérivent de données à très larges échelles et ne reflètent pas les conditions locales qui affectent directement la croissance, la germination et les traits fonctionnels. Cela soulève d'ailleurs aussi une question de compromis entre échelle de l'étude et qualité des modèles d'habitat mis en œuvre. Dans leur étude Elmendorf & Moore (2008) ont calibrés les modèles d'habitat à partir de variables très fines (sol, radiation) mesurées localement. Il est vrai que cela semble donner des résultats satisfaisants et que cette approche est intéressante, mais elle n'est pas vraiment généralisable à l'échelle de la distribution entière d'une espèce. De plus dans notre site d'étude nous avons montré que les données régionales semblent bien refléter les gros contrastes climatiques en termes de température et radiation. (3) C'est alors peut-être le type de données utilisées qui ne reflète pas complètement les conditions environnementales locales qui ne se résument pas aux températures et radiations. Des données de sol par exemple (ex. pH, texture) ou de ressource (ex. lumière, humidité, Fig. 3.2

Encadré 3.2. Evaluation des modèles d'habitat (Pearce & Ferrier 2001)

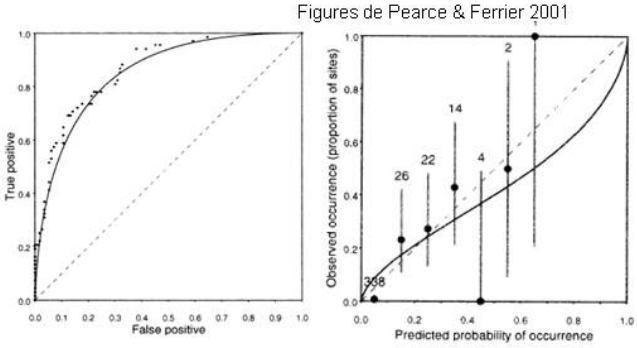
Discrimination (Swets 1988)

Les probabilités issues des modèles d'habitat peuvent être converties en prédiction binaires à l'aide d'un seuil (présence au dessus du seuil et absence en dessous), un seuil optimal permettant de discriminer au mieux présences et absences. Le pouvoir prédictif (discriminant) du modèle peut-être évaluer grâce à l'aire sous la courbe ROC («receiver operating characteristic »), l'AUC («area under the curve») qui représente le pourcentage de présences correctement prédites contre le pourcentage d'absences incorrectement prédites calculés pour différents seuils (Fig. gauche). Classiquement il est considéré qu'un modèle ayant une AUC entre 0.9 et 1 est excellent, entre 0.8 et 0.9 est bon, entre 0.7 et 0.8 est correcte, entre 0.6 et 0.7 est mauvais et entre 0.5 et 0.6 est très mauvais (équivalent à un modèle aléatoire). Il existe d'autres méthodes telles que celle du Kappa (Cohen, 1960).

Calibration

La calibration décrit l'accord entre les probabilités prédites et la proportion de sites occupés correspondant à chaque classe de prédictions (Fig. droite). Elle se décompose en biais et étalement; le biais décrit une surestimation ou sous-estimation systématique des probabilités d'occurrence de l'espèce tandis que l'étalement décrit l'éloignement de la courbe de calibration à la bissectrice.

Figures de Pearce & Ferrier 2001



lien a) pourraient permettre d'améliorer la pertinence des modèles et les rapprocher d'une caractérisation des conditions subies localement par les individus (Coudun & Gegout 2007 ; Elmendorf & Moore 2008) et donc de leur performance. Mais pour le moment ces données ne sont pas disponibles à large échelle.

5. Perspectives

Confronter directement la qualité de l'habitat donnée par les modèles d'habitat avec la réponse des estimateurs de la performance (paramètres démographiques, traits fonctionnels) aux gradients n'était finalement peut-être pas la meilleure manière de procéder (Fig. 3.2). En choisissant cette démarche, nous avons comparé deux estimateurs différents de la niche (une en fréquence d'occurrence et une fonctionnelle) et une absence de résultat ne permet pas trancher entre la défaillance de l'un ou l'autre des estimateurs. Je pense que pour poursuivre l'évaluation du sens écologique des modèles d'habitat il serait désormais nécessaire de se pencher d'avantage sur deux aspects complémentaires :

Aspects écologiques – Cela concerne la partie droite de la Fig. 3.2, c.à.d. comment estimer la fitness ? Et quels sont les liens existants entre les différents estimateurs de cette fitness ? Il serait essentiel en effet de comprendre comment sont liés fréquences d'occurrence, abondance, traits fonctionnels, paramètres démographiques et fitness aux niveaux inter- et intraspécifique. A l'échelle interspécifique, Poorter (2006 & 2008) ont par exemple montré que les traits foliaires étaient de bons estimateurs des paramètres démographiques (Croissance liée à la SLA, à la densité du bois et à Hmax et Survie liée à la durée de vie des feuilles) ; Cingolani *et al.* (2007) ont montré que ce n'était pas les mêmes traits qui influaient sur les patrons locaux de présence et d'abondance. Elmendorf & Moore (2008) rapportent aussi que des relations positives, négatives ou inexistantes ont été décrites entre abondance et fitness au sein d'une espèce. Maintenant il faudrait continuer d'améliorer notre connaissance de ces relations, en particulier au niveau intraspécifique. Cela passe par la recherche de traits fonctionnels, d'ensemble de traits ou d'un indice synthétique de ces traits (vision de type stratégie fonctionnelle dans un espace multidimensionnel de traits, voir Chapitre 2) représentant les différents aspects de la performance et donnant une quantification fonctionnelle de la niche, comme suggéré par Violle (2009). D'autre part il faudrait aussi se pencher sur la question particulière de la fréquence d'occurrence et de ce qu'elle représente. Pour cela je pense qu'il serait nécessaire de revenir en arrière et d'utiliser des modèles simples purement descriptifs (et non prédictifs), pour estimer la fréquence d'occurrence. Cela permettrait de comparer réellement les patrons d'occurrence connus avec des patrons d'abondance ou des courbes de réponse de traits. S'il s'avère alors que la fréquence d'occurrence n'est pas liée à la fitness ou à ses estimateurs, alors il n'y aura plus de raison de penser que les modèles d'habitat (qui font le lien entre fréquence d'occurrence et environnement) sont une estimation de la niche (qui fait le lien entre fitness et gradients de

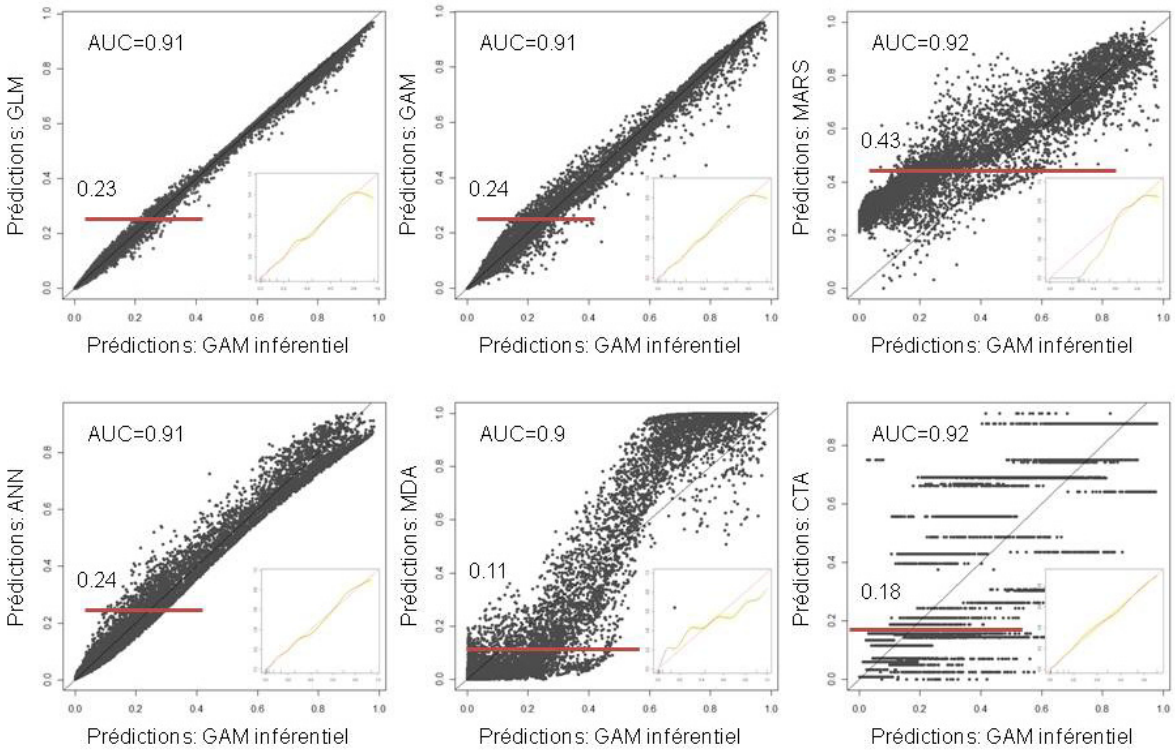


Fig. 3.13. Qualité de l'habitat donnée pour P. abies par différents algorithmes.
Ces modèles ont été calibrés sur 14741 relevés contenant 3098 présences. Les qualités de l'habitat données par les différents algorithmes sont présentés en fonction des qualités de l'habitat données par la méthode du moyennage de modèle appliquée à un GAM, d'AUC=0.91 (Encadré 2.2). Pour chaque algorithme est donné son AUC ainsi que sa courbe de calibration (Encadré 3.2) ainsi que le seuil optimal de prédiction binaire (ligne rouge).

ressources). Par contre, s'il s'avère que la fréquence d'occurrence est liée à la fitness ou à ses estimateurs, alors il faudra trouver des modèles d'habitat efficaces qui permettent de conserver ce lien et de quantifier la niche des espèces.

Modélisation de l'habitat – Comme nous l'avons vu en introduction de ce Chapitre, les modèles d'habitat ont fait l'objet de nombreuses études comparatives (Elith *et al.* 2006) et de bricolage statistique (Real *et al.* 2006 *vs* Article 3). Austin (2007) avait critiqué une telle accumulation d'information sur ces modèles (et les divers algorithmes disponibles) sans réflexion écologique préalable et sans test de compatibilité avec la théorie écologique. Il avait également soulevé la question de l'évaluation des modèles d'habitat qui repose sur des méthodes purement statistiques et non pas sur leur réalisme écologique. Une étude très récente s'est enfin penchée sur le pourquoi des différences observées entre modèles (Elith & Graham 2009). Cette étude montre en particulier des différences flagrantes entre les courbes de réponse estimées par les différents modèles et les conséquences sur la projection des distributions dans l'espace géographique. Je pense en effet que ce genre d'études cherchant à comprendre le fonctionnement fin des modèles devrait se généraliser et qu'il faudrait également mettre au point des méthodes d'évaluation des modèles en lien avec les grandeurs écologiques qu'ils sont censés représenter. Comme nous l'avions suggéré dans l'Article 3 et illustré dans la Fig. 3.13 et dans la Fig. 5 d'Elith (2009), il est particulièrement étrange de constater que différents algorithmes calibrés sur un même jeu de données produisent des valeurs de qualité d'habitat très différentes. Ces algorithmes (Fig. 3.13) sont pourtant tous bien calibrés et ont un bon pouvoir prédictif (Encadré 3.2), cela signifie que les valeurs hautes (resp. basses) de qualité de l'habitat représentent plutôt bien les présences (resp. absences). Cependant s'ils sont censés donner une mesure de la même grandeur (la niche ?), lequel croire ? Et pourquoi ? A l'origine, les modèles d'habitat ont été pensés pour décrire des distributions et prédire des présences potentielles sous une forme binaire. Dans ce contexte, des algorithmes ayant une bonne capacité à discriminer les conditions environnementales correspondant plutôt à des présences ou à des absences et des méthodes d'évaluations de ce pouvoir de discrimination de type AUC (Encadré 3.2) étaient appropriées. Aujourd'hui, si l'on souhaite réellement obtenir de bons modèles quantifiant la niche il faudrait rechercher les algorithmes les plus appropriés et ayant un bon pouvoir de quantification et non plus de calibration. C'est à mon sens un pré-requis à l'utilisation des sorties de ces modèles sous leur forme brute (probabilités de présence). Sans un effort particulier dans ce sens, il faut simplement admettre que les modèles d'habitat sont en fait de simples modèles de distribution et les utiliser en transformant leurs sorties sous forme binaires.

Chapitre 4

Prise en compte de la variabilité fonctionnelle intraspécifique et implications

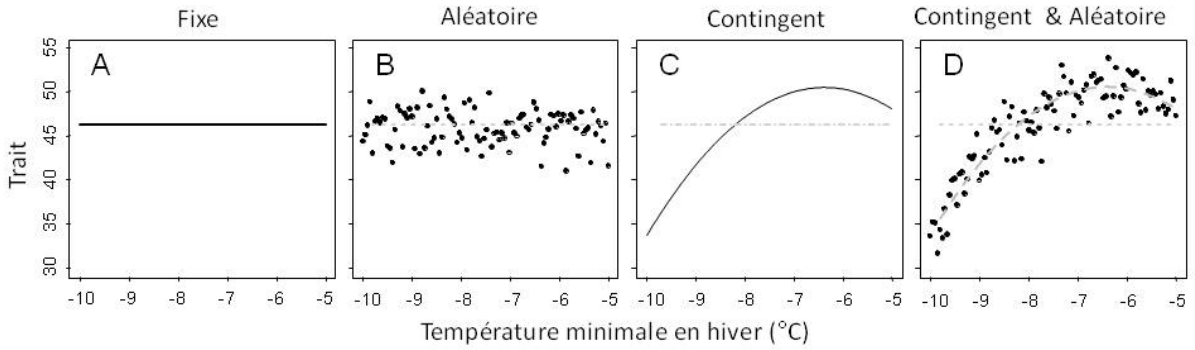


Fig. 4.1. Les différents types de variabilité intraspécifique pour une espèce unique.

*A. Sans variabilité ; B. avec une variabilité uniquement aléatoire suivant une loi normale de moyenne= trait moyen et d'écart-type=CV*trait moyen ; C. avec une variabilité uniquement contingente à l'environnement, c'est-à-dire qui a la forme d'une fonction de l'environnement ; D. avec une combinaison des deux types de variabilité : contingente à l'environnement et aléatoire. NB : Notons que ce que j'appelle ici de la « variabilité aléatoire » n'est pas réellement aléatoire. C'est la variabilité qui n'est pas expliquée directement par l'environnement. Cette variabilité peut être due par exemple à la micro-hétérogénéité ou à la stochasticité de l'environnement, aux interactions biotiques, à l'âge des individus, à la diversité génétique (Chapitre 2). Comme on ne connait pas l'ensemble des facteurs qui la gouverne ni leurs influences respectives on ne pourra la représenter ici que sous une forme aléatoire.*

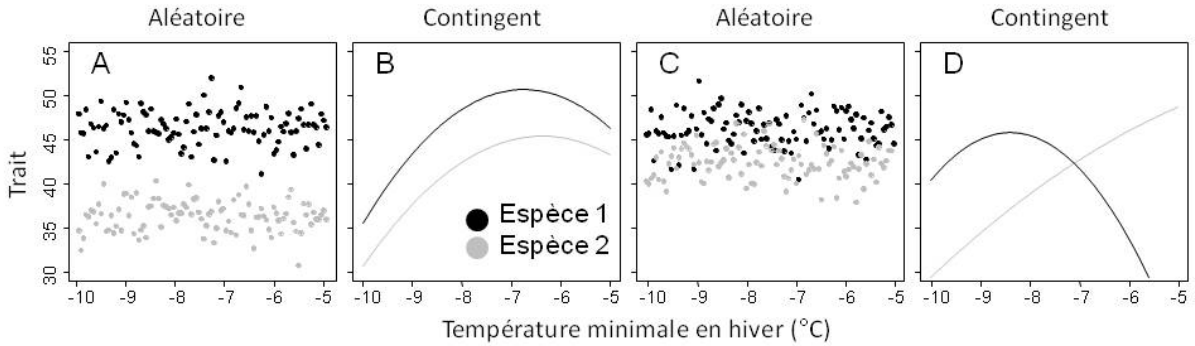


Fig. 4.2. Les différents types de variabilité intraspécifique pour deux espèces.

A. Les deux espèces présentent une variabilité aléatoire, mais leurs distributions ne se chevauchent pas ; B. Les deux espèces présentent une variabilité contingente et les fonctions décrivant le lien entre trait et environnement ne se croisent pas ; C. Les deux espèces présentent une variabilité aléatoire et leurs distributions se chevauchent ; D. Les deux espèces présentent une variabilité contingente et les fonctions décrivant le lien entre trait et environnement se croisent. NB : voir légende de la Fig. 4.1.

1. Pourquoi et comment prendre en compte la variabilité fonctionnelle intraspécifique?

Comme je l'ai déjà évoqué, une grande partie des utilisations des traits fonctionnels se fait à partir de moyennes de traits définies sur quelques individus d'une ou plusieurs populations, entrainant une classification fixe des espèces dans le temps et dans l'espace. Or nous avons montré au Chapitre 2 qu'il existait une variabilité fonctionnelle intraspécifique non négligeable. Nous avons également montré que cette variabilité pouvait entrainer des modifications de classifications lorsque les distances fonctionnelles entre espèces étaient faibles. De plus quelques études récentes ont pointé du doigt l'importance de la variabilité fonctionnelle intraspécifique dans divers domaines de l'écologie. Par exemple concernant le fonctionnement des écosystèmes, Lecerf (2008) a mis en évidence une forte variabilité des traits foliaires d'*Alnus glutinosa* à l'échelle Européenne. Cette variabilité, due aux conditions environnementales très contrastées d'un bout à l'autre de la distribution de l'espèce, entraîne de fortes variations du taux de décomposition des feuilles et par conséquent joue un rôle sur les cycles de l'azote et du carbone. Il y aurait donc autant de variabilité du taux de décomposition des feuilles au sein d'une espèce qu'entre espèces. Il a aussi été montré que la variabilité intraspécifique des traits foliaires pouvait avoir un fort impact sur la résistance à l'herbivorie (Boege & Dirzo 2004). D'autres études concernant l'évaluation de la diversité fonctionnelle (Leps *et al.* 2006; Cianciaruso *et al.* 2009) ou encore la modélisation de la végétation (Kleidon *et al.* 2007; Savage *et al.* 2007 ; Article 1) prônent aussi une prise en compte de la variabilité fonctionnelle intraspécifique. Par conséquent il apparait aujourd'hui essentiel de tester les effets de l'inclusion de cette variabilité dans les modèles écologiques et les études empiriques. Cela pourrait nous permettre d'affiner notre compréhension de la biodiversité et du fonctionnement des écosystèmes notamment pour prédire leur réponse aux changements globaux (McGill *et al.* 2006). Si cela n'a pas été fait jusqu'à présent c'est que (i) les données sur la variabilité fonctionnelle intraspécifique manquent (Chapitre 2) et que (ii) collecter des données permettant de tester son impact dans tous les domaines d'utilisation des traits fonctionnels n'est ni réaliste, ni compatible avec les moyens humains et financiers généralement disponibles (Chapitre 1). Il est cependant possible de tester l'effet de la variabilité fonctionnelle intraspécifique, sans mesurer tout et partout, en faisant de l'écologie virtuelle, c.à.d. en utilisant nos connaissances sur cette variabilité, pour la simuler dans les jeux de données existants et les modèles. Mettre en place des simulations réalistes de sa prise en compte nécessite de disposer d'un cadre conceptuel et quantitatif permettant de déterminer l'ampleur, la structure, les vecteurs de la variabilité fonctionnelle intraspécifique, ainsi que les proportions de variabilité inter- et intraspécifique dans le système étudié. C'est ce que nous avons tenté d'apporter dans les Chapitres 2 et 3. Le cadre quantitatif ne fait pas tout et simuler de la variabilité fonctionnelle intraspécifique dans

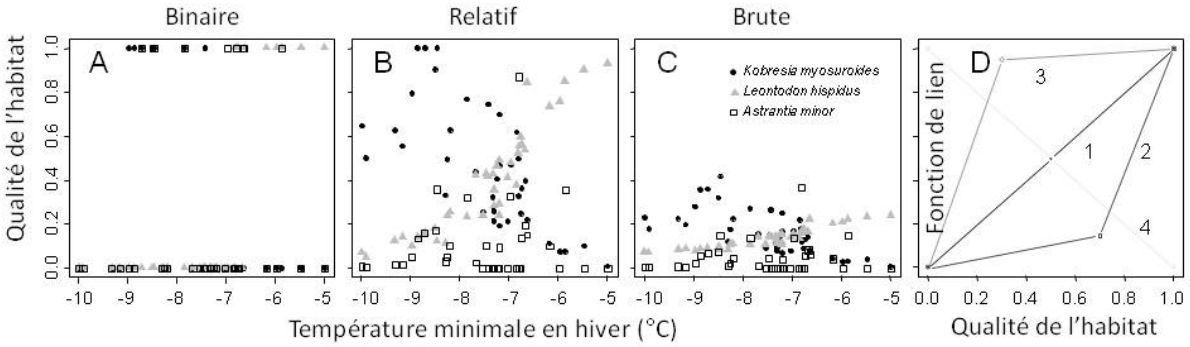


Fig. 4.3. Qualité de l'habitat (issue des modèles d'habitat) et fonctions de lien par rapport à celle-ci.
A. Qualité de l'habitat sous forme binaire après conversion à l'aide d'un seuil. Ici on a pris arbitrairement un seuil correspondant au quantile à 80% des sorties pour chacune des espèces, soit 0.26 pour K. myosuroides, 0.16 pour L. hispidus et 0.06 pour A. minor ; B. Qualité de l'habitat sous forme relative, c'est-à-dire rééchelonnées entre 0 et 1 pour chacune des espèces ; C. Sorties des modèles sous leur forme brute ; D. Les courbes 1 à 4 représentent différentes fonction de lien possible permettant de contraindre des paramètres de modèles ou des valeurs de traits par la qualité de l'habitat.

	Hmax			LDMC		
Espèce	Moyenne	SD	CV	Moyenne	SD	CV
CS	11,2	3,9	0,35	37,5	5,2	0,14
DG	44,3	16,2	0,37	26,7	5,2	0,19
DO	5,5	1,7	0,31	41,2	6,7	0,16
FP	42,7	8,2	0,19	34,1	4,3	0,13
GM	6,1	2,6	0,43	27,2	3,9	0,14
J				48,6	7,4	0,15
LD				44,1	6,7	0,15
LV	25,6	9,1	0,36	14,3	2,1	0,15
P				51,8	4,1	0,08
PV	12,5	6,1	0,49	21,6	2,9	0,13
RF	35,5	13	0,36	46,6	5,5	0,12
SC	12,2	3,8	0,31	36,3	9	0,25
SH	2,4	0,7	0,27	36,3	5,1	0,14
SN	22,5	7,5	0,33	18,3	2,8	0,15
TA	5	2,3	0,46	28,5	2,9	0,1
VM	17,3	5,4	0,31	42	5,3	0,13
Moyenne	18,7	6,2	0,35	34,7	4,9	0,14
SD	14,4	4,6	0,08	11,1	1,9	0,04

Tableau 4.1. Variabilité des traits pour les espèces de la base Guisane 2080 par espèce.
Ici sont donnés les moyennes, écart-type (SD) et coefficients de variation (CV) de Hmax et LDMC pour chacune des 16 espèces de la base Guisane 2080 (Chapitres 1 & 2). Les deux dernières lignes donnent les moyennes et écart-type (SD) de chaque colonne. Le CV résulte ici directement du rapport des deux colonnes précédentes : CV=SD/Moyenne

un jeu de données, par exemple pour étudier l'impact de celle-ci sur des calculs de diversité fonctionnelle, soulève un certain nombre de questions.

1.1. Quoi faire changer ?

Tout d'abord il est essentiel de se demander quels traits fonctionnels ou paramètres démographiques faire varier (dans un modèle on pourra se poser la même question sur les paramètres de celui-ci). Doit-on les faire varier tous ? Tous autant ? Tous de la même manière ? Quels traits sont variables au sein d'une espèce? Les traits liés à la fécondité (ex. nombre de graine) ou au succès d'installation (ex. masse des graines, taux de germination), ceux liés à la croissance (ex. taux de croissance) et à la morphologie (ex. hauteur), ceux liés à la réponse aux perturbations (ex. résistance *vs* tolérance à l'herbivorie) ou encore les paramètres démographiques (ex. taux de croissance de la population) ? Nos résultats (Chapitres 2 et 3) ne permettent de répondre à ces questions que partiellement, les traits fonctionnels que nous avons mesurés sont variables, mais n'ont pas tous la même amplitude de variation; les paramètres démographiques semblent eux aussi variables mais nous n'avons pas trouvé de loi quand à cette variabilité.

1.2. Comment faire pour une espèce ?

Il est ensuite nécessaire de se demander comment inclure la variabilité des traits fonctionnels pour une espèce donnée. Quels sont l'ampleur, la structure, les vecteurs de celle-ci ? Faut-il remplacer les valeurs fixes (Fig. 1.1-A) et uniques par des distributions (décrivant une variabilité aléatoire, Fig. 1.1-B), par une fonction de l'environnement (décrivant la contingence des traits aux conditions environnementales, Fig. 1.1-C), ou par les deux et en quelles proportions (Fig. 1.1-D) ? Dans le dernier cas la variabilité aléatoire au sein des populations est-elle constante dans l'environnement ? Notons que ce que j'appelle ici de la « variabilité aléatoire » n'est pas réellement aléatoire. C'est la variabilité qui n'est pas expliquée directement par les grands gradients environnementaux. Cette variabilité peut être due par exemple à la micro-hétérogénéité ou à la stochasticité de l'environnement, aux interactions biotiques, à l'âge des individus ou encore à la diversité génétique (Chapitre 2). Comme on ne connait pas l'ensemble des facteurs qui la gouverne ni leurs influences respectives on ne pourra la représenter ici que sous une forme aléatoire. Nos résultats donnent de nombreuses indications permettant de répondre à ces questions; ils vont dans le sens d'une combinaison des deux types de variabilité (contingence à l'environnement et variabilité aléatoire au sein des populations), la proportion des deux étant dépendante des traits (Chapitre 2); de plus nous n'avons pas observé de variation de la variabilité le long des gradients (Tables 4.1 et 4.2, analyses non présentées).

Trait	Hmax						LDMC					
Population	Moyenne		SD		CV		Moyenne		SD		CV	
Espèce	moyenne	SD	moyenne	SD	moyenne	SD	moyenne	SD	moyenne	SD	moyenne	SD
CS	11,2	3,2	2,4	0,8	0,22	0,07	37,5	3,4	3,7	2,2	0,1	0,06
DG	44,3	13,8	9,5	3,1	0,22	0,07	26,7	3,1	4,2	1,7	0,16	0,07
DO	5,5	1,4	1	0,4	0,18	0,04	41,2	3,8	3,4	2,8	0,08	0,05
FP	41,1	5,2	6,8	1,8	0,17	0,04	34,1	3,1	3,1	1	0,09	0,03
J	80,5	78,5	32,5	29,7	0,4	0,14	48,5	3,9	3,4	1,4	0,07	0,03
GM	5,9	1,7	2	0,7	0,35	0,1	27,2	2,3	2,9	1,9	0,1	0,06
LD	8,4	3,7	1,6	0,6	0,21	0,07	43,9	3,8	3,1	1,8	0,07	0,04
LV	25,6	6,7	6,4	2,2	0,26	0,08	14,3	1,1	1,8	0,6	0,13	0,05
P	7,6	2	3,2	1,4	0,41	0,11	51,8	1,7	1,7	0,9	0,03	0,02
PV	12,5	5,3	3,1	1,4	0,26	0,09	21,6	2,1	2	0,6	0,09	0,03
RF	35,5	7	10,6	5,5	0,3	0,15	46,6	3,2	2,8	1,4	0,06	0,03
SC	12,2	2,8	2,6	1,2	0,21	0,07	36,3	7,3	5,3	2,5	0,14	0,05
SH	2,6	0,6	0,7	0,5	0,24	0,12	36,1	3,6	2,4	0,8	0,07	0,02
SN	22,5	4,5	6,2	1,8	0,28	0,06	18,3	1,7	2,4	0,6	0,13	0,04
TA	5,3	1,9	1,6	0,7	0,29	0,06	28,5	1,5	2,6	0,6	0,09	0,02
VM	17,3	4,2	3,5	1,4	0,2	0,06	42	2,2	3	1,5	0,07	0,03
Moyenne	**21,13**	**8,91**	**5,86**	**3,33**	**0,26**	**0,08**	**34,66**	**2,99**	**2,99**	**1,39**	**0,09**	**0,04**

Table 4.2. Variabilité des traits pour les espèces de la base Guisane 2080 par population.
Ici sont donnés les moyennes et écart-types (SD) des moyennes, écart-types (SD) et coefficients de variation (CV) obtenus dans les différentes populations mesurées pour chaque espèce pour Hmax et LDMC.

1.3. Comment faire pour plusieurs espèces ?

Lorsque plusieurs espèces sont considérées en même temps, il est enfin indispensable de se demander comment s'organise la variation simultanée des traits des différentes espèces dans l'environnement. Toutes les espèces sont-elles autant variables ? Ont-elles les mêmes proportions de variabilité contingente à l'environnement et aléatoire au sein des populations? Les distributions ou courbes de réponse des traits des différentes espèces peuvent-elles ou doivent-elles se superposer ou se croiser en fonction des conditions environnementales (Fig. 4.2) ? Nous n'avons pas trouvé de patron de variabilité entre les différentes espèces, par rapport à leurs formes de vie ou à leurs exigences environnementales. Par contre, nous avons montré que la classification des espèces dans l'espace des traits pouvait ne pas être conservée dans certains cas et que les traits répondaient certainement aux gradients environnementaux en fonction des exigences environnementales des espèces.

1.4. Quel lien avec les modèles d'habitat ?

Nos résultats suggèrent l'utilisation des modèles d'habitat pour simuler la variabilité contingente à l'environnement. Demeurent alors les questions évoquées au chapitre 3; faut-il prendre en compte les sorties des modèles d'habitat sous une forme binaire (indiquant simplement une présence ou absence de l'espèce et donc du paramètre, Fig. 4.3-A) ? Sous une forme relative (indiquant si les conditions sont plus ou moins favorables, Fig. 4.3-B) s'accompagnant d'un ré-échelonnage des sorties entre 0 et 1 pour l'ensemble des espèces ? Ou encore sous forme brute (Fig. 4.3-C), avec tous les problèmes qui y sont liés (Chapitre 3)? Quelle fonction de lien (Fig. 4.3-D) doit être utilisée pour relier le trait fonctionnel à la qualité de l'habitat?

Dans la suite de ce chapitre, nous avons utilisé le cadre conceptuel et quantitatif développé dans les Chapitres 2 & 3 et ci-dessus afin de tester les effets d'une inclusion de la variabilité fonctionnelle intraspécifique dans deux grands domaines d'utilisation des traits fonctionnels. Dans une première partie je développerais une étude sur la prise en compte la variabilité fonctionnelle intraspécifique dans les calculs d'indices de diversité fonctionnelle, qui est l'une des utilisations majeurs des traits fonctionnels, permettant de lier diversité et fonctionnement des écosystèmes. Dans une deuxième partie je développerais une étude sur la prise en compte de la variabilité fonctionnelle intraspécifique dans un modèle de dynamique du paysage.

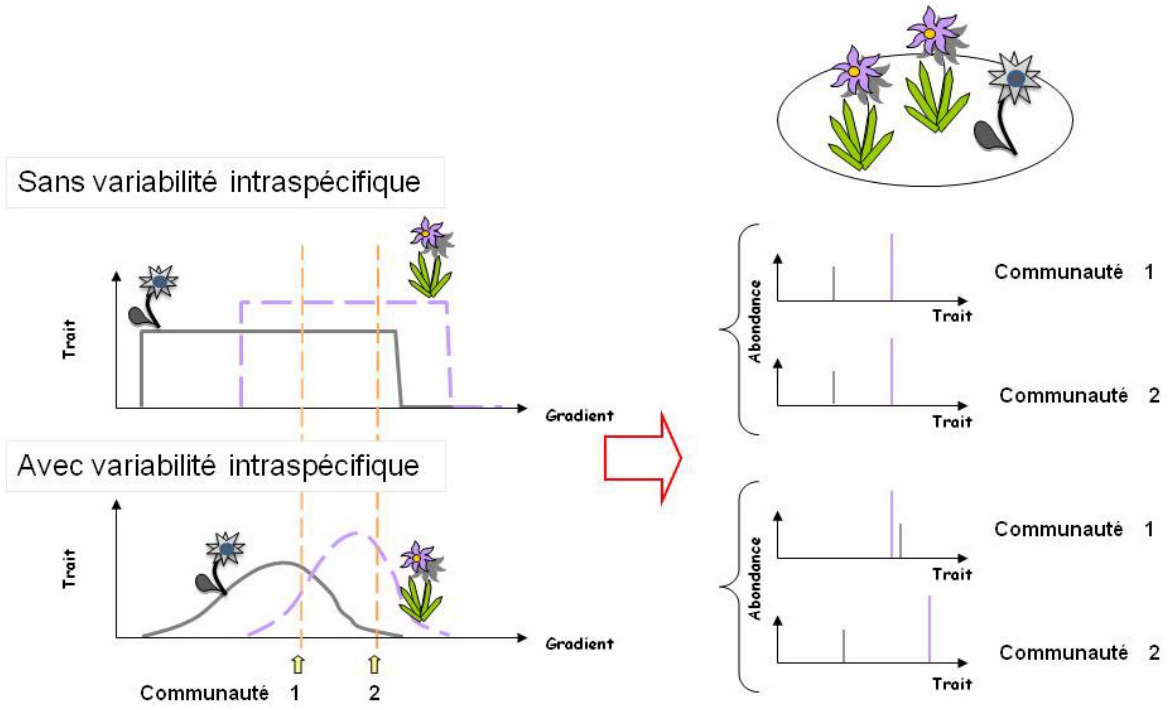


Fig. 4.4. Effet de la prise en compte de la variabilité fonctionnelle intraspécifique sur la diversité fonctionnelle.
Considérons une communauté végétale constituée de deux espèces (mauve et grise), l'espèce mauve étant deux fois plus représentée que la grise. 1. (en haut) Sans variabilité intraspécifique, la composition fonctionnelle de la communauté sera la même dans différentes conditions environnementales. 2. (en bas) Avec de la variabilité intraspécifique contingente à l'environnement, la composition fonctionnelle de la communauté pourra être très différente selon les conditions environnementales dans lesquelles elle se trouve.

2. Implication dans l'évaluation de la diversité fonctionnelle

2.1. Le concept de diversité fonctionnelle

Historiquement, la diversité des systèmes biologiques était donnée par le nombre d'espèces différentes qui les constituaient, toutes les espèces étant considérées comme équivalentes. De nombreuses études se sont attelées à montrer que cette diversité spécifique influait positivement sur les processus écosystémiques tels que la productivité ou la résistance aux invasions (Loreau *et al.* 2001; Hooper *et al.* 2005), mais ces résultats sont restés controversés du fait d'un possible effet confondant de l'identité des espèces, appelé «effet d'échantillonnage» (Aarssen 1997; Hector *et al.* 2002). Le concept de diversité fonctionnelle s'est développé plus récemment, permettant de lier l'environnement, la structure des communautés et les processus écosystémiques tout en considérant l'identité des espèces (Diaz & Cabido 2001; Tilman 2001); de nombreuses études ont mis en évidence un lien entre diversité fonctionnelle et fonctionnement des écosystèmes (Hooper *et al.* 2005 ; Naeem & Wright 2003), en particulier en ce qui concerne les cycles biogéochimiques (Diaz *et al.* 2007a) ou les régimes de perturbations (Lavorel & Garnier 2002). Il est donc essentiel de quantifier la diversité fonctionnelle pour comprendre le lien entre diversité et fonctionnement des écosystèmes, mais aussi pour comprendre la réponse de la diversité végétale aux gradients et aux changements environnementaux (Leps *et al.* 2006). Pourtant il n'existe pas réellement de consensus sur la manière de définir cette diversité fonctionnelle et encore moins sur la manière de l'évaluer (Petchey & Gaston 2006).

Mason *et al.* (2005) ont proposé de définir la diversité fonctionnelle d'une communauté comme la distribution des valeurs de traits et des abondances associées et de l'évaluer selon trois composantes orthogonales: (a) la richesse correspond à l'espace fonctionnel (gamme de trait en une dimension) occupé par la communauté, plus l'espace est grand plus la communauté est riche; (b) la régularité décrit la régularité de la distribution des abondances dans l'espace fonctionnel (la gamme de trait), plus les abondances sont similaires et les valeurs de traits équidistantes dans l'espace fonctionnel plus la régularité est grande; (c) la divergence caractérise la dispersion des valeurs de traits et les abondances associées autour de la valeur moyenne des traits dans la communauté, plus les espèces ayant des valeurs de traits extrêmes sont abondantes plus la divergence est forte (Villéger *et al.* 2008).

Les dernières années ont vu se mettre en place de nombreuses avancées conceptuelles et méthodologiques, la mise en place de critères pour sélectionner des indices efficaces (Mason *et al.* 2003) et le développement d'indices intégrant simultanément plusieurs traits (Petchey & Gaston 2002 ; Botta-Dukat 2005). Cependant de nombreuses questions se posent encore et en particulier il reste à définir quelles méthodes sont adaptées à quelles questions de recherche et dans quel contexte. Faut-il ou non prendre en compte les abondances des

Encadré 4.1. Propriétés des communautés et diversité fonctionnelle

1. Moyenne pondérée d'un trait au niveau de la communauté, avec p_i l'abondance relative de l'espèce i et x_i la valeur moyenne de trait de l'espèce i dans la communauté

$$Moyenne.pondérée = \sum_{i=1}^{S} p_i x_i$$

2. Richesse fonctionnelle relative de la communauté j par rapport à l'ensemble des communautés étudiées , avec Max_j (resp. Min_j) la valeur maximale (resp. minimale) du trait dans la communauté j, et Max (resp. Min) la valeur maximale (resp. minimale) du trait observé dans l'ensemble des communautés étudiées

$$Richesse_j = \frac{Max_j - Min_j}{Max - Min}$$

3. Entropie quadratique de Rao de la communauté, c'est une moyenne des dissimilarités (d_{ij}) entre les espèces i et j , pondérée par les abondances relatives des deux espèces (p_i et p_j)

$$Entropie\ .Rao = \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} d_{ij}\, p_i\, p_j$$

espèces suivant l'hypothèse du ratio de biomasse (Grime 1998; Petchey & Gaston 2006)? Faut-il calculer les distances ou dissimilarités fonctionnelles entre les espèces comme la somme des longueurs des branches d'un dendrogramme établi sur les différences fonctionnelles (Petchey & Gaston 2002), comme la surface de non recouvrement des distributions de traits de deux espèces (Leps *et al.* 2006), ou encore comme une distance Euclidienne ou de Mahalanobis dans l'espace des traits (Botta-Dukat 2005; Podani & Schmera 2006) ? Combien de traits prendre en compte et lesquels (Petchey & Gaston 2006 ; Villéger *et al.* 2008) ? Faut-il prendre en compte la variabilité intraspécifique (Mason *et al.* 2005; Leps *et al.* 2006) et comment (Cianciaruso *et al.* 2009) ?

C'est à ce dernier point que nous nous sommes intéressés. L'analyse suivante propose (1) d'étudier la stabilité des indices de diversité fonctionnelle à la prise en compte de la variabilité fonctionnelle intraspécifique et (2) d'analyser les effets de sa structure et de ses sources (aléatoire, contingente, Fig. 4.1) sur cette stabilité. Evaluer la diversité fonctionnelle des communautés fait en effet partie des nombreuses applications des traits fonctionnels et a été jusqu'à présent essentiellement menée à l'aide de moyennes de traits par espèce, même si la question de la prise en compte de la variabilité intraspécifique a été soulevée à plusieurs reprises (Mason *et al.* 2005 ; mais voir Leps *et al.* 2006 pour une première prise en compte effective). D'une communauté à une autre on s'attend en effet à ce que la diversité fonctionnelle soit modifiée du fait de plusieurs composantes changeant simultanément : la composition spécifique, l'abondance des différentes espèces (quant elle est prise en compte), mais aussi la variabilité des traits de chaque espèce dans l'environnement; deux communautés ayant la même composition en termes d'espèces et d'abondance pouvant ainsi avoir des diversités fonctionnelles très différentes (Fig. 4.4).

2.2. Effet de la variabilité fonctionnelle intraspécifique, de sa structure et de ses sources sur la stabilité des indices de diversité fonctionnelle

2.2.1. Principe

Encore une fois nous avons utilisé une expérience semi-virtuelle (Chapitre 1). L'étude a été menée sur les 41 communautés (prairies et landes, 115 espèces) de la vallée de la Guisane décrites au Chapitre 1 (voir 2.5.2.4). Ces communautés se situent entre 1800 et 2800 m et comportent entre 22 et 50 espèces. Comme nous avons vu au Chapitre 2 que les différents traits n'avaient ni la même ampleur, ni la même structure de variabilité, nous nous sommes restreints à des indices calculés sur un unique trait et nous avons testé deux traits différents: Hmax et LDMC (Chapitre 1, 2.5). Nous avons calculé trois indices décrivant les propriétés et la diversité des communautés (Encadré 4.1): (i) une moyenne des valeurs de traits de la communauté, pondérée par les abondances relatives de chaque espèce (CWM pour «community weighted mean», Violle *et al.* 2007). Elle représente la valeur de trait moyen

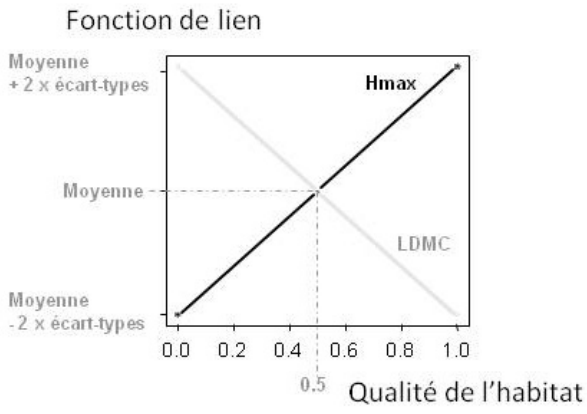
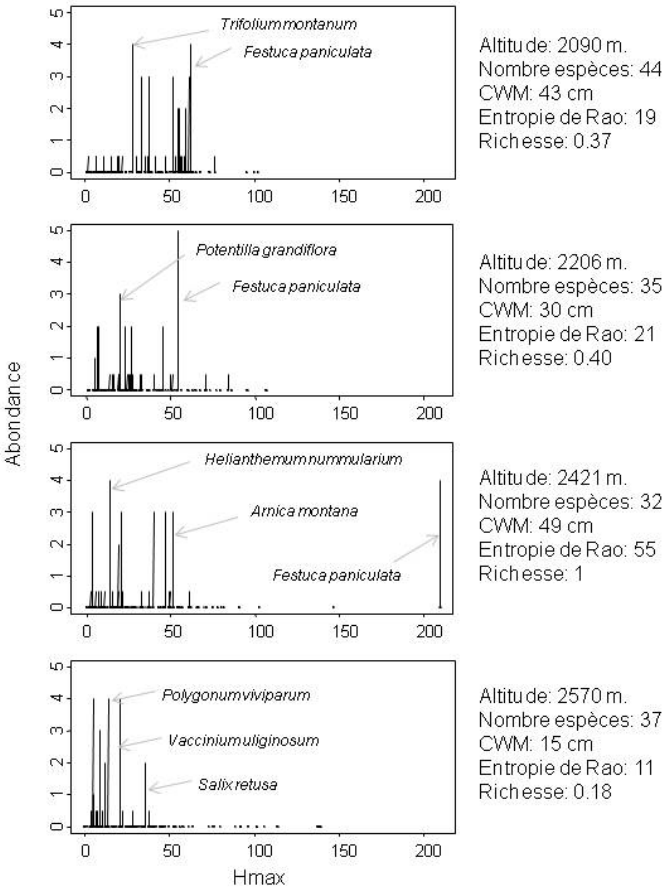


Fig. 4.5. Fonctions de lien choisies pour Hmax et LDMC dans le cas d'une variabilité intraspécifique de type « Niche ».
Conformément aux résultats des Chapitres 2 & 3, Hmax (resp. LDMC) suit une fonction de lien linéaire croissante (resp. décroissante), donnant des valeurs de Hmax maximales (resp. LDMC minimales) dans les zones d'habitat favorable et des valeurs de Hmax minimales (resp. LDMC maximales) dans les zones d'habitat peu favorable. Les valeurs de Hmax et LDMC varient de deux écart-types autour de la moyenne pour que les résultats soient comparables au cas d'une variabilité de type «Identique».

Fig. 4.6. Exemples de distributions de Hmax obtenues pour quatre communautés avec une variabilité intraspécifique de type « Identique» et un CV de 0.45 (la plus grande variabilité) .
Pour chacune des communautés sont donnés son altitude, le nombre d'espèces qu'elle contient, sa moyenne de trait pondérée (CWM), son entropie de Rao et sa richesse. Dans chacune on a indiqué le nom des espèces les plus abondantes. Dans les deux communautés du haut, F. paniculata présente une hauteur de l'ordre de 60 cm, dans la troisième communauté en partant du haut, cette même espèce présente une hauteur de plus de 200 cm, conduisant à de fortes valeurs des indices de diversité (CWM, Entropie de Rao et Richesse). Il est à noter qu'une telle hauteur pour cette espèce est peu réaliste, mais nous avons utilisé ce cas de variabilité extrême (CV=0.45), assez peu réaliste au vu des données de terrain (CV moyen observé de 0.35) dans un but d'illustration.



attendue lors d'un échantillonnage aléatoire dans cette communauté (Leps *et al.* 2006). Selon l'hypothèse du ratio de biomasse (Grime 1998) –qui propose que le fonctionnement des écosystèmes est essentiellement déterminé par les espèces dominantes et leurs caractéristiques- les processus écosystémiques devraient être prédictibles à l'aide des CWM de traits (Garnier *et al.* 2004; Hooper *et al.* 2005; Garnier *et al.* 2007). Par exemple dans les communautés herbacées, la CWM de la teneur des feuilles en azote est fortement corrélée à la productivité primaire (Garnier *et al.* 2004). (ii) un indice de richesse, donné par la gamme relative de trait de chaque communautés par rapport à l'ensemble des communautés étudiées (Mason *et al.* 2005). (iii) l'entropie quadratique de Rao (Rao 1982; Botta-Dukat 2005) qui caractérise la dissimilarité fonctionnelle moyenne des espèces de la communauté ; elle représente la divergence de la communauté.

A partir de ces données réelles, nous avons fait varier « artificiellement » les valeurs de traits autour des moyennes des espèces, les présences et abondances des espèces étant données par les observations *in situ* et variant donc « naturellement » entre les différentes communautés. Nous avons simulé de la variabilité intraspécifique pour les 115 espèces présentes dans les communautés selon trois structures: (a) *Identique*: toutes les espèces ont le même coefficient de variation (CV), celui-ci variant par pas de 0,05 entre 0 et 0,45 (Fig. 2.6). La valeur moyenne de trait pour l'espèce i dans la communauté j est donc une réalisation de la loi normale de moyenne *trait moyen de l'espèce i* et d'écart-type CV x *trait moyen de l'espèce i* (Fig. 4.6). (b) *Espèce*: chaque espèce a son propre coefficient de variation, invariant entre les communautés (CVi). Ces coefficients de variation sont établis à partir des données du Chapitre 2 (Tableau 4.1), ils suivent pour Hmax une loi normale de moyenne 0,35 et d'écart type 0,08, se retrouvant donc entre 0,2 et 0,55 ; et pour LDMC ils suivent une loi normale de moyenne 0,14 et d'écart-type 0,04, se situant donc entre 0,05 et 0,22. La valeur moyenne de trait pour l'espèce *i* dans la communauté *j* est donc une réalisation de la loi normale de moyenne *trait moyen de l'espèce i* et d'écart-type CVi x *trait moyen de l'espèce i*. (c) *Niche*: l'ampleur de la variabilité est donnée par un coefficient de variation unique pour toutes les espèces (CV) variant par pas de 0,05 entre 0 et 0,45 et les valeurs de traits sont contraintes par les modèles d'habitat (HS) projetés pour chaque espèce dans chaque communauté (Article 6). En accord avec les résultats des Chapitres 2 et 3, Hmax (resp. LDMC) suit une fonction monotone croissante (resp. décroissante) de HS, la valeur moyenne du trait étant obtenue pour HS=0,5. Nous avons choisi des fonctions de lien linéaires (Fig. 4.5) et nous avons déterminé les paramètres de celles-ci (pente et ordonnée à l'origine) de manière à avoir des variabilités similaires au cas (a) à CV donné ; c.à.d. des valeurs étalées jusqu'à plus ou moins deux écart-types autour de la moyenne de l'espèce (Fig. 4.5). La valeur moyenne de trait pour l'espèce i dans la communauté j ($Trait_{ij}$) est donc donnée par :

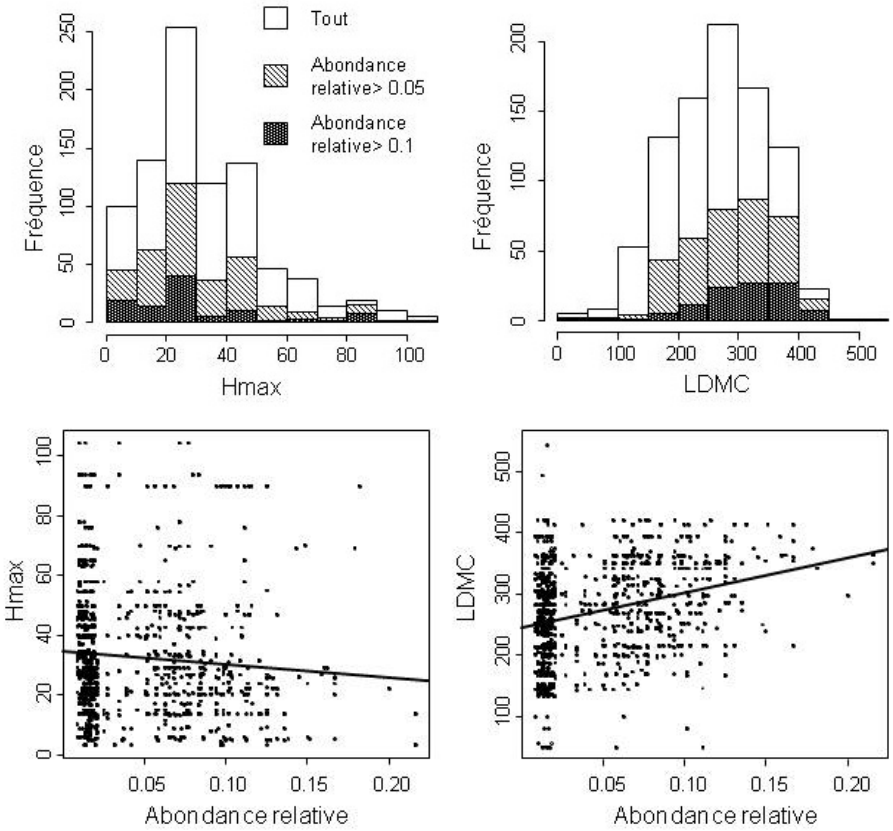


Fig. 4.7. Distributions de Hmax et LDMC dans les communautés.
(En haut) Les histogrammes blancs donnent les distributions des deux traits dans l'ensemble des communautés, chaque espèce étant comptée autant de fois qu'elle est présente au sein de celles-ci. En gris clair, on a limité le comptage aux espèces ayant une abondance supérieure à 0,05 (40% des occurrences) et en gris clair à celles ayant une abondance supérieure à 0,1 (10% des occurrences). (En bas) Relations entre les traits et les abondances relatives de chaque espèce dans toutes les communautés. Les droites représentent les régressions linéaire entre les deux, mais ces relations n'ont aucun pouvoir explicatif (R^2=0), elles ne font qu'indiquer des tendances : décroissante pour Hmax, les espèces les plus abondantes sont les plus petites et croissante pour LDMC, les espèces les plus abondantes ont les plus fortes LDMC.

$Trait_{ij} = Trait.moyen.espèce_i \times (1 - 2 \times CV(1 + 2 \times HS_{ij}))$, avec HS_{ij} la valeur du modèle d'habitat de l'espèce i dans la communauté j (Fig. 4.5).

Pour les cas (a) et (b), cent simulations ont été effectuées. Afin de mesurer l'effet de la variabilité intraspécifique simulée, nous avons comparé les indices calculés dans chaque cas à ceux calculés sans variabilité (appelés *indices moyens*) ; le R^2 de la régression linéaire entre les indices calculés avec la variabilité et les *indices moyens* donne le pourcentage d'information donné par les *indices moyens* par rapport aux indices calculés avec la variabilité. Les abondances des espèces dans chaque communauté sont celles enregistrées sur le terrain (Chapitre 1) et n'ont donc pas été modifiées dans les simulations.

2.2.2. Résultats

Afin de s'assurer que le calcul des indices de diversité fonctionnelle n'est pas biaisé par l'utilisation d'abondances relatives des espèces en classes de Braun-Blanquet (1952) nous avons vérifié que les indices mesurés étaient similaires à ceux calculés avec les médianes de classes de couverture (+: 1%, 1: 5%, 2: 17%, 3: 37%, 4: 62%, 5: 87%) : R^2=0,96 entre CWM avec les classes vs avec les médianes.

Au sein des communautés, la moyenne pondérée (CWM) pour Hmax avait tendance à décroître avec l'altitude (r=-0,5) et la moyenne pondérée pour LDMC avait tendance à augmenter avec l'altitude (r=0,24). Les deux traits n'avaient pas la même distribution, avec une plus grande représentation des espèces ayant une petite Hmax et une distribution de LDMC centrée (Fig. 4.7). De plus les deux traits n'avaient pas la même répartition dans les communautés: les espèces dominantes avaient plutôt des LDMC fortes et elles avaient également plutôt des Hmax basses (Fig. 4.7), ce qui est une caractéristique des communautés végétales de montagne.

Les trois indices de diversité fonctionnelle se sont révélés sensibles à la prise en compte d'une variabilité fonctionnelle intraspécifique avec un éloignement par rapport aux *indices moyens* d'autant plus grand que la variabilité incluse (CV) était forte (Fig. 4.8 & 4.9). Les indices calculés sur les deux traits n'avaient pas la même sensibilité à la variabilité intraspécifique. Les indices calculés pour LDMC étaient beaucoup plus sensibles (R^2 descendant jusqu'à 0,1 pour les moyennes pondérées) que ceux calculés pour Hmax (R^2 descendant à 0,8 pour les moyennes pondérées). Cela peut s'interpréter à l'aide des résultats ci-dessus sur les distributions et répartitions des traits. Les espèces dominantes ayant des LDMC fortes, leurs valeurs de LDMC étaient très variables entre les communautés, puisque la variabilité de chaque espèce était une fonction de sa moyenne de trait (moyenne x CV). Cela a entrainé une forte modification des distributions de valeurs de LDMC au sein des communautés et donc des indices. Au contraire les espèces dominantes ayant des Hmax

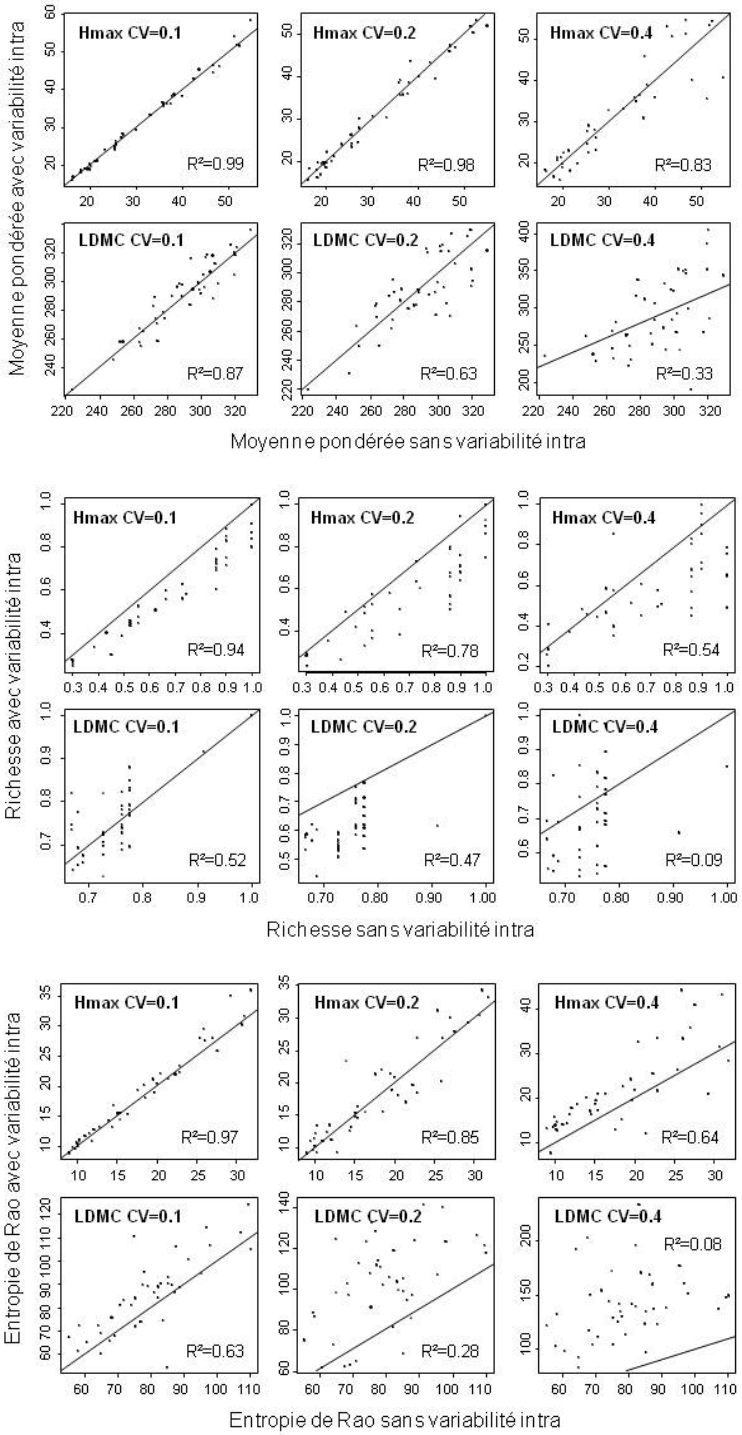


Fig. 4.8. Indices de diversité fonctionnelle calculés avec et sans variabilité fonctionnelle intraspécifique (type «Identique »). *(En haut) les moyennes des traits pondérées, (Au milieu) la richesse des communautés, (En bas) l'Entropie de Rao. Chacun des indices est calculé pour Hmax (ligne du haut) et pour LDMC (ligne du bas) avec des variabilités intraspécifiques correspondant à des CV de 0.1, 0.2 et 0.4 (de gauche à droite). Dans chaque graphique, chacun des points représente une communauté ; ces graphiques donnent les indices calculés avec de la variabilité intraspécifique en fonction des indices calculés sans cette variabilité, c'est-à-dire sur les traits moyens. La droite représente la bissectrice du plan. Le R² est celui de la régression linéaire entre les indices avec et sans variabilité.*

faibles, elles étaient peu variables en valeur absolue, entrainant de faibles modifications des distributions de Hmax et donc des indices.

Les trois indices se sont comportés de manière assez similaire face à la prise en compte de la variabilité intraspécifique en termes d'éloignement aux *indices moyens* (Fig. 4.9); le caractère informatif des *indices moyens* diminuant légèrement moins vite face à l'augmentation de la variabilité pour les moyennes pondérées que pour les deux autres indices, sauf pour la variabilité de type *Niche*. Il est par contre à noter une différence intéressante entre ces trois indices (Fig. 4.8): les moyennes pondérées avaient tendance à se disperser autour des valeurs des *indices moyens*, les entropies de Rao avaient tendance à être plus grands que les *indices moyens* avec la variabilité et les indices de richesse avaient tendance à être plus petits que les indices *moyens* avec la variabilité (Encadré 4.2).

Nos résultats indiquent également que la structure de la variabilité pourrait avoir un rôle dans le caractère informatif des *indices moyens*. En effet il n'y a pas de différence significative pour Hmax entre les indices calculés selon la méthode *Identique* (CV=0,35) et la méthode *Espèce*. Par contre pour LDMC les R^2 sont plus forts ($p<0,05$; Fig. 4.9) pour la moyenne pondérée et l'entropie de Rao selon la méthode *Espèce* que selon la méthode *Identique*. Pour ce qui est de la méthode *Niche*, elle donne des R^2 plus bas (en dessous du quantile 0,1) que la méthode identique, sauf pour Rao sur Hmax et richesse sur LDMC (Fig. 4.9).

Sur le terrain, nous avons trouvé des CV moyens de 0,14 pour LDMC et 0,35 pour Hmax (Tableau 4.1). Dans cette tranche de variabilité, les moyennes pondérées de Hmax présentent avec les trois structures différentes des R^2 supérieurs à 0,8, ce qui veut dire que les *indices moyens* apportent déjà en eux-mêmes plus de 80% de l'information sur la diversité fonctionnelle (Fig. 4.9). Les moyennes pondérées de LDMC montrent elles des R^2 entre 0,6 et 0,9 (0,75 en moyenne) et seulement 0,35 pour la structure *Niche*. Par contre pour les deux autres indices, les R^2 restent autour de 0,8 pour Hmax, mais sont en dessous de 0,65 pour LDMC (Fig. 4.9).

2.2.3. Discussion & perspectives

Nos résultats montrent qu'il y a un fort effet de l'ampleur de variabilité intraspécifique et de sa structure sur le calcul des indices de diversité fonctionnelle. Evaluer l'importance de cet effet, c'est-à dire décider que l'on peut se contenter des *indices moyens* et ne pas prendre en compte la variabilité intraspécifique ne peut pas se décider *a priori*. Cela dépend du trait étudié (distribution, répartition par rapport aux abondances), de sa variabilité, de la structure de celle-ci et de l'indice calculé et du but de l'analyse. Par exemple, dans notre cas en admettant que 80% d'information nous suffisent, on pourrait négliger la variabilité intraspécifique dans les calculs de moyennes pondérées sur Hmax (CWM), mais pas dans les autres calculs.

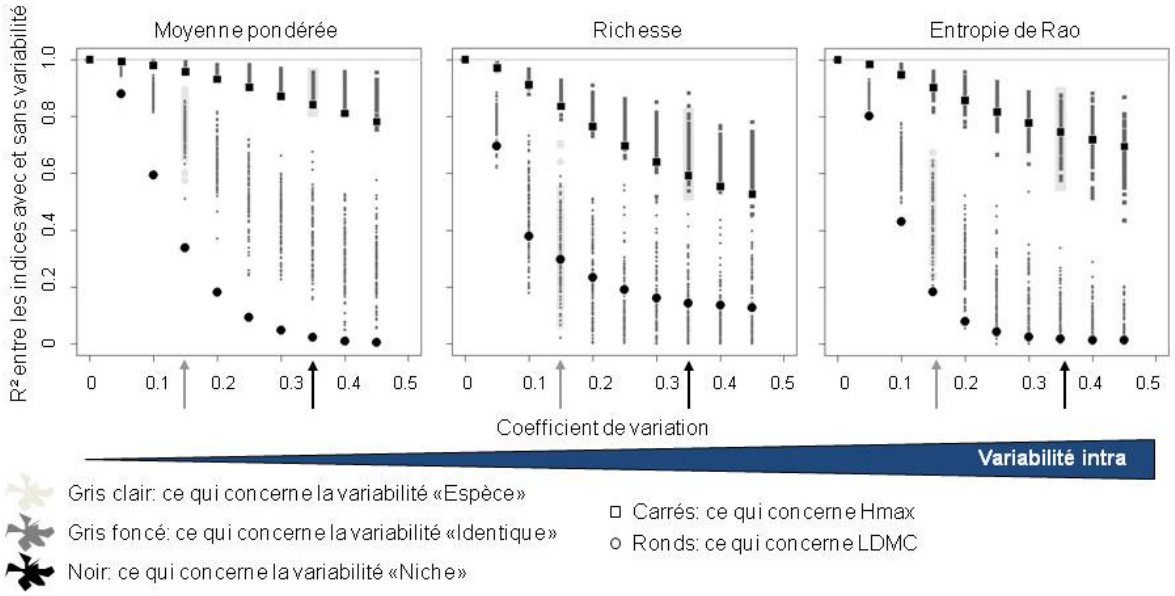


Fig. 4.9. Effet de la variabilité intraspécifique sur le calcul des indices de diversité fonctionnelle.
Ici sont présentés les résultats pour les trois indices, de gauche à droite : les moyennes de trait pondérées, la richesse fonctionnelle et l'entropie de Rao et pour les deux traits : Hmax les carrés et LDMC les ronds. Chaque graphique donne les R² des régressions linéaires entre les indices calculés avec de la variabilité intraspécifique et sans en fonction de la quantité de variabilité intraspécifique caractérisée par les coefficients de variation (CV). Le R² de la régression linéaire entre les indices calculés avec la variabilité et les « indices moyens » donne le pourcentage d'information donné par les « indices moyens » par rapport aux indices calculés avec la variabilité.
En Noir sont donnés les résultats correspondant à la variabilité intraspécifique de type « Niche », il y a un unique calcul par niveau de variabilité. En gris foncé sont donnés les résultats correspondant à la variabilité de type « Identique» pour lequel toutes les espèces ont la même variabilité : on a fait dans ce cas 100 simulations pour chaque niveau de variabilité. En gris clair sont donnés les résultats correspondant à la variabilité de type « Espèce », chaque espèce ayant son propre niveau de variabilité, donné par une loi normale de moyenne 0,35 pour Hmax et 0,15 pour LDMC et d'écart-type 0,02 pour Hmax et 0,03 pour LDMC, correspondant aux valeurs observées sur le terrain. Les flèches grises (resp. noires) au niveau des axes des abscisses indiquent 0,15 (resp. 0,35), le CV moyen observé pour LDMC (resp. Hmax) sur le terrain.

Nous avons bien fait ici de nous contenter d'indices calculés à partir d'un unique trait. En effet les deux traits étudiés donnent des résultats très différents et leur combinaison résulterait probablement en un résultat également différent. De plus tester une telle combinaison de traits demanderait dans notre cas de combiner deux traits qui ne varient pas de la même manière *in situ* (CV de 0,14 et 0,35, Table 4.1), ce qui risque de compliquer les calculs et les interprétations.
Il est intéressant de noter que la structure *Niche* donne les plus mauvais résultats (R^2 faibles, Fig. 4.9). Cela pourrait être du à une interaction entre les effets des changements de composition spécifique (changements naturels de présence et d'abondance) entre les communautés et les effets de la variabilité (artificielle) des traits. En effet dans la structure *Niche* la valeur des traits suit la qualité de l'habitat des espèces (ex. Hmax plus faible dans un habitat moins adéquate), c.à.d. varie de manière systématique et non plus aléatoire comme dans les autres cas ; or la distribution et l'abondance des espèces dans les communautés pourraient aussi dépendre de cette qualité de l'habitat (Chapitre 3). Ainsi la variation simultanée de ces deux grandeurs (variation naturelle de l'abondance et artificielle des traits) le long des gradients environnementaux pourrait être responsable de la forte baisse du pouvoir informatif des *indices moyens*, cela reste à tester. De plus l'effet de cette structure *Niche* dépend de la fonction de lien (résultats non montrés), il faudrait donc tester différentes possibilités.

Cette étude va à l'encontre des résultats présentés par Cianciaruso *et al.* (2009) qui annoncent un effet très faible de la variabilité intraspécifique sur le calcul du FD de Petchey et Gaston (2002). Une partie des résultats présentés dans cette étude (Cianciaruso *et al.* 2009) sont émis à partir de communautés purement virtuelles et de distributions virtuelles de traits standardisées (loi normales de moyenne 0 et d'écart-type zéro), tous les traits variant aléatoirement avec la même amplitude (CV fixé). Ces choix sont peu réalistes et moins propices à ce genre de simulations que des communautés et des distributions de traits réelles, en particulier à cause des interactions qui peuvent exister comme nous l'avons dit entre les changements dans la composition des espèces, les abondances et la variabilité des traits d'une communauté à une autre. De plus il est à noter que les coefficients de variations qu'ils utilisent se situent entre 0 et 5, alors que nous n'avons observé sur le terrain des coefficients de variation qu'entre 0 et 0,5 tous traits confondus. Cela montre qu'un cadre quantitatif comme celui que nous avons mis au point dans les Chapitres 2 et 3 est véritablement nécessaire à la mise en place de telles simulations, afin que celles-ci soient réalistes et utiles en termes d'aide à la décision (la variabilité intraspécifique a-t-elle ou non un effet dans mon cas d'étude ?). Il faut cependant noter que l'indice FD est très différent de ceux que nous avons calculés ici et qu'il serait intéressant de reproduire notre analyse sur cet indice particulier, c'est en cours.

Encadré 4.2. Effet de la variabilité intraspécifique sur la dispersion des valeurs de traits

Supposons que la variabilité intraspécifique est aléatoire (Fig. 4.1) et identique pour toutes les espèces et que la valeur de trait d'une espèce dans une communauté est une réalisation de la loi normale de moyenne M égale à la moyenne de trait de l'espèce et d'écart-type SD=CVxM, où CV est la variabilité intraspécifique du trait. Observons ce qui se passe à l'extrémité haute de la distribution des valeurs de trait dans une communauté.

a. Une seule espèce très différente des autres, l'espèce noire (cas de gauche) de paramètres M1 & CV, est à l'extrémité de la distribution. Il y a alors une chance sur deux que la valeur de trait observée dans la communauté (Tmax) soit plus grande que la valeur moyenne de trait M1: Pr(Tmax>M1)=0.5.

b. Deux espèces se situent à l'extrémité de la distribution (cas du milieu et de droite), l'espèce noire et l'espèce grise de paramètres M2 et CV. Dans une loi normale Pr(M+/-4xSD)~1. Pour avoir plus d'une chance sur deux d'avoir Tmax>M1, il faut qu'une partie de la distribution de l'espèce grise soit plus grande que M1, c.à.d. qu'il faut 4xCVxM2>|M1-M2| (1). On pose Δ=|M1-M2|, la condition (1) devient: 4xCVxM1-4xCVxΔ>Δ, puis 4xCVxM1/(1+4xCV)>Δ.
Pour Hmax, avec un CV de 0.1 on remplit la condition (1) dans 16 des 41 communautés, pour un CV de 0.2 on la remplit pour 27 communautés et pour un CV de 0.3 pour 32 d'entre elles.
Pour LDMC, avec un CV de 0.1 on remplit la condition (1) dans 35 des 41 communautés, pour une CV de 0.2 on la remplit pour 40 communautés et pour un CV de 0.3 pour toutes.

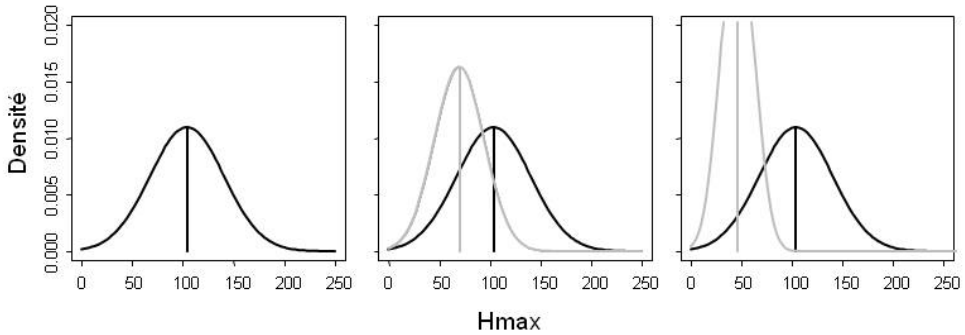
- Lorsqu'il y a de la variabilité intraspécifique, la plupart des communautés étudiées sont donc telles que Pr(Tmax>M1)>0.5.

- Ce raisonnement tient pour l'extrémité inférieure de la distribution. Pour Hmax (resp LDMC) avec un CV de 0.3 seulement 14 (resp 28) des communautés valident la condition (1) et donc Pr(Tmin<m1)>0.5.

- Pour 32 des 42 comm, des CV de 0.35 pour Hmax et 0.15 pour LDMC, correspondant aux moyennes de CV observés *in situ* suffisent à remplir la condition (1) aux deux extrémités de la distribution.

- Lorsque l'on observe la distribution de l'ensemble des traits, toutes communautés confondues, pour Hmax il suffit d'avoir CV=0.11 (resp. CV=0.37) pour que la condition (1) soit validée à l'extrémité haute (resp. basse) et pour LDMC il suffit d'avoir CV =0.09 (resp. CV=0.15) pour que la condition (1) soit validée à l'extrémité haute (resp. basse). Par rapport aux valeurs de CV observée *in situ*, cela implique que la condition (1) est toujours remplie. Il y a donc plus d'une chance sur deux que l'amplitude de la distribution augmente à chaque extrémité quand on introduit la variabilité intraspécifique.

- Au final dans la plupart des communautés, la Richesse absolue calculée avec de la variabilité intra (Tmax-Tmin) est plus grande que la Richesse absolue calculée sans (M1-m1). Cela entraine: (1) une dissimilarité moyenne entre les espèces plus grande avec de la variabilité intra que sans et donc une augmentation de l'entropie de Rao quand on prend en compte la variabilité intra; (2) une Richesse absolue qui augmente avec la variabilité intra, mais moins et moins souvent que l'amplitude de la distribution de traits sur l'ensemble des communautés; et donc la Richesse relative des communautés qui est le rapport entre la Richesse absolue et l'amplitude de la distribution totale diminue quand on prend en compte la variabilité intra



Dans l'analyse présentée ci-dessus, nous n'avons pris en compte que de la variabilité entre populations (aléatoire dans le cas *Identique* ou contingente dans le cas *Niche*, Fig. 4.1), mais pas de variabilité au sein des populations. Comme nous avons vu qu'il existait une forte variabilité des traits au sein des populations (Chapitre 2), il serait intéressant de la prendre en compte dans les calculs d'indices de diversité fonctionnelle ; cela nécessiterait cependant des calculs plus complexes et plus nombreux; en effet dans ce cas on ne peut plus se contenter de calculer des distances (ex. Euclidienne) entre les moyennes des espèces dans chaque condition pour évaluer par exemple l'entropie de Rao, il faut utiliser la méthode des « recouvrements » proposée par Leps *et al.* (2006) pour calculer les dissimilarités entre espèces ; de plus il faudrait tester différents niveaux de variabilité au sein des populations (*in situ:* moyenne des CV par population entre 0,15 et 0,45 pour Hmax et 0,05 et 0,16 pour LDMC selon les espèces) pour chaque niveau de variabilité entre populations et chaque type de structure. Cela est une prochaine étape indispensable pour comprendre l'importance de la variabilité intraspécifique dans les utilisations des traits fonctionnels.

3. Implication dans la modélisation de la dynamique d'un paysage

3.1. Vers des modèles « hybrides »

Il existe aujourd'hui de nombreux modèles permettant d'étudier la dynamique de la végétation. Ces modèles sont implémentés à l'échelle des individus (ex. modèles basés sur les individus, DeAngelis & Mooij 2005), des populations (Jeltsch *et al.* 2008), des peuplements forestiers (ex. modèles de trouée, Bugmann 2001), des paysages (Baker 1985), des régions (ex. modèles d'habitat, Guisan & Thuiller 2005), des biomes ou de la planète entière (ex. DGVM, Sitch *et al.* 2003). Tous sont des simplifications volontaires et orientées de la réalité, c'est-à-dire que dans chacun l'accent est mis sur un (ou quelques) processus ou patron d'intérêt qui est décrit en détail, tandis que les autres sont simplifiés au maximum ou exclus. Par exemple, les modèles d'habitat (Chapitre 3) décrivent des patrons de distribution sans prendre en compte explicitement les processus biologiques liés à la dynamique des populations, la dispersion ou la compétition. Ces modèles sont donc simplistes et critiquables (Heikkinen *et al.* 2006) mais permettent de caractériser la qualité de l'habitat pour de très nombreuses espèces à de larges échelles spatiales (Thuiller *et al.* 2005). En contrepartie, les modèles de paysage (Baker 1985) décrivent le plus souvent les grands types de végétation ou groupes fonctionnels (PFT) qui structurent le paysage. Ils s'attachent à prendre en compte explicitement les processus de dispersion et de perturbation tout en incluant des représentations simplifiées de la démographie et des interactions biotiques. Ils supposent que la qualité de l'habitat constante est influençable par les perturbations et la charge utile, mais non directement par le climat.

L'une des solutions envisagée aujourd'hui pour améliorer la qualité des prédictions faites en modélisation de la végétation est de coupler différents modèles déjà existant afin de tirer le meilleur parti des points forts de chaque modèle et de pallier au mieux leurs limites respectives (Keith *et al.* 2008 ; Article 1, Article 2, Article 11). Cela permettrait en particulier de combiner des processus n'ayant pas lieu à la même échelle.

Envisager une telle hybridation pourrait en particulier permettre d'inclure de la variabilité intraspécifique et de pallier au fait que nombre de modèles décrivent encore la végétation sous forme de PFT grossiers (ex. herbacées, arbres décidus, arbres sempervirents) et uniformes, tous les individus, cohortes, populations d'un PFT donné étant considérés comme similaires (même fécondité, stature, croissance, taux de recrutement…). Dans certains modèles (Treemig, modèle de trouée, Lischke *et al.* 2006), la croissance des individus ou cohortes est déjà affectée par la qualité de l'habitat à travers une fonction reliant la croissance potentielle à trois variables directes (disponibilité en eau, lumière, énergie thermique, Liebig's "Law of the Minimum"). Cependant paramétrer ces courbes de croissance potentielle nécessite une grande quantité de données de terrain (données de croissance dans différentes conditions) et ne peut se faire que sur les espèces arborées. Utiliser des modèles d'habitat pour décrire la qualité de l'habitat et contraindre la performance des espèces, en accord avec les résultats du Chapitre 3, pourrait permettre d'inclure facilement (techniquement parlant) de la variabilité intraspécifique dans les modèles de paysage pour de nombreuses espèces ou PFT. Cela est d'autant plus intéressant à tester que trouver une manière plus réaliste (que des traits moyens) de paramétrer espèces et groupes fonctionnels est aujourd'hui au cœur d'un débat actif (Lavorel *et al.* 2007). Ces avancées sont désormais la clé qui permettra de lier la dynamique des individus, espèces, groupes fonctionnels et communautés végétales au fonctionnement des écosystèmes (ex. cycles biogéochimiques, boucles de rétroaction sol-biosphère-atmosphère, Suding *et al.* 2008).

3.2. L'exemple de Lamos/Biomove

Le workshop intitulé "*Simulating the dynamic response of species to environmental change? Lamos, a hybrid platform*", organisé par W. Thuiller avec entre autres I. Davies (Australian National University, Australie), G. Midgley (South African National Biodiversity Institute, Afrique du Sud) & L. Hannah (Conservation International, Etats-Unis), s'est tenu a Grenoble du 4 au 8 Février 2008. Au cours de celui-ci j'ai participé à la finalisation des aspects conceptuels d'un couplage entre un modèle de dynamique de succession végétale (FATE) et de qualité de l'habitat (Modèle d'habitat) nommé Biomove. Le développement de ce modèle hybride en tant que tel est mené par Ian Davies et est implémenté au sein de la plateforme de modélisation LAMOS. L'Article 5 est issu de ce workshop.

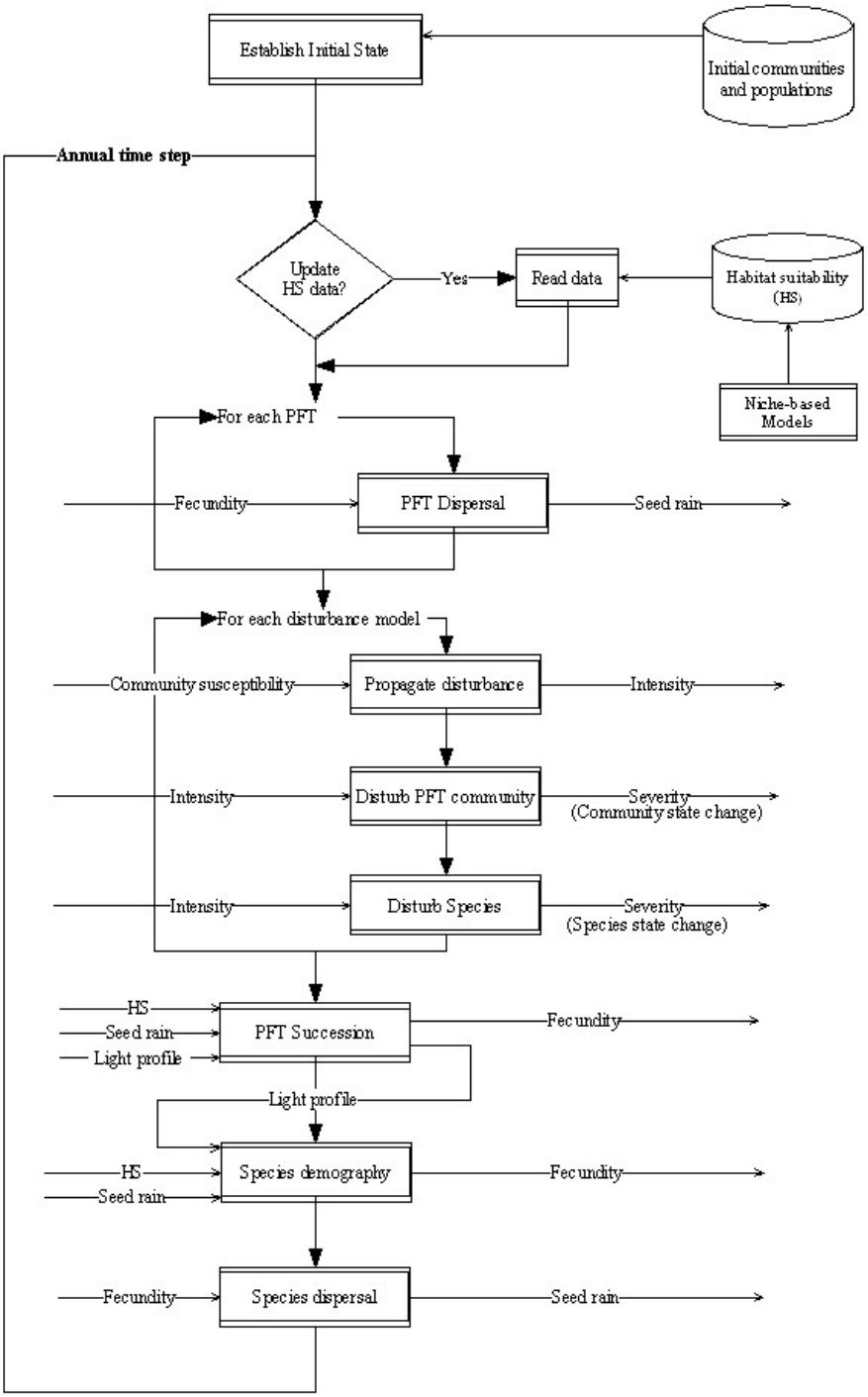


Fig. 4.10. Diagramme des principales routines de Biomove.
Les flèches verticales indiquent l'ordre d'appel des différentes routines et les flèches horizontales indiquent les principales entrées et sorties de chacune de ces routines. (Figure issue de l'Article 5)

3.2.1. Lamos, une plateforme de modélisation

Lamos (pour « LAndscape MOdelling Shell ») est une plateforme de modélisation spatialement et temporellement explicite qui inclut les processus basiques de la dynamique de végétation résultant des interactions entre PFT ou espèces, des perturbations et des patrons spatiaux (Lavorel 2001, Fig. 4.10). Cette plateforme propose de décrire la végétation sous forme de PFT (ou espèces structurantes) et son utilisation suppose que les conditions abiotiques sont adéquates pour chaque PFT implémenté dans la zone étudiée (ex : pas d'influence du climat). Au sein de Lamos, le paysage est décrit comme une grille de pixels, dans laquelle la dynamique de la végétation est modélisée par trois modules interdépendants (Fig. 4.10): (1) un modèle de succession gouverne la dynamique de la succession végétale à l'intérieur de chaque pixel et à chaque pas de temps ; (2) un modèle de dispersion simule la dispersion des graines à travers le paysage. Le patron de distribution des graines peut suivre (i) une loi exponentielle négative non directionnelle par rapport au pixel d'origine de la graine, (ii) une distribution locale, les graines restant dans leur pixel d'origine, ou (iii) une dispersion uniforme dans le paysage. Cette dispersion initiale peut être couplée à un modèle de dispersion longue distance; une partie limitée des graines est alors dispersée jusqu'à une distance maximale (L) de manière inversement proportionnelle à la distance à l'origine. (3) un modèle de perturbation établi à partir des sorties du modèle de succession et pour chaque PFT en fonction de sa sensibilité, la part de biomasse (abondance) détruite (mort) ou perturbée (régénération). Il peut s'agir de différents types de perturbations tels que des feux, de la fauche ou du pâturage.

3.2.2. Biomove : un modèle de succession hybride implémenté dans Lamos.

Biomove a été conçu comme un modèle de succession au sein de la plateforme Lamos. C'est un modèle novateur qui permet de modéliser simultanément la dynamique des PFT déterminant la structure de la végétation, et la démographie d'une espèce particulière au sein de cette végétation. De plus, ces deux dynamiques couplées dépendent de la qualité de l'habitat pour chaque espèce/PFT impliqué (ex : issue d'un modèle d'habitat), cette qualité de l'habitat pouvant être spatialement hétérogène et pouvant changer dans le temps. La manière dont la dynamique de la végétation est implémentée dans Biomove dérive du modèle de succession FATE (Moore & Noble 1990). Cela donne un modèle de succession semi-quantitatif reposant sur la compétition pour la lumière et assez simple à paramétrer avec des connaissances de terrains ou des données bibliographiques ou de bases de données. Les PFT sont représentés par des cohortes d'âge et d'abondance données. En fonction de son âge, chaque cohorte se situe dans une strate de hauteur donnée (basse, intermédiaire ou haute) et en fonction de son abondance, chacune opère une interception partielle de la lumière, limitant la lumière disponible dans les strates inférieures (niveau de lumière faible,

Traits de vie		
Durée de vie		200 ans
Age de maturité		30 ans
Durée de vie du pool de graines actives		5 ans
Strate de hauteur des immatures		Intermédiaire
Strate de hauteur des matures		Haute
Abondance maximale		1300 arbres/ha extinction intermédiaire de lumière
Fécondité potentielle		20 000 graines/ pixel /an
Tolérance des différents âges aux différents niveaux de lumière disponible:		
	Germinants	Intermédiaire
	Juvéniles	Fort
	Matures	Fort
Taux de germination pour les différents niveaux de lumière disponible:		
	Faible lumière	Fort
	Lumière intermédiaire	Intermédiaire
	Forte lumière	Intermédiaire
Paramètres de dispersion		
Dispersion longue distance		oui
Fraction de graines concernées		0.001
Distance moyenne de dispersion (kernel)		400 mètres
Distance maximale de dispersion (longue distance)		1.8 kilomètre

Table 4.3. Principales valeurs de paramètres utilisées pour les simulations sous Biomove.

intermédiaire ou fort). Le recrutement et la survie d'autres cohortes dans les strates inférieures du même pixel dépendent de la tolérance des différents PFT à l'ombrage.

Ce qui rend Biomove novateur c'est le couplage qu'il propose entre la dynamique des PFT dominants, déterminant la structure de la végétation, et la démographie d'une espèce particulière (par ex. une espèce rare à forte valeur patrimoniale ou une espèce menacée). Ce couplage est unidirectionnel (effet des PFT sur l'espèce) et suppose que l'espèce particulière reste non dominante et n'affecte pas ou peu son environnement. Il permet d'inclure l'effet des interactions biotiques sur la dynamique d'une espèce particulière dont on connait la démographie, sans pour autant mettre en place des modèles démographiques complexes pour l'ensemble de la végétation. La dynamique de l'espèce est donnée par un modèle démographique mélangeant les approches par âge et par stade. Les étapes juvéniles sont décrites par des âges (entre 0 et l'âge de maturité) et les étapes adultes par des stades (par ex. mature, sénescent), le passage d'un stade à l'autre étant caractérisé dans une matrice de transition. Paramétrer ce modèle nécessite de connaitre les taux démographiques clés de l'espèce, ainsi que la durée de vie des individus, l'âge de leur maturité, la réponse des différents stades à la perturbation et la capacité de dispersion de l'espèce. Il est aussi nécessaire de connaitre la tolérance à l'ombrage pour différentes étapes de la vie des individus (germinant, immature, mature), la disponibilité en lumière résultant de la dynamique du paysage et étant incorporée dans le modèle démographique pour induire de la compétition pour la lumière. Je ne parlerai plus ci-après de ce couplage qui est assez éloigné de mon propos.

Ce qui rend Biomove novateur, c'est également son couplage avec un modèle d'habitat, qui permet de répercuter la qualité de l'habitat sur la performance des cohortes (variabilité intraspécifique contingente à l'environnement, Chapitre 3 et Fig. 4.1) et donc de simuler l'influence des changements environnementaux (par ex. réchauffement climatique) sur la structure de la végétation. La calibration, validation et projection des modèles d'habitat pour chaque PFT/espèce sur l'aire étudiée, se font de manière indépendante à Lamos, par exemple à l'aide de Biomod (Thuiller *et al.* 2009); il est essentiel que l'utilisateur choisisse sciemment la méthode qu'il utilise (Chapitre 3). L'hétérogénéité de la qualité de l'habitat dans l'espace est donnée par la projection des modèles d'habitat sur l'ensemble de la zone traitée dans Biomove. La modification de la qualité de l'habitat dans le temps est donnée par la projection des modèles d'habitat dans le futur grâce aux scénarii climatiques futurs disponibles (IPCC 2007). Ces projections peuvent être effectuées de manières annuelles ou pluriannuelles (ex : tous les dix ans). Une fois importées dans Biomove, les projections des modèles d'habitat sont utilisées pour pondérer les processus démographiques (recrutement, fécondité) et ont pour effet d'exclure le recrutement des PFT et de l'espèce dans les habitats non favorables (Keith *et al.* 2008). Le choix des paramètres devant être pondérés par

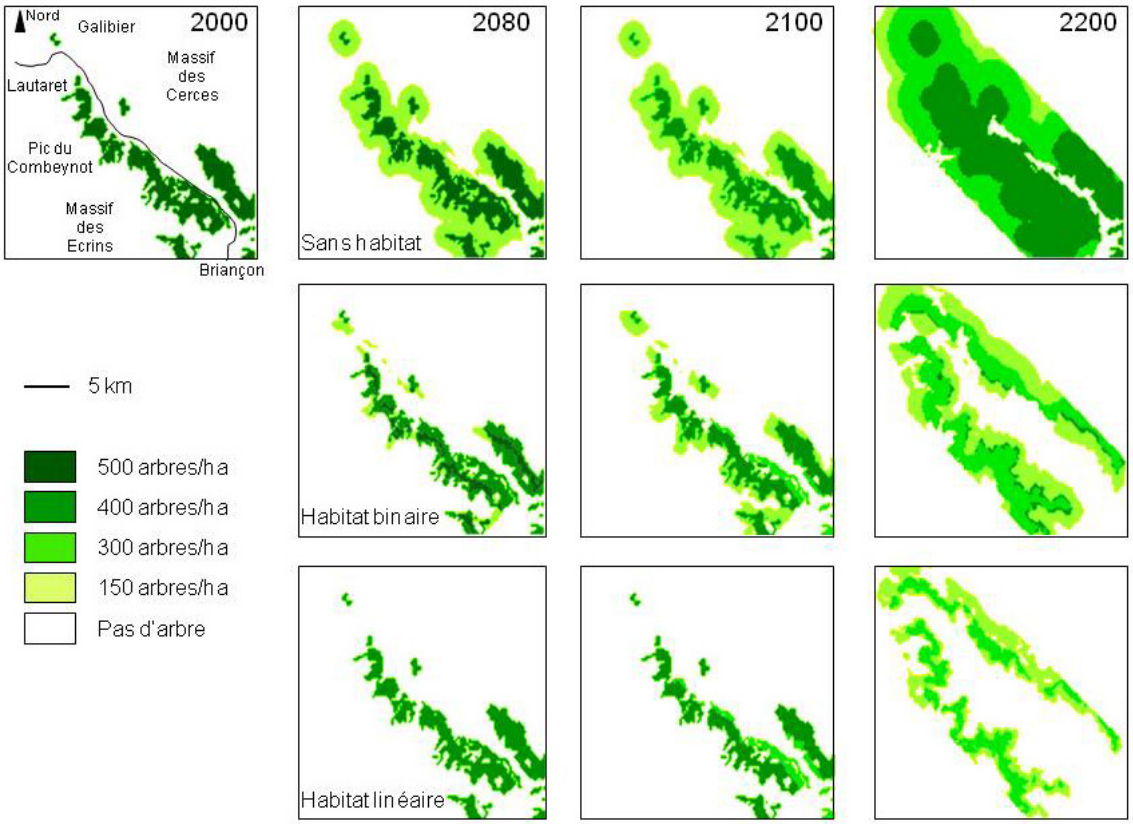


Fig. 4.11. Dynamique de la distribution de Larix decidua dans la vallée de la Guisane.
Les résultats représentent ici l'abondance mature de Larix decidua (niveau de vert). (En haut à gauche) les conditions initiales en 2000. Puis de gauche à droite les distributions en 2080, 2100 et 2200. De haut en bas les dynamiques observées lors des simulations sans effet de la qualité de l'habitat (ligne du haut), avec un effet de la qualité de l'habitat par une fonction de lien binaire (ligne du milieu) et avec un effet de la qualité de l'habitat par une fonction de lien linéaire croissante (ligne du bas).

l'environnement et les choix des fonctions de lien (Fig. 4.3) entre le modèle d'habitat et les paramètres doivent être décidés avant de mettre en place les simulations avec Biomove.

3.2.3. Cas d'étude: dynamique de Larix decidua dans la vallée de la Guisane

Principe

Afin d'illustrer le couplage entre modèle de succession et modèle d'habitat dans Biomove, j'ai mené quelques simulations simples visant à décrire la dynamique de *Larix decidua Mill* seul (sans compétition) dans la vallée de la Guisane entre 2000 et 2080. Le mélèze d'Europe est une espèce arborée sub-alpine pionnière et héliophile, son installation est souvent suivie de celle d'autres arbres tels que *Pinus cembro* ou *P. uncinata*. Les individus de cette espèce atteignent la maturité vers 20-30 ans et peuvent vivre jusqu'à 400 ans (200 ans dans les simulations). Les individus matures produisent plusieurs milliers de graines mais pas tous les ans (~10 ans), c'est le phénomène de masting ; ces graines ont une durée de vie de 3 à 7 ans. Pour simplifier, les arbres produisent ici des graines tous les ans avec une fécondité maximale de 10^6 graines par pixel. Comme modèle de dispersion, j'ai utilisé une loi exponentielle négative (distance moyenne de dispersion=400 m, elle est exagérée ici et devrait plutôt être de 200, la largeur des pixels), couplée à une dispersion aléatoire longue distance (distance max=1,8 km) pour 0,1% des graines. L'ensemble des paramètres de cette espèce, modélisée dans Biomove comme un PFT est donné dans le Tableau 4.3. J'ai utilisé les données de couverture végétale (Corine land cover à 200 x 200 m, Tableau 1.1) en Guisane pour définir la répartition actuelle des arbres dans la vallée et j'ai admis que la grande majorité des boisements étaient des mélézins. Comme modèle d'habitat j'ai choisi d'utiliser un modèle linéaire généralisé de famille binomiale (GLM, Encadré 2.2) liant la présence du mélèze à la Température minimale en hiver (« Température »), selon la formule :

$$\text{Pr}\,\textit{ésence_Mélèze} \approx \textit{Température} + \textit{Température}^2 .$$

Pour simuler un réchauffement climatique entre 2000 et 2080, j'ai augmenté la température uniformément de 0,5 °C tous les dix ans, soit une augmentation totale de 4°C. J'ai ensuite projeté le modèle d'habitat sur ces données de température modifiées afin d'obtenir une carte de qualité de l'habitat tous les dix ans. J'ai ensuite effectué trois types de simulations (sans inclure de perturbation): (1) sans effet de la qualité de l'habitat ; (2) avec un effet binaire de la qualité de l'habitat (seuil arbitraire à 0,6, il faudrait par la suite en tester d'autres) sur le taux de recrutement et (3) avec une fonction de lien linéaire entre qualité de l'habitat et taux de recrutement (recrutement nul dans les sites non favorables et maximal dans les sites de qualité de l'habitat maximal). J'ai également regardé l'évolution des simulations jusqu'à 2100 et 2200 (la qualité de l'habitat restant identique à celle de 2070), 2200 étant la date à laquelle les arbres de la distribution initiale (2000) meurent de mort « naturelle ».

Résultats

La distribution initiale du mélèze est globalement limitée aux altitudes en dessous de 2500 m avec une abondance moyenne de l'ordre de 500 arbres/ha (Fig. 4.11). L'évolution de la distribution des arbres diffère fortement entre les trois types de simulations. Dans le cas (1) où la dynamique n'est pas contrainte par la qualité de l'habitat, la distribution des arbres s'étend jusqu'à recouvrir des sommets comme le Pic du Combeynot (Nord Ouest de la carte) à plus de 3000 m d'altitude (Fig. 4.11). Dans les deux autres cas pour lesquels la distribution est contrainte par la qualité de l'habitat, des nuances d'abondance se créent entre les sites plus ou moins favorables du fait de l'effet de la qualité de l'habitat sur le recrutement. La distribution du mélèze ne s'étend pas mais « se déplace » spatialement, jusqu'à former des peuplements denses (environ 250 arbres/ha) entre 2200 et 2500 m d'altitude et des peuplements plus lâches (environ 100 arbres/ha) autour. Les aires les plus chaudes (et basses en altitude) qui étaient peuplées en 2000 ne le sont plus en 2200 (Fig. 4.11). Il y a donc une influence très forte de la qualité de l'habitat sur la dynamique du mélèze en présence d'un réchauffement climatique et sans compétition ni perturbation. Il y a également une influence de la manière dont cette qualité de l'habitat est prise en compte (fonction de lien binaire ou linéaire).

3.2.4. Discussion sur l'exemple et questions ouvertes

Je n'ai pas cherché ici à montrer des simulations réalistes, ni a prédire une quelconque dynamique future du mélèze dans la vallée de la Guisane. J'ai plutôt cherché à mettre en évidence les éléments manquants pour pouvoir faire des prédictions réalistes à l'avenir. Suivent quelques points de discussion et des questions encore ouvertes pour atteindre ce but :

Signification des modèles d'habitat

Nous avons vu au Chapitre 3 que l'utilisation des modèles d'habitat rencontrait encore un certain nombre de limites techniques et conceptuelles et que le débat était loin d'être clos quand à la signification des sorties de ces modèles. Tout d'abord, afin que le couplage entre modèle d'habitat et modèle de succession soit conceptuellement valide, il faudrait être certain que le modèle d'habitat pris en compte donne une estimation de la niche fondamentale et non de la niche réalisée. En effet on ne peut pas inclure les effets compétitifs deux fois : indirectement avec le modèle d'habitat (si celui-ci représente la niche réalisée) et directement avec le modèle de succession. Ensuite projeter les modèles d'habitat dans le futur nécessite de faire deux hypothèses fortes: (1) il faut supposer que le modèle d'habitat donne une estimation de la niche fondamentale de l'espèce. Il faut aussi supposer que sa distribution présente n'est pas déjà contrainte par des effets compétitifs qui peuvent changer dans le temps: l'erreur de prédiction dans le futur sera d'autant plus grande que l'espèce n'occupe pas la totalité de sa niche au temps présent (Thuiller *et al.* 2004a); (2) il faut

également supposer que la relation entre la présence de l'espèce modélisée et les gradients environnementaux est invariante dans le temps, c'est-à dire qu'il n'y a pas de plasticité phénotypique, ni d'adaptation génétique des populations de l'espèce aux nouvelles conditions environnementales qu'elles subissent (Article 2). Enfin comme nous l'avons vu dans l'exemple, la forme de la fonction de lien (ex. binaire, linéaire : x, ou d'autres formes comme $x^{1/2}$ ou x^3 résultats non présentés : Boucher 2008) change considérablement les résultats pour la dynamique de la distribution d'une espèce seule. Or nous avons montré au Chapitre 3 qu'il n'était pas évident de déceler une relation entre les sorties des modèles d'habitat et performance, trait fonctionnel ou paramètre démographique des populations d'une espèce. Si l'on ne dispose pas de suffisamment d'information pour décider des fonctions de lien à mettre en place, on peut tester la sensibilité de ces fonctions pour chacun des PFT/espèce à l'aide de simulations suivant un dispositif factoriel. On peut également utiliser une approche simplifiée avec des fonctions de lien binaires qui permettent déjà d'éliminer les zones d'habitat non favorables et de n'autoriser la dynamique de succession que dans les zones favorables (Article 1). On pourrait supposer qu'en utilisant une fonction binaire, on introduit moins d'information tout en limitant le biais, cela est à tester. Il faut rajouter aussi que Biomove ne nécessite pas la mise en œuvre d'un modèle d'habitat et que les cartes d'habitat (binaire ou continue) peuvent résulter d'autre modèles ou de connaissance d'experts.

Problème lié au caractère semi-quantitatif de Biomove

Comme décrit ci-dessus, la partie modèle de succession de Biomove est un modèle semi-quantitatif, c'est-à-dire que les principales grandeurs (hauteur des cohortes, disponibilité en lumière, taux de recrutement) sont implémentées sous forme de variables catégorielles ordonnées (faible, intermédiaire, forte). De ce fait, la pondération de certains processus par la qualité de l'habitat résulte en des « sauts » d'une catégorie à une autre, par exemple pour une qualité de l'habitat de 0.20 on pourra avoir un taux de recrutement faible et pour une qualité de l'habitat de 0.21 un recrutement intermédiaire. Pour une espèce/PFT donné, on observe donc des discontinuités dans la distribution des abondances (Fig. 4.11), conduisant à des discontinuités dans l'extinction de la lumière par cette espèce/PFT et donc des discontinuités dans son effet compétitif et dans toute la dynamique qui en résulte. Cela pose problème en particulier par rapport au choix des fonctions de lien, puisque le changement d'une fonction donnée à une fonction très légèrement différente peut entrainer un saut dans la répercussion de la qualité de l'habitat sur les processus. Utiliser une fonction de lien binaire permettrait de limiter ces discontinuités.

Bord de progression et bord de traine (« leading» et « trailing edges »)

Introduire de la variabilité intraspécifique dans un modèle de dynamique de végétation permet en particulier de distinguer les zones d'habitat favorable, dans lesquelles les cohortes ont les paramètres optimaux (taux de fécondité, de recrutement… maximaux), des zones d'habitat moins favorable, dans lesquelles les cohortes sont dotées de paramètres «sub-optimaux» (taux de fécondité, de recrutement… restreints). Dans les zones d'habitat moins favorables ont distingue en particulier : (1) les zones qui n'étaient pas favorables et sont en train de le devenir (bord de progression), ces zones sont potentiellement colonisables par l'espèce si sa fécondité, sa capacité de dispersion et ses interactions avec les autres espèces le permettent et (2) les zones qui étaient favorables et sont en train de devenir non-favorables (bord de traine), les populations qui y sont présentes vont potentiellement s'éteindre, sauf si elles deviennent des puits, fournis en propagules par des populations viables. Modéliser la dynamique de ces deux « bords » de la distribution de l'espèce est l'un des enjeux majeurs du moment (Article 2). Cependant le bord de traine demeure peu étudié (Hampe & Petit 2005). Modéliser la dynamique du bord de traine soulève des questions importantes quand à la mortalité des populations (Hampe & Petit 2005). Les populations présentes dans le bord de traine doivent-elles s'éteindre dès que l'habitat devient non-favorable ? Progressivement avec une diminution de la durée de vie? Ou doivent-elles persister et mourir de « mort naturelle » à la fin de la durée de vie maximale? On pourrait par exemple s'attendre à ce que les espèces les plus longévives et en particulier les arbres soient moins rapidement affectées par une diminution de la qualité de l'habitat.

Des interactions complexes

Dans l'exemple présenté ci-dessus tout a été simplifié au maximum, mais on pourrait ajouter d'autres espèces/PFT et de la perturbation. Aux modifications de la distribution dues aux changements climatiques s'ajouteraient alors des modifications dues à des changements de perturbation (changement d'utilisation des terres) et des modifications dues à la compétition et à la modification de la distribution des espèces/PFT compétiteurs, également soumis à tout ces changements. Nous avons vu que modéliser une espèce seule soumise à une contrainte unique n'était déjà pas trivial. Ajouter d'autres processus complique encore les choses, d'autant plus que les effets de l'ensemble des processus ne s'additionnent pas linéairement mais interagissent pour façonner la distribution finale de chaque espèce/PFT (résultats non présentés : Boucher 2008). Ainsi les résultats obtenus au final pourraient être assez éloignés de ce que l'on aurait pu prévoir connaissant les effets de chaque processus séparément. Je ne présente aucun résultat ici n'étant pour le moment pas à même de les expliquer. Comprendre et prédire la nature et les effets des interactions possibles entre démographie, compétition, dispersion, perturbation (ou modification du régime de perturbation) et changement « climatiques » (qui pourraient aussi bien être d'autres types de

modification de l'habitat) est l'un des enjeux de demain.

4. Conclusion

L'objectif de ce chapitre était de tester la prise en compte de la variabilité fonctionnelle intraspécifique dans deux cas d'étude très différents. Cet objectif nous a d'abord amené à soulever des questions génériques sur la manière de prendre en compte cette variabilité. Et en particulier comment la prendre en compte sans la mesurer, mais simplement en la simulant dans des jeux de données réels (semi-virtuel). Nous avons ensuite utilisé ce cadre conceptuel et l'ensemble des résultats des Chapitres 2 & 3 pour tester la prise en compte de la variabilité fonctionnelle intraspécifique dans deux grands champs d'application des traits fonctionnels: l'évaluation de la diversité fonctionnelle et la modélisation de la dynamique de végétation. Dans chacun des cas d'étude nous avons montré que la prise en compte de la variabilité fonctionnelle intraspécifique pouvait avoir un impact fort sur les résultats obtenus. Les questions et perspectives soulevées ont été différentes dans les deux cas et laissent penser que l'inclusion de la variabilité intraspécifique doit être considérée en détail dans chaque champ d'utilisation des traits fonctionnels.

Ce Chapitre dans son ensemble montre que malgré notre contribution (Chapitre 2 & 3) pour améliorer la connaissance quantitative sur la variabilité intraspécifique, il manque encore un certain nombre d'informations et d'ordres de grandeurs pour pouvoir inclure tous ses aspects dans des analyses semi-virtuelles ou de modélisation. En particulier dans aucune des deux analyses nous n'avons pu prendre en compte la variabilité au sein des populations, ni d'autres sources de variabilité structurée (à l'opposée d'aléatoire) que celle induite par l'environnement. Il serait par conséquent intéressant de quantifier les autres sources de variation des traits telles que la micro-hétérogénéité ou à la stochasticité spatiale et temporelle de l'environnement, les interactions biotiques ou la diversité génétique (Grassein 2009). Cela permettrait de prendre en compte cette variabilité et de l'intégrer sous une forme autre que purement aléatoire. D'autant plus que nous avons montré que la structure de la variabilité et non pas simplement son ampleur pouvait jouer un rôle.

Il faudrait maintenant effectuer les mêmes tests sur d'autres modèles, d'autres indices de diversité, d'autres systèmes d'étude que les milieux subalpins de la Guisane et dans d'autres champs d'application des traits fonctionnels (ex. définition de PFT). Dans les études régionales ou les contrastes fonctionnels entre espèces ne sont pas nécessairement très forts (Chapitre 2), il faudrait tester systématiquement la prise en compte de la variabilité fonctionnelle intraspécifique. Cela peut-être fait par des analyses similaires aux nôtres reposant sur des données de base de traits, sur les ordres de grandeur que nous avons montrés ou sur d'autres mesures *in situ* de la variabilité. Cela permettrait de quantifier l'erreur faite lorsqu'elle est négligée et de décider de la négliger ou non en connaissance de cause.

Conclusion générale

Les objectifs de cette thèse était de (1) quantifier la variabilité fonctionnelle intraspécifique in situ, (2) comprendre la manière dont elle était structurée et (3) étudier son impact dans diverses applications des traits fonctionnels. Pour remplir ces objectifs, nous avons entre autre mené une quantification *in situ* de la variabilité des traits fonctionnels sur seize espèces dans une vallée alpine. J'ai déjà longuement discuté les résultats, limites et perspectives des différentes études présentées dans ce manuscrit, je me contenterais donc ici de résumer les avancées principales par rapport aux attentes données en introduction :

(1) Grâce à l'échantillonnage mis en place, les données nous ont permis de mettre en évidence une forte variabilité des traits fonctionnels au sein des espèces. Une telle variabilité suggère de décrire les espèces non pas par de simples moyennes de traits, mais par des moyennes et des variances. De plus elle suggère de considérer les mesures de terrain sous forme de syndrome de traits au niveau des individus (Chapitre 2).

(2) A l'échelle interspécifique, nous avons retrouvé à partir des données individuelles le compromis de traits classique et les grandes stratégies le long de l'axe d'acquisition-conservation des ressources tiré par la corrélation négative entre SLA et LDMC. A l'échelle intraspécifique, les compromis étaient légèrement différents entre espèces et moins forts qu'à l'échelle interspécifique, mais néanmoins présents (Chapitre 2).

(3) Par une comparaison des variabilités inter- et intraspécifique, nous avons montré (selon la sélection d'espèces) que la variabilité intraspécifique demeurait moindre mais tout de même non négligeable (~30%). Cette variabilité conduit à la formation d'un continuum fonctionnel. Nous pensons qu'il n'est pas raisonnable à l'échelle régionale, lorsque les contrastes fonctionnels entre espèces sont faibles de considérer la variabilité intraspécifique comme négligeable. Au contraire il semble raisonnable de la négliger dans les études à large échelle qui mettent en jeu de très gros contrastes fonctionnels (Chapitre 2).

(4) Nous avons montré que la variabilité fonctionnelle intraspécifique résultait de différences entre populations dues en partie aux gradients environnementaux régionaux et à d'autres facteurs non mesurés. Elle résultait en outre d'une variabilité forte au sein des populations et au sein des individus. Dans notre étude les différents traits avaient des variabilités, des structures de variabilités et des réponses aux gradients différentes. De plus nous n'avons trouvé aucun patron de variabilité ou de structure de variabilité entre les différentes formes de vie ou les espèces ayant les mêmes exigences environnementales. Par ailleurs nous avons mis en évidence que chaque trait pourrait suivre une courbe de réponse générique le long de la distribution de l'espèce. Dans ce cas la relation trait-gradient observée pour chaque espèce

dépendrait de la portion de gradient observée par rapport à la distribution de l'espèce (partie haute, intermédiaire ou basse). Cela a été partiellement confirmé par la confrontation des mesures de traits (quantification fonctionnelle de la niche) et de la qualité de l'habitat donnée par les modèles d'habitat (quantification de la niche en fréquence d'occurrence) (Chapitres 2 & 3).

(5) Nous avons testé un couplage entre un modèle d'habitat et un modèle de dynamique du paysage qui permet de prendre en compte la part de variabilité fonctionnelle intraspécifique contingente aux gradients. Cette prise en compte était fortement sensible au type de relation entre qualité de l'habitat et les traits fonctionnels (pas de relation, relation binaire, relation linéaire). Compte tenu des résultats de la thèse et donc de la nature idiosyncratique des relations, une réponse binaire semblerait une stratégie intéressante à poursuivre (Chapitres 3 & 4).

(6) Notre étude quantitative de la variabilité fonctionnelle intraspécifique nous a permis de tester son effet sur des calculs d'indices de diversité fonctionnelle par l'intermédiaire de simulations semi-virtuelles. Nous avons ainsi montré que différents indices calculés à partir de différents traits n'étaient pas tous aussi robustes à la prise en compte de la variabilité intraspécifique. Nous avons notamment montré que l'ampleur de la variabilité et sa structure jouaient un rôle sur le caractère informatif des indices calculés à partir des traits moyens (Chapitre 2 & 4).

De plus ce travail de thèse :

(7) Nous a amené à mettre en place une stratégie d'échantillonnage efficace. Nous avons mis en évidence l'importance de prévoir un échantillonnage dans le bon espace d'étude, et nous avons par conséquent privilégié l'espace climatique aux espaces géographique et topographique. Nous avons également montré l'importance d'utiliser des expériences virtuelles ou semi-virtuelles pour calibrer un échantillonnage et choisir l'effort d'échantillonnage et la stratégie les plus appropriés à mettre en évidence le patron ou la relation recherché. Nous avons réfléchi à (et testé) des mesures de précision pour estimer non pas des moyennes ou des variances –comme c'est le cas usuellement en théorie de l'échantillonnage-, mais plutôt des courbes de réponse (Chapitre 1).

(8) Nous a amené à travailler sur la signification biologique/écologique des modèles d'habitat. Nous avons montré que ceux-ci pouvaient potentiellement caractériser la niche fonctionnelle de l'espèce (Chapitre 3).

Il a été pour moi très enrichissant d'étudier de multiples aspects d'un même concept en utilisant des approches variées: du terrain à la modélisation en passant par des approches statistiques sur données empiriques et des expériences semi-virtuelles combinant l'ensemble.

L'ensemble de ce travail est bien sur perfectible et connait de nombreuses limites discutées dans l'ensemble du manuscrit. Ce travail apporte toutefois un premier cadre conceptuel et quantitatif sur la variabilité fonctionnelle intraspécifique. Il apporte aussi des premiers exemples d'implications et donne quelques grandes lignes à suivre pour décider s'il faut prendre en compte ou non cette variabilité dans les différents champs d'utilisation des traits fonctionnels; et si oui comment.

Ainsi à l'échelle régionale, la variabilité fonctionnelle intraspécifique pourrait être vue comme une source de diversité au même titre que les patrons de distribution et d'abondance des différentes espèces et que les différences fonctionnelles entre espèces. La variabilité fonctionnelle intraspécifique résulte de nombreux facteurs intrinsèques, biotiques et abiotiques et apparait à différentes échelles (population, individu). Elle reflète les différences de performance entre les individus et les différentes populations d'une même espèce. Vue à l'échelle de l'espèce, elle a un rôle essentiel à jouer dans la manière dont celle-ci peut répondre à des changements environnementaux. Vue à l'échelle des communautés végétales, elle peut influer sur sa dynamique en entraînant des inversements dans le comportement relatif des espèces et les interactions biotiques -par la non conservation des classifications fonctionnelles le long des gradients. De plus elle peut fortement contribuer à la diversité fonctionnelle d'une communauté et aux différences de diversité entre communautés. Vue à l'échelle de l'écosystème, elle peut jouer un rôle sur les interactions biotiques (Boege *et al.* 2008) et les propriétés de celui-ci (Lecerf *et al.* 2008).

Un grand progrès a été fait lors de l'émergence de l'écologie fonctionnelle et du remplacement de la dénomination taxonomique des espèces en présence par leurs caractéristiques fonctionnelles (Calow 1987). Un autre grand progrès a été fait lorsque la présence des espèces a été remplacée par leur abondance dans les études liant les propriétés des communautés végétales au fonctionnement des écosystèmes (ratio de biomasse, Grime 1998). Un nouveau progrès est en train d'être fait avec la prise en compte progressive de la variabilité fonctionnelle intraspécifique. Cela apportera d'autant plus qu'abondance et traits fonctionnels varient tout deux dans l'environnement (effets abiotiques et biotiques) et pourraient avoir des effets synergiques ou antagonistes sur la dynamique des communautés et le rôle de chaque espèce dans le fonctionnement des écosystèmes.

Bibliographie

Aarssen L.W. (1997) High productivity in grassland ecosystems: effected by species diversity or productive species? *Oikos*, 80, 183-184

Ackerly D.D. & Cornwell W.K. (2007) A trait-based approach to community assembly: partitioning of species trait values into within- and among- community components. *Ecology Letters*, 10, 135-145

Ackerly D.D., Schwilk D.W. & Webb C.O. (2006) Niche evolution and adaptive radiation: Testing the order of trait divergence. *Ecology*, 87, S50-S61

Adinarayana (2002) Response surface optimisation of the critical medium components for the production of alkaline protease by a newly isolated bacillus species. *Journal of Pharmaceutical Science*, 5, 272-278

Akaike H. (1974) A new look at statistical model identification. *IEEE Transactions on Automatic Control*, AU-19, 716-722

Al Haj Khaled R., Duru M., Theau J.P., Plantureux S. & Cruz P. (2005) Variation in leaf traits through seasons and N-availability levels and its consequences for ranking grassland species. *Journal of Vegetation Science*, 16, 391-398

Allouche O., Tsoar A. & Kadmon R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223-1232

Araujo M.B. & Guisan A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677-1688

Araujo M.B. & Luoto M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16, 743-753

Araújo M.B. & Pearson R.G. (2005) Equilibrium of species' distributions with climate. *Ecography*, In press

Araújo M.B. & Williams P. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, 96, 331-345

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

Austin M.P. (1987) Models for the analysis of species' response to environment gradients. *Vegetatio*, 69, 35-45

Austin M.P., Nicholls A.O. & Margules C.R. (1990) Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monographs*, 60, 161-177

Baker W.L. (1985) A review of models of landscape change. *Landscape Ecology*, 2, 111-133

Bart J., Hofschen M. & Peterjohn B.G. (1995) Reliability of the breeding bird survey: Effects of restricting surveys to roads. *Auk*, 112, 758-761

Baskin C.C. & Baskin J.M. (1998) *Seeds: ecology, biogeography, and evolution of dormancy and germination*. Acad.Pr., San Diego.

Benichou P. & Le Breton O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie*, 7, 23-24

Benzécri J.P. (1983) Analyse de l'inertie intra-classe par l'analyse d'un tableau de correspondances. *Les Cahiers de l'Analyse des données*, 8, 351-358

Bjorklund M. & Bergek S. (2009) On the relationship between population differentiation and sampling effort: is more always better? *Oikos*, 118, 1127-1129

Boege K. & Dirzo R. (2004) Intraspecific variation in growth, defense and herbivory in *Dialium guianense* (Caesalpiniaceae) mediated by edaphic heterogeneity. *Plant Ecology*, 175, 59-69

Bonnier G. (1887) Note sur des cultures comparées de mêmes espèces à diverses altitudes. *Bull. Soc. Bot. Fr*, t xxxiv, 467-469

Botta-Dukat Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, 16, 533-540

Boucher F. (2008) Réponse de la végétation aux changements climatiques et aux changements d'utilisation des terres dans la vallée de la Guisane (05): paramétrage et test de sensibilité d'un modèle de paysage (LAMOS). In. Ecole Poytechnique (M1), Paris

Bowman W.D. (1995) Altitudinal variation in leaf N and P concentrations in *frasera speciosa* (green gentian). In: *Research/Monitoring Program* (ed. report F). Open Space and Moutain Parks, Boulder

Box G.E.P. & Draper N.R. (1959) A basis for the selection of a response surafce design. *Journal of American Statistical Association*, 54, 622-654

Braun-Blanquet J., Roussine N. & Negre R. (1952) *Les groupements vegetaux de la France Mediterraneenne.* Publications CNRS, Montpellier, France.
Breiman L. (2001) Random forests. *Machine Learning*, 45, 5-32
Breiman L., Friedman J.H., Olshen R.A. & Stone C.J. (1984) *Classification and regression trees*. Chapman and Hall, New York.
Brito J.C., Crespo E.G. & Paulo O.S. (1999) Modelling wildlife distributions: logistic multiple regression vs overlap analysis. *Ecography*, 22, 251-260
Brooker R.W., Maestre F.T., Callaway R.M., Lortie C.L., Cavieres L.A., Kunstler G., Liancourt P., Tielborger K., Travis J.M.J., Anthelme F., Armas C., Coll L., Corcket E., Delzon S., Forey E., Kikvidze Z., Olofsson J., Pugnaire F., Quiroz C.L., Saccone P., Schiffers K., Seifan M., Touzard B. & Michalet R. (2008) Facilitation in plant communities: the past, the present, and the future. *Journal of Ecology*, 96, 18-34
Bugmann H. (2001) A review of forest gap models. *Climatic Change*, 51, 259-305
Burnham K.P. & Anderson D.R. (2004) Multimodel inference - understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261-304
Byars S.G., Papst W. & Hoffmann A.A. (2007) Local adaptation and cogradient selection in the alpine plant, *Poa hiemata*, along a narrow altitudinal gradient. *Evolution*, 61, 2925-2941
Calow P. (1987) Towards a definition of functional ecology. *Functional Ecology*, 1, 57-61
Canham C.D., LePage P.T. & Coates K.D. (2004) A neighborhood analysis of canopy tree competition: effects of shading versus crowding. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 34, 778-787
Chase J.M. & Leibold M.A. (2003) *Ecological niches*. Chicago University Press, Chicago.
Chave J. (2004) Neutral theory and community ecology. *Ecology Letters*, 7, 241-253
Choler P., Michalet R. & Callaway R.M. (2001) Facilitation and competition on gradients in alpine plant communities. *Ecology*, 82, 3295-3308
Chuine I., Cour P. & Roussau D.D. (1999) Selecting models to predict the timing of flowering of temperate trees: implications for tree phenology modelling. *Plant, Cell and Environment*, 22, 1-13
Cianciaruso M.V., Batalha M.A., Gaston K.J. & Petchey O.L. (2009) Including intraspecific variability in functional diversity. *Ecology*, 90, 81-89
Cingolani A.M., Cabido M., Gurvich D.E., Renison D. & Diaz S. (2007) Filtering processes in the assembly of plant communities: Are species presence and abundance driven by the same traits? *Journal of Vegetation Science*, 18, 911-920
Cohen J. (1960) A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.*, 20, 37-46
Coomes D.A. & Allen R.B. (2007) Effects of size, competition and altitude on tree growth. *Journal of Ecology*, 95, 1084-1097
Cordell S., Goldstein G., Mueller-Dombois D., Webb D. & Vitousek P.M. (1998) Physiological and morphological variation in *Metrosideros polymorpha*, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. *Oecologia*, 113, 188-196
Cornelissen J.H.C., Cerabolini B., Castro-Diez P., Villar-Salvador P., Montserrat-Marti G., Puyravaud J.P., Maestro M., Werger M.J.A. & Aerts R. (2003a) Functional traits of woody plants: correspondence of species rankings between field adults and laboratory-grown seedlings? *Journal of Vegetation Science*, 14, 311-322
Cornelissen J.H.C., Lavorel S., Garnier E., Díaz S., Buchmann N., Gurvich D.E., Reich P.B., ter Steege H., Morgan H.D., van der Heijden M.G.A., Pausas J.G. & Poorter H. (2003b) Handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 51, 335-308
Cornwell W.K., Schwilk D.W. & Ackerly D.D. (2006) A trait-based test for habitat filtering: Convex hull volume. *Ecology*, 87, 1465-1471
Coudun C. & Gegout J.C. (2007) Quantitative prediction of the distribution and abundance of Vaccinium myrtillus with climatic and edaphic factors. *Journal of Vegetation Science*, 18, 517-524
Courbois J.-Y., Katz S.L., Isaak D.J., Steel E.A., Thurow R.F., Wargo Rub A.M., Olsen T. & Jordan C.E. (2008) Evaluating probability sampling strategies for estimating redd counts: an example with Chinook salmon (Oncorhynchus tshawytscha). *Canadian Journal of Fisheries and Aquatic Sciences*, 65, 1814-1830

Cox P.A. & Elmqvist T. (2000) Pollinator extinction in the Pacific Islands. *Conservation Biology*, 14, 1237-1239

Craine J.M., Lee, W.G. (2003) Covariation in leaf and root traits for native and non-native grasses along and altitudinal gradient. *Oecologia*, 134, 471-478

Dangasuk O.G. & Panetsos K.P. (2004) Altitudinal and longitudinal variations in *Pinus brutia* (Ten.) of Crete Island, Greece: some needle, cone and seed traits under natural habitats. *New Forest*, 27, 269-284

de Bello F., Thuiller W., Leps J., Choler P., Clement J.C., Macek P., Sebastia M.T. & Lavorel S. (2009) Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. *Journal of Vegetation Science*, 20, 475-486

de Groot W.J., Bothwell P.M., Carlsson D.H. & Logan K.A. (2003) Simulating the effects of future fire regimes on western Canadian boreal forests. *Journal of Vegetation Science*, 14, 355-364

DeAngelis D.L. & Mooij W.M. (2005) Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology Evolution And Systematics*, 36, 147-168

Debeljak M., Dzerosky S., Jerina K., Kobler A. & Adamic M. (2001) Habitat suitability modelling for red deer (*Cervus elaphus* L.) in South-central Slovenia with classification trees. *Ecological Modelling*, 138, 321-330

DeWitt T.J., Sih A. & Wilson D.S. (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, 13, 77-81

Diaz S. & Cabido M. (2001) Vive la difference: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, 16, 646-655

Díaz S. & Cabido M. (1997) Plant functional types and ecosystem function in relation to global change. *Journal of Vegetation Science*, 8, 463-474

Diaz S., Hodgson J.G., Thompson K., Cabido M., Cornelissen J.H.C., Jalili A., Montserrat-Marti G., Grime J.P., Zarrinkamar F., Asri Y., Band S.R., Basconcelo S., Castro-Diez P., Funes G., Hamzehee B., Khoshnevi M., Perez-Harguindeguy N., Perez-Rontome M.C., Shirvany F.A., Vendramini F., Yazdani S., Abbas-Azimi R., Bogaard A., Boustani S., Charles M., Dehghan M., de Torres-Espuny L., Falczuk V., Guerrero-Campo J., Hynd A., Jones G., Kowsary E., Kazemi-Saeed F., Maestro-Martinez M., Romo-Diez A., Shaw S., Siavash B., Villar-Salvador P. & Zak M.R. (2004) The plant traits that drive ecosystems: evidence from three continents. *Journal of Vegetation Science*, 15, 295-304

Diaz S., Lavorel S., Chapin F.S., Tecco P.A., Gurvich D.E. & Grigulis K. (2007a) In: *Terrestrial Ecosystem in a Changing World* (ed. Canadell J PD, Pitelka LF), pp. 79-91. Springer, Berlin

Diaz S., Lavorel S., de Bello F., Quetier F., Grigulis K. & Robson M. (2007b) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 20684-20689

Dodélec S. & Chessel D. (1991) Recent developments in linear ordination methods for environmental sciences. *Advances in Ecology, India*, 1, 133-155

Dormann C.F., McPherson J.M., Araujo M.B., Bivand R., Bolliger J., Carl G., Davies R.G., Hirzel A., Jetz W., Kissling W.D., Kuhn I., Ohlemuller R., Peres-Neto P.R., Reineking B., Schroder B., Schurr F.M. & Wilson R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30, 609-628

Droege S. (1990) The North American Breeding Bird Survey. In: *Survey Designs and Statistical Methods for the Estimation of Avian Population Trends* (eds. Sauer JR & Droege S), p. 1–4. USDI, Fish and Wildlife Service, Biological Reports 90

Dyer A.R., Goldberg D.E., Turkington R. & Sayre C. (2001) Effects of growing conditions and source habitat on plant traits and functional group definition. *Functional Ecology*, 15, 85-95

Elith J. & Graham C.H. (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32, 66-77

Elith J., Graham C.H., Anderson R.P., Dudik M., Ferrier S., Guisan A., Hijmans R.J., Huettman F., Leathwick J.R., Lehmann A., Li J., Lohmann L., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.M., Peterson A.T., Phillips S., Richardson K., Schachetti Pereira R., Schapire R.E., Soberón J., Williams S.E., Wisz M. & Zimmermann N.E. (2006) Novel methods improve predictions of species' distributions from occurrence data. *Ecography*, in press

Elmendorf S.C. & Moore K.A. (2008) Use of community-composition data to predict the fecundity and abundance of species. *Conservation Biology*, 22, 1523-1532

Elton C.S. (1927) *Animal ecology*. Sedgwick and Jackson, London.
Engler R., Guisan A. & Rechsteiner L. (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41, 263-274
Eriksson O. (2002) Ontogenetic niche shifts and their implications for recruitment in three clonal Vaccinium shrubs: Vaccinium myrtillus, Vaccinium vitis-idaea, and Vaccinium oxycoccos. *Canadian Journal of Botany-Revue Canadienne De Botanique*, 80, 635-641
Fielding A.H. & Bell J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38-49
Fonseca C.R., Overton J.M., Collins B. & Westoby M. (2000) Shifts in trait-combinations along rainfall and phosphorus gradients. *Journal of Ecology*, 88, 964-977
Friedman J.H. (1991) Multivariate Adaptive Regression Splines. *Annals of Statistics*, 19, 1-67
Garnier E., Cortez J., Billès G., Navas M.-L., Roumet C., Debussche M., Laurent G., Blanchard A., Aubry D., Bellmann A., Neill C. & Toussaint J.-P. (2004) Plant functional markers capture ecosystem properties during secondary succession. *Ecology*, 85, 2630-2637
Garnier E., Laurent G., Bellmann A., Debain S., Berthelier P., Ducout B., Roumet C. & Navas M.-L. (2001) Consistency of species ranking based on functional leaf traits. *New Phytologist*, 152, 69-83
Garnier E., Lavorel S., Ansquer P., Castro H., Cruz P., Dolezal J., Eriksson O., Fortunel C., Freitas H., Golodets C., Grigulis K., Jouany C., Kazakou E., Kigel J., Kleyer M., Lehsten V., Leps J., Meier T., Pakeman R., Papadimitriou M., Papanastasis V.P., Quested H., Quetier F., Robson M., Roumet C., Rusch G., Skarpe C., Sternberg M., Theau J.P., Thebault A., Vile D. & Zarovali M.P. (2007) Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: A standardized methodology and lessons from an application to 11 European sites. *Annals of Botany*, 99, 967-985
Gaston K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.
Gates S. & Donald P.F. (2000) Local extinction of British farmland birds and the prediction of further loss. *Journal of Applied Ecology*, 37, 806-820
Gause G.F. (1934) *The struggle for existence*. Dover publications, New York.
Gerdol R. (2005) Growth performance of two deciduous *Vaccinium* species in relation to nutrient status in a subalpine heath. *Flora*, 200, 168-174
Graham C.H., Ferrier S., Huettmann F., Moritz C. & Peterson A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19, 497-503
Grassein F. (2009) Mécanismes de variations des traits fonctionnels dans les praires alpines. In. Université Joseph Fourier, Grenoble
Grime J.P. (1977) Evidence for existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *American Naturalist*, 111, 1169-1194
Grime J.P. (1998) Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *Journal of Ecology*, 86, 902-910
Grime J.P., Thompson K., Hunt R., Hodgson J.G., Cornelissen J.H.C., Rorison I.H., Hendry G.A.F., Ashenden T.W., Askew A.P., Band S.R., Booth R.E., Bossard C.C., Campbell B.D., Cooper J.E.L., Davison A.W., Gupta P.L., Hall W., Hand D.W., Hannah M.A., Hillier S.H., Hodkinson D.J., Jalili A., Liu Z., Mackey J.M.L., Matthews N., Mowforth M.A., Neal A.M., Reader R.J., Reiling K., RossFraser W., Spencer R.E., Sutton F., Tasker D.E., Thorpe P.C. & Whitehouse J. (1997) Integrated screening validates primary axes of specialisation in plants. *Oikos*, 79, 259-281
Grinnell J. (1917) The niche-relationships of the California Trasher. *Auk*, 34, 131-135
Gross N., Suding K.N., Lavorel S. & Roumet C. (2007) Complementarity as a mechanism of coexistence between functional groups of grasses. *Journal of Ecology*, 95, 1296-1305
Guisan A., Broennimann O., Engler R., Vust M., Yoccoz N.G., Lehmann A. & Zimmermann N.E. (2006) Using niche-based models to improve the sampling of rare species. *Conservation Biology*, 20, 501-511
Guisan A. & Thuiller W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993-1009

Guisan A. & Zimmermann N.E. (2000) Predictive habitat distribution models in Ecology. *Ecological Modelling*, 135, 147-186

Haddad N.M., Hudgens B., Damiani C., Gross K., Kuefler D. & Pollock K. (2008) Determining optimal population monitoring for rare butterflies. *Conservation Biology*, 22, 929-940

Hampe A. (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology & Biogeography*, 13, 469-476

Hampe A. & Petit R.J. (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, 8, 461-467

Harper J.L. (1964) The individual in the population. *Journal of Ecology*, 52, 149-158

Harper J.L. (1977) *Population biology of plants*. Academic press, London.

Hastie T.J. & Tibshirani R. (1990) *Generalized additive models*. Chapman and Hall, London.

He J.S., Wang Z., Wang X., Schmid B., Zuo W., Zhou M., Zheng C., Wang M. & Fang J. (2006) A test of the generality of leaf trait relationships on the Tibetan Plateau. *New Phytologist*, 170, 835-848

Hector A., Bazeley-White E., Loreau M., Otway S. & Schmid B. (2002) Overyielding in grassland communities: testing the sampling effect hypothesis with replicated biodiversity experiments. *Ecology Letters*, 5, 502-511

Heikkinen R.K., Luoto M., Araujo M.B., Virkkala R., Thuiller W. & Sykes M.T. (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress In Physical Geography*, 30, 751-777

Higgins S.I., Richardson D.M., Cowling R.M. & Trinder-Smith T.H. (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology*, 13, 303-313

Hirzel A.H., Hausser J., Chessel D. & Perrin N. (2002) Ecological-niche factor analysis: How to compute habitat- suitability maps without absence data? *Ecology*, 83, 2027-2036

Hooper D.U., Chapin III F.S., Ewel J.J., Hector A., Inchausti P., Lavorel S., Lawton J.H., Lodge D.M., Loreau M., Naeem S., Schmid B., Setälä H., Symstad A.J., Vandermeer J. & Wardle D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, 75, 3-35

Hovenden M.J. & Vander Schoor J.K. (2003) Nature vs nurture in the leaf morphology of Southern beech, *Nothofacus cunninghamii* (Nothofagaceae). *New Phytologist*, 161, 585-594

Hughes R.M. & Peck D.V. (2008) Acquiring data for large aquatic resource surveys: the art of compromise among species, logistics, and reality. *Journal of the North American Benthological Society*, 27, 837-859

Hultine K.R. & Marshall J.D. (2000) Altitude trends in conifer leaf morphology and stable carbon isotope composition. *Oecologia*, 123, 32-40

Hutchinson G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 145-159

Inouye (2001) response surface experimental designs for investigating interspecific competition. *Ecology*, 82, 2696-2706

IPCC (2007) Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Asssessment. In: (ed. Press CU). Report of the Intergovernmental Panel on Climate Change., Cambridge, UK

Jeltsch F., Moloney K.A., Schurr F.M., Kochy M. & Schwager M. (2008) The state of plant population modelling in light of environmental change. *Perspectives in Plant Ecology Evolution and Systematics*, 9, 171-189

Joshi J., Schmid B., Caldeira M.C., Dimitrakopoulos P.G., Good J., Harris R., Hector A., Huss-Danell K., Jumpponen A., Minns A., Mulder C.P.H., Pereira J.S., Prinz A., Scherer-Lorenzen M., Siamantziouras A.S.D., Terry A.C., Troumbis A.Y. & Lawton J.H. (2001) Local adaptation enhances performance of common plant species. *Ecology Letters*, 4, 536-544

Jouglet J.P. (1999) *Les végétations des alpages des Alpes françaises du sud.* Cemagref éditions, Antony, France.

Kattge J., Knorr W., Raddatz T. & Wirth C. (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology*, 15, 976-991

Kattge J., Knorr W., Raddatz T. & Wirth C. (in prep) The nature of plant traits as a basis for a generic structure of trait databases.

Kearney M. (2006) Habitat, environment and niche: what are we modelling? *Oikos (opinion)*
Kearney M. & Porter W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12, 334-350
Keddy P.A. (1992) Assembly and response rules - 2 goals for predictive community ecology. *Journal of Vegetation Science*, 3, 157-164
Keith D.A., Akcakaya H.R., Thuiller W., Midgley G.F., Pearson R.G., Phillips S.J., Regan H.M., Araujo M.B. & Rebelo T.G. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, 4, 560-563
Kergélen M. (1993) *Index synonymique de la flore de France*. Muséum National d'Histoire Naturelle, Secrétariat Faune-Flore, Paris.
Kitijama K. & Fenner M. (2000) Ecology of seedling regeneration. In: *Seeds: The Ecology of Regeneration in Plant Communities*, pp. 331-359. CABI Publishing, New York
Knevel I.C., Bekker R.M., Bakker J.P. & Kleyer M. (2003) Life-history traits of the Northwest European flora: The LEDA database. *Journal of Vegetation Science*, 14, 611-614
Körner C. (1989) The nutritional status of plants from high altitudes. A worldwide comparison. *Oecologia*, 81, 379–391
Körner C. (1998) A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, 115, 445-459
Körner C. (1999) *Alpine plant life*. Springer-Verlag, Berlin.
Körner C. (2007) The use of "altitude" in ecological research. *Trends in Ecology & Evolution*, 22, 570-574
Körner C., Bannister P. & Mark A.F. (1986) Altitudinal variation in stomatal conductance, nitrogen content and leaf anatomy in different plant life forms in New Zealand. *Oecologia*, 69, 577-588
Kozak K.H. & Wiens J.J. (2006) Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, 60, 2604-2621
Kühn I., Durka W. & Klotz S. (2004) BiolFlor - a new plant-trait database as a tool for plant invasion ecology. *Diversity and Distributions*, 10, 363-365
Lake J.K. & Ostling A. (2009) Comment on "Functional Traits and Niche-Based Tree Community Assembly in an Amazonian Forest". *Science*, 324, 1015c-U31
Lavorel S. (2001) LAMOS: A LAndscape MOdelling Shell for studying the role of landscape scale processes in global change effects. *GCTE News*, 17, 5-6
Lavorel S., Diaz S., Cornelissen J.H.C., Garnier E., Harrison S.P., McIntyre S., Pausas J., Pérez-Harguindeguy N., Roumet C. & Urcelay C. (2007) Plant functional types: are we getting any closer to the Holy Grail? In: *Terrestrial Ecosystems in a Changing World* (eds. Canadell J, Pitelka LF & Pataki D), pp. 171-186. Springer-Verlag
Lavorel S., Gachet S., Sahl A., Gaucherand S. & Bonet R. (2009) A plant functional traits data base for the Alps - Understanding functional effects of changed grassland management. In: *Data mining for global trends in mountain biodiversity* (eds. Körner C & Spehn E), pp. 106-123. CRC Press / Taylor and Francis, Boca Raton
Lavorel S. & Garnier E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology*, 16, 545-556
Lavorel S., McIntyre S., Landsberg J. & Forbes T.D.A. (1997) Plant functional classifications: from general groups to specific groups based on response to disturbance. *Trends in Ecology & Evolution*, 12, 474-478
Lawler J.J., White D., Neilson R.P. & Blaustein A.R. (2006) Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, 12, 1568-1584
Lawton J.H. (1999) Are there general laws in ecology? *Oikos*, 84, 177-192
Leathwick J.R. & Austin M.P. (2001) Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, 82, 2560-2573
Lecerf A. & Chauvet E. (2008) Intraspecific variability in leaf traits strongly affects alder leaf decomposition in a stream. *Basic and Applied Ecology*, 9, 598-605
Lehsten V. & Kleyer M. (2007) Turnover of plant trait hierarchies in simulated community assembly in response to fertility and disturbance. *Ecological Modelling*, 203, 270-278
Lemeshow S. & Levy P.S. (1999) *Sampling of Populations: Methods and Applications*. 3 edition (Feb 12 1999) edn. Wiley-Interscience.

Leps J., de Bello F., Lavorel S. & Berman S. (2006) Quantifying and interpreting functional diversity of natural communities: practical considerations matter. *Preslia*, 78, 481-501

Link W.A. & Barker R.J. (2006) Model weights and the foundations of multimodel inference. *Ecology*, 87, 2626-2635

Lischke H., Zimmermann N.E., Bolliger J., Rickebusch S. & Loffler T.J. (2006) TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling*, 199, 409-420

Liu C., Berry P.M., Dawson T.E. & Pearson R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385-393

Lopez-Darias M., Lobo J.M. & Gouat P. (2008) Predicting potential distributions of invasive species: the exotic Barbary ground squirrel in the Canarian archipelago and the west Mediterranean region. *Biological Invasions*, 10, 1027-1040

Loreau M., Naeem S., Inchausti P., Grime J.P., Hector A., Hooper D.U., Raffaelli D., Schmid B., Tilman D. & Wardle D.A. (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, 294, 804-808

Luo T., Luo J. & Pan Y. (2005) Leaf traits and associated ecosystem characteristics across subtropical and timberline forests in the Gongga Moutains, Eastern Tibetan Plateau. *Oecologia*, 142, 261-273

Mac Arthur R.H. (1972) *Geographical ecology: patterns in the distribution of species*. Princeton University Press.

Mack R.N., Simberloff D., Lonsdale W.M., Evans H., Clout M. & Bazzaz F.A. (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, 10, 689-710

Mackenzie D.I., Bailey L.L. & Nichols J.D. (2004) Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology*, 73, 546-555

Madin J.S., Bowers S., Schildhauer M.P. & Jones M.B. (2008) Advancing ecological research with ontologies. *Trends in Ecology & Evolution*, 23, 159-168

Maguire D.A. & Bennett W.S. (1996) Patterns in vertical distribution of foliage in young coastal Douglas-Fir. *Canadian journal of forest research*, 26, 1991-2005

Mahalanobis P.C. (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2, 49-55

Mason N.W.H., MacGillivray K., Steel J.B. & Wilson J.B. (2003) An index of functional diversity. *Journal of Vegetation Science*, 14, 571-578

Mason N.W.H., Mouillot D., Lee W.G. & Wilson J.B. (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos*, 111, 112-118

McCain C.M. (2009) Vertebrate range sizes indicate that mountains may be 'higher' in the tropics. *Ecology Letters*, 12, 550-560

McCullagh P. & Nelder J.A. (1989) *Generalized linear models*. Chapman & Hall.

McGill B.J., Enquist B.J., Weiher E. & Westoby M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21, 178-185

McGraw J.B. & Antonovics J. (1983) Experimental ecology of *Dryas octopetala* ecotypes. 1. Ecotypic differentiation and life-cycle stages of selection. *Journal of Ecology*, 71, 879-897

Menke W.W. (1973) Response surface methodology Applications to Biological Data from field Measurements. *Ecology*, 54, 920-923

Meziane D. & Shipley B. (1999) Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. *Plant Cell And Environment*, 22, 447-459

Milla R., Escudero A. & Iriondo J.M. (2009) Inherited variability in multiple traits determines fitness in populations of an annual legume from contrasting latitudinal origins. *Annals Of Botany*, 103, 1279-1289

Milsom T.P., Langton S.D., Parkin W.K., Peel S., Bishop J.D., Hart J.D. & Moore N.P. (2000) Habitat models of bird species' distribution: an aid to the management of coastal grazing marshes. *Journal of Applied Ecology*, 37, 706-727

Mohler C.L. (1983) Effect of sampling pattern on estimation of species distributions along gradients. *Vegetatio*, 54, 97-102

Mooney H.A. & Hobbs R.J. (2000) *Invasive species in a changing world*. Island Press, Washington, D.C, California.

Moore A.D. & Noble I.R. (1990) An individualistic model of vegetation stand dynamics. *Journal of Environmental Management*, 31, 61-81
Naeem S. & Wright J.P. (2003) Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem. *Ecology Letters*, 6, 567-579
Norberg J., Swaney D.P., Dushoff J., Lin J., Casagrandi R. & Levin S.A. (2001) Phenotypic diversity and ecosystem functioning in changing environments: A theoretical framework. *Proceedings of The National Academy of Sciences of The United States of America*, 98, 11376-11381
Olser A.R., Sedransk J., Edwards D., Gotway C.A., Liggett W., Rathbun S., Reckhow K.H. & Young L.J. (1999) Statistical issues for monitoring ecological and natural resources in the United States. *Environmental Monitoring and Assessment*, 54, 1-45
Pachepsky E., Bown J.L., Eberst A., Bausenwein U., Millard P., Squire G.R. & Crawford J.W. (2007) Consequences of intraspecific variation for the structure and function of ecological communities Part 2: Linking diversity and function. *Ecological Modelling*, 207, 277-285
Paulsen J., Weber U.M. & Korner C. (2000) Tree growth near treeline: Abrupt or gradual reduction with altitude? *Arctic Antarctic and Alpine Research*, 32, 14-20
Pearce J. & Ferrier S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225-245
Pearce J.L. & Boyce M.S. (2006) Modelling distribution and abundance with presence-only data. *Journal Of Applied Ecology*, 43, 405-412
Pearman P.B., Guisan A., Broennimann O. & Randin C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution*, 23, 149-158
Pearson R.G., Dawson T.P., Berry P.M. & Harrison P.A. (2002) SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. *Ecological Modelling*, 154, 289-300
Peratoner G., Joergensen R.G. & Spatz G. (2007) Growth of Trifolium alpinum: effects of soil properties, symbionts and pathogens. *Ecological engineering*, 30, 349-355
Perrings C. (2003) The economics of abrupt climate change. *Philosophical Transactions of the Royal Society of London B*, 361, 2043-2059
Petchey O.L. & Gaston K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, 5, 402-411
Petchey O.L. & Gaston K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741-758
Peterjohn B.G. (1994) The North American breeding bird survey. *Journal of American Birding Association*, 26, 386-398
Peterson A.T., Lash R.R., Carroll D.S. & Johnson K.M. (2006) Geographic potential for outbreaks of Marburg hemorrhagic fever. *American Journal of Tropical Medicine And Hygiene*, 75, 9-15
Phillips S.J., Anderson R.P. & Schapire R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231-259
Phillips S.J., Dudik M., Elith J., Graham C.H., Lehmann A., Leathwick J. & Ferrier S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19, 181-197
Pierce S., Ceriani R.M., de Andreis R., Luzzaro A. & Cerabolini B. (2007) The leaf economics spectrum of Poaceae reflects variation in survival strategies. *Plant Biosystems*, 141, 337-343
Pinheiro J.C. & Bates D.M. (2000) *Mixed-effects models in S and S-Plus*. Springer, New York.
Podani J. & Schmera D. (2006) On dendrogram-based measures of functional diversity. *Oikos*, 115, 179-185
Poorter L. & Bongers F. (2006) Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology*, 87, 1733-1743
Poorter L., Wright S.J., Paz H., Ackerly D.D., Condit R., Ibarra-Manriques G., Harms K.E., Licona J.C., Martinez-Ramos M., Mazer S.J., Muller-Landau H.C., Pena-Claros M., Webb C.O. & Wright I.J. (2008) Are functional traits good predictors of demographic rates? Evidence from five Neotropical forests. *Ecology*, 89, 1908-1920
Pornon A. & Lamaze T. (2007) Nitrogen resorption and photosynthetic activity over leaf life span in an evergreen shrub, Rhododendron ferrugineum, in a subalpine environment. *New Phytologist*, 175, 301-310
Potts J.M. & Elith J. (2006) Comparing species abundance models. *Ecological Modelling*, 199, 153-163
Pulliam H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, 3, 349-361

Randin C.F., Dirnbock T., Dullinger S., Zimmermann N.E., Zappa M. & Guisan A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, 33, 1689-1703

Rao C.R. (1964) The use and interpretation of principal correspondence analysis in applied research. *Sankhya A*, 26, 329-359

Rao C.R. (1982) Diversity and dissimilarity coefficients - a unified approach. *Theoretical Population Biology*, 21, 24-43

Raxworthy C.J., Martinez-Meyer E., Horning N., Nussbaum R.A., Schneider G.E., Ortega-Huerta M.A. & Peterson A.T. (2004) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, 426, 837-841

Real R., Barbosa A.M., Rodriguez A., Garcia F.J., Vargas J.M., Palomo L.J. & Delibes M. (2009) Conservation biogeography of ecologically interacting species: the case of the Iberian lynx and the European rabbit. *Diversity and Distributions*, 15, 390-400

Real R., Marcia Barbosa A. & Vargas J.M. (2006) Obtaining environmental favourability functions from logistic regression. *Environmental and ecological statistics*, 13, 237-245

Reich P.B., Ellsworth D.S., Walters M.B., Vose J.M., Gresham C., Vollin J.C. & Bowman W.D. (1999) Generality of leaf trait relationships: a test across six biomes. *Ecology*, 80, 1955-1969

Reich P.B., Walters M.B. & Ellsworth D.S. (1997) From tropics to tundra: Global convergence in plant functioning. *Proceedings of The National Academy Of Sciences Of The United States Of America*, 94, 13730-13734

Reich P.B., Wright I.J., Cavender-Bares J., Craine J.M., Oleksyn J., Westoby M. & Walters M.B. (2003) The evolution of plant functional variation: Traits, spectra, and strategies. *International Journal of Plant Sciences*, 164, S143-S164

Ridgeway G. (1999) The state of boosting. *Computing Science and Statistics*, 31, 172-181

Robertson G.P., Coleman D.C., Bledsoe C.S. & Sollins P. (1999) *Standard soil methods for long-term ecological research*. Oxford University Press, New York Oxford.

Roche P., Diaz-Burlinson N. & Gachet S. (2004) Congruency analysis of species ranking based on leaf traits: which traits are the more reliable? *Plant Ecology*, 174, 37-48

Rome M. (2009) Traits fonctionnels végétaux et fonctionnement d'un paysage agropastoral à l'étage sub-alpin. In. Université Joseph Fourier (Master 2), Grenoble

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

Ryser P. & Aeschlimann U. (1999) Proportional dry-matter content as an underlying trait for the variation in the relative growth rate among 22 Eurasian populations of *Dactylis glomerata*. *Functional Ecology*, 13, 473-482

Savage V.M., Webb C.T. & Norberg J. (2007) A general multi-trait-based framework for studying the effects of biodiversity on ecosystem functioning. *Journal of Theoretical Biology*, 247, 213-229

Schröter D., Cramer W., Leemans R., Prentice I.C., Araújo M.B., Arnell N.W., Bondeau A., Bugmann H., Carter T., Garcia C.A., de la Vega-Leinert A.C., Erhard M., Ewert F., Glendining M., House J.I., Kankaanpää S., Klein R.J.T., Lavorel S., Lindner M., Metzger M.J., Meyer J., Mitchell T., Reginster I., Rounsevell M., Sabaté S., Sitch S., Smith B., Smith J., Smith P., Sykes M.T., Thonicke K., Thuiller W., Tuck G., Zaehle S. & Zierl B. (2005) Ecosystem service supply and human vulnerability to global change in Europe. *Science*, 310, 1333-1337

Scott E.M., Dixon P., Voigt G. & Whicker W. (2008) Choice and criteria for selection of sampling strategies in environmental radioactivity monitoring. *Applied Radiation and Isotopes*, 66, 1575-1581

Segurado P. & Araújo M.B. (2004) Evaluation of methods for modelling species probabilities of occurrence. *Journal of Biogeography*, 31, 1555-1568

Seo C., Thorne J.H., Hannah L. & Thuiller W. (2009) Scale effects in species distribution models: implications for conservation planning under climate change. *Biology Letters*, 5, 39-43

Shea K. & Chesson P. (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology & Evolution*, 17, 170-175

Shipley B. (1995) Structured interspecific determinants of specific leaf-area in 34 species of herbaceous angiosperms. *Functional Ecology*, 9, 312-319

Shipley B. (2000) Plasticity in relative growth rate and its components following a change in irradiance. *Plant Cell and Environment*, 23, 1207-1216

Shipley B. (2003) Interspecific consistency and intraspecific variability of specific leaf area with respect to irradiance and nutrient availability. *Ecoscience*, 10, 74-79

Sitch S., Smith B., Prentice I.C., Arneth A., Bondeau A., Cramer W., Kaplan J.O., Levis S., Lucht W., Sykes M.T., Thonicke K. & Venevsky S. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ Dynamic Global Vegetation Model. *Global Change Biology*, 9, 161-185

Soberon J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10, 1115-1123

Sparks J.P. & Ehleringer J.R. (1997) Leaf carbon isotope discrimination and nitrogen content for riparian trees along elevational transects. *Oecologia*, 109, 362-367

Steffen W.L. (1996) A periodic table for ecology? A chemist's view of plant functional types. *Journal of Vegetation Science*, 7, 425-430

Stockwell D.R.B. & Peters D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *Int. J. Geographical Information Science*, 13, 143-158

Stockwell D.R.B. & Peterson A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148, 1-13

Strauss B. & Biedermann R. (2007) Evaluating temporal and spatial generality:How valid are species–habitat relationship models? *Ecological Modelling*, 204, 104-114

Suding K.N., Lavorel S., Chapin F.S., Cornelissen J.H.C., Diaz S., Garnier E., Goldberg D., Hooper D.U., Jackson S.T. & Navas M.L. (2008) Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology*, 14, 1125-1140

Swets K.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, 240, 1285 - 1293

Teixeira J., Ferrand N. & Arntzen J.W. (2001) Biogeography of the golden-striped salamander *Chioglossa lusitanica*: a field survey and spatial modelling approach. *Ecography*, 24, 618-624

Thomas P.A., El-Barghathi M. & Polwart A. (2007) Biological Flora of the British Isles: *Juniperus communis L. Journal of Ecology*, 95, 1404-1440

Thuiller W., Brotons L., Araujo M.B. & Lavorel S. (2004a) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27, 165-172

Thuiller W., Lafourcade B., Engler R. & Araujo M.B. (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373

Thuiller W., Lavorel S., Araújo M.B., Sykes M.T. & Prentice I.C. (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences, USA*, 102, 8245-8250

Thuiller W., Lavorel S., Midgley G., Lavergne S. & Rebelo T. (2004b) Relating plant traits and species distributions along bioclimatic gradients for 88 Leucadendron taxa. *Ecology*, 85, 1688-1699

Thuiller W., Lavorel S., Sykes M.T. & Araújo M.B. (2006) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, 12, 49-60

Tilman D. (2001) Functional diversity. In: *Encyclopedia of Biodiversity* (ed. Levin SA), p. 109–120. Academic Press, San Diego, CA

Tonne F. (1954) *Besser bauen mit besonnungs- und tageslicht-planung*, Hofmann, Schorndorf.

Treves A. & Karanth K.U. (2003) Human-carnivore conflict and perspectives on carnivore management worldwide. *Conservation Biology*, 17, 1491-1499

Tscharntke T., Klein A., M., Kruess A., Steffan-Dewenter I. & Thies C. (2005) Landscape perspectives on agriculturla intensification and biodiversity-ecosystem service management. *Ecology Letters*, 8, 857-874

Vaieretti M.V., Diaz S., Vile D. & Garnier E. (2007) Two measurement methods of leaf dry matter content produce similar results in a broad range of species. *Annals of Botany*, 99, 955-958

Van Horn B. (2002) Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. In: *Predicting species occurrences: issues of accuracy and scale* (eds. Scott JM, Heglund PJ, Morrison ML, Haufler JB, Raphael MG, Wall WA & Samson FB). Island Press, Covelo, California

VanderWal J., Shoo L.P., Graham C. & William S.E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, 220, 589-594

Vieilledent G. (2009) Structurer l'incertitude et la variabilité dans les modèles de dynamique forestière - Application à la coexistence du Sapin et de l'Epicéa en forêt de montagne. In. AgroParisTech-ENGREF, Grenoble

Vile D., Shipley B. & Garnier E. (2006) A structural equation model to integrate changes in functional strategies during old-field succession. *Ecology*, 87, 504-517

Villéger S., Mason N.W.H. & Mouillot D. (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290-2301

Violle C. & Jiang L. (2009) Towards a trait-based quantification of species niche. *Journal of Plant Ecology-Uk*, 2, 87-93

Violle C., Navas M.L., Vile D., Kazakou E., Fortunel C., Hummel I. & Garnier E. (2007) Let the concept of trait be functional! *Oikos*, 116, 882-892

Wegner C., Wunderlich M. & Kessler M. (2003) Foliar C:N ratio of ferns along an Andean Elevational Gradient. *Biotropica*, 35, 486-490

Weiher E., van der Werf A., Thompson K., Roderick M., Garnier E. & Eriksson O. (1999) Challenging Theophrastus: a common core list of plant traits for functional ecology. *Journal of Vegetation Science*, 10, 609-620

Welker J.M., Molau U., Parsons A.N., Robinson C.H. & Wookey P.A. (1997) Responses of Dryas octopetala to ITEX environmental manipulations: a synthesis with circumpolar comparisons. *Global Change Biology*, 3, 61-73

Westman W.E. (1991) Measuring realized niche spaces: climatic response of chaparral and coastal sage scrub. *Ecology*, 72, 1678-1684

Westoby M. (1998) A leaf-heigh-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, 199, 213-227

Westoby M., Falster D.S., Moles A.T., Vesk P.A. & Wright I.J. (2002) Plant ecological strategies: Some leading dimensions of variation between species. *Annual Review of Ecology and Systematics*, 33, 125-159

Wiens J.J. & Graham C.H. (2005) Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution and Systematics*, 36, 519-539

Wilson P.J., Thompson K. & Hodgson J.G. (1999) Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. *New Phytologist*, 143, 155-162

Woodward F.I. (1986) Ecophysiological studies on the shrub *Vaccinium myrtillus* L. taken from a wide altitudinal range. *Oecologia*, 70, 580-586

Woodward F.I. & Diament A.D. (1991) Functional approaches to predicting ecological effects of global change. *Functional Ecology*, 5, 202-212

Wright I.J., Reich P.B., Cornelissen J.H.C., Falster D.S., Groom P.K., Hikosaka K., Lee W., Lusk C.H., Niinemets U., Oleksyn J., Osada N., Poorter H., Warton D.I. & Westoby M. (2005) Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography*, 14, 411-421

Wright I.J., Reich P.B., Westoby M., Ackerly D.D., Baruch Z., Bongers F., Cavender-Bares J., Chapin T., Cornelissen J.H.C., Diemer M., Flexas J., Garnier E., Groom P.K., Gulias J., Hikosaka K., Lamont B.B., Lee T., Lee W., Lusk C., Midgley J.J., Navas M.L., Niinemets U., Oleksyn J., Osada N., Poorter H., Poot P., Prior L., Pyankov V.I., Roumet C., Thomas S.C., Tjoelker M.G., Veneklaas E.J. & Villar R. (2004) The worldwide leaf economics spectrum. *Nature*, 428, 821-827

Wright I.J. & Westoby M. (2001) Understanding seedling growth relationships through specific leaf area and leaf nitrogen concentration: generalisations across growth forms and growth irradiance. *Oecologia*, 127, 21-29

Wright J.W., Davies K.F., Lau J.A., McCall A.C. & McKay J.K. (2006) Experimental verification of ecological niche modeling in a heterogeneous environment. *Ecology*, 87, 2433-2439

Wright S.J. & Muller-Landau H.C. (2006) The future of tropical forest species. *Biotropica*, 38, 287-301

Xu R.H. (2003) Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, 22, 3527-3541

Zimmermann N.E. & Kienast F. (1999) Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *Journal of Vegetation Science*, 10, 469-482

Zvereva E.L., Toivonen E. & Kozlov M.V. (2008) Changes in species richness of vascular plants under the impact of air pollution: a global perspective. *Global Ecology and Biogeography*, 17, 305-319

Annexes

Liste des annexes

Article 1
Albert C.H., Thuiller W., Lavorel S., Davies I., Garbolino E., *Land-use change and subalpine tree dynamics: colonization of Larix decidua in French subalpine grasslands*. Journal of Applied Ecology, 2008, 45: 659-669.

Article 2
Thuiller W., **Albert C**., Araujo M.B. , Berry P.M., Cabeza M., Guisan A., Hickler T., Midgley G. , Paterson J., Shurr F.M., Sykes M.T., Zimermann N.E., *Predicting global change impacts on plant species' distributions: Future challenges.* Perspectives in Plant Ecology, Evolution and Systematic, 2008, 9 (3-4):137-152.

Article 3
Albert C.H. & Thuiller W., *Favourability functions versus probability of presence: advantages and misuses.* Ecography, 2008, 31: 417-422.

Article 5
Midgley G. F., Davies I. D., **Albert C. H.,** Altwegg R., Hannah L., Hughes G. O., Ries L. P. & Thuiller W. *BioMove – an integrated platform simulating the dynamic response of species to environmental change*. Sous presse à Ecography

Article 6
Thuiller W.*, **Albert C.H.*,** Dubuis A., Randin C., Guisan A., *Variation in habitat suitability models does not always relate to variation in species' plant functional traits.* Sous presse à Biology Letters: doi:10.1098/rsbl.2009.0669). *Contributions équivalentes des deux premiers auteurs.

Article 7
Albert C.H., Thuiller W., Yoccoz N.G., Soudan A., Boucher F., Saccone P., Lavorel S. *Intraspecific functional variability: extent, structure and sources of variation within an alpine valley.* (En revision mineure à Journal of Ecology)

Article 8
Albert C.H., Thuiller W., Dubuis A., Guisan A., Yoccoz N.G., Aubert S., Douzet R., Lavorel S., *Intraspecific functional variability: on its relative importance in respect to interspecific variability.* (Soumis à Ecology)

Article 9
Albert C.H., Yoccoz N.G, Graham C., Edwards T., Zimmermann N.E., Thuiller W. *Sampling in ecology and evolution - integrating questions and designs.* (Soumis à Ecography)

Journal of Applied Ecology 2008, **45**, 659–669 doi: 10.1111/j.1365-2664.2007.01416.x

Land-use change and subalpine tree dynamics: colonization of *Larix decidua* in French subalpine grasslands

Cécile H. Albert[1,2]*, Wilfried Thuiller[1,2], Sandra Lavorel[1,2], Ian D. Davies[3] and Emmanuel Garbolino[4]

[1]*Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France;* [2]*Station Alpine Joseph Fourier, CNRS UMS 2925, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France;* [3]*School of Resources, Environment and Society, Australian National University, Canberra ACT 0200, Australia;* [4]*Ecole des Mines de Paris – Pole Cindyniques BP 207 – F06904 Sophia Antipolis, France*

Summary

1. In many places in Europe, trees are currently far below their theoretical altitudinal limit because of past land use. However, under the current crisis affecting mountain agriculture, trees are likely to recolonize the subalpine belt. Addressing the societal issues associated with such landscape change requires predictive tools to assess vegetation dynamics in relation to management strategies. This study aims to analyse the factors determining tree dynamics at the subalpine ecotone and to evaluate the impact of land-use change on landscape vegetation patterns.

2. We developed a hierarchical scaling approach and applied it to a European larch (*Larix decidua* Mill.) colonization in the French Alps. Our landscape case study focused on subalpine south-facing grasslands of Villar d'Arêne (France). First, we used a habitat-suitability model to delineate the climatic suitability habitats of larch in the French Alps. Then we used a landscape model (LaMoS) to determine whether larch is able to colonize these grasslands predicted to be climatically suitable, considering land use and local vegetation. The sensitivity of landscape colonization patterns to land-use scenarios (abandonment, delayed mowing or traditional mowing) and biological attributes of larch were analysed with a factorial simulation experiment.

3. The accuracy of the habitat-suitability model at the French Alps scale allowed for the prediction that the grasslands of Villar d'Arêne were highly suitable for larch.

4. Simulations highlighted the effect of land use on larch establishment at the local but not at the landscape scale. On the other hand, larch attributes such as dispersal capacity, juvenile tolerance to light interaction (competition and facilitation) and its capacity to survive on infertile soil were shown as essential for tree spatial dynamics.

5. *Synthesis and applications*. Combining a habitat-suitability model with a spatially and temporally explicit landscape model enhances the capacity to account for environmental factors acting at different overlapping scales. This modelling strategy increases the robustness and accuracy of predictions, a prerequisite for landscape management in a global change context. We conclude that ongoing and future agri-environmental policies have to be quickly adapted to protect biodiversity and ecosystem services provided by subalpine grasslands.

Key-words: climate, habitat-suitability modelling, landscape dynamics model, landscape management, mowing

Introduction

The reality of rapid anthropogenic global changes such as increasing temperatures, elevated carbon dioxide concentration, nitrogen deposition and land-use change is nowadays accepted, and these changes are expected to affect plant and animal species. Global warming appears to have already affected the ranges of species (Parmesan & Yohe 2003), with loss, expansion or relocation of habitats sometimes leading to extinctions, and many recent studies have forecast its widespread effects

*Correspondence author. E-mail: cecile.albert@m4x.org

© 2007 The Authors. Journal compilation © 2007 British Ecological Society

on biodiversity and ecosystem functioning (Diaz *et al.* 2005). However, land-use change is expected to be the main influence on terrestrial ecosystems for the coming decades, followed by climate change (Sala *et al.* 2000). The direct effect of these two drivers on species distributions is not yet well understood, and interactions between them represent one of the largest uncertainties in projections of future biodiversity change (Sala *et al.* 2005).

Among the various biogeographical regions, arctic and alpine ecosystems are certainly the most temperature-constrained and thus are expected to be particularly sensitive to climate warming (Grabherr 2003). With increasing temperatures and changed patterns in snow cover, plants are expected to adapt their life cycles or to shift their ranges to more suitable habitats upwards and polewards (Walther *et al.* 2005), which is supported by accumulating observations. Walther *et al.* (2005) showed an increase in species richness on Swiss high summits between 1912 and 2003, with a mean increase in the number of species of 138%. Woody colonization at the tree line is an expected response to global changes. This climatically constrained limit (Körner 1999), separating subalpine forests from alpine meadows (≈2300 m in the Alps), is expected to be shifted upslope as global warming provides an increase in suitable habitats toward higher altitudes (Kullman 2002; Holtmeier & Broll 2005; Coop & Goivnish 2007). However, to the best of our knowledge there is no clear trend at the landscape scale of trees moving upslope (Holtmeier & Broll 2005). As land use is the key driver of tree colonization into man-made grasslands below the tree line (Rickebusch 2006), it is difficult to distinguish climate-related signals from a response to land-use change, beyond land-use legacies (Didier 2001) and atmospheric CO_2 concentration (Körner 2003). Tree colonization into grasslands is also influenced by dispersal and biotic processes (competition and facilitation) (Li & Wilson 1998; Dullinger *et al.* 2004; Kunstler *et al.* 2006).

In Europe, trees were cut long ago to make way for agricultural use, and grasslands have been prevalent on south-facing slopes since at least the 16th century (Quétier *et al.* 2007), with land-use practices precluding tree establishment. Today, decreased management intensity and land abandonment foster quick grassland recolonization (Didier 2001). Landscape changes brought on by tree encroachment into former hay meadows threaten the remarkable diversity of plants, birds, insects and mammals hosted by these habitats (Vandvik *et al.* 2005), and also alter the delivery of many other ecosystem services (Tappeiner *et al.* 2003): negatively, as in the case of recreation or water availability, or positively, for example through carbon sequestration or slope and snow stability. Management choices to move landscapes towards a desirable future, whether forested or not, therefore require forecasts of the future ranges of species under different scenarios of environmental change (Milsom *et al.* 2000).

Three broad types of model are used to assess the likely impacts of global changes on vegetation dynamics. At a global-to-continental scale, dynamic vegetation models are based on averaged processes and are therefore of limited use for investigating species dynamics at the landscape scale (Sitch *et al.* 2003). At a regional scale, statistical habitat-suitability models are useful to quantify species–environment relationships, but do not directly include ecological processes, and are static (Guisan & Thuiller 2005). Finally, at the landscape scale, mechanistic forest-gap models or succession models are usually able to cope with spatial dynamics (dispersal), biotic interactions and survival, but do not usually include the direct influence of environmental gradients such as climate or CO_2 on the dynamics (Cousins *et al.* 2003; Schumacher *et al.* 2004).

Here we aim to analyse the factors determining tree dynamics at the subalpine ecotone and to evaluate the impact of land-use change on landscape vegetation patterns. Combining the strength of existing models at the regional and landscape scales, we developed an original approach combining hierarchically two types of model, in the case study of the European larch (*Larix decidua* Mill.) in the French Alps, and a focus on subalpine south-facing grasslands of Villar d'Arêne (France). First assuming that macroclimate drives the regional distribution of plant species (Woodward & Beerling 1997), we defined the climatically suitable habitats for European larch in the French Alps using a habitat-suitability model (Guisan & Zimmermann 2000). This model gave a potential distribution representing Hutchinson's realized niche for larch without directly including biotic and biological factors that might prevent a species from occupying its entire suitable habitat (e.g. competitive exclusion, dispersal limitation). Then, scaling down to the landscape level in one of these climatically suitable sites, Villar d'Arêne south-facing grasslands, we determined whether larch could colonize those grasslands under three land-use scenarios, considering the dispersal ability of larch and its ability to compete against the local vegetation. For this purpose we used a landscape dynamic succession model (LaMoS Landscape Modelling Shell; Cousins *et al.* 2003), parameterized using both literature and experimental data. The sensitivity of landscape colonization patterns to land-use scenarios (abandonment, delayed mowing or traditional mowing) and biological attributes of larch was analysed using a factorial simulation experiment. The results are discussed considering the relevance of such a modelling approach for guiding management in the context of global change.

Materials and methods

STUDY SPECIES

Larix decidua (European larch) is a deciduous tree of montane forests and montane-subalpine grasslands (500–2600 m). European larch is often described as a heliophile pioneer species (Rameau *et al.* 1993). Its establishment is quickly followed by trees such as *Pinus cembro* and *Pinus uncinata* in the high subalpine belt, *Picea abies* and *Abies alba* in the low subalpine belt, and *Fagus sylvatica* below.

The life span of larch can reach 400 years (Carrer & Urbinati 2004), but we used a mean value of 200 years, which is enough to contrast with the life span of graminoids. Larch becomes mature after 20–30 years and has mast seeding years approximately every

Table 1. Eleven environmental variables: 10 bioclimatic variables from the meteorological model Aurelhy at a resolution of 200 m × 200 m and one soil texture variable from the *European soil database*

Type	Variable	Five PCA-selected variables
Bioclimatic (200 × 200 m)	Relative humidity	×
	Solar radiation	×
	Vapour pressure deficit	
	Degree days above 5 °C (°C)	×
	Number of frost events	×
	Total precipitations of the year (mm)	
	Summer precipitations (mm)	×
	Average maximal temperature (°C)	
	Average minimal temperature (°C)	
	Minimal temperature in winter (°C)	
Soil (1 × 1 km)	Soil texture	

10 years, with seed life span ranging from 3 to 7 years. Seeds are mostly dispersed by wind and birds (Rameau *et al.* 1993).

A HIERARCHICAL APPROACH AT TWO SCALES

Assuming that climate is driving the regional distribution of plant species, and that dispersal, biotic interactions and land-use change are the most important factors determining landscape-scale distribution, we developed an original approach based on a hierarchical approach at two scales. The first model was applied at the regional scale (French Alps) to quantify the ecological niche of larch, and then by GIS mapping to identify areas of high habitat suitability. The second model investigated at the landscape scale whether larch could migrate and establish in interaction with the current vegetation (grass-dominated) according to different scenarios of land-use change.

Regional-scale suitability

To delineate the niche space of larch, we used habitat-suitability modelling (Guisan & Zimmermann 2000), which links species occurrence data with selected environmental factors (Austin 2002).

We used presence and absence data for larch from the botanical database SOPHY (http://sophy.u-3mrs.fr/sophy.htm; Brisse *et al.* 1995), which contains more than 3 million observations of plants in France for 4500 taxa. The 2156 plots of *L. decidua* (375 observed presences and 1781 observed absences) in the French Alps were used for this study. Each plot is characterized by its geographical coordinates with an accuracy between 100 m and 1 km, the altitude of the plot and the cover-abundance index of the plant (from 1 to 6), according to the Braun–Blanquet methodology (Mueller-Dombois & Ellenberg 1974).

To describe the environmental niche of larch, we used 11 variables (Table 1): a set of bioclimatic variables obtained from a digital elevation model and from the meteorological model Aurelhy (Benichou & Le Breton 1987), based on interpolated measurements at a resolution of 200 m × 200 m and soil texture data from the *European soil database* (http://eusoils.jrc.it/ESDB/index.htm) at a resolution of 1 km × 1 km. A condensed set of five variables (relative humidity, summer precipitation, number of frost events, solar radiation, number of degree days above 5 °C) was determined using principal components analysis to limit the collinearity within the environmental drivers.

Among the species distribution models commonly used, generalized linear models (GLM) and generalized additive models (GAM) have proven to be one of the best compromises between interpretability and predictability (Guisan & Thuiller 2005).

Both GLM and GAM models were implemented with a LOGIT link function, specifically for binomial variables such as presence/absence of a species. We first used GAM in an exploratory way, to identify the most probable shapes of the response curves, which were then fitted within a GLM. The shapes of the response curves were identified using a cubic spline smoother and a two-sided stepwise procedure based on the Akaike information criterion (AIC; Akaike 1974) to choose the most parsimonious model. We then estimated the mathematical formula (linear, quadratic or polynomial) corresponding to the response curve of each significant explanatory variable, and built a GLM with those terms. The goodness of fit was estimated by a measure of deviance reduction (adjusted D^2) given by (Guisan & Zimmermann 2000):

$$\text{adjusted } D^2 = 1 - [(n-1)/(n-p)] \times [\text{residual deviance/null deviance}], \quad \text{eqn 1}$$

where n is the number observations and p is the number of parameters.

The model was calibrated on a random fraction of the original data (70%) and its predictions were tested on the remaining data (30%), assuming the two random fractions to be quasi-independent. In order to assess the predictive capacity of the model, model predictions were compared with real observations using the area under the curve (AUC) of a receiver-operating characteristics plot (Swets 1988; Fielding & Bell 1997). This AUC measure allows us to test whether the obtained predictions differ significantly from a random prediction. We used the following rough guide (Swets 1988): poor predictions for AUC were contained in the range 0·5–0·7; useful predictions for AUC in the range 0·7–0·9; and good to excellent predictions when AUC is >0·9. A threshold was then calculated to optimize the separation in a contingency table of the true presences and true absences, and the outputs of the model were then transformed with this threshold into binary results for presences and absences.

The final model gave the habitat suitability index of the species as a function of abiotic conditions and allowed us to determine geographically potentially highly suitable locations. We projected the model over the French Alps, with zeros where the index was below the selected threshold and the nough value of the index where

it was above the threshold. Zooming down allowed us to observe the local habitat-suitability patterns and to determine high-suitability areas for larch.

Landscape dynamics and land-use change scenarios

Case study site

The Villar d'Arêne study area (45·04°N, 6·34°E; altitude 1650–2500 m) is located in the central French Alps, on a south-facing slope near the Lautaret pass. This area covers 1292 ha characterized by subalpine climatic conditions with mean temperatures reaching –7·4 °C in February and +19·5 °C in July, total annual precipitation averaging 956 mm, and prevailing winds mainly coming from the west.

This site was selected from the outputs of the regional model as a typical location within the subalpine belt with high habitat suitability for larch, but still devoid of trees. Its vegetation patterns, floristic composition and land-use dynamics are well documented, making it an ideal site for detailed landscape modelling.

Until the 12th century, the area used to be covered with a dominant pine species (*P. uncinata*) before trees were removed for agriculture and grazing (Ali *et al*. 2003). The site was then terraced and ploughed in its lower section and mown and/or grazed in its upper section. Cropping was abandoned after World War II, and management then involved summer mowing with or without fertilization on both terraced and unterraced fields. More recently, mowing has been abandoned in the steeper and least accessible parts of the landscape. Nowadays grazing pressure has decreased and a smaller area is still mown for conservation purposes (NATURA 2000 directive; Quétier *et al*. 2007).

Current land cover is exclusively grasslands, dominated by four grass functional types with high local abundance and, respectively, represented by the most common species: *Bromus erectus*, *Dactylis glomerata*, *Festuca paniculata* and *Sesleria caerulea.* Their landscape-scale distribution and dominance patterns are determined by past and current land use (Quétier *et al*. in press). The main tree species in Villar d'Arêne is *L. decidua*, which is colonizing the north-facing slope; however, few individual trees currently occur in the grassland area with a patchy distribution.

Landscape dynamics model

Once a climatically suitable site had been chosen, a second approach at the landscape scale aimed to study the potential establishment of larch in the existing communities according to various scenarios of land-use change. To analyse larch dynamics on a landscape scale, including biotic and biological processes, we used a landscape dynamics model: LaMoS (Cousins *et al*. 2003). LaMoS is based on the concept of plant functional type (PFT; Lavorel *et al*. 1997), defined as a group of species with similar responses to and/or effects on their environment. It assumes that abiotic conditions are suitable for each PFT in the area where they are studied.

LaMoS is a spatially and temporally explicit model that accounts for basic vegetation dynamic processes resulting from the interactions between plant functional types (or species), habitat conditions, disturbances and spatial patterns (Cousins *et al*. 2003). Within LaMoS, the landscape is described as a raster grid, in which landscape dynamics is modelled by three interacting modules. First, a succession model drives within-pixel yearly successional dynamics. We used a modification of the FATE model (Moore & Noble 1990), which determines the abundance of competing PFTs represented as age cohorts based on a simple set of traits relating to plant life history, tolerance to light interaction at different life stages, and recruitment (for details see Cousins *et al*. 2003; Quétier *et al*. in press) (Table 2). Second, a dispersal model distributes seeds across the landscape. We used a negative exponential model (Table 2) coupled with a random long-distance dispersal function (Table 2), which is also exponentially correlated with distance to seed source but allows a very small (10^{-4}) fraction of seeds to be dispersed much further (1·8 km). Third, a disturbance model establishes the potential disturbed fraction of abundance (output from the succession model) for each PFT and affects each PFT specifically, leading to death, resprouting or no effect on the different defined age classes. We used a uniform mowing disturbance within fields still in use under various land-use scenarios.

Villar d'Arêne grasslands have already been modelled with LaMoS, including four leaf-height-seed (LHS)-based PFTs, classifying the dominant graminoids in contrasting nutrient economies and competitive effects and responses after Westoby's (1998) LHS plant strategy scheme. This method uses three plant traits (leaf area, plant height and seed weight) to determine species' strategies for growth, persistence and regeneration. The four LHS-based PFTs resulting from field work, named after the archetype species in each cluster, are Bromus PFT, representing medium-sized, disturbance-tolerant species; Dactylis PFT, representing tall, acquisitive species; Festuca PFT, representing tall, competitive and conservative species; and Sesleria PFT, representing small-sized, stress-tolerant species (unpublished data). The model including the four contrasting grass PFTs was calibrated under past conditions, run until the present with historical land use, and validated under current conditions.

The four PFTs are thought to have contrasting nutrient economies and contrasting competitive effects and responses, and were parameterized with both field and experiment data.

Future land-use scenarios

We added to this model a tree PFT, parameterized to represent larch. Most of the necessary parameters were found in the literature, but germination rates under different light levels, representing light interaction (competition and facilitation), were determined with a controlled experiment (see below). Three sensitive parameters remained undefined and were tested with simulations: dispersal ability; resource-uptake efficiency, which determines the demographic response to pixel fertility; and juvenile response to light interaction.

To test the impact of land-use change on tree dynamics, we used three scenarios of land-use change representing local land-use projections by 2050, developed by Quétier (2006). These were based on local projections of Europe-wide scenarios of land-use change (Rounsevell *et al*. 2006), formulated on the basis of the Intergovernmental Panel on Climate Change (IPCC) global storylines (Nakicenovic & Swart 2000). The first scenario represents a global, economically growing world, leading to land abandonment because isolated mountain lands are no longer profitable. The second scenario represents a global, sustainably developing world. It leads to delayed mowing, maintained in isolated regions only for environmental purposes as a way of protecting biodiversity, but this non-profitable activity is delayed until the end of summer (August). The third scenario represents a regional, sustainably developing world, leading to the persistence of traditional agriculture. The environmental value and regional identity of Villar d'Arêne allow protection of the strong local agriculture, including the early (June or July) mowing of grasslands. This early mowing disturbance prevents the Festuca PFT from producing seeds (Quétier *et al*. in press).

Table 2 LaMoS parameterization for larch species

Parameter	Value
(a) Life parameters	
Low resource threshold*	0·20 or 0·60
Maximum biomass	10 000 t ha^{-1}
Maximum abundance	8000 trees ha^{-1}
Life of active pool	5 years
Immature stratum	Medium
Mature stratum	High
Time to maturity	30 years
Maximum abundance	Leads to a medium extinction of light
Potential fecundity	200 000 seeds per tree every 10 years
Life span	200 years
Stage tolerance to available light:	
Germinants	Medium and high light level
Juveniles*	High light level or medium and high light level
Matures	High light level
Germination rate for different light availability**:	
Low light level	High
Medium light level	Medium
High light level	Low
(b) Dispersal and disturbance response parameters	
Long-distance dispersal*	Yes or no
Long-distance dispersal fraction	0·0001
Mean dispersal distance	0·4
Maximum long-distance throw	1·8 km
Type of disturbance	Constant severe harvest (60% of biomass)
Frequency	Once a year
Response of larch:	
Germinants	All killed
Juveniles	All killed
Matures	All killed

*The three parameter tested with the simulations.
**Parameter of germination rate determined with the controlled experiment.

Simulations

The LaMoS model has already been shown to be sensitive to several parameters, including seed dispersal, germination rate, resource-uptake efficiency, and juvenile tolerance to light availability (Cousins *et al.* 2003; Grigulis *et al.* 2005; Quétier *et al.* 2007). Among these key parameters only germination was estimated by experiments, the others being tested with simulation runs. We used a factorial simulation design to test the sensitivity of larch colonization to three of the critical auto-ecological parameters: with vs. without long-distance seed dispersal ability (Table 2); shade-tolerant vs. intolerant juveniles (Table 2); and high vs. low resource-uptake efficiencies (Table 2), under the three scenarios of land-use change (A1 = abandonment, B1 = late mowing, B2 = early mowing) (Fig. 2b), yielding a total of 24 different simulations. Simulations were run on a 43 × 75-pixel (90 × 90-m) landscape representing the altitudinal fertility gradient at the site (five values from 1 at the bottom to 0·2 at the top), with a maximal value (fertility = 1) for formerly cropped and fertilized grasslands (Fig. 2a) (Quétier *et al.* in press). Land-use scenarios were translated into disturbance maps, contrasting areas with annual mowing and abandoned areas where successional development proceeded undisturbed. Each simulation was run for 230 years post-land-use change to observe at least one complete generation of trees, using current conditions as the initial state (Fig. 2c), with a line of trees with maximum abundance introduced on the north-facing slope of Villar d'Arêne (initial conditions for larch; personal observations). We recorded maps of mature abundance of each PFT every 30 years. Simulations were not repeated because the model was stable and deterministic.

LARCH SEED GERMINATION UNDER DIFFERENT LIGHT CONDITIONS

We quantified the effect of light conditions on seed germination using a controlled experiment. Our experiment was set up under unlimited conditions of water supply, because both light and water are usually expected to influence the germination rate (Broncano *et al.* 1998). We used seeds of *L. decidua* from the northern internal Alps, provided by the French Office National des Forêts. After the seeds had been stored in cold conditions (2–3 °C) from 8 March to 4 April 2006, on 5 April they were placed in five pots, which received three different light treatments (measured light: 100, 36 or 0% of total full sunlight) applied with shade cloth to (partially) shade the pots containing 50 seeds. The seeds were monitored every 2 days until 8 May, and the first root's appearance was recorded as the sign of a successful germination.

All statistical models and analyses were carried out with the free-source software R (ver. 2·3·1; Ihaca & Gentleman 1996). All map treatments were realized using ARCGIS ver. 9·1 (ESRI, Redlands, CA, USA).

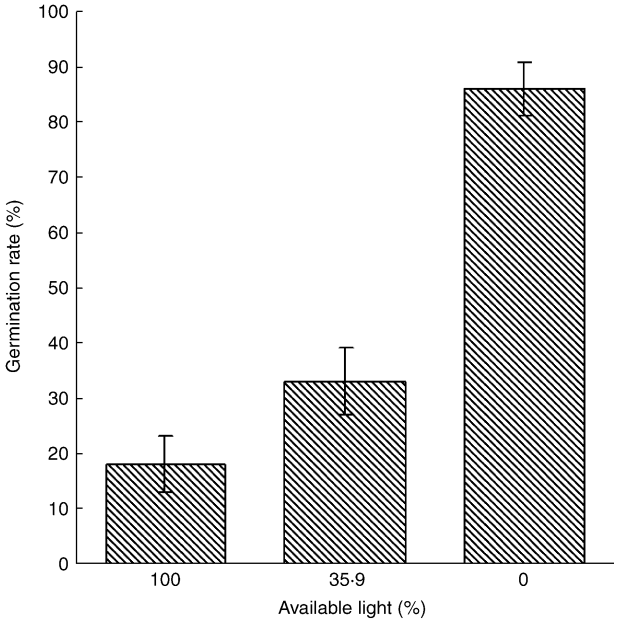


Fig. 1. Effects of light availability on germination rate of *Larix decidua*.

Results

PARAMETERIZATION OF GERMINATION WITH EXPERIMENTAL RESULTS

There was a negative relationship between light availability and germination (Fig. 1), showing the relatively high ability of larch seeds to germinate under a completely dark cover, although saplings needed light to grow subsequently. This experiment allowed us to parameterize the germination rate in LaMoS under three different light levels (low, medium, high) as required by the LaMoS model (Table 2).

HABITAT SUITABILITY FOR LARCH IN THE FRENCH ALPS

The stepwise GAM identified that the response of species to the selected environmental variables was quadratic. Soil type was not selected by the stepwise procedure as a pertinent variable to explain larch distribution.

We then built a GLM using the variables selected by the stepwise GAM and the predetermined parametric shapes. The explanatory and predictive powers of the final model were fairly high (D^2 = 43%, AUC evaluation = 0·92, AUC calibration = 0·92).

The geographical projection (Fig. 3) showed that under current climatic conditions, larch was predicted potentially to occupy a relatively narrow altitudinal range within mountainous areas. A zoom of the projection identified the south-facing grasslands of Villar d'Arêne as fairly suitable for larch (Fig. 2d,e). The binary-converted results (threshold = 0·189) led us to consider the high part of the grasslands as unsuitable for larch.

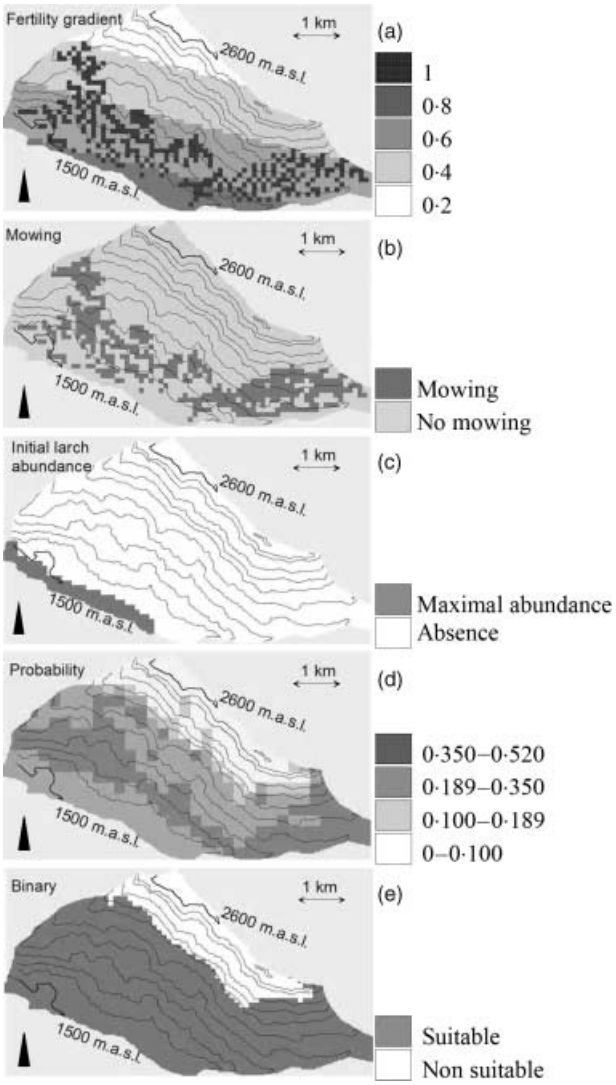


Fig. 2. Map layers of the grasslands of Villar d'Arêne along an altitudinal gradient: (a) relative soil fertility from 0·2 to 1. (formerly fertilized grasslands = dark patches); (b) current mowing disturbance; (c) current location of *L. decidua*; (d) larch habitat suitability (200 m× 200 m); (e) binary larch potential distribution (0 below the threshold of 0·189 and 1 above).

EFFECTS OF AUTOECOLOGY AND LAND-USE CHANGE ON LARCH DYNAMICS AT THE LANDSCAPE SCALE

The low part of Villar d'Arêne grasslands was potentially suitable for larch. We then determined whether larch could colonize grasslands by dispersing seeds from its current distribution area on the north-facing slope, germinating and recruiting in local vegetation under the effects of soil-fertility patterns and different land-management regimes (mowing vs. abandonment) (Fig. 4). The LaMoS simulations allowed us to highlight key drivers for the ability of larch to establish in Lautaret grassland. Juveniles appeared outside the initial population source after 20 years, and mature trees about 50 years after land-use change.

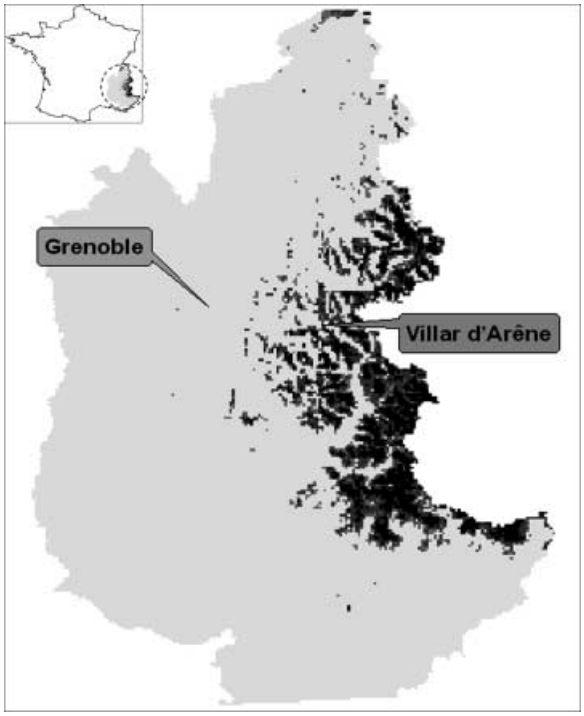


Fig. 3. Larch suitability map at the regional scale (French Alps). Dark to light, from high to low suitability index.

Simulations showed that land use partly drives larch distribution. In particular, mown grasslands were devoid of larch because the mowing disturbance directly prevented trees from growing. Further, if larch had a low dispersal ability (no long-distance dispersal), mowing dramatically reduced colonization by larch (results not presented here), potentially leading to local extinction (see Appendix S1 in Supplementary material).

However, the ecological attributes of larch appear to be of crucial importance in predicting its potential colonization into the subalpine landscape. Maps of mature tree abundance showed three main landscape patterns. (1) When larch had high resource-uptake efficiency, shade-tolerant juveniles and a long-distance dispersal capacity, colonization was efficient over the whole landscape, even in poor soils at higher altitude, and was nearly independent of land use (Fig. 4b,e,h). (2) With shade-tolerant juveniles, but no long-distance dispersal ability, colonization was patchy, with two different patterns. Either larch had low resource-uptake efficiency, so that patches coincided with fertilized meadows (see Appendix S1), or larch had high resource-uptake efficiency and was limited only by biotic interactions (Fig. 4a,d,g). (3) With shade-intolerant juveniles and no long-distance dispersal, larch colonization was limited to a few individuals and a small spatial extent (see Appendix S1), or led to local extinction; on the other

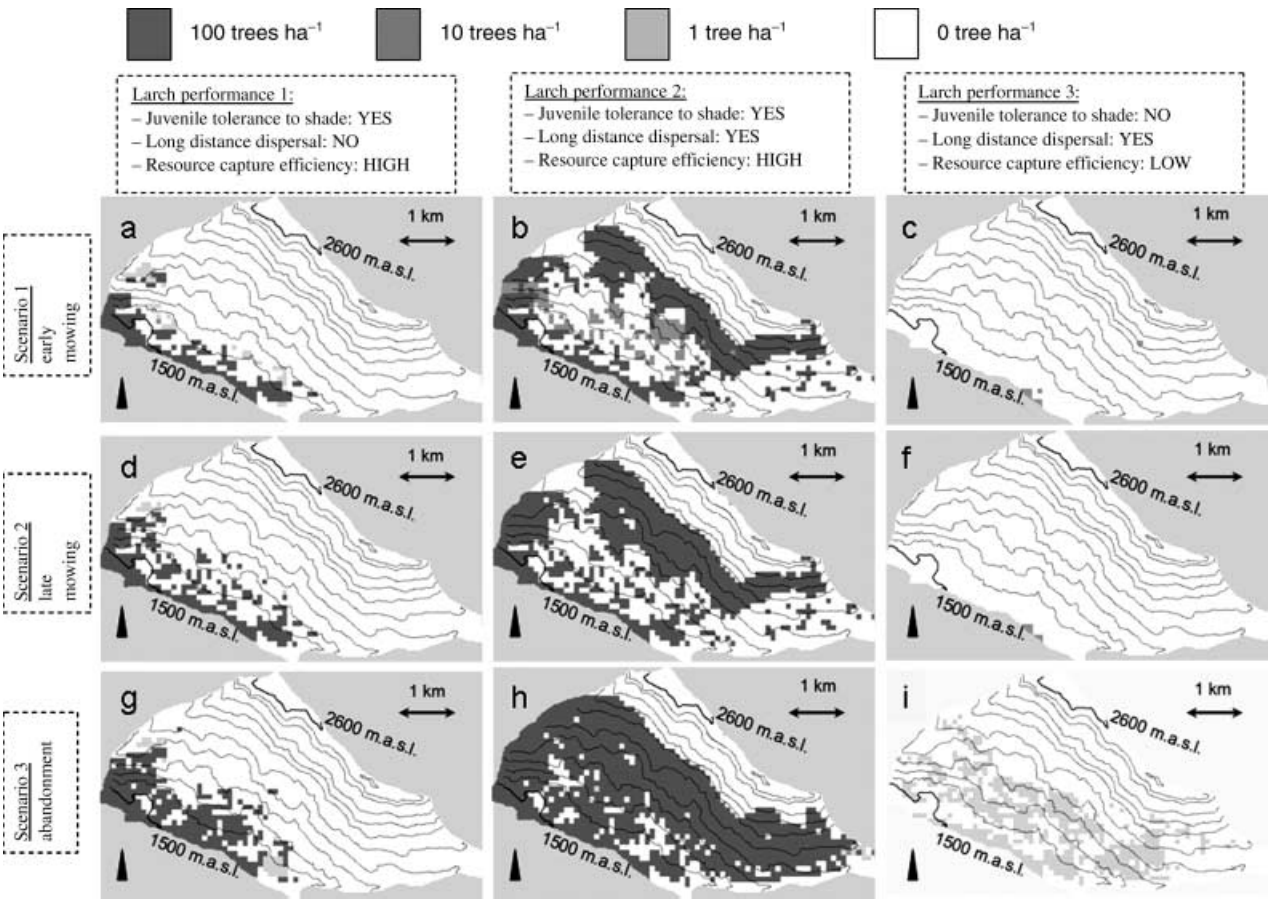


Fig. 4. Maps of abundance of mature larch trees and patterns of grassland colonization by *Larix decidua* along the altitudinal gradient. Only nine maps from 24 are presented here (see Appendix S1 for the remainder).

© 2007 The Authors. Journal compilation © 2007 British Ecological Society,

hand, with shade-intolerant juveniles and a long-distance dispersal ability, colonization was spatially broad but limited to few individuals, or also led to local extinction (Fig. 4c,f,i).

Discussion

MODELLING THE SIMULTANEOUS EFFECTS OF CLIMATE AND LAND-USE CHANGE ON LANDSCAPES

The new approach proposed here, relying on combined models at the regional and landscape scales, offers several assets for predicting future landscape dynamics. First, the use of habitat-suitability modelling as a first step fulfils the most important assumption of LaMoS, which is the suitability of the study site for modelled species. Including both abiotic and biotic conditions in landscape dynamics thereby increases the realism of projections. Second, this method combines two different study scales, and uses state-of-the-art models at each scale. As a first step, habitat-suitability models can be established at regional or global scales, given a suitable sample size of presence records to train the model reliably (Stockwell & Peterson 2002). Under these conditions one of the most important assumptions of these models can be fulfilled: pseudo-equilibrium between the species of interest and the current environment (Guisan & Thuiller 2005). This does not mean that the species occupies its whole niche (Pulliam 2000), because habitat-suitability models provide a geographical representation of the entire niche, including truly occupied locations and those that could be occupied but are not currently because of competitive exclusion, land use or low dispersal ability (Pulliam 2000). The Villar d'Arêne landscape is an example of such a climatically suitable but unoccupied locality. As a second step, landscape models such as LaMoS are appropriate for a local analysis on a well known site identified as climatically suitable. This approach is more parsimonious than attempting to run landscape models at large scales, a time-consuming exercise with numerous parameters that are difficult to obtain at a large scale (e.g. former land use). The proposed approach appears promising for addressing scaling problems, including limitations of models at each scale (Peters *et al.* 2004). It may also reduce uncertainties in trying to address climate and land-use changes, and their interactions, simultaneously. As far as we know, few model combinations have already been implemented in such a way (Guisan & Thuiller 2005). They are some examples for animal species where landscape models are used as variables for the niche-based models (Wintle *et al.* 2005). Concerning plants, there are very few studies presenting such a combination. Dullinger *et al.* (2004) carried out a similar analysis without explicitly using a hierarchical approach, but by building a spatially explicit model incorporating both climate, biotic interactions and dispersal. Recent developments of forest gap models implicitly combined these niche-based models and landscape models, but they are only applicable to trees and ignore grasslands or open habitats (Schumacher 2004).

IDENTIFYING DRIVERS OF LARCH EXPANSION IN SUBALPINE GRASSLAND LANDSCAPES

Our simulations suggest that European larch is likely to establish viable populations in the Villar d'Arêne grasslands within the next 50 years under the various scenarios implemented. This potential colonization by trees of abandoned or less-used grasslands is not surprising, as similar colonization patterns have already been observed in old fields from mountain (Didier 2001) and lowland regions (Debussche & Lepart 1992).

The patterns of colonization depend on both future land management and larch ecological characteristics. Our simulations suggest that, under current and projected management practices, the landscape-scale distribution of trees is ultimately driven by the ecological characteristics of larch rather than by the ability of management to contain colonization. This result confirms a previous simulation study that demonstrated that colonization of the Villar d'Arêne landscape by the undesirable grass *F. paniculata* was primarily driven by constraints on its recruitment into formerly ploughed grasslands, rather than by current management (Quétier *et al.* in press). Likewise, Tasser & Tappeiner (2002) demonstrated that current tree colonization patterns are determined primarily by dispersal (e.g. distance to seed source) and secondarily by current management.

Although mowing could prevent tree establishment locally, its effects appeared limited at the landscape scale because of its small extent, even under the most intensive scenarios (scenarios 2 and 3, 15% of landscape area mown). Under such patchy land use, trees were shown to establish between mown patches. None of our scenarios considers an extension of the mown area that could contain tree establishment; at best, mowing is conserved to its current, relatively small extent.

Under these land-use conditions, we show that the most important traits determining the colonization of subalpine grasslands such as those of Villar d'Arêne are related to the three main phases of colonization: dispersal of seeds to reach a new location (dispersal ability); ability of the species to develop in this new location (resource-uptake efficiency); and ability of the species to sustain competition from colonized communities (juvenile tolerance to light interaction). These results corroborate those of many experimental and modelling studies (Bleher *et al.* 2002; Dullinger *et al.* 2004) showing the importance of dispersal ability (mean distance and distribution) and inter- and intraspecific competition in tree colonization patterns. Likewise, Quétier *et al.* (in press) showed that the response of the Villar d'Arêne landscape to land-use scenarios depended on the ability of the Festuca PFT to disperse seeds. These traits are not always easy to quantify, but experiments such as the one we conducted for germination could be used to reduce uncertainty in their estimation.

Finally, our simulations are informative in estimating the time-scale of colonization. For instance, abandoned grasslands could turn into woody communities within a few decades. This study also shows the importance of land-use legacies

(Donohue *et al*. 2000; Quétier *et al*. 2007), and how the knowledge of former land use could aid the prediction of future landscape dynamics.

IMPLICATIONS FOR FUTURE LANDSCAPES AND ECOSYSTEM SERVICES

Results from simulation studies such as this can offer guidance for management of semi-natural mountain grasslands in the context of global change. At the landscape scale, agricultural land use can have either a positive or a negative effect on biodiversity and associated goods and services (Tscharntke *et al*. 2005; Bolliger *et al*. 2007). Mowing, as carried out for several centuries, plays a crucial role in maintaining a high level of biodiversity (Maurer *et al*. 2006) by creating an intermediate disturbance intensity that prevents competitive exclusion by late-successional species (shrubs, trees) or large perennial tussock grasses (Quétier *et al*. 2007). This diversity, and the corresponding functional composition of the vegetation, are associated with the delivery of multiple services that are important to local stakeholders. Our results suggest that current patterns of agricultural use may not be able to sustain desirable levels of biodiversity associated with open landscapes, and we highlight some possible alternatives.

We examined three plausible scenarios of land use that lead to three possible futures for the next few decades, depending mostly on uncertainties about larch autoecology. Knowledge of these futures should inform management decisions that are required to adapt to future changes or mitigate them if they are found to be undesirable.

(1) Persistence of open and species-rich grasslands in Villar d'Arêne and similar localities. The current management regime does not seem intense enough for this, and more intense and/or widespread disturbances are needed to resist colonization by larch, as well as by undesirable grassland species (Quétier *et al*. in press). Maintaining the current open landscape would require mowing a larger area, as was done until World War II. The introduction of targeted grazing management, focusing more intensively on unmown grasslands and applied early in the season, would also be required to restore species-rich grasslands (Jouglet 1999).

(2) Establishment and management of a larch forest. This alternative scenario might be favoured for multifunctional agriculture, as the light cover provided by larch forest sustains a high diversity of plants and animals (Rameau *et al*. 1993). Larch forests already occur in the Guisane valley, southward from the Lautaret pass, and provide useful protection against landslides and avalanches. They can also provide shelter and some fodder for stock. The downside is that this type of forest needs to be managed because it is not a stable successional stage, and larch would quickly be superseded by other tree species such as some pines (Ozenda & Borel 1995). Under this scenario, the effects of forest establishment on current grassland communities should be investigated further. Landscape simulations predict the disappearance of the Sesleria and Dactylis PFTs in favour of the dominant and abundant Festuca PFT when land is abandoned (Quétier *et al*. in press) and when larch colonization occurs. Such a change in grassland functional composition would inevitably lead to some fundamental changes in ecosystem function and associated goods and services, such as the availability of fresh water, carbon sequestration or soil conservation. For example, dominance by the Festuca PFT is associated with an increase in biomass production, but a decrease in fodder quality and strong litter accumulation, with a loss of flower diversity and thus of aesthetic value (Quétier *et al*. in press).

(3) Establishment of mountain forest to prevent soil erosion and avalanches. In this case, no specific management is required and trees might be left to expand their range. This change may come at the cost of biodiversity and water availability (Tscharntke *et al*. 2005), but with positive consequences for carbon storage.

The effects of each of these futures on tourism should also be considered, as tourism is currently, and is likely to remain, a key sector sustaining the mountain economy.

Acknowledgements

We would like to acknowledge Marie-Pascale Colace for her help for the controlled experiment, Christian Piedallu and Jean-Claude Gégout for providing a digital version of larch distribution, and also Sophie Rickebusch for her thoughtful and detailed comments on the manuscript. This work was funded by the Agence National pour la Recherche through the QDiv project (Quantifying the Effects of Global Environmental Change on Terrestrial Plant Diversity). We received support from European Commission's FP6 MACIS (Minimization of and Adaptation to Climate change Impacts on biodiversity, N°044399) and FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe) projects.

References

Akaike, H. (1974) A new look at statistical model identification. *IEEE Transactions on Automatic Control*, **AU-19**, 716–722.

Ali, A.A., Carcaillet, C., Guendon, J.-L., Quinif, Y., Roiron, P. & Terral, J.-F. (2003) The Early Holocene tree line in the southern French Alps: new evidence from travertine formations. *Global Ecology and Biogeography*, **12**, 411–419.

Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101–118.

Benichou, P. & Le Breton, O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *la Météorologie*, **7**, 23–24.

Bleher, B., Oberrath, R. & Bohning-Gaese, K. (2002) Seed dispersal, breeding system, tree density and the spatial pattern of trees – a simulation approach. *Basic and Applied Ecology*, **3**, 115–123.

Bolliger, J., Kienast, F., Soliva, R. & Rutherford, G. (2007) Spatial sensitivity of species habitat patterns to scenarios of land use change (Switzerland). *Landscape Ecology*, **22**, 773–789.

Brisse, H., De Ruffray, P., Grandjouan, G. & Hoff, M. (1995) European vegetation survey. La banque de données phytosociologiques ≪SOPHY≫. *Annali di Botanica L.*, **III**, 191–223.

Broncano, M.J., Riba, M. & Retana, J. (1998) Seed germination and seedling performance of two Mediterranean tree species, holm oak (*Quercus ilex* L.) and Aleppo pine (*Pinus halepensis* Mill.): a multifactor experimental approach. *Plant Ecology*, **138**, 17–26.

Carrer, M. & Urbinati, C. (2004) Age-dependent tree-ring growth responses to climate in *Larix decidua* and *Pinus cembra*. *Ecology*, **85**, 730–740.

Coop, J.D. & Goivnish, T.J. (2007) Spatial and temporal patterns of recent forest encroachment in montane grasslands of the Valles Caldera, New Mexico, USA. *Journal of Biogeography*, **34**, 914–927.

Cousins, S.A.O., Lavorel, S. & Davies, I. (2003) Modelling the effects of landscape pattern and grazing regimes on the persistence of plant species with

high conservation value in grasslands in south-eastern Sweden. *Landscape Ecology*, **18**, 315–332.

Debussche, M. & Lepart, J. (1992) Establishment of woody plants in mediterranean old fields: opportunity in space and time. *Landscape Ecology*, **6**, 133–145.

Diaz, S., Tilman, D., Fargione, J. *et al.* (2005) Biodiversity regulation of ecosystem services. *Ecosystems and Human Well-being: Current State and Trends* (eds R. Hassan, R. Scholes and N. Ash), pp. 297–329. Island Press, Washington, DC, USA.

Didier, L. (2001) Invasion patterns of European larch and Swiss stone pine in subalpine pastures in the French Alps. *Forest Ecology and Management*, **145**, 67–77.

Donohue, K., Foster, D.R. & Motzkin, G. (2000) Effects of the past and the present on species distribution: land-use history and demography of wintergreen. *Journal of Ecology*, **88**, 303–316.

Dullinger, S., Dirnbock, T. & Grabherr, G. (2004) Modelling climate change-driven tree line shifts: relative effects of temperature increase, dispersal and invasibility. *Journal of Ecology*, **92**, 241–252.

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

Grabherr, G. (2003) Alpine vegetation dynamics and climate change – a synthesis of long-term studies and observations. *Alpine Biodiversity in Europe – A Europe-Wide Assessment of Biological Richness and Change* (eds L. Nagy, G. Grabherr, C. Körner and D.B.A. Thompson), pp. 399–409. Springer, Dordrecht, the Netherlands.

Grigulis, K., Lavorel, S., Davies, I.D., Dossantos, A., Lloret, F. & Vila, M. (2005) Landscape-scale positive feedbacks between fire and expansion of the large tussock grass, *Ampelodesmos mauritanica*. Catalan shrublands. *Global Change Biology*, **11**, 1042–1053.

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

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

Holtmeier, F.-K. & Broll, G. (2005) Sensitivity and response of northern hemisphere altitudinal and polar tree lines to environmental change at landscape and local scales. *Global Ecology and Biogeography*, **14**, 395–410.

Ihaca, R. & Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–314.

Jouglet, J.P. (1999) *Les végétations des alpages des Alpes françaises du sud*. Cemagref éditions, Antony Cedex, France.

Körner, C. (1999) *Alpine Plant Life*. Springer, Dordrecht, the Netherlands.

Körner, C. (2003) Ecological impacts of atmospheric CO_2 enrichment on terrestrial ecosystems. *Philosophical Transactions of the Royal Society of London B*, **361**, 2023–2041.

Kullman, L. (2002) Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, **90**, 68–77.

Kunstler, G., Curt, T., Bouchaud, M. & Lepart, J. (2006) Indirect facilitation and competition in tree species colonization of sub-mediterranean grasslands. *Journal of Vegetation Science*, **17**, 379–388.

Lavorel, S., McIntyre, S., Landsberg, J. & Forbes, T.D.A. (1997) Plant functional classifications: from general groups to specific groups based on response to disturbance. *Trends in Ecology and Evolution*, **12**, 471–478.

Li, X. & Wilson, S.D. (1998) Facilitation among woody plants establishing in an old field. *Ecology*, **79**, 2694–2705.

Maurer, K., Weyand, A., Fischer, M. & Stöcklin, J. (2006) Old cultural traditions, in addition to land use and topography, are shaping plant diversity of grasslands in the Alps. *Biological Conservation*, **130**, 438–446.

Milsom, T.P., Langton, S.D., Parkin, W.K. *et al.* (2000) Habitat models of bird species' distribution: an aid to the management of coastal grazing marshes. *Journal of Applied Ecology*, **37**, 706–727.

Moore, A.D. & Noble, I.R. (1990) An individualistic model of vegetation stand dynamics. *Journal of Environmental Management*, **31**, 61–81.

Mueller-Dombois, D. & Ellenberg, H. (1974) *Aims and Methods of Vegetation Ecology*. Wiley, New York.

Nakicenovic, N. & Swart, R. (2000) *Emissions Scenarios*. Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Ozenda, P. & Borel, J.-L. (1995) Possible response of mountain vegetation to a global climatic change: the case of the Western Alps. *Ecological Impacts of Climate Change in the Alps and Fennoscandian Mountains* (eds A. Guisan, J.I. Holten, R. Spichiger and L. Tessier), pp. 137–144. Conserv. Jardin Botanique, Genève.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Peters, D.P.C., Herrick, J.E., Urban, D.L., Gerdner, R.H. & Breshears, D.D. (2004) Strategies for ecological extrapolation. *Oikos*, **106**, 627–636.

Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.

Quétier, F. (2006) *La vulnérabilité des écosystèmes semi-naturels européens aux changements d'utilisation des terres. Application aux prairies subalpines de Villar d'Arène, France*. PhD thesis, l'Agro-Montpellier, Montpellier, France.

Quétier, F., Lavorel, S., Thuiller, W. & Davies, I. (in press) Plant trait-based assessment of ecosystem service sensitivity to land-use change in mountain grasslands. *Ecological Applications*, in press.

Quétier, F., Thébault, A. & Lavorel, S. (2007) Linking vegetation and ecosystem response to complex past and present land use changes using plant traits and a multiple stable state framework. *Ecological Monographs*, in press.

Rameau, J.C., Mansion, D., Dume, G. *et al.* (1993) *Flore Forestière Française: Guide Écologique Illustré, 2*. Institut pour le Développement Forestier, Paris.

Rickebusch, S. (2006) *Modelling tree population dynamics at the alpine and boreal tree-line ecotones in response to climate and land-use change*. PhD thesis, Université de Lausanne, Lausanne, France.

Rounsevell, M.D.A., Reginster, I., Araujo, M.B. *et al.* (2006) A coherent set of future land use change scenarios for Europe. *Agriculture, Ecosystems and Environment*., **114**, 57–68.

Sala, O.E., Chapin, S.F. III, Armesto, J.J. *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.

Sala, O.E., Van Vuuren, D., Pereira, H.M. *et al.* (2005) Biodiversity across scenarios. *Ecosystems and Human Well-Being: Scenarios* (eds S.R. Carpenter, P.L. Pingali, E.M. Bennett and M. Zurek), pp. 375–408. Island Press, Washington, DC, USA.

Schumacher, S. (2004) *The role of large scale disturbances and climate for the dynamics of forested landscapes in the European Alps.* PhD thesis, Swiss Federal Institute of Technology, Zurich.

Schumacher, S., Bugmann, H. & Mladenoff, D.J. (2004) Improving the formulation of tree growth and succession in a spatially explicit landscape model. *Ecological Modelling*, **180**, 175–194.

Sitch, S., Smith, B., Prentice, I.C. *et al.* (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ Dynamic Global Vegetation Model. *Global Change Biology*, **9**, 161–185.

Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.

Swets, K.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.

Tappeiner, U., Tappeiner, G., Hilbert, A. & Mattanovich, E. (2003) *The EU Agricultural Policy and the Environment. Evaluation of the Alpine Region*. Blackwell Science, Berlin.

Tasser, E. & Tappeiner, U. (2002) Impact of land use changes on mountain vegetation. *Applied Vegetation Science*, **5**, 173–184.

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I. & Thies, C. (2005) Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management. *Ecology Letters*, **8**, 857–874.

Vandvik, V., Heedgaard, E., Maren, I.E. & Aarrestad, P.A. (2005) Managing heterogeneity: the importance of grazing and environmental variation on post-fire succession in heathlands. *Journal of Applied Ecology*, **42**, 139–149.

Walther, G.-R., Beibner, S. & Burga, C.A. (2005) Trends in the upward shift of alpine plants. *Journal of Vegetation Science*, **16**, 541–548.

Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.

Wintle, B.A., Bekessy, S.A., Venier, L.A., Pearce, J.L. & Chisholm, R.A. (2005) Utility of dynamic-landscape metapopulation models for sustainable forest management. *Conservation Biology*, **19**, 1930–1943.

Woodward, F.I. & Beerling, D.J. (1997) The dynamics of vegetation change: health warnings for equilibrium 'dodo' models. *Global Ecology and Biogeography Letters*, **6**, 413–418.

Received 29 January 2007; accepted 27 August 2007
Handling Editor: Rob Freckleton

Supplementary material

The following supplementary material is available for this article.

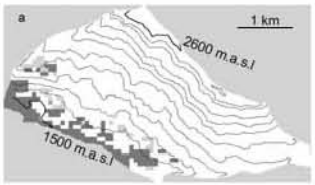
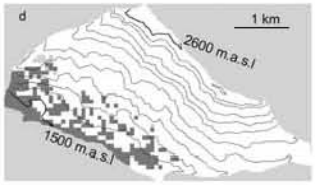
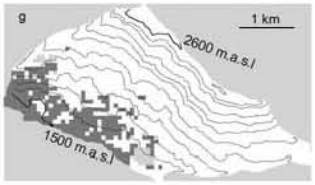
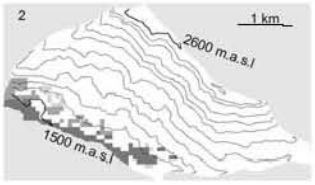
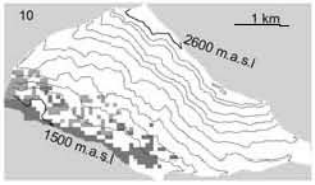
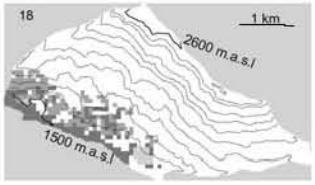
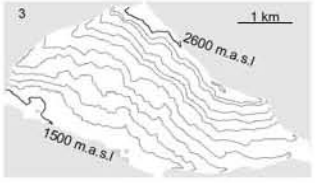
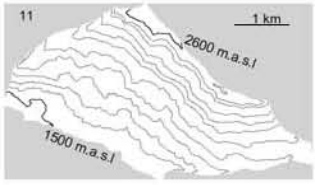
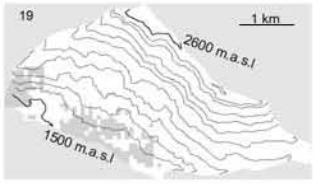
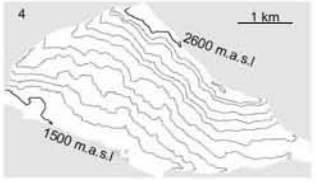
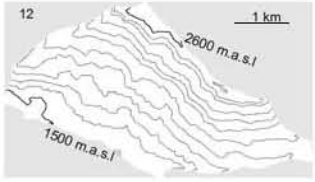
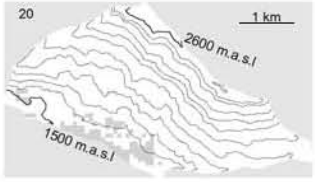
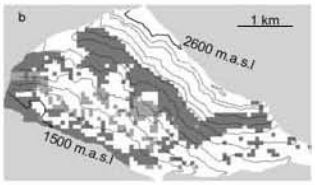
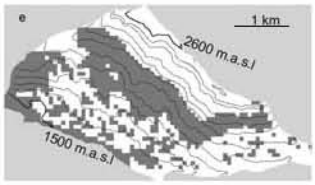
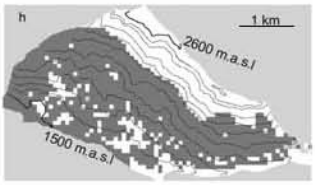
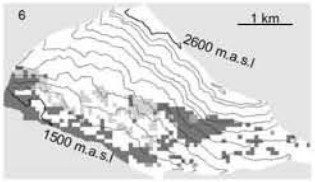
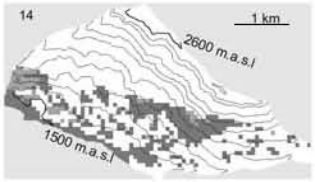
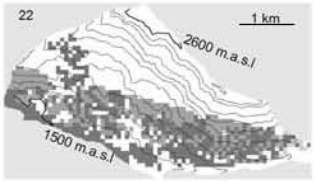
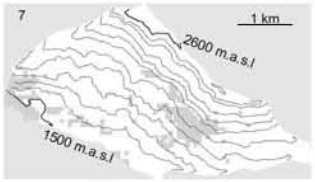
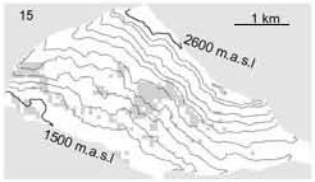
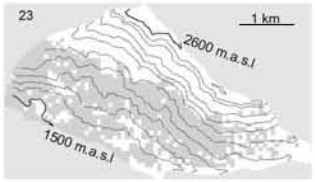
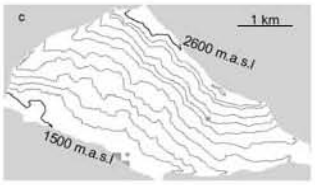
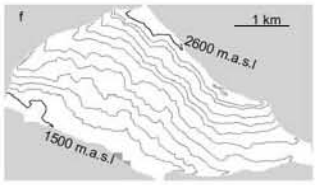
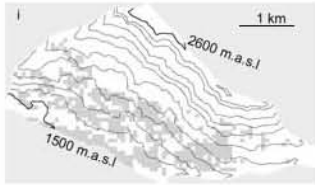
Appendix S1. Maps of the abundance of mature larch trees and patterns of grassland colonization by *Larix decidua*

along the altitudinal gradient for the 24 simulations combining land-use scenarios and larch autoecology

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2664.2007.01416.x.
(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

100 Trees/ha | 10 Trees /ha | 1 Tree /ha | 0 Tree /ha

Disturbance			Generalist	Tolerant juveniles	Long distance dispersal
Early Mowing	Late Mowing	Abandonment			
			Y	Y	
				Y	
			Y		
					
			Y	Y	Y
				Y	Y
			Y		Y
					Y

Available online at www.sciencedirect.com
ScienceDirect
Perspectives in Plant Ecology, Evolution and Systematics 9 (2008) 137–152
Perspectives in Plant Ecology, Evolution and Systematics
www.elsevier.de/ppees

Predicting global change impacts on plant species' distributions: Future challenges

Wilfried Thuiller[a,*], Cécile Albert[a,1], Miguel B. Araújo[b,1], Pam M. Berry[c,1], Mar Cabeza[d,1], Antoine Guisan[e,1], Thomas Hickler[f,1], Guy F. Midgley[g,1], James Paterson[c,1], Frank M. Schurr[h,1], Martin T. Sykes[f,1], Niklaus E. Zimmermann[i,1]

[a]*Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France*
[b]*Departmento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC, Cl José Gutiérrez Abascal 2, Madrid 28006, Spain*
[c]*Environmental Change Institute, Oxford University Centre for the Environment, Dyson Perrins Building, South Parks Road, Oxford OX1 3QY, UK*
[d]*Department of Biological and Environmental Sciences, University of Helsinki, P.O. Box 65, Finland*
[e]*Laboratoire de Biologie de la Conservation, Département d'Ecologie et d'Evolution, Université de Lausanne, Bâtiment de Biologie, CH-1015 Lausanne, Switzerland*
[f]*Department of Physical Geography and Ecosystems Analysis, Lund University, Sölvegatan 12, 223 62 Lund, Sweden*
[g]*Climate Change Research Group, Kirstenbosch Research Center, South African National Biodiversity Institute, P/Bag x7, Claremont 7735, Cape Town, South Africa*
[h]*Plant Ecology and Nature Conservation, University of Potsdam, Maulbeerallee 2, D-14469 Potsdam, Germany*
[i]*Land-use Dynamics, Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland*

Received 3 March 2007; received in revised form 11 September 2007; accepted 18 September 2007

Abstract

Given the rate of projected environmental change for the 21st century, urgent adaptation and mitigation measures are required to slow down the on-going erosion of biodiversity. Even though increasing evidence shows that recent human-induced environmental changes have already triggered species' range shifts, changes in phenology and species' extinctions, accurate projections of species' responses to future environmental changes are more difficult to ascertain. This is problematic, since there is a growing awareness of the need to adopt proactive conservation planning measures using forecasts of species' responses to future environmental changes.

There is a substantial body of literature describing and assessing the impacts of various scenarios of climate and land-use change on species' distributions. Model predictions include a wide range of assumptions and limitations that are widely acknowledged but compromise their use for developing reliable adaptation and mitigation strategies for biodiversity. Indeed, amongst the most used models, few, if any, explicitly deal with migration processes, the dynamics of population at the "trailing edge" of shifting populations, species' interactions and the interaction between the effects of climate and land-use.

In this review, we propose two main avenues to progress the understanding and prediction of the different processes occurring on the leading and trailing edge of the species' distribution in response to any global change phenomena.

*Corresponding author. Tel.: +33 4 76 63 54 53; fax: +33 4 76 51 42 79.
E-mail address: wilfried.thuiller@ujf-grenoble.fr (W. Thuiller).
[1]Contributed equally to the paper. They are sorted alphabetically.

1433-8319/$ - see front matter © 2007 Rübel Foundation, ETH Zürich. Published by Elsevier GmbH. All rights reserved.
doi:10.1016/j.ppees.2007.09.004

Deliberately focusing on plant species, we first explore the different ways to incorporate species' migration in the existing modelling approaches, given data and knowledge limitations and the dual effects of climate and land-use factors. Secondly, we explore the mechanisms and processes happening at the trailing edge of a shifting species' distribution and how to implement them into a modelling approach. We finally conclude this review with clear guidelines on how such modelling improvements will benefit conservation strategies in a changing world.
© 2007 Rübel Foundation, ETH Zürich. Published by Elsevier GmbH. All rights reserved.

Keywords: Species distribution modeling; Habitat models; Process-based models; Global change; Conservation planning

Introduction

There is now compelling evidence that species are already shifting their ranges in response to on-going changes in regional climates (Parmesan and Yohe, 2003; Root et al., 2005; Walther et al., 2005; Lavergne et al., 2006), that species are altering their phenology (Menzel and Fabian, 1999; Visser and Holleman, 2001; White et al., 2003; Zavaleta et al., 2003) and that some species are facing extinction, or have become extinct (Parmesan, 2006; Pauli et al., 2006; Pounds et al., 2006; Foden et al., 2007). Given the contemporary biodiversity crisis, effective conservation strategies that offset the climate change threats to species persistence will be critical in maintaining species and genetic diversity. To date, conservation planning has mostly focused on preserving pattern and has acted reactively. However, scientists and stakeholders recognize that the dynamic nature of biodiversity requires a paradigm shift in approach (Cabeza and Moilanen, 2001; Pressey et al., 2007). This should include proactive planning and focusing on processes rather than patterns. However, proactive responses are limited without forecasting and the implementation of advanced models to do this. Responses of biodiversity to land-use and climate change are receiving increasing attention in conservation planning science, with decision tools being developed to incorporate such forecasts (Hannah et al., 2007; Pressey et al., 2007).

The Millennium Ecosystem Assessment (MEA, 2005) mentions climate change as the largest forthcoming threat to biodiversity across most biomes, and in a recent review on systematic conservation planning, Pressey et al. (2007) recognize accounting for climate change to be an urgent issue and one of the major challenges for the future. However, Pressey et al. (2007) also emphasize the added problems due to the uncertainties associated with both climate change and projections of associated species' range shifts. Although methodological exercises have attempted to design protected areas while accounting for species' shifts (e.g. Hannah et al., 2007), conservation practitioners prefer to avoid the uncertainties in projections and concentrate on protecting current patterns of biodiversity. Nevertheless, several studies evaluating the future performance of protected areas (e.g. Araújo et al., 2004; Thuiller et al., 2006a) call for a need to include the dynamic nature of biodiversity in conservation planning.

To increase confidence in model projections, methodologies must acknowledge clearly the uncertainties involved and try to obtain conditional measurements of confidence in the forecasts made (Thuiller et al., 2004; Schröter et al., 2005; Berry et al., 2006; Araújo and New, 2007; Thuiller, 2007). Whereas confidence in model projections is likely to increase as the realism of key assumptions within a particular modelling approach are improved (Sutherst et al., 2000), tractability often decreases, due to more demanding parameterizations. Thus a trade-off exists between complexity and tractability in modelling species and identifying the most reliable and unbiased solution is not a trivial task (Thuiller, 2007). A second trade-off in modelling species' responses to environmental changes is between generality and specificity of predictions. At the generality end of the gradient, process-based dynamic global vegetation models (DVMs) are generalized to such an extent that they can simulate global patterns of vegetation distribution, as well as carbon, nutrient and water cycling from existing knowledge of the mechanisms driving these processes (e.g. Sitch et al., 2003; Woodward and Lomas, 2004; Thuiller et al., 2006c). The drawback is that primary producers are classified into a small number of "plant functional types" (PFTs), which provide a coarse classification of biodiversity for several conservation planning applications. Tree species-specific responses have been extensively explored with "forest gap models" (Botkin et al., 1972; Shugart, 1984; Sykes and Prentice, 1995, 1996; Bugmann, 2001). These models have been criticized for being highly parameterized for particular species and sites (Badeck et al., 2001), but there has been considerable progress in the development of generalized forest gap models, which can now be applied across different regions, at least in the temperate zone, and account for population demographics, species' interactions and physiological processes (Bugmann and Solomon, 2000; Hickler et al., 2004). At the specificity end of the gradient, habitat models (Guisan and Thuiller, 2005), which fit species environmental niches explicitly, ignore all mechanisms driving species'

demography (but see Sutherst, 2000; Kriticos et al., 2003) and species' interactions (but see Leathwick and Austin, 2001). They assume that the fitted relationship between the presence/absence or abundance of a given species and the environmental conditions measured in a site is a good surrogate for such demographic processes (Araújo and Guisan, 2006; Austin, 2007). These habitat models are simplistic but they focus explicitly at the species level, allowing the modelling of numerous species, and can thus be used to estimate patterns of current and future diversity (Box 1; Peterson et al., 2002; Thomas et al., 2004; Thuiller, 2004; Ferrier and Guisan, 2006).

Most efforts in the past have attempted to understand and predict how the distribution of a selected species or group of species as a whole will be modified in response to a given environmental change (mostly climate and land-use changes). However, little attention has been given to simulating processes at the leading and trailing edges of the distribution where range change happens (migration, persistence, extinction, Hampe and Petit, 2005). The leading edge is fundamentally important under global change, as it is commonly accepted that range expansions depend mostly on populations at the colonization front. The leading edge is also seen as controlled by rare long-dispersal dispersal events followed by exponential population growth (Hewitt, 1993; Hampe and Petit, 2005). Alternatively, the trailing edge remains largely understudied, but recent reviews have demonstrated their crucial role by maintaining long-term stores of species genetic diversity and places of speciation (Hampe and Petit, 2005). The aim of this review is therefore to highlight and discuss two research avenues that should provide better ecological understanding and better predictions. First we review, discuss and propose new avenues to understand and predict the processes happening at the leading edge of species' distributions. New directions are proposed to include population dynamics, land-use and biotic interactions more explicitly in the models. Then, following the same scheme, we analyse the processes happening at the trailing edge of species' distributions (persistence, mortality, genetic adaptation) and propose avenues to include them in a modelling framework.

Because it was impossible to cover all organisms and associated processes, we explicitly deal with plant models in this review, but most of the concepts discussed also apply to other organisms.

Box 1. Species' distribution model definition.

It is relatively difficult to classify models as they all share some theories, concepts or assumptions. In this review, we follow Kearney (2006):

We refer to "*habitat models*" as any model which relates the physical nature (biotic and abiotic) of a region with respect to a species, with no direct mechanistic links necessarily occurring between those descriptors and the species. Typical habitat models are reviewed in Austin (2002, 2007), Guisan and Thuiller (2005) and Guisan and Zimmermann (2000). They are purely descriptive and relate to a particular space and time frame.

We refer to "*process-based models*" as any model that uses mechanistic links between the growth and fitness of species, or more abstract plant functional types (PFTs), and a range of environmental or biological (e.g. competing species or PFTs) variables. Examples range from dynamic vegetation model (based on PFTs, Woodward, 1992), population viability analysis (based on population dynamics, Possingham and Davies, 1995), plant population modelling (Jeltsch et al., 2008) phenological models (based on phenology; Chuine et al., 2000), or diffusion/spread models (With, 2002).

Predicting the leading-edge response – forecasting species migration

Background

Evidence from paleo-ecological studies (Huntley, 1991; Prentice et al., 2000; Davis and Shaw, 2001) and multiple observations of current species' range shifts (Walther et al., 2005; Parmesan, 2006) show that migration (i.e. a directional shift in a species' range) will be a very likely response of plant species to ongoing and future climate change. In this paper, we define migration as the result of four processes: fecundity, dispersal, recruitment and population growth. Understanding migration holistically is important because: (i) it is clear that the rate of change in local conditions threatens to outpace the migration capabilities of many species with limited dispersal abilities (Solomon and Kirilenko, 1997; Pounds et al., 1999); and (ii) current levels of habitat modification and fragmentation further reduce the ability of most plant species' propagules to disperse across landscapes (Collingham and Huntley, 2000). This has obvious negative implications for species' persistence, including range reduction and possible extinction (Solomon and Kirilenko, 1997; Midgley et al., 2007).

Despite the acknowledged importance of plant migration in response to global change (Higgins et al., 2003a; Neilson et al., 2005; Midgley et al., 2007), few

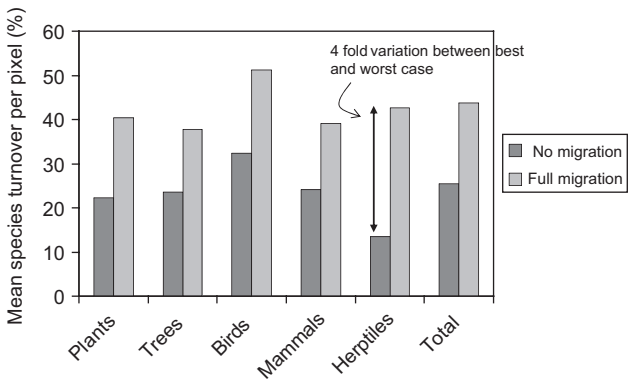


Fig. 1. Uncertainties in projected mean species turnover in Europe by 2050 associated with two extremes in migration assumptions (from Thuiller et al., 2005a, 2006b; Araújo et al., 2006, and Araujo and Thuiller, unpublished data). Two extreme assumptions, no or unlimited migration, were used to account for the migration ability of each species. The histogram highlights the uncertainty entering through these assumptions (see herptiles for the most striking evidence, showing almost four times the turnover with unlimited migration).

modelling studies explicitly include migration processes when simulating geographical plant response (but see Collingham et al., 1996; Lischke et al., 2006; Midgley et al., 2006; Albert et al., 2007). DVMs do not consider migration either, the plant functional type or species of interest being expected to grow wherever the conditions are suitable. Habitat models usually assume either no migration or unlimited migration (Peterson et al., 2002; Thuiller et al., 2005b; Schwartz et al., 2006, Fig. 1), although most published future projections have assumed unlimited dispersal. As a consequence, seed dispersal and the associated patterns of plant migration are considered to be among the most significant uncertainties in projecting climate change impacts on plant species' ranges (Fig. 1, Higgins et al., 2003a; Neilson et al., 2005; Pearson, 2006; Midgley et al., 2007).

Key issues

The simplest way to incorporate migration into models of species or plant functional type distributions is to assume that all populations of a species migrate at identical rates (Williams et al., 2005; Broennimann et al., 2006; Midgley et al., 2006). Estimates of such migration rates can be derived from observed rates of contemporary migration (Cain et al., 2000; Bullock et al., 2006; Kunstler et al., 2007), from the paleo-records (Davis, 1981; Clark et al., 1998), from genetic isolation by distance slopes (Hamrick and Loveless, 1989; Petit et al., 2001; Kinlan and Gaines, 2003), from expert knowledge or from a mix of these sources. For instance, Midgley et al. (2006) assumed the rate of plant migration to be a maximum of 1 km per decade for ant- and rodent-dispersed Proteaceae species of the Cape region, and 3 km per decade for wind-dispersed species. Such an approach can be easily combined with habitat models and may be used in global change analyses to assess risks (e.g. Williams et al., 2005) and to reduce the uncertainty of projections related to dispersal (e.g. Broennimann et al., 2006).

One shortcoming of this simple approach is that there are still few available data on dispersal distances, let alone migration rates of non-invasive plant species in present-day landscapes (for a summary see Parmesan and Yohe, 2003). Moreover, the assumption of identical migration rates across a species' range is likely to be untenable in many cases since the migration capacity of a population depends on its local dynamics (e.g. Skellam, 1951). Furthermore, external factors such as the land-use change, landscape fragmentation and the connectivity of patches of suitable habitat are bound to complicate the modelling of dispersal since probabilities of dispersal in heterogeneous habitats can become anisotropic (e.g. Pitelka et al., 1997; Higgins et al., 2003b; Pearson and Dawson, 2005). Similarly, as shown with the rapid spread of plant invasive alien species throughout the globe (Wadsworth et al., 2000; Kowarik, 2003; Levin et al., 2003), human-mediated dispersal plays a major role in long-distance dispersal events although it remains difficult to actually measure and integrate this into a modelling framework (Schmidt, 1989; Hodkinson and Thompson, 1997; Tikka et al., 2001; Zwaenepoel et al., 2006).

These problems can be partly avoided through the combination of a migration model with projections from habitat models. For example, Iverson et al. (2004) coupled a habitat model with a simple model of habitat colonization to predict the future distribution of North American tree species. The colonization model uses a cellular automaton that predicts the colonization of suitable cells based on habitat quality, a correlate of local abundance in occupied cells, and the distance between occupied and potentially colonized cells. For all of the studied species, Iverson et al. (2004) parameterized the dispersal component of their colonization model to produce a migration rate of 50 km/century in fully forested landscapes, a value inspired by maximum estimates of average Holocene tree migration in North America. Using the estimated dispersal parameter, the colonization model was then used to simulate the spread of species in present-day landscapes with lower habitat quality. This approach is computationally efficient and innovative in that it uses paleo-ecological information. However, it also has a number of limitations including the lack of consideration of population dynamics, the use of ad-hoc dispersal kernels (e.g. the modelling of dispersal with an inverse power law that does not integrate to 1), the assumption of identical Holocene migration rates for all studied species, and indeed the

assumption that migration estimates based on the paleorecord are correct, and relevant for today's fragmented landscapes (for discussion see Clark et al., 1998; Pearson, 2006).

More complex hybrid approaches entail models of local population dynamics and mathematically adequate descriptions of dispersal and recruitment (Jongejans et al., 2008). Classical metapopulation models represent recruitment processes by a single function (Hanski, 1994), whereas sophisticated spatially explicit gap models like LANDIS/LANDCLIM (He and Mladenoff, 1999; Schumacher and Bugmann, 2006), or landscape models like LAMOS (Cousins et al., 2003; Albert et al., 2007) model components of the recruitment process, such as seed production, seed dispersal, and seedling recruitment, separately. The spatially explicit forest gap model TREEMIG (Lischke et al., 2006) includes probably the most complete representation of seed dispersal and subsequent regeneration processes, by including density-dependent antagonists and intra- and inter-specific seedling competition.

Perspectives

Given that complex gap and metapopulation models are computationally very expensive when they are run at large spatial scales, it seems more promising to incorporate results of migration theory into species-distribution models. There is a well-developed body of theory on the migration of species, but so far only little pieces of the theory have been used to improve range shift projections. Migration theory has led to a suite of models for population spread (e.g. Skellam, 1951; Kot et al., 1996; Neubert and Caswell, 2000; Clark et al., 2001). While the classical model of Skellam (1951) described migration as a diffusion process characterized by mean dispersal distance, later models have emphasized the importance of kernel shape and long-distance dispersal (Kot et al., 1996; Higgins and Richardson, 1999; Clark et al., 2001). In their simplest form, these models require information on a population's finite rate of growth, λ (or on net reproductive rate, R_0, and generation time, T), and on the dispersal kernel of offspring as a function of distance from the mother plant ($f(x)$). In the following, we point out how better models and data on dispersal and demography may help us improve migration forecasts.

The understanding and quantification of long-distance seed dispersal has undergone considerable progress in recent years (Nathan, 2006; Kunstler et al., 2007). Mechanistic models for wind dispersal now make reliable predictions at scales relevant for the long-distance movement of seeds and can be parameterized for large numbers of species (Nathan et al., 2002; Tackenberg et al., 2003; Schurr et al., 2005). Moreover, these models seem to be informative for the biogeographical dynamics of plant species. For South African Proteaceae, Schurr et al. (2007) showed that mechanistically derived estimates of wind dispersal potential (Tackenberg et al., 2003; Schurr et al., 2005, Fig. 2) explain inter-specific variation in the ratio between realized and potential range size (the latter estimated with habitat models by Midgley et al., 2003). This suggests that for these predominantly wind-dispersed species, the incorporation of mechanistic models for wind dispersal may considerably improve distribution forecasts.

Mechanistic models for other dispersal vectors are less advanced than models for wind dispersal. In the case of seed dispersal by animals this is partly due to the fact that the environmental and behavioural factors influencing animal movement are so varied that models are often highly case-specific (e.g. Westcott and Graham, 2000). However, Powell and Zimmermann (2004) have demonstrated how mathematical homogenization techniques allow combining information on the behaviour of animals across different spatial scales with simple plant population models. The results are promising in that (a) considerably better historical migration estimates result from these analyses and (b) a general method to derive dispersal kernels is provided, which can be extended to other dispersal vectors that operate at multiple spatial scales. The further development of such generic models for seed dispersal by animals and abiotic factors other than wind will greatly improve our ability to forecast range dynamics (Nathan, 2006).

Statistical techniques also provide a promising avenue for estimating dispersal kernels from easily available information such as parent and seedling distributions (Ribbens et al., 1994; Clark et al., 1998; Clark, 2005) or molecular data (Cain et al., 2000; Godoy and Jordano, 2001). In the case of molecular methods, however, it remains unclear how widely applicable they are (Rousset, 2001) and whether they can be used to infer the shape of dispersal kernels.

While the quantification of long-distance dispersal is a comparatively new field, data on population growth rates have been collected by population ecologists over decades (see e.g. Franco and Silvertown, 2004). It should be noted, however, that these data typically stem from dense populations in the "core" of a species' range (Franco and Silvertown, 2004). The dynamics of such dense populations may differ in important ways from the dynamics of sparse populations at the spreading front that drive species expansion (Kot et al., 1996; Sagarin et al., 2006). For example, reduced intraspecific competition at the spreading front may result in increased population growth rates, thereby promoting migration. On the other hand, sparse populations at the expanding edge of a moving population may suffer from

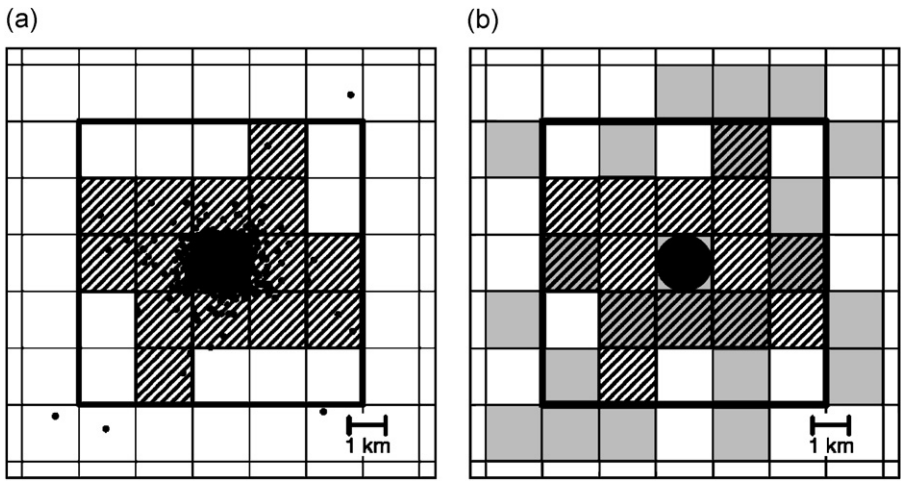


Fig. 2. A protocol used to estimate the dispersal ability of 37 species of South African Proteaceae. The protocol combines species-specific simulations of a mechanistic model for seed dispersal by wind with species-specific estimates of habitat distribution: (a) for each of 10,000 different dispersal environments, the dispersal model simulated the post-dispersal locations of 10,000 seeds (black dots) starting from the central source cell. The neighbouring cells reached by at least one seed (hatched cells) were determined from this seed shadow. To represent barriers to Proteaceae seed dispersal that arise from limited fire extent, only the 5×5 cell neighbourhood (delimited by the bold line) was considered. (b) The distribution of cells reached (hatched) was overlaid on the distribution of suitable habitat (grey), yielding the number of suitable cells that can be reached from a potentially suitable source cell (black circle) (from Schurr et al., 2007).

Allee effects that can substantially slow down spread rates (e.g. Kot et al., 1996; Taylor and Hastings, 2005). So far, only a few empirical studies have looked for Allee effects at expanding range edges, but for invasive plant species there is some evidence that Allee effects indeed occur at the invasion front (Cappuccino, 2004; Davis et al., 2004). This calls for further empirical research on the importance of Allee effects for the spread of native and invasive plant species. Moreover, we currently lack data on the post-dispersal survival of long-distance dispersed seeds (Nathan, 2006).

We suggest that empirical observations and experiments should urgently be combined with habitat modelling approaches to allow the spatial simulation of climate change impacts on demographic rates. A number of modelling approaches exist for providing such a mechanistic link. Metabolic theory makes very simple and generic predictions on the temperature dependence of population growth rates and carrying capacities (Brown et al., 2004). However, these predictions are probably too coarse to be useful in models for the range dynamics of individual species and independent tests using a broad range of taxa across different regions have found limited support for some of the general predictions in the theory (Hawkins et al., 2007). To develop and test more refined models of climate effects on demography, data are required on demographic rates collected across climatic gradients. More complex climatic effects on population growth rates are embedded in DVMs (e.g. Sitch et al., 2003; Woodward and Lomas, 2004 ♯2633), but species-specific predictions cannot be made from these, and the formulations of population dynamics are very simplified (Moorcroft et al., 2001). Some forest gap models represent many of the processes that are thought to influence climatic effects on population growth rates, including competition and indirect effects on dispersal through earlier reproduction or seed numbers (Lischke et al., 2006). In these models, tree growth is commonly assumed to be maximal if the climate corresponds with the current climate in the centre of a species' distribution (e.g. Bugmann, 2001; Schumacher et al., 2004, 2006).

Towards the edge of the distribution, growth and establishment is reduced, while growth efficiency-related mortality increases. Plant species are limited not only by their absolute limits of survival but also through competition from other species, which might grow better in a given climate. Parameterizing these models for large numbers of plants other than trees in the temperate and boreal zones or PFTs, however, might not be possible in the near future. Instead, process-based vegetation models can be used to predict general changes in vegetation structure, such as canopy density or vegetation type, which implies information on changes in habitat type and quality for most species.

By altering demographic rates, inter-specific interactions may speed up or slow down spread rates, as well as decrease or increase the area a species can occupy (e.g. Travis et al., 2005). Modelling the effect of inter-specific interactions on spread rates is particularly challenging since, at the spreading front, a species is likely to interact with species it has not encountered before. These novel

interactions may not be captured by the extrapolation of habitat models and call for the incorporation of process-based interaction modules into range shift models. As an example, the range shift model of Lischke et al. (2006) contains a submodel describing intra- and inter-specific competition for light. For more complex interactions such as competition for multiple resources, facilitation, pollination and herbivory we lack generic models. While the development of such generic models is highly complex, it also seems highly rewarding: generic models for species' interactions may greatly expand our ability to (1) replace descriptions of potential ranges by mechanistic range models, (2) assess the impact of invasive species on biodiversity, and (3) predict the dynamics of novel communities arising under global change.

The combination of habitat models for potential ranges with mechanistic models of species spread holds great promise for the improvement of range shift projections. It should be noted, however, that such hybrid models provide no general solution for the incorporation of migration processes into range shift projections. Consider, for instance, a species that is expanding poleward in response to past climatic changes, as seems to be the case for many European tree species whose range still appears to be limited by post-glacial expansion (Svenning and Skov, 2004). A habitat model fitted to the present-day distribution of such species would tend to underestimate its cold tolerance and hence the extent of its current and future potential range. Consequently, a process-based migration model based on these bioclimatic projections is likely to underestimate the future range of this species, thus overestimating extinction risks from models of range. This problem may be somewhat reduced by not using the "best fit" between a species' distribution and climate; but instead the most extreme value of any bioclimatic variable found within a species' geographic range (Bugmann and Solomon, 2000).

An approach that completely circumvents this problem is the joint statistical fitting of a mechanistic range dynamics model and a habitat model to species records. A step in this direction has been taken by Wikle (2003), who used a hierarchical Bayesian approach to fit a Skellam-type model of population migration to a time series of House Finch records in the eastern US. Wikle's model allowed for random spatial variation in both population growth rate and mean dispersal distance. While he did not model these parameters as a function of climatic variables, the hierarchical Bayesian framework employed is flexible enough to incorporate functional relationships between demographic parameters and climatic variables. Limits to the broad application of this approach may be posed by the availability of high-quality data on the spatio-temporal dynamics of species and by the computational challenges encountered in the statistical fitting of complex models (Clark, 2005). However, the hierarchical Bayesian framework is a powerful way of linking process-based models and data, and this potential should be exploited for the improvement of range shift projections.

Predicting the trailing edge response – forecasting species probability of persistence

Background

To predict the future distribution of a species, one has to consider not only the ability of the species to colonize new sites, but also the ability to persist in current sites or face localized population extinction. The effects of local extinction may be seen at the level of entire species' ranges: in South African Proteaceae, species with life histories that are more vulnerable to population extinction tend to fill less of their potentially suitable range (Schurr et al., 2007). Extinction is also likely to be seen in small and isolated populations driven by reduced intra-population genetic variation and reduced fitness stemming from a greater degree of environmental stress and inbreeding depression (Willi et al., 2006). These small and isolated populations are often found at the rear edge of a range; indeed, the fossil record shows that extinction commonly occurred at rear-edge populations, particularly during periods of rapid climate change (Davis and Shaw, 2001).

Recent empirical evidence suggests that climate change already drives the extinction of rear-edge plant populations leading to a distribution with a 'trailing' edge (Lesica and McCune, 2004; Hampe, 2005; Lavergne et al., 2005, 2006; Wilson et al., 2005). For the desert tree *Aloe dichotoma* in southern Africa, Foden et al. (2007) found that trailing edge populations showed clearly negative demographic rates, and strongly positive rates were observed at the leading edge of the range, making population growth rate a sensitive and useful indicator of incipient change in range.

However, for the majority of species detecting any decline at the rear edge is hampered by a lack of available survey data (Thomas et al., 2006) and there is also a danger that the larger amount of evidence for leading-edge expansion may result in a lack of research on climate change effects on the rear edge (Hampe and Petit, 2005). Moreover, this lack of research is granted more urgency as it has been argued recently that population extinction at the rear edge of shifting ranges may be of disproportional importance for conservation because of the high level of regional genetic diversity between isolated populations (Hampe and Petit, 2005).

The short-term response of plant populations to environmental change partly depends on basic aspects of a species' life history: populations with long generation time, high reproductive rates, and persistent life cycle stages are likely to persist longer in the face of environmental change. Additionally, phenotypic plasticity is important for short-term reactions to environmental change, since it may allow a species to buffer short-term environmental change and maintain viable populations (Ghalambor et al., 2007).

For the long-term persistence of natural populations facing directional changes in environmental conditions, evolutionary adaptation becomes crucial (Etterson and Shaw, 2001). Ultimately, populations persist only if the rate of adaptive evolution at least matches the rate of environmental change (Bürger and Lynch, 1995). The two main ingredients necessary to predict how phenotypic traits in a population might change in response to environmental change are the magnitude of plasticity in the population (which determines the potential of the population to respond over the short term), and the speed at which the population can evolve (which determines the potential of the population to respond over the longer term). However, there is little data to quantify the speed of evolution in the face of climate change (but see Berteaux et al., 2004). This lack of hard data was emphasized by Hedrick (2001), who highlighted that "determining the potential for adaptation of a population in a changing environment is a critical unanswered question in most cases". Alternatively, the fossil record favours the hypothesis of niche conservatism for most species in a climate change context (Prinzing et al., 2001; Ackerly, 2003; Wiens and Graham, 2005; but see Broennimann et al., 2007 for the case of biological invasions); this may be true for isolated populations at the rear edge of a shifting range because the relatively poor *intra*-population diversity will reduce the evolutionary potential in the face of rapid environmental change (Davis and Shaw, 2001; Parmesan, 2006; Willi et al., 2006).

Also, given that microevolutionary processes occur across generations, generation time is an important factor affecting the evolutionary potential of species if environmental change is rapid (Rosenheim and Tabashnik, 1991). From that, species with a long generation time such as perennial herbs, shrubs and trees should have a lower capability to respond to new selective pressures than species with a short generation time such as annual herbs (Rosenheim and Tabashnik, 1991). On the other hand, species with a long generation time have a higher probability of population persistence, at least over short timescales (see above). Thus, to predict the long-term fate of populations exposed to environmental change, one has to quantify the effects of both ecological and evolutionary processes on demographic rates (Botkin et al., 2007).

Key issues

The understanding and quantification of demographic processes is key to the modelling of population extinction under environmental change. The link between demography and population extinction has been studied extensively in the field of population viability analysis (PVA, see Menges, 2000; Reed et al., 2002 for reviews of PVAs of plants), leading to a substantial theory on population extinction (e.g. Lande, 1993; Frank and Wissel, 2002; Grimm and Wissel, 2004). A minimal spatial model for population extinction contains the growth rate and carrying capacity of individual populations, and the probability of dispersal between populations (the latter is relevant since immigration may rescue small populations from extinction, Brown and Kodric-Brown, 1977). The basic demographic quantities necessary for predicting population extinction are thus easily identified; the challenge to distribution modelling lies in predicting how these quantities will be modified by environmental conditions, inter-specific interactions, phenotypic plasticity and evolutionary change. In the following, we provide perspectives on how these modifications can be represented in distribution models.

Perspectives

Environmental effects on demographic processes can be easily incorporated into distribution models if environmental change causes an abrupt change in demographic rates, leading to rapid population extinction. In the context of land-use change, such abrupt changes have already been incorporated into distribution forecasts by overlaying habitat models' forecasts with a post-modelling mask (Broennimann et al., 2006). Climate change and invasive species are, however, likely to induce gradual changes in demographic parameters and the probability of local extinction. Such gradual changes pose a greater challenge to distribution modelling because their description requires quantitative information on the relationship between demography and environment. In the context of species' migration, we have already discussed how distribution models can incorporate gradual effects of *mean environmental conditions* on demographic rates. In the context of population extinction, however, it seems particularly important to additionally consider the effects of *temporal and spatial variability in environmental conditions* (e.g. Higgins et al., 2000; Menges, 2000), especially extreme events (Parmesan et al., 2000).

Temporal variability in the environment is commonly believed to increase the probability of population extinction (e.g. Menges, 2000), and its consideration in distribution forecasts is particularly important if environmental variability increases due to climate change.

The more temporally variable the environment becomes, the more the persistence of a population will depend on its ability to maintain reproductive potential during unfavourable periods (e.g. Higgins et al., 2000). Plant populations can store reproductive potential in seed and sapling banks or in long-lived adults (Bond and Midgley, 2001). Higgins et al. (2000) noted that species with a high storage reproductive potential may even benefit from an environment becoming more variable, as this may lead to a higher frequency of conditions beneficial for recruitment. Their simple model for local population dynamics includes the storage effect as well as temporal variability in the environment, and could be embedded into process-based distribution models.

Broad-scale spatial variability in the environment is usually accounted for in habitat models and this information can be linked to demographic rates in a process-based approach. Small-scale spatial variability, however, poses a problem to distribution modelling. Existing distribution models are typically grid based and assume that environmental conditions within large grid cells are constant. Yet, small-scale environmental variability within grid cells may be crucially important for predicting species' responses to climate change. In a heterogeneous grid cell, a species may be able to escape to another habitat type or to higher altitudes and therefore persist longer than in a homogeneous environment (Jackson and Overpeck, 2000). Species' distribution models can account for small-scale variability in environmental conditions by considering the area covered by different habitat types or the altitudinal range within each grid cell. A similar approach has already been taken to describe habitat transformation through the incorporation of a data layer that gives the proportion of area transformed inside each grid cell (e.g. Latimer et al., 2006).

Biotic interactions are also likely to alter the demographic determinants of population persistence or extinction. Competition for resources is undoubtedly important for determining range size (Case et al., 2005; Brooker, 2006; Morin and Chuine, 2006). Leathwick and Austin (2001) showed how competition by *Nothofagus* shaped the distribution of other tree genera in New Zealand. Facilitation can have similar controls on rear-edge boundaries: Castro et al. (2004) have shown that Scot's pine (*Pinus sylvestris*) seedlings at their southern limit in Spain have higher survival rates under shrubs (that effectively shelter the pine seedlings from drought, frost, hail and trampling) than on bare ground or under woodland canopies. Both these examples highlight how interactions refine the broad-scale control that the environment has on a species' distribution and suggest that interactions may become increasingly important for range dynamics in the future. Approaches that can be used to model the effects of interactions on species' distribution and population spread may also be useful for modelling the consequences of interactions for population extinction.

Phenotypic plasticity and evolutionary adaptation in response to climate change are currently being studied (e.g. Willi et al., 2006; Bridle and Vines, 2007), and may find their way into process-based models over the next few years. Indeed, understanding the processes and patterns of gene flow and local adaptation requires a detailed knowledge of how geographic space and landscape characteristics structure populations (Antonovics et al., 2002). We foresee that such information will be transferable into distribution modelling in the near future. One promising research direction would thus be to simulate adaptive genetic responses in spatially explicit models of species' distribution, a new task for the field of landscape genetics (Manel et al., 2003; Metcalf and Pavard, 2007). Tools for jointly simulating the dynamics of plant populations and genotypes are now becoming available: the modelling framework AMELIE (Kuparinen and Schurr, 2007) links the spatio-temporal dynamics of plant populations and genotypes, and is flexible in its description of life histories, seed and pollen dispersal, reproductive systems, and demographic and environmental stochasticity. Individual life cycle stages can be described by a wide range of alternative submodels, allowing, for instance, plant growth to be modelled with a physiological model. The AMELIE framework can thus be used for including evolutionary processes into distribution modelling, given that sufficient demographic and genetic data are available.

Alternatively, when no data are available, simple adaptive scenarios may be drawn up and included into distribution modelling outputs (Berteaux et al., 2004). For instance, for a quantitative trait such as response to climate (influenced by multiple genes, each having a small effect), we can describe several situations depending on the magnitude of phenotypic plasticity (PP) and the speed of evolution. In one extreme situation, contemporary evolution is slow and there is virtually no PP for the trait considered. Sources of change in the average trait value are weak, and if the selection gradient acting on the trait (the slope of the relationship between the trait and fitness) is steep, the average fitness of the population decreases under the pressure of climate change (Berteaux et al., 2004). In the opposite case PP is high and contemporary evolution is fast, so that sources of change in the average trait value are strong, and the average fitness of the population remains stable. A particular situation may occur, where antagonistic plastic and genetic responses to environmental change will lead to apparent evolutionary stasis (no change or contradictory change in the phenotype with time), despite selection and genetic variation of the trait (Merilä et al., 2001). Depending on the magnitude of PP and the speed of evolution, one can draw optimistic or

pessimistic scenarios when predicting changes in species' distribution under environmental change (Berteaux et al., 2004).

To conclude, we argue that range gain and loss are two sides of the same coin: range loss is caused by the extinction of existing populations, whereas migration is driven by newly founded populations that escape extinction. The process-based modelling of range dynamics will thus greatly benefit from better data on the factors causing the extinction or persistence of small populations.

Conclusion

Models are simplifications of reality and often begin life by helping researchers to formalize their understanding of a particular process or pattern of interest. Models are thus primarily important aids to research. Difficulties may therefore arise when such theoretical models are used to guide conservation planning, management and to support the formulation of policy decisions (e.g. IPCC). The magnitude of uncertainties in species' range modelling is currently so great that it might lead conservation planners, policy makers and other stakeholders to question the overall usefulness of science as an aid to solve real world problems. Bridging the perceived gap between science and societal needs is of paramount importance if we want to make progress and contribute meaningfully, as scientists, to solving the global environmental change crises.

The conservation agenda is now moving on to consider adaptation to climate change and here a landscape approach is more applicable for testing strategies such as habitat re-creation, corridors and ecosystem resilience. Habitat models may contribute to this in providing a broader framework for understanding and predicting climate change impacts, delineating potential corridors for species migration (Williams et al., 2005) or testing the effectiveness of reserve networks to protect biodiversity (Burns et al., 2003; Araújo et al., 2004; Bomhard et al., 2005). Perhaps we should endorse Sutherland's suggestion (Sutherland, 2006) and ask conservationists to identify desired future states and then use models to backcast in order to identify strategies for achieving the desired state. The time is ripe for modellers to explore how far models can be taken to both fulfil the requirements of and answer the critical questions for conservation in a dynamic future.

One key point is still puzzling to us and generates lively discussions among the authors of this paper: Are more complex models (i.e. including a large number of parameters) likely to be more accurate than simple models (i.e. including few parameters)? When should we rely on complex versus simple models to aid policy decisions? This is a reflection of an older debate over the merits of accuracy versus realism, or between simplicity versus complexity. We could of course argue that adding dispersal and physiology makes models more realistic and closer to the truth, but are these models necessarily more accurate? This is a complex issue as habitat models have proved to be very accurate under current conditions even though they ignore species physiology and other properties (Hijmans and Graham, 2006). Nevertheless, some conclusions can be tentatively made: complex models are likely to be more accurate at finer resolutions, whereas simple models are likely to offer useful and parsimonious solutions at broader scales; complex models are probably more accurate under novel situations and for species with non-equilibrium distributions with current environmental conditions (Peng, 2000); the development of complex models is necessary not only as an academic exercise to help understanding of the importance of different drivers of change on the distribution of species, but also as an aid to identify the optimal trade-offs between precision and generality. Furthermore, future research may help develop models that are still simple yet have better predictive power (e.g. based on more advanced environmental predictors or including simple interactions; Leathwick and Austin, 2001; Robertson et al., 2003; Araújo et al., 2005; Guisan et al., 2006; Heikkinen et al., 2006).

In recent years we have seen a new generation of models, termed "hybrid" models (TreeMig (Lischke et al., 2006), LPJ-GUESS (Smith et al., 2001), Hybrid (Friend et al., 1997)). These models try to achieve a compromise between realism–accuracy and complexity–simplicity. Such developments are opening new pathways for research and are bound to provide many answers for the questions raised in this review. One of the limitations of these approaches is that they are highly sophisticated, data hungry, and require detailed knowledge of ecological processes that is usually unavailable for large numbers of species and regions. Nevertheless, the literature abounds with experimental and observational studies carried out at local scales which are only rarely used by modellers. These studies have been accumulating over at least two centuries and offer a powerful source of information. This, if appropriately synthesized and digested, could be extremely useful in helping to select and fine tune the parameters entering process-based models across a variety of species, functional types and regions (Midgley and Thuiller, 2005). A systematic use of the information available in the literature, and directed pairing of experimentalists with distribution modellers, would help in testing model performance against observations recorded with a moderate level of climate change already observed, and thus help to identify the gaps in

knowledge that need filling in order to efficiently set the targets for conducting large-scale experiments required to obtain meaningful information for the parameterization of distribution models. Existing global biological traits databases are an example of accumulated meta-information that will prove to be extremely helpful in a more realistic parameterization of models.

The future holds great promise for researchers in global change biology: there is indeed much to be done, but there is also real hope that integrative programmes and projects will help filling in gaps between distinctive fields of research, thus contributing to the emergence of the new generation of forecasting methodologies that are more useful for decision making.

Acknowledgments

We would like to thank Kirk Moloney and Florian Jeltsch for having invited us to participate to this special issue. The ideas leading to this paper were developed as part of the kick-off meeting of the EU-funded FP6 MACIS (Minimization of and Adaptation to Climate change Impacts on biodiverSity, No. 044399) project, chaired by Martin Sykes at Lund University (Sweden) in November 2006. A.G., M.B.A., M.T.S., N.E.Z. and W.T. also receive support from the EU-funded FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, No. 066866 GOCE) integrated project. W.T. and C.A. were also supported by the "Agence National pour la Recherche" through the QDiv project (Quantifying the Effects of Global Environmental Change on Terrestrial Plant Diversity). F.M.S. acknowledges support from the European Union through Marie Curie Transfer of Knowledge Project FEMMES (MTKD-CT-2006-042261). We thank C. Dormann, A.T. Peterson and on anonymous reviewer for their constructive comments on the manuscript.

References

Ackerly, D.D., 2003. Community assembly, niche conservatism, and adaptive evolution in changing environments. Int. J. Plant Sci. 164, S165–S184.

Albert, C., Thuiller, W., Lavorel, S., Davies, I.D., Garbolino, E., 2007. Land use change and sub-alpine tree dynamics: colonisation of *Larix decidua* in French sub-alpine grasslands. Journal of Applied Ecology in press, doi:10.1111/j.1365-2664.2007.01416.x.

Antonovics, J., Newman, T.J., Best, B.J., 2002. Spatially explicit studies on the ecology and genetics of population margins. In: Silvertown, J., Antonovics, J. (Eds.), Integrating Ecology and Evolution in a Spatial Context. Blackwell, Oxford, pp. 97–117.

Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. J. Biogeogr. 33, 1677–1688.

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22, 42–47.

Araújo, M.B., Cabeza, M., Thuiller, W., Hannah, L., Williams, P.H., 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. Global Change Biol. 10, 1618–1626.

Araújo, M.B., Thuiller, W., Williams, P.H., Reginster, I., 2005. Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. Global Ecol. Biogeogr. 14, 17–30.

Araújo, M.B., Thuiller, W., Pearson, R.G., 2006. Climate warming and the decline of amphibians and reptiles in Europe. J. Biogeogr. 33, 1712–1728.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Model. 157, 101–118.

Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol. Model. 200, 1–19.

Badeck, F.-W., Lischke, H., Bugmann, H., Hickler, T., Hönninger, K., Lasch, P., Lexer, M.J., Mouillot, F., Schaber, J., Smith, B., 2001. Tree species composition in European pristine forests. Comparison of stand data to model predictions. Climatic Change 51, 307–347.

Berry, P.M., Rounsevell, M.D.A., Harrison, P.A., Audsley, E., 2006. Assessing the vulnerability of agricultural land use and species to climate change and the role of policy in facilitating adaptation. Environ. Sci. Policy 9, 189–204.

Berteaux, D., Réale, D., McAdam, A.G., Boutin, S., 2004. Keeping pace with fast climate change: can arctic life count on evolution? Integr. Comp. Biol. 44, 140–151.

Bomhard, B., Richardson, D.M., Donaldson, J.S., Hughes, G.O., Midgley, G.F., Raimondo, D.C., Rebelo, A.G., Rouget, M., Thuiller, W., 2005. Potential impacts of future land use and climate change on the red list status of the Proteaceae in the Cape Floristic region, South Africa. Global Change Biol. 11, 1452–1468.

Bond, W.J., Midgley, J.J., 2001. The persistence niche: ecology of sprouting in woody plants. Trends Ecol. Evol. 16, 45–51.

Botkin, D.B., Janak, J.F., Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. J. Ecol. 60, 849–872.

Botkin, D., Saxe, H., Araújo, M.B., Betts, R., Bradshaw, R., Cedhagen, T., Chesson, P., Davis, M.B., Dawson, T., Etterson, J., Faith, D.P., Guisan, A., Ferrier, S., Hansen, A.S., Hilbert, D., Kareiva, P., Margules, C., New, M., Skov, F., Sobel, M.J., Stockwell, D., 2007. Forecasting effects of global warming on biodiversity. Bioscience 57, 227–236.

Bridle, J.R., Vines, T.H., 2007. Limits to evolution at range margins: when and why does adaptation fail? Trends Ecol. Evol. 22, 140–147.

Broennimann, O., Thuiller, W., Hughes, G.O., Midgley, G.F., Alkemade, J.R.M., Guisan, A., 2006. Do geographic distribution, niche property and life form explain plants' vulnerability to global change? Global Change Biol. 12, 1079–1093.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T.P., Guisan, A., 2007. Evidence of

climatic niche shift during biological invasion. Ecol. Lett. 10, 701–709.

Brooker, R.W., 2006. Plant–plant interactions and environmental change. New Phytol. 171, 271–284.

Brown, J.H., Kodric-Brown, A., 1977. Turnover rates in insular biogeography – effect of immigration on extinction. Ecology 58, 445–449.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., West, G.B., 2004. Toward a metabolic theory of ecology. Ecology 85, 1771–1789.

Bugmann, H., 2001. A review of forest gap models. Climatic Change 51, 259–305.

Bugmann, H.K.M., Solomon, A.M., 2000. Explaining forest composition and biomass across multiple biogeographical regions. Ecol. Appl. 10, 95–114.

Bullock, J.M., Shea, K., Skarpass, O., 2006. Measuring plant dispersal: an introduction to field methods and experimental design. Plant Ecol. 186, 217–234.

Bürger, R., Lynch, M., 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49, 151–163.

Burns, C.E., Johnson, K.M., Schmitz, O.J., 2003. Global climate change and mammalian species diversity in US national parks. Proc. Natl. Acad. Sci. USA 100, 11474–11477.

Cabeza, M., Moilanen, A., 2001. Design of reserve networks and the persistence of biodiversity. Trends Ecol. Evol. 16, 242–248.

Cain, M.L., Milligan, B.G., Strand, A.E., 2000. Long-distance seed dispersal in plant populations. Am. J. Bot. 87, 1217–1227.

Cappuccino, N., 2004. Allee effect in an invasive alien plant, pale swallow-wort *Vincetoxicum rossicum* (Asclepiadaceae). Oikos 106, 3–8.

Case, T.J., Holt, R.D., McPeek, M.A., Keitt, T.H., 2005. The community context of species' borders: ecological and evolutionary perspectives. Oikos 108, 28–46.

Castro, J., Zamora, R., Hodar, J.A., Gomez, J.M., 2004. Seedling establishment of a boreal tree species (*Pinus sylvestris*) at its southernmost distribution limit: consequences of being in a marginal Mediterranean habitat. Journal of Ecology 92, 266–277.

Chuine, I., Cambon, G., Comtois, P., 2000. Scaling phenology from the local to the regional levels: advances from species-specific phenological models. Global Change Biol. 6, 1–11.

Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. Ecol. Lett. 8, 2–14.

Clark, J.S., Fastie, C., Hurtt, G.H., Jackson, S.T., Johnson, C., King, G.A., Lewis, M., Lynch, J., Pacala, S., Prentice, I.C., Schupp, E.W., Webb III, T., Wyckoff, P., 1998. Reid's paradox of rapid plant migration. BioScience 48, 13–24.

Clark, J.S., Lewis, M., Horvath, L., 2001. Invasion by extremes: population spread with variation in dispersal and reproduction. Am. Nat. 157, 537–554.

Collingham, Y.C., Huntley, B., 2000. Impacts of habitat fragmentation and patch size upon migration rates. Ecol. Appl. 10, 131–144.

Collingham, Y., Hill, M.O., Huntley, B., 1996. The migration of sessile organisms: a simulation model with measurable parameters. J. Veg. Sci. 7, 831–846.

Cousins, S.A.O., Lavorel, S., Davies, I., 2003. Modelling the effects of landscape pattern and grazing regimes on the persistence of plant species with high conservation value in grasslands in south-eastern Sweden. Landscape Ecol. 18, 315–332.

Davis, M.B., 1981. Quaternary history and the stability of forest communities. In: West, D.C., Shugart, H.H., Botkin, D.B. (Eds.), Forest Succession. Springer, New York, pp. 132–153.

Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptative responses to quaternary climate change. Science 292, 673–679.

Davis, H.G., Taylor, C.M., Civille, J.C., Strong, D.R., 2004. An Allee effect at the front of a plant invasion: Spartina in a Pacific estuary. J. Ecol. 92, 321–327.

Etterson, J.R., Shaw, R.G., 2001. Constraint to adaptive evolution in response to global warming. Science 294, 151–154.

Ferrier, S., Guisan, A., 2006. Spatial modelling of biodiversity at the community level. J. Appl. Ecol. 43, 393–404.

Foden, W., Midgley, G.F., Hughes, G.O., Bond, W.J., Thuiller, W., Hoffman, M.T., Kaleme, P., Underhill, L.G., Rebelo, A.G., Hannah, L., 2007. A changing climate is eroding the geographical range of the Namib Desert tree Aloe through population declines and dispersal lags. Divers. Distrib. 13, 645–653.

Franco, M., Silvertown, J., 2004. A comparative demography of plants based upon elasticities of vital rates. Ecology 85, 531–538.

Frank, K., Wissel, C., 2002. A formula for the mean lifetime of metapopulations in heterogeneous landscapes. Am. Nat. 159, 530–552.

Friend, A.D., Stevens, A.K., Knox, R.G., Cannell, M.G.R., 1997. A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). Ecol. Model. 95, 249–287.

Ghalambor, C.K., McKay, J.K., Carroll, S.P., Reznick, D.N., 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Funct. Ecol. 21, 394–407.

Godoy, J.A., Jordano, P., 2001. Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. Mol. Ecol. 10, 2275–2283.

Grimm, V., Wissel, C., 2004. The intrinsic mean time to extinction: a unifying approach to analysing persistence and viability of populations. Oikos 105, 501–511.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M.P., Overton, J.M.C., Aspinall, R., Hastie, T.J., 2006. Making better biogeographical predictions of species' distributions. J. Appl. Ecol. 43, 386–392.

Hampe, A., 2005. Fecundity limits in *Frangula alnus* (Rhamnaceae) relict populations at the species' southern range margin. Oecologia 143, 377–386.

Hampe, A., Petit, R.J., 2005. Conserving biodiversity under climate change: the rear edge matters. Ecol. Lett. 8, 461–467.

Hamrick, J.L., Loveless, M.D., 1989. The genetic structure of tropical tree populations: associations with reproductive biology. In: Bock, J.H., Linhart, Y.B. (Eds.), The Evolutionary Ecology of Plants. Westview Press, Boulder, CO, USA.

Hannah, L., Midgley, G.F., Andelman, S., Araújo, M.B., Hughes, G.O., Martinez-Meyer, E., Pearson, R.G., Williams, P.H., 2007. Protected area needs in a changing climate. Frontiers Ecol. Environ. 5, 131–138.

Hanski, I., 1994. A practical model of metapopulation dynamics. J. Anim. Ecol. 63, 151–162.

Hawkins, B.A., Araújo, M.B., Cabrero-Sañudo, F.J., Castro Parga, I., Diniz-Filho, J.A.F., Ferrer-Castán, D., Field, R., Gómez, J.F., Qian, H., Hortal, J., Kerr, J.T., Lobo, J.M., Montoya, D., Moreno, J.C., Olalla-Tárraga, M.À., Pausas, J.G., Rahbek, C., Rodríguez, M.À., Sanders, N.J., Albuquerque, F.S., Williams, P.H., 2007. A global evaluation of metabolic theory as an explanation of diversity gradients. Ecology.

He, H.S., Mladenoff, D.J., 1999. Spatially explicit and stochastic simulation of forest-landscape fire disturbance and succession. Ecology 80, 81–99.

Hedrick, P.W., 2001. Evolutionary conservation biology. In: Fox, C.W., Roff, D.A., Fairbairn, D.J. (Eds.), Evolutionary Ecology: Concepts and Case Studies. Oxford University Press, Oxford, UK, pp. 371–383.

Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W., Sykes, M.T., 2006. Methods and uncertainties in bioclimatic envelope modeling under climate change. Prog. Phys. Geogr. 30, 751–777.

Hewitt, G.M., 1993. Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In: Lees, D.R., Edwards, D. (Eds.), Evolutionary Patterns and Processes. Academic Press, London, pp. 97–123.

Hickler, T., Smith, B., Sykes, M.T., Davis, M., Sugita, S., Walker, K., 2004. Using a generalized vegetation model to simulate vegetation dynamics in northeastern USA. Ecology 85, 519–530.

Higgins, S.I., Richardson, D.M., 1999. Predicting plant migration rate in a changing World: the role of long-distance dispersal. Am. Nat. 153, 464–475.

Higgins, S.I., Pickett, S.T.A., Bond, W.J., 2000. Predicting extinction risks for plants: environmental stochasticity can save declining populations. Trends Ecol. Evol. 15, 516–520.

Higgins, S.I., Clark, J.S., Nathan, R., Hovestadt, T., Schurr, F., Fragasos, J.M.V., Aguiar, M.R., Ribbens, E., Lavorel, S., 2003a. Forecasting plant migration rates: managing uncertainty for risk assessment. J. Ecol. 91, 341–347.

Higgins, S.I., Lavorel, S., Revilla, E., 2003b. Estimating plant migration rates under habitat loss and fragmentation. Oikos 101, 354–366.

Hijmans, R.J., Graham, C.H., 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. Global Change Biol. 12, 2272–2281.

Hodkinson, D.J., Thompson, K., 1997. Plant dispersal: the role of man. J. Appl. Ecol. 34, 1484–1496.

Huntley, B., 1991. How plants respond to climate change: migration rates, individualism and the consequences for plant communities. Ann. Bot. 67 (Suppl 1), 15–22.

Iverson, L.R., Schwartz, M.W., Prasad, A.M., 2004. Potential colonization of newly available tree species habitat under climate change: an analysis for five eastern US species. Landscape Ecol. 19, 787–799.

Jackson, S.T., Overpeck, J.T., 2000. Responses of plant populations and communities to environmental changes of the late quaternary. Paleobiology 26, 194–220.

Jeltsch, F., Moloney, K.A., Schurr, F.M., Köchy, M., Schwager, M., 2008. The state of plant population modelling in light of environmental change. Perspect. Plant Ecol. Evol. Systematics 9, 171–190.

Jongejans, E., Skarpaas, O., Shea, K., 2008. Dispersal, demography and spatial population models for conservation and control management. Perspect. Plant Ecol. Evol. Systematics 9, 153–170.

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

Kinlan, B.P., Gaines, S.D., 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84, 2007–2020.

Kot, M., Lewis, M.A., van den Driesshe, P., 1996. Dispersal data and the spread of invading organisms. Ecology 77, 2027–2042.

Kowarik, I., 2003. Human agency in biological invasions: secondary releases foster naturalisation and population expansion of alien plant species. Biol. Invasions 5, 293–312.

Kriticos, D.J., Sutherst, R.W., Brown, J.R., Adkins, S.W., Maywald, G.F., 2003. Climate change and the potential distribution of an invasive alien plant: *Acacia nilotica* ssp. indica in Australia. J. Appl. Ecol. 40, 111–124.

Kunstler, G., Thuiller, W., Curt, T., Bouchaud, M., Jouvie, R., Deruette, F., Lepart, J., 2007. *Fagus sylvatica* L. recruitment across a fragmented Mediterranean landscape, importance of long distance effective dispersal, abiotic conditions and biotic interactions. Divers. Distrib., in press.

Kuparinen, A., Schurr, F.M., 2007. A flexible modelling framework linking the spatio-temporal dynamics of plant genotypes and populations: application to gene flow from transgenic forests. Ecol. Model. 202, 476–486.

Lande, R., 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am. Nat. 142, 911–927.

Latimer, A.M., Wu, S.S., Gelfand, A.E., Silander, J.A., 2006. Building statistical models to analyze species distributions. Ecol. Appl. 16, 33–50.

Lavergne, S., Thuiller, W., Molina, J., Debussche, M., 2005. Environmental and human factors influencing rare plant local occurrence, extinction and persistence: a 115 year study in the Mediterranean region. J. Biogeogr. 32, 799–811.

Lavergne, S., Molina, J., Debussche, M., 2006. Fingerprints of environmental change on the rare Mediterranean flora: a 115-year study. Global Change Biol. 12, 1466–1478.

Leathwick, J.R., Austin, M.P., 2001. Competitive interactions between tree species in New Zealand's old-growth indigenous forests. Ecology 82, 2560–2573.

Lesica, P., McCune, B., 2004. Decline of arctic-alpine plants at the southern margin of their range following a decade of climate warming. J. Veg. Sci. 15, 679–690.

Levin, S.A., Muller-Landau, H.C., Nathan, R., Chave, J., 2003. The ecology and evolution of seed dispersal: a theoretical perspective. Annu. Rev. Ecol. Evol. Syst. 34, 575–604.

Lischke, H., Zimmermann, N.E., Bolliger, J., Rickebusch, S., Löffler, T.J., 2006. TreeMig: a forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. Ecol. Model. 199, 409–420.

Manel, S., Schwartz, M.K., Luikart, G., Taberlet, P., 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. 18, 189–197.

MEA, 2005. Ecosystems and Human Well-Being: Biodiversity Synthesis. World Resources Institute, Washington, DC.

Menges, E.S., 2000. Population viability analyses in plants: challenges and opportunities. Trends Ecol. Evol. 15, 51–55.

Menzel, A., Fabian, P., 1999. Growing season extended in Europe. Nature 397, 659.

Merilä, J., Sheldon, B.C., Kruuk, L.E.B., 2001. Explaining statis: microevolutionary studies in natural populations. Genetica 112–113, 199–222.

Metcalf, C.J.E., Pavard, S., 2007. Why evolutionary biologists should be demographers. Trends Ecol. Evol. 22, 205–212.

Midgley, G.F., Thuiller, W., 2005. Global environmental change and the uncertain fate of biodiversity. New Phytol. 167, 638–641.

Midgley, G.F., Hannah, L., Millar, D., Thuiller, W., Booth, A., 2003. Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic region. Biol. Conserv. 112, 87–97.

Midgley, G.F., Hughes, G.O., Thuiller, W., Rebelo, A.G., 2006. Migration rate limitations on climate change induced range shifts in Cape Proteaceae. Divers. Distrib. 12, 555–562.

Midgley, G.F., Thuiller, W., Higgins, S.I., 2007. Plant species migration as a key uncertainty in predicting future impacts of climate change on ecosystems: progress and challenges. In: Canadell, J.G., Pataki, D.E., Pitelka, L.F. (Eds.), Terrestrial Ecosystems in a Changing World. Springer, New York, pp. 129–137.

Moorcroft, P.R., Hurtt, G.C., Pacala, S.W., 2001. A method for scaling vegetation dynamics: the ecosystem demography model (ED). Ecol. Monogr. 74, 557–586.

Morin, X., Chuine, I., 2006. Niche breadth, competitive strength and range size of tree species: a trade-off based framework to understand species distribution. Ecol. Lett. 9, 185–195.

Nathan, R., 2006. Long distance dispersal of plants. Science 313, 786–788.

Nathan, R., Katul, G., Horn, H.S., Thomas, S.M., Oren, R., Avissar, R., Pacala, S.W., Levin, S.A., 2002. Mechanisms of long-distance dispersal of seeds by wind. Nature 418, 409–413.

Neilson, R.P., Pitelka, L.F., Solomon, A.M., Nathan, R., Midgley, G.F., Fragoso, J., Lischke, H., Thompson, K., 2005. Forecasting regional to global plant migration in response to climate change: challenges and directions. Bioscience 55, 749–759.

Neubert, M., Caswell, H., 2000. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. Ecology 81, 1613–1628.

Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. Annu. Rev. Ecol. Evol. Syst. 37, 637–669.

Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421, 37–42.

Parmesan, C., Root, T.L., Willig, M.R., 2000. Impacts of extreme weather and climate on terrestrial biota. Bull. Am. Meteorol. Soc. 81, 443–450.

Pauli, H., Gottfried, M., Reiter, K., Klettner, C., Grabherr, G., 2006. Signals of range expansions and contractions of vascular plants in the high Alps: observations (1994–2004) at the GLORIA*master site Schrankogel, Tyrol, Austria. Global Change Biol. 13, 147–156.

Pearson, R.G., 2006. Climate change and the migration capacity of species. Trends Ecol. Evol. 21, 111–113.

Pearson, R.G., Dawson, T.P., 2005. Long-distance plant dispersal and habitat fragmentation: identifying conservation targets for spatial landscape planning under climate change. Biol. Conserv. 123, 389–401.

Peng, C., 2000. From static biogeographical model to dynamic global vegetation model: a global perspective on modelling vegetation dynamics. Ecol. Model. 135, 33–54.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H., Stockwell, D.R.B., 2002. Future projections for Mexican faunas under global climate change scenarios. Nature 416, 626–629.

Petit, C., Freville, H., Mignot, A., Colas, B., Riba, M., Imbert, E., Hurtrez-Bousses, S., Virevaire, M., Olivieri, I., 2001. Gene flow and local adaptation in two endemic plant species. Biol. Conserv. 100, 21–34.

Pitelka, L.F., Gardner, R.H., Ash, J., Berry, S., Gitay, H., Noble, I.R., Saunders, A., Bradshaw, R.H.W., Brubaker, L., Clark, J.S., Davis, M.B., Sugita, S., Dyer, J.M., Hengeveld, R., Hope, G., Huntley, B., King, G.A., Lavorel, S., Mack, R.N., Malanson, G.P., McGlone, M., Prentice, I.C., Rejmanek, M., 1997. Plant migration and climate change. Am. Sci. 85, 464–473.

Possingham, H.P., Davies, I., 1995. Alex: a model for the viability analysis of spatially structured populations. Biol. Conserv. 73, 143–150.

Pounds, J.A., Fogden, M.P.L., Campbell, J.H., 1999. Biological response to climate change on a tropical mountain. Nature 398, 611–615.

Pounds, A.J., Bustamante, M.R., Coloma, L.A., Consuegra, J.A., Fogden, M.P.L., Foster, P.N., La Marca, E., Masters, K.L., Merino-Viteri, A., Puschendorf, R., Ron, S.R., Sanchez-Azofeifa, G.A., Still, C.J., Young, B.E., 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. Nature 439, 161–167.

Powell, J.A., Zimmermann, N.E., 2004. Multi-scale analysis of seed dispersal and the resolution of Reid's paradox. Ecology 85, 490–506.

Prentice, I.C., Jolly, D., BIOME.6000.participants, 2000. Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. J. Biogeogr. 27, 507–519.

Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M., Wilson, K.A., 2007. Conservation planning in a changing world. Trends Ecol. Evol., in press.

Prinzing, A., Durka, W., Klotz, S., Brandl, R., 2001. The niche of higher plants: evidence for phylogenetic conservatism. Proc. Roy. Soc. London B 268, 2383–2389.

Reed, J.M., Mills, L.S., Dunning, J.B., Menges, E.S., McKelvey, K.S., Frye, R., Beissinger, S.R., Anstett, M.C., Miller, P., 2002. Emerging issues in population viability analysis. Conserv. Biol. 16, 7–19.

Ribbens, E., Silander, J.A., Pacala, S.W., 1994. Seedling recruitment in forests: calibrating models to predict patterns of tree seedling dispersion. Ecology 76, 1794–1806.

Robertson, M.P., Peter, C.I., Villet, M.H., Ripley, B.S., 2003. Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modelling techniques. Ecol. Model. 164, 153–167.

Root, T.L., MacMynowski, D.P., Mastrandrea, M.D., Schneider, S.H., 2005. Human-modified temperatures induce species changes: joint attribution. Proc. Natl. Acad. Sci. USA 102, 7465–7469.

Rosenheim, J.A., Tabashnik, B.E., 1991. Influence of generation time on the rate of response to selection. Am. Nat. 137, 527–541.

Rousset, F., 2001. Genetic approaches to the estimation of dispersal rates. In: Clobert, J., Danchin, E., Dhondt, A.A., Nichols, J.D. (Eds.), Dispersal. Oxford University Press, New York, pp. 18–28.

Sagarin, R.D., Gaines, S.D., Gaylord, B., 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. Trends Ecol. Evol. 21, 524–530.

Schmidt, W., 1989. Plant dispersal by motor cars. Vegetatio 80, 147–152.

Schröter, D., Cramer, W., Leemasn, R., Prentice, I.C., Araújo, M.B., Arnell, N.W., Bondeau, A., Bugmann, H., Carter, T.R., Gracia, C.A., de la Vega-Leinert, A.C., Erhard, M., Ewert, F., Glendining, M., House, J.I., Kankaanpää, S., Klein, R.J.T., Lavorel, S., Lindner, M., Metzger, M.J., Meyer, J., Mitchell, T.D., Reginster, I., Rounsevell, M., Sabaté, S., Sitch, S., Smith, B., Smith, J.F., Smith, P., Sykes, M.T., Thonicke, K., Thuiller, W., Tuck, G., Zaehle, S., Zierl, B., 2005. Ecosystem service supply and vulnerability to global change in Europe. Science 310, 1333–1337.

Schumacher, S., Bugmann, H., 2006. The relative importance of climatic effects, wildfires and management for future forest landscape dynamics in the Swiss Alps. Global Change Biol. 12, 1435–1450.

Schumacher, S., Bugmann, H., Mladenoff, D.J., 2004. Improving the formulation of tree growth and succession in a spatially explicit landscape model. Ecol. Model. 180, 175–194.

Schumacher, S., Reineking, B., Sibold, J., Bugmann, H., 2006. Modeling the impact of climate and vegetation on fire regimes in mountain landscapes. Landscape Ecol. 21, 539–554.

Schurr, F.M., Bond, W.J., Midgley, G.F., Higgins, S.I., 2005. A mechanistic model for secondary seed dispersal by wind and its experimental validation. J. Ecol. 93, 1017–1028.

Schurr, F.M., Midgley, G.F., Rebelo, A.G., Reeves, G., Poschlod, P., Higgins, S.I., 2007. Colonization and persistence ability explain the extent to which plant species fill their potential range. Global Ecol. Biogeogr. 16, 449–459.

Schwartz, M.W., Iverson, L.R., Prasad, A.M., S.N., M., O'Connor, R.J., 2006. Predicting extinctions as a result of climate change. Ecology 87, 1611–1615.

Shugart, H.H., 1984. A Theory of Forest Dynamics. The Ecological Implications of Forest Succession Models. Springer, New York.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W., Sykes, M.T., Thonicke, K., Venevsky, S., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biol. 9, 161–185.

Skellam, I.G., 1951. Random dispersal in theoretical populations. Biometrika 38, 196–218.

Smith, B., Prentice, I.C., Sykes, M.T., 2001. Representation of vegetation dynamics in modelling of European ecosystems: comparison of two contrasting approaches. Global Ecol. Biogeogr. 10, 621–637.

Solomon, A.M., Kirilenko, A.P., 1997. Climate change and terrestrial biomass: what if trees do not migrate!. Global Ecol. Biogeogr. Lett. 6, 139–148.

Sutherland, W.J., 2006. Predicting the ecological consequences of environmental change: a review of the methods. J. Ecol. 43, 599–616.

Sutherst, R.W., 2000. Climate change and invasive species – a conceptual framework. In: Mooney, H.A., Hobbs, R.J. (Eds.), Invasive Species in a Changing World. Island Press, Washington, DC, pp. 211–240.

Sutherst, R.W., Maywald, H.F., Russel, B.L., 2000. Estimating vulnerability under global change: modular modelling of pests. Agric. Ecosyst. Environ. 82, 303–319.

Svenning, J.-C., Skov, F., 2004. Limited filling of the potential range in European tree species. Ecol. Lett. 7, 565–573.

Sykes, M.T., Prentice, I.C., 1995. Boreal forest futures: modelling the controls on tree species range limits and transient responses to climate change. Water Air Soil Pollut. 82, 415–428.

Sykes, M.T., Prentice, I.C., 1996. Climate change, tree species distributions and forest dynamics: a case study in the mixed conifer/northern hardwoods zone of northern Europe. Climatic Change 34, 161–177.

Tackenberg, O., Poschold, P., Bonn, S., 2003. Assessment of wind dispersal potential in plant species. Ecol. Monogr. 73, 191–205.

Taylor, C.M., Hastings, A., 2005. Allee effects in biological invasions. Ecol. Lett. 8, 895–908.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., Ferreira de Siquiera, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarseveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L., Williams, S.E., 2004. Extinction risk from climate change. Nature 427, 145–148.

Thomas, C.D., Franco, A.M.A., Hill, J.K., 2006. Range retractions and extinction in the face of climate warming. Trends Ecol. Evol. 21, 415–416.

Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. Global Change Biol. 10, 2020–2027.

Thuiller, W., 2007. Biodiversity-climate change and the ecologist. Nature 448, 550–552.

Thuiller, W., Araújo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L., Lavorel, S., 2004. Biodiversity conservation-uncertainty in predictions of extinction risk. Nature 430, 1 (p following 33; discussion following 33 doi:10.1038/nature02716).

Thuiller, W., Lavorel, S., Araújo, M.B., 2005a. Niche properties and geographic extent as predictors of species sensitivity to climate change. Global Ecol. Biogeogr. 14, 347–357.

Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T., Prentice, I.C., 2005b. Climate change threats to plant diversity in Europe. Proc. Natl. Acad. Sci. USA 102, 8245–8250.

Thuiller, W., Broennimann, O., Hughes, G.O., Alkemade, J.R.M., Midgley, G.F., Corsi, F., 2006a. Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. Global Change Biol. 12, 424–440.

Thuiller, W., Lavorel, S., Sykes, M.T., Araújo, M.B., 2006b. Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. Divers. Distrib. 12, 49–60.

Thuiller, W., Midgley, G.F., Hughes, G.O., Bomhard, B., Drew, G., Rutherford, M.C., Woodward, F.I., 2006c. Endemic species and ecosystem vulnerability to climate change in Namibia. Global Change Biol. 12, 759–776.

Tikka, P.M., Hogmander, H., Koski, P.S., 2001. Road and railway verges serve as dispersal corridors for grassland plants. Landscape Ecol. 16, 659–666.

Travis, J.M.J., Brooker, R.W., Dytham, C., 2005. The interplay of positive and negative species interactions across an environmental gradient: insights from an individual-based simulation model. Biol. Lett. 1, 5–8.

Visser, M.E., Holleman, L.J.M., 2001. Warmer springs disrupt the synchrony of oak and winter moth phenology. Proc. Roy. Soc. London B 268, 289–294.

Wadsworth, R.A., Collingham, Y.C., Willis, S.G., Huntley, B., Hulme, P.e., 2000. Simulating the spread and management of alien riparian weeds: are they out of control? J. Appl. Ecol. 37, 28–38.

Walther, G.-R., Berger, S., Sykes, M.T., 2005. An ecological "footprint" of climate change. Proc. Roy. Soc. London B 272, 1427–1432.

Westcott, D.A., Graham, D.L., 2000. Patterns of movement and seed dispersal of a tropical frugivore. Oecologia 122, 249–257.

White, M.A., Brunsell, N., Schwartz, M.D., 2003. Vegetation phenology in global change studies. In: Schwartz, M.D. (Ed.), Phenology: An Integrative Environmental Science. Kluwer Academic Publishers, Dordrecht, pp. 453–466.

Wiens, J.J., Graham, C.H., 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annu. Rev. Ecol. Evol. Syst. 36, 519–539.

Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84, 1382–1394.

Willi, Y., Van Busrirk, J., Hoffmann, A.A., 2006. Limits to the adaptive potential of small populations. Annu. Rev. Ecol. Evol. Syst. 37, 433–458.

Williams, P.H., Hannah, L., Andelman, S., Midgley, G.F., Araújo, M.B., Hughes, G.O., Manne, L., Martinez-Meyer, E., Pearson, R.G., 2005. Planning for climate change: identifying minimum-dispersal corridors for the Cape Proteaceae. Conserv. Biol. 19, 1063–1074.

Wilson, R.J., Gutiérrez, D., Gutiérrez, J., Martinez, D., Agudo, R., Monserrat, V.J., 2005. Changes to the elevational limits and extent of species ranges associated with climate change. Ecol. Lett. 8, 1138–1146.

With, K.A., 2002. The landscape ecology of invasive spread. Conserv. Biol. 16, 1192–1203.

Woodward, F.I., 1992. Predicting plant responses to global environmental change. New Phytol. 122, 239–251.

Woodward, F.I., Lomas, M.R., 2004. Vegetation dynamics-simulating responses to climatic change. Biol. Rev. 79, 643–670.

Zavaleta, E.S., Shaw, M.R., Chiariello, N.R., Thomas, B.D., Cleland, E.E., Field, C.B., Mooney, H.A., 2003. Grassland responses to three years of elevated temperature, CO_2, precipitation, and N deposition. Ecol. Monogr. 73, 585–604.

Zwaenepoel, A., Roovers, P., Hermy, M., 2006. Motor vehicles as vectors of plant species from road verges in a suburban environment. Basic Appl. Ecol. 7, 83–93.

Forum

Ecography 31: 417–422, 2008
doi: 10.1111/j.2008.0906-7590.05221.x
© 2008 The Authors. Journal compilation © 2008 Ecography
Subject Editor: Jane Elith. Accepted 13 February 2008

Favourability functions versus probability of presence: advantages and misuses

Cécile H. Albert and Wilfried Thuiller

C. H. Albert (cecile.albert@m4x.org) and Wilfried Thuiller, Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Univ. Joseph Fourier, BP 53, FR-38041 Grenoble Cedex 9, France.

Logistic regression has been used to predict species distribution since the early age of predictive biogeography (Austin et al. 1984, Guisan and Thuiller 2005, Austin 2007). This is a parametric generalisation of usual linear regression that deals with a non-normally distributed response, and it has been utilized successfully to explain and model the distribution of various organisms over local to large spatial scales (Guisan et al. 1998, Teixeira et al. 2001, Fleishman et al. 2003, Thuiller et al. 2006a, b). The probability outputs of these models are commonly used under two forms: either they are summarised into binary results by choosing a threshold (Liu et al. 2005) or they are used un-summarised as raw probabilities or indexes (Real et al. 2006) where high values are taken as characterising highly suitable sites for the modelled species. The raw probability outputs are calibrated so they depend on the prevalence (ratio between the number of presences and the total number of records in the dataset) of the training dataset: the mean fitted value across all sites is equal to the prevalence in the training dataset. Because of this, rare species tend to have many low predicted values at sites across a region compared with more common species, so comparing suitability of sites across species in a conservation planning context can be a complex task. That is, the common species can effectively swamp the rare species in inter-species comparisons. In a recent paper, Real et al. (2006) proposed a method to circumvent this potential drawback by introducing a favourability function. They introduced the idea that a biased dataset is one where sampling prevalence (as deduced from samples) differs from 50%. Rescaling the probability outputs of logistic regression with reference to a so called ideal sample prevalence of 50%, the favourability function is expected to improve predictions in terms of description of the favourable areas for the species and to facilitate the comparison between species (and so that 0.5 would become the threshold for all species). Real et al. (2006) also claim that this function makes the model probability outputs independent of the sampling prevalence of the training dataset. The proposed formula for favourability F is the following:

$$F = \frac{\dfrac{P}{1-P}}{\dfrac{n_1}{n_0} + \dfrac{P}{1-P}},$$

with P the probability output of the logistic regression and n_1 and n_0 the respective number of presences and absences in the dataset.

To test this favourability function, Real et al. (2006) used virtual species that respond to a single environmental factor, providing "true" species distributions that can be sampled in different ways, and used to evaluate model outcomes (Hirzel et al. 2001, Reineking and Schroder 2006, Austin et al. 2006, Jimenez-Valverde and Lobo 2007, Meynard and Quinn 2007). The species all used the same environmental variable, and the same logistic relationship, but had varying population prevalence. A first test was realized with a single virtual species constructed with a 50% population prevalence to test the effect of sampling prevalence on the probability and the favourability outputs. Three sub-samples were extracted with respective sampling prevalences of 20% (presences are under represented compared to the true distribution), 50% (full distribution) and 80% (presences are over represented compared to the true distribution). Each sample was used to fit a logistic regression and results were compared: probability outputs were very different (maps) and favourability outputs were said to be qualitatively similar. Then on this specific case the favourability function was able to improve outputs resulting from unbalanced training datasets. The favourability function was then applied to two other virtual species with respective population prevalences of 20 and 80% to show that the favourability function makes it easier to compare species with different population prevalences. Both fully sampled (sampling prevalence equal to population

prevalence), the two virtual species obviously presented very different logistic regression outputs. However, rescaling these probability outputs, the favourability function allowed direct comparison of the two species in the overall landscape.

We found this method appealing for its modelling approach and innovative in its use of virtual species. In particular, Real et al. (2006) have clearly indicated why they require model probability outputs to be independent of the sampling prevalence and therefore comparable across prevalences. The method (Real et al. 2006) has merit in dealing with this problem but relies on a problematic definition of bias (sampling prevalence other than 50%) not ecologically justified. We propose that bias be defined as a sampled prevalence different from the "true" prevalence (the real probability of finding the species in the landscape), because it seems meaningful to consider as unbiased a sample which represents the ecological reality. Predictions from a logistic regression model will be calibrated in such a way that they accurately reflect the prevalence of a species throughout a region. We have to deal carefully with the word "bias" because it is commonly used but has several meanings; outputs of logistic regressions are not biased in the sense that they reflect the input data, but if built on biased samples they can be difficult to interpret.

Let us then move on without focussing on the word "bias", and consider what the favourability function (Real et al. 2006) does. We propose to test the favourability function ("Favourability 50") and compare it with a modified function ("Modified favourability") defined as follows:

$$F = \frac{\dfrac{P}{1-P}}{\dfrac{Pr_{Sample}}{Pr_{True}} + \dfrac{P}{1-P}},$$

where Pr_{Sample} and Pr_{True} are respectively the prevalence of the sampled dataset and the true prevalence of the virtual species using simulations with a range of both sample prevalence and population prevalence.

We first built a virtual landscape of 32 per 20 pixels (each pixel representing a 10×10 km area) based on a mean temperature parameter. This mean temperature parameter was distributed along a realistically varying gradient (0–15°C) and slightly modified with the addition of a random normal value. This led to a landscape with a temperature parameter ranging from −5 to 20°C and a mean of 7.6°C (Fig. 1a). We then established a logistic response to this single environmental factor. The response was higher for higher mean temperatures. This response to the environment was then used to create three virtual species distributions with respective population prevalence of 20% (VS20), 50% (VS50) and 80% (VS80) (Fig. 1b, c, d).

For each species, samples of 80 pixels with prevalence of 20, 50 and 80% were randomly extracted. They represented random partial samplings in the field, with perfect detection. This sampling strategy was chosen to cross-test the effect of the true and the sampling prevalence on the logistic regression, the Favourability 50 and Modified favourability outputs. For VS20, three more samples of 200 pixels were also selected with prevalences of 5, 10 and 15%, which are

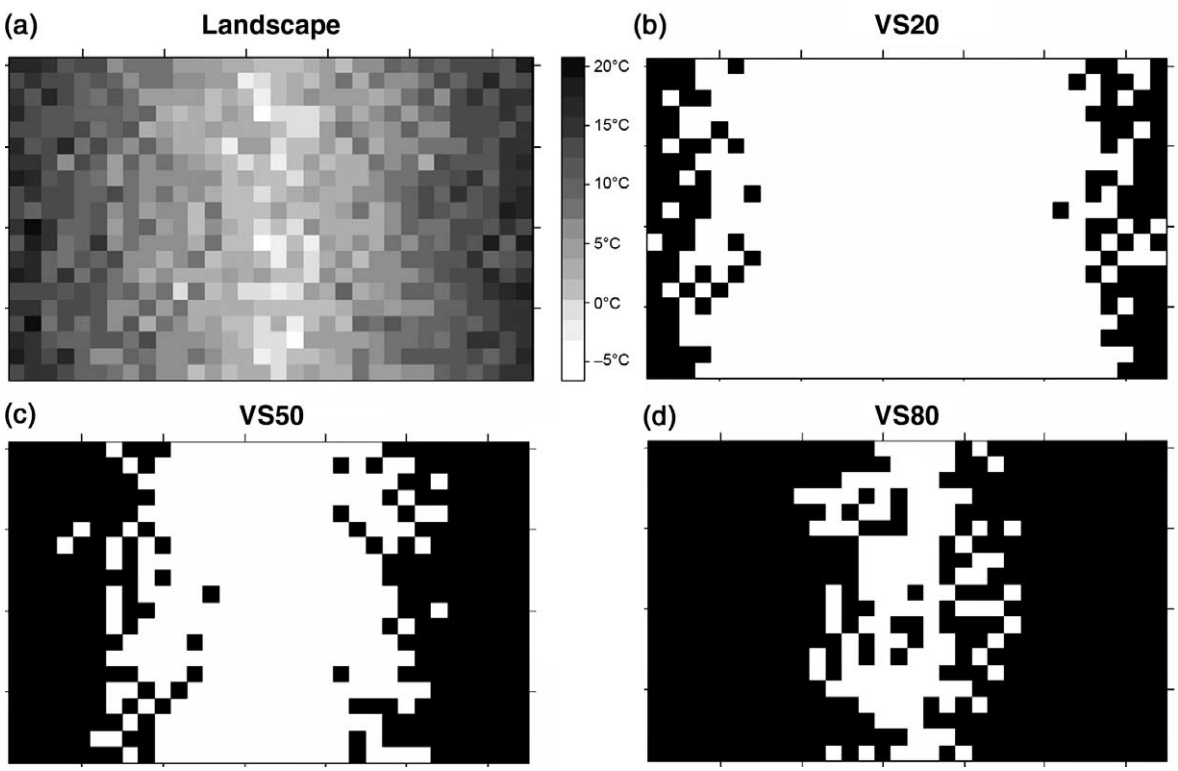


Figure 1. Map of the virtual landscape: mean annual temperature (from −5 to 20°C) (a), VS20 species in presence (black) vs absence (white) (b), VS50 species in presence vs absence (c), VS80 species in presence vs absence (d).

sample prevalences close to those observed in datasets obtained in natura.

Logistic regressions were performed on the sampled datasets and on the true distributions (full distribution) of the three species. The probability outputs obtained with the sampled dataset were then transformed with the Favourability 50 and the Modified favourability functions. For each model, the raw predictions as well as the Favourability 50 and Modified favourability predictions were evaluated using the area under the receiver operating characteristic curve (AUC) (Swets 1988) for the whole landscape.

We then compared each output dataset (probability or favourability) to the model fitted with the full dataset to explore the performance of each model and formula outputs for the different sampling and population prevalence. To determine if predictions for a single species were consistent and independent from the sample prevalence, we also compared the outputs obtained from samples with a prevalence of 20% (and 10% for VS20) and those with a prevalence of 80%.

To quantify this we used a weighed Kappa coefficient (Cohen 1960) with multiple modalities. We used 5 classes: 0–0.2 (very low suitable environment), 0.2–0.4 (low suitable), 0.4–0.6 (medium suitable), 0.6–0.8 (quite suitable) and 0.8–1 (suitable) to discriminate the outputs values. N being size of the datasets, the n_{ij} were defined as the counts in the resulting confusion table and a weight matrix W (5, 5) was chosen to classify errors by their seriousness: $Wij = 1 - \frac{|i-j|}{5-1}$. Then the Kappa coefficient was calculated as follows:

$$\text{with } Po = \sum_{i=1}^{5}\sum_{j=1}^{5} W_{ij} \times \frac{n_{ij}}{N} \text{ and}$$

$$Pe = \sum_{i=1}^{5}\sum_{j=1}^{5} W_{ij} \times \frac{n_{i.}}{N} \times \frac{n_{.j}}{N},\ K = \frac{Po - Pe}{1 - Pe}.$$

We finally used a rough guideline to classify the kappa values: above 0.8, the two models were considered as equivalent, between 0.6 and 0.8, the two models were considered as nearly different and under 0.6 the two models were considered as very different.

Obviously, the raw regression and the Favourability 50 were identical for a sampling prevalence of 50% and the raw regression and the Modified favourability were identical for a sampling prevalence equal to the population one, leading to the same values of the Kappa coefficient (K) (Fig. 2).

For the three species the best predictions (higher Kappa values, Kappa >0.88) with the raw model were obtained when sampling and population prevalences coincide (for instance with a sampling prevalence of 20 for the species VS20). However, as soon as the sampling prevalence was different from the population one, raw predictions were very different from the model performed on the exhaustive distribution of the species (Fig. 2). The Favourability 50 led to better predictions than the raw model (for example Kappa $=0.9$ instead of 0.47) when the sampling prevalence was very different from the population one. Nevertheless predictions became worse than the raw ones when the sampling prevalence was close to the population one (for example Kappa $=0.25$ instead of 0.74). As for the Modified

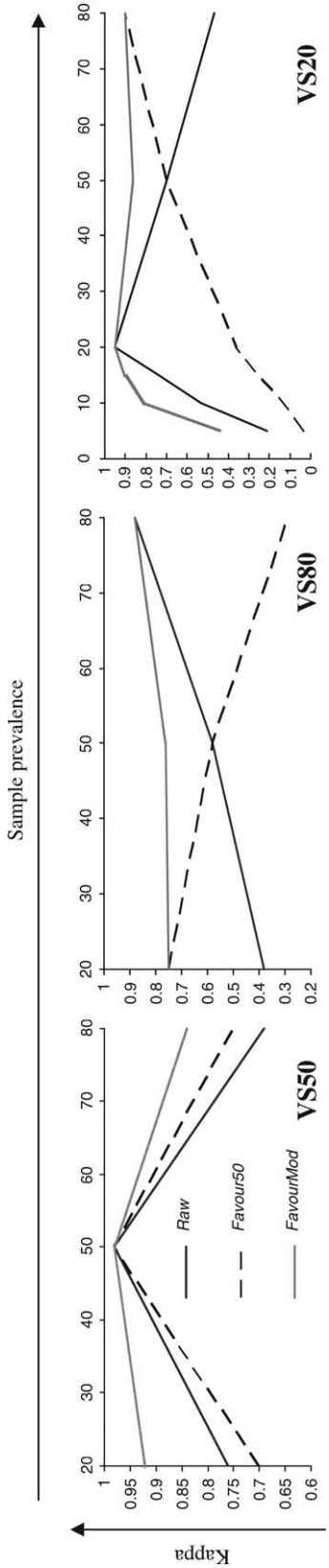


Figure 2. Cohen's Kappa values between each model/formula outputs and the ones of the model performed on the true distributions, for each sample prevalence and for each species. For Kappa values under 0.8 both predictions sets can be considered as different.

Table 1. Cohen's weighted Kappa coefficients for each species and each model/formula between two outputs resulting from samples with prevalences of 20 and 80% for the three species and 10 and 80% for VS20.

	Species	Raw	Favourability 50	Modified favourability
VS20	Sample prevalence 20 vs 80	0.43	0.38	0.88
VS20	Sample prevalence 10 vs 80	0.20	0.16	0.72
VS50	Sample prevalence 20 vs 80	0.49	0.45	0.87
VS80	Sample prevalence 20 vs 80	0.33	0.39	0.72

favourability, it gave predictions ($0.81 < \text{Kappa} < 0.98$ except for the sampling prevalence of 5%) as good as or better than the raw ones in all cases (Fig. 2).

The Kappa coefficients calculated between two models outputs resulting from two samples of the same species differing by their sampling prevalence are presented in the Table 1. As expected the raw regression led to higher probability outputs when the sampling prevalence was high (80%) compared to when it was low (20%) (Table 1, Fig. 3). This results in relatively low Kappas for the three species (0.33–0.49, Table 1). The Favourability 50 function reversed these tendencies (Fig. 3) and still led to low Kappas (0.38–0.45, Table 1). Modified favourability properly adjusted the outputs (Fig. 3), leading to higher values of kappa (0.72–0.88, Table 1).

All AUC values (predictions tested against observations, Table 2) were very high (over 0.98), showing that all the performed models discriminate a presence from an absence almost perfectly. As the Favourability 50 and Modified favourability are rescaling formulas they are changing only calibration and not discrimination, so their use leads necessarily to the same AUC values and to rescaled thresholds, (Jimenez-Valverde and Lobo 2007).

This research sought to improve the favourability formula proposed by Real et al. (2006) with a formula relying on the ratio between the sampling and the true prevalence of a virtual species. The improved formula has proven to be efficient in reducing sampling-induced error in the logistic regression outputs, even for consequent bias of the training sample (sampling prevalence of 80% instead of 20% or the contrary) and for datasets close to those obtained in natura (sample prevalence of 5 or 10% for VS20). It also yielded similar predictions for a given species, even for samples with different prevalences. We believe that this formula is more realistic than the one proposed by Real et al. (2006) because it does not suppose that a particular prevalence of 50% is the reference one. However, this formula will be difficult to apply to real species for which true prevalences are poorly known. In these cases, one should investigate the origin of the low prevalence problem and the possible solutions for getting around it.

The possible origins of a low prevalence in a sample are multiple. First, such a sample might properly represent the true distribution of a rare species because most environments are unsuitable for it (Guisan and Thuiller 2005, Jimenez-Valverde and Lobo 2007, Meynard and Quinn 2007). Secondly, it can result from a sampling-induced bias due to either low detectability or poor sampling strategy (Mackenzie et al. 2004, Wintle et al. 2005). Thirdly, it can result from the fact that the considered area is at the edge of the species environmental niche or does not include its whole niche (Hernandez et al. 2006, Randin et al. 2006, Vanreusel et al. 2007). Fourth, the sample area may be too large in relation to the whole niche of the species as it is often the case for restricted and rare species (Martin et al. 2005). These four cases differ both in the origin of the low prevalence of the sample they induce and in the relevant solutions. The first case is not an error as such, because the modelled probabilities truly reflect the species prevalence. In fact, given relevant predictors, a suitable site, if it exists, can be modelled with a high probability – one would simply not expect many of them overall. If, alternatively, there are no predictors that identify highly suitable sites, many probabilities may be low, and sensible strategies for dealing with this in the light of higher predictions for more common species will be required. The second case is a substantial problem, highly mobile or small vagile species are often difficult to detect and good sampling strategies are always difficult to design in such a way as to be both statistically and ecologically sound. Quantifying the detection probability for a given species could be a solution to obtain a more statistically conclusive statement about possible reasons for not detecting a given species as opposed to a simple "none were found" conclusion (Wintle et al. 2004, Royle et al. 2005, Mackenzie 2006). The third and the fourth cases are modelling problems. The choice of the scale and of the resolution in statistical habitat modelling is one of the greatest challenges in this field (Guisan and Thuiller 2005). To obtain reliable probability outputs, input data should include the whole niche (the core, the edge and the outside of the niche) of the species (Austin and

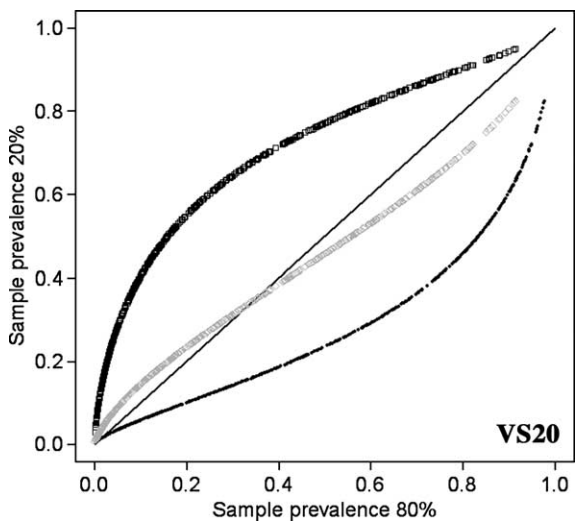


Figure 3. VS20 species. Comparison between outputs resulting from a training data with a sample prevalence of 20% against one of 80%. Black dots are representing raw regression outputs, black squares the Favourability 50 outputs and grey squares the Modified favourability. This comparison shows the sample prevalence dependence of the models.

Table 2. Values of the area under the curve of the receiver operating characteristic curve for the different species and the different sampling prevalences. The three model/formula lead in each case to the same AUC value.

Species	Prevalence	AUC
VS20	5	1
VS20	10	1
VS20	15	1
VS20	20	1
VS20	50	1
VS20	80	1
VS50	20	1
VS50	50	1
VS50	80	1
VS80	20	0.98
VS80	50	0.98
VS80	80	0.98

Meyers 1996, Thuiller et al. 2004, Strauss and Biedermann 2007, Tsoar et al. 2007). However, the study area must not be too large compared to the extent of occurrence of the species to avoid the inclusion of a large number of inappropriate absences and lower the probability outputs.

It therefore seems that a single formula aiming at getting around this sampling prevalence problem cannot be universal and that the situation has to be considered case by case with an accurate ecological knowledge of the species to make the right modelling choices (Pearce et al. 2001, McPherson and Jetz 2007).

Our analysis also shows that innovative mathematical formula have to be developed and tested carefully before being used by decision makers. Models and in particular habitat models, as simplifications of reality and predictive tools are more and more used to guide conservation and restoration planning, management and support the formulation of policy decisions (Bayliss et al. 2005, Thompson et al. 2006, Titeux et al. 2007, Vanreusel et al. 2007). For this reason, predictions and models have to be as reliable and consistent as possible (Mech 2006, Nicholson and Possingham 2007) and need to be used carefully. Sampling strategies have to be statistically and ecologically sound (Guisan and Thuiller 2005) and supported by a strong ecological knowledge of the studied species (Pearce et al. 2001, McPherson and Jetz 2007). Habitat models should also satisfy the underlying assumptions: the studied species should be in quasi-equilibrium with its environment and not, for example, colonising a new area, the studied area should include the whole niche of the species without being too large (Guisan and Thuiller 2005) and environmental explanatory variables should be as proximal as possible (Austin 2002). Users of habitat suitability models have to bear in mind that the models are correlative and static in nature and usually focus on abiotic conditions (but see Leathwick and Austin 2001). Habitat suitability models should not thus be applied to species whose distributions are constrained for example by the occurrence of a specific prey or a competitor if these are not accounted for. Moreover, the outputs of habitat suitability models have to be interpreted with the clarification from expert opinion.

In species distribution modelling two kinds of results are commonly used: the first one is the binary transformation of logistic regression outputs (Liu et al. 2005, Jimenez-Valverde and Lobo 2007); this is based on the choice of a threshold which aims to separate occupied from unoccupied sites. The selection of this threshold and the possible ways to calculate it have been extensively discussed during the last ten years (Fielding and Bell 1997, Pearce and Ferrier 2000, Vayssières et al. 2000, Manel et al. 2001, Fielding 2002, Liu et al. 2005, Allouche et al. 2006, Jimenez-Valverde and Lobo 2007). The resulting binary outputs and their predictive performance can be influenced by the training sample prevalence and also by the validation sample prevalence per se (Manel et al. 2001, Liu et al. 2005), but are not depending on a rescaling formula like the Favourability 50 (rescaling results only lead to a rescaling in the threshold) (Jimenez-Valverde and Lobo 2007). The threshold must then be chosen carefully and according to the final goal of the modelling study (Liu et al. 2005).

The second kind of results is raw probability outputs. Those results can be interpreted in two different ways; they can be seen as an absolute measure of probability of occurrence but must be validated by calibration plots (Pearce and Ferrier 2000) or as a relative measure of the likelihood of species occurrence (Pearce and Ferrier 2000). If there is bias in the training sample, it is more robust to use them as an ordinal suitability index.

Finally, when binary results are not sufficient, it is possible to avoid the drawback of unbalanced species by considering the results as relative values and using them for example as a percentage of the maximum of the raw probability outputs. It is also possible to weight the logistic regression (Teixeira et al. 2001) in a way to balance the sample without losing information and to avoid a biased rescaling afterward, but this again relies either on the knowledge of the true prevalence of the species or on the assumption of an ideal 0.5 prevalence. Another way to use those models outputs would be to confront them systematically with expert opinion and eventually rescale them by including this field knowledge. For example the maximum, minimum and medium values of the probability outputs could be compared with in situ measures of abundance, fitness or other ecological performances.

Acknowledgements – This work was funded by the "Agence National pour la Recherche" through the QDiv project (Quantifying the Effects of Global Environmental Change on Terrestrial Plant Diversity). WT received support from European Commission's FP6 MACIS (Minimisation of and Adaptation to Climate change Impacts on biodiverSity No. 044399) and ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe) projects. We thank three anonymous reviewers for their very helpful comments and critics on the manuscript and Jane Elith for all her comments and her major help.

References

Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – J. Appl. Ecol. 43: 1223–1232.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. – Ecol. Model. 200: 1–19.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. – Ecol. Model. 157: 101–118.

Austin, M. P. and Meyers, J. A. 1996. Current approaches to modelling the environmental niche of eucalypts: implication for management of forest biodiversity. – For. Ecol. Manage. 85: 95–106.
Austin, M. P. et al. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. – Vegetatio 55: 11–27.
Austin, M. P. et al. 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. – Ecol. Model. 199: 197–216.
Bayliss, J. L. et al. 2005. The use of probabilistic habitat suitability models for biodiversity action planning. – Agricult. Ecosyst. Environ. 108: 228–250.
Cohen, J. 1960. A coefficient of agreement for nominal scales. – Educ. Psychol. Meas. 20: 37–46.
Fielding, A. H. 2002. What are the appropriate characteristics of an accuracy measure. – In: Scott, J. M. et al. (eds), Predicting species occurrences – issues of accuracy and scale. Island Press, pp. 303–313.
Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – Environ. Conserv. 24: 38–49.
Fleishman, E. et al. 2003. Validation tests of predictive models of butterfly occurrence based on environmental variables. – Conserv. Biol. 17: 806–817.
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Guisan, A. et al. 1998. Predicting the potential distribution of plant species in an alpine environment. – J. Veg. Sci. 9: 65–74.
Hernandez, P. A. et al. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. – Ecography 29: 773–785.
Hirzel, A. H. et al. 2001. Assessing habitat-suitability models with a virtual species. – Ecol. Model. 145: 111–121.
Jimenez-Valverde, A. and Lobo, J. M. 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. – Acta. Oecol. 31: 361–369.
Leathwick, J. R. and Austin, M. P. 2001. Competitive interactions between tree species in New Zealand's old-growth indigenous forests. – Ecology 82: 2560–2573.
Liu, C. et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. – Ecography 28: 385–393.
Mackenzie, D. I. 2006. Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. – J. Wildl. Manage. 70: 368–374.
Mackenzie, D. I. et al. 2004. Investigating species co-occurrence patterns when species are detected imperfectly. – J. Anim. Ecol. 73: 546–555.
Manel, S. et al. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. – J. Appl. Ecol. 38: 921–931.
Martin, T. G. et al. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. – Ecol. Lett. 8: 1235–1246.
McPherson, J. M. and Jetz, W. 2007. Effects of species' ecology on the accuracy of distribution models. – Ecography 30: 135–151.
Mech, D. L. 2006. Prediction failure of a wolf landscape model. – Wildl. Soc. Bull. 34: 874–877.
Meynard, C. N. and Quinn, J. F. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. – J. Biogeogr. 34: 1455–1469.
Nicholson, E. and Possingham, H. P. 2007. Making conservation decisions under uncertainty for the persistence of multiple species. – Ecol. Appl. 17: 251–265.
Pearce, J. and Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. – Ecol. Model. 133: 225–245.
Pearce, J. L. et al. 2001. Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution. – J. Appl. Ecol. 38: 412–424.
Randin, C. F. et al. 2006. Are niche-based species distribution models transferable in space? – J. Biogeogr. 33: 1689–1703.
Real, R. et al. 2006. Obtaining environmental favourability functions from logistic regression. – Environ. Ecol. Stat. 13: 237–245.
Reineking, B. and Schroder, B. 2006. Constrain to perform: regularization of habitat models. – Ecol. Model. 193: 675–690.
Royle, A. et al. 2005. Modelling occurence and abundance of species when detection is imperfect. – Oikos 110: 353–359.
Strauss, B. and Biedermann, R. 2007. Evaluating temporal and spatial generality: how valid are species–habitat relationship models? – Ecol. Model. 204: 104–114.
Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. – Science 240: 1285–1293.
Teixeira, J. et al. 2001. Biogeography of the golden-striped salamander *Chioglossa lusitanica*: a field survey and spatial modelling approach. – Ecography 24: 618–624.
Thompson, L. M. et al. 2006. A spatial modeling approach to identify potential butternut restoration sites in Mammoth Cave National Park. – Restor. Ecol. 14: 289–296.
Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – Ecography 27: 165–172.
Thuiller, W. et al. 2006a. Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. – Global Change Biol. 12: 424–440.
Thuiller, W. et al. 2006b. Endemic species and ecosystem vulnerability to climate change in Namibia. – Global Change Biol. 12: 759–776.
Titeux, N. et al. 2007. Fitness-related parameters improve presence-only distribution modelling for conservation practice: the case of the red-backed shrike. – Biol. Conserv. 138: 207–223.
Tsoar, A. et al. 2007. A comparative evaluation of presenceonly methods for modelling species distribution. – Divers. Distrib. 13: 397–405.
Vanreusel, W. et al. 2007. Transferability of species distribution models: a functional habitat approach for two regionally threatened butterflies. – Conserv. Biol. 1: 201–212.
Vayssières, M. P. et al. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. – J. Veg. Sci. 11: 679–694.
Wintle, B. A. et al. 2005. Estimating and dealing with detectability in occupancy surveys for forest owls and arboreal marsupials. – J. Wildl. Manage. 69: 905–917.
Wintle, B. A. et al. 2004. Precision and bias of methods for estimating point survey detection probabilities. – Ecol. Appl. 14: 703–712.

Accepté par Ecography: 10.1111/j.1600-0587.2009.06000.x

BioMove – an integrated platform simulating the dynamic response of species to environmental change

Guy F. Midgley[1], Ian. D. Davies[2], Wilfried Thuiller[3] Cécile H. Albert[3], Res Altwegg[1, 4], Lee Hannah[5], Gregory O. Hughes[6], Lydia R. O'Halloran[7], Changwan Seo[8], James H. Thorne[9]

1 Climate Change Research Group, Kirstenbosch Research Center, South African National Biodiversity Institute, P/Bag x7, Claremont 7735, Cape Town, South Africa
2 Fenner School of Environment and Society, Australian National University, Canberra ACT 0200, Australia;
3 Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France
4 Animal Demography Unit, Department of Zoology, University of Cape Town, Rondebosch 7701, Republic of South Africa
5 Conservation International, based at Bren School of Environmental Science and Management, University of California, Santa Barbara, CA, USA
6 Cambridge Environmental Assessments, Woodthorne, Wergs Road, Wolverhampton, WV6 8TQ, UK
7 Department of Zoology, Oregon State University, Corvallis, OR 97331 USA
8 Department of Geoinformatics, University of Seoul, Seoul 130-743, Korea
9 Dept. Environmental Science and Policy, University of California, Davis, CA, USA

Corresponding author: Ian. D. Davies. ian.davies@anu.edu.au / Wilfried Thuiller: wilfried.thuiller@ujf-grenoble.fr

BioMove simulates plant species' geographic range shifts in response to climate, habitat structure and disturbance, at annual time steps. This spatially explicit approach integrates species' bioclimatic suitability and population-level demographic rates with simulation of landscape-level processes (dispersal, disturbance, species' response to dynamic dominant vegetation structure). Species population dynamics are simulated through matrix modelling that includes scaling demographic rates by climatic suitability. Dispersal functions simulate population spread. User-specified plant functional types (PFTs) provide vegetation structure that determines resource competition and disturbance. PFTs respond annually through dispersal, inter-PFT competition and demographic shifts. BioMove provides a rich framework for dynamic range simulations.

One of the core needs in conservation biology is the assessment of the long term viability of populations of a species under a range of possible management interventions and environmental changes. Several approaches for assessing population viability have been proposed and implemented. These population viability approaches (PVA) provide a useful tool to assess population viability in a given context (Lacy 1993). However, climate change that is likely to cause changes in species home range and in population sizes within home ranges, minimizes the value of these approaches as they are currently designed (Thuiller et al. 2008). They are usually not spatially explicit and do not include the direct effect of climate change on demographic parameters. They also ignore the interactive dynamics between environmental change and landscape process (e.g. succession, shift of biome boundary) and are often difficult to parameterise for a large number of species.

Alternatively, landscape models have been traditionally used to explore the influence of disturbances on vegetation dynamics and look for emerging patterns over long time spans (e.g., hundreds of years, Baker 1989). They typically do not focus on one subject species. Rather, they address biotic interactions between groups of species commonly called plant functional types that represent different components of a community (e.g. Albert et al. 2008). Landscape models explicitly focus on processes such as dispersal, fire and human disturbances, and for this reason include simplified representation of the demography and biotic interactions. These simplifications of vegetation dynamics are essential to represent landscape processes where local-scale processes are over-ridden by larger-scale disturbances and where habitat suitability is expected to be constant. For this reason, they are traditionally not used in a climate change context.

Finally, climate impact assessments on species are mostly carried out with habitat suitability models characterising the suitability of a given environment for a species, based on climate and other inputs (such as soil type or land cover) (Heikkinen et al. 2006). These approaches allow climate change driven spatial shifts in species suitability to be mapped. While these attempts can explore impacts on a wide range of species, and thus provide useful, even if preliminary, estimates of risk of extinction for endemic faunas and floras, they are criticised for a range of reasons. Firstly, they derive the likelihood of occurrence in any particular site by correlating occurrence data with selected environmental variables – an approach that has been much discussed and is currently considered pragmatic but problematic (e.g. Dormann 2007). Secondly, they do not realistically account for population-level processes especially at the leading and trailing edges of geographic ranges, including dispersal and establishment of populations in new geographic ranges and persistence in sub-optimal conditions. Thirdly, they are blind to community interactions (e.g., competition) and disturbances (e.g., fire or grazing).

The models described above all offer unique contributions to conservation biology modelling. We propose here that a hybrid approach, combining the advantages of PVA approaches, landscape models and habitat suitability models, would permit more realistic estimates of potential population change rates. This would also allow the interactive effects of landscape dynamics and environmental changes to account for idiosyncratic changes that may be occurring throughout different parts of a species' home range (Keith et al. 2008).

Our proposition has led us to develop BioMove, a novel modelling tool designed initially to assess global change risk to species persistence. With its roots in ecological niche theory (Hutchinson 1957), BioMove couples landscape and population modelling to provide a more robust simulation of the demographic and community processes that affect persistence and realization in a newly available range.
BioMove has been designed inside the LAMOS framework (Lavorel 2001) which uses plant functional traits to integrate different ecological processes operating at the landscape scale. Within each pixel of a geo-referenced grid, community dynamics are modelled by applying a succession model that uses demographic, regeneration and morphological traits at the functional type or species level. BioMove is implemented as a succession model in LAMOS and uses its dispersal and disturbance methods. We believe BioMove offers a novel approach in that it provides a method to simultaneously model plant function types (PFT) that determine vegetation structure and the demography of particular species with regard to dynamic habitat suitability in response to environmental and landscape changes (Fig. 1).

To model the vegetation dynamics, BioMove uses a modified version of the FATE succession model (Moore and Noble 1990). It is relatively easy to parameterise with the knowledge of field biologists and relevant data from the current literature.

The PFTs are chosen based on their relevance to the ecosystem studied and the ecological questions at hand. For instance, to investigate the ability of the European larch to colonize semi-abandoned subalpine grasslands, Albert et al. (2008) defined four leaf-height-seed (LHS)-based PFTs, classifying the dominant graminoids in contrasting nutrient economies and competitive effects and responses after Westoby's (1998) LHS plant strategy scheme and one tree PFT representing the larch. These four grass PFTs and one tree PFT were sufficient to represent the vegetation structure of the study area. In another study, Hannah et al. (2008) used BioMove to simulate the interaction of a target pine, *Pinus lambertiana*, with

a single PFT representing the oak-dominated vegetation expected to displace the pine under climate change.

BioMove requires semi-quantitative parameters to simulate the PFTs, including their competitive ability at different life stages, their dispersal ability (via both short and long distance dispersal functions), their mortality and regeneration responses to fire, grazing and other disturbances relevant to the study area (disturbance model). This information can be derived from literature, trait databases, field measures or expert knowledge. In its current configuration, only competition for light is simulated in the FATE model. The FATE module in BioMove handles the vegetation and seed pool dynamics by tracking annual cohort abundance for juvenile and mature age classes.

BioMove simulates the dynamics of a given plant species within this dynamic landscape. We believe this decoupling between species and PFTs is necessary as few models include biotic interactions and the resulting mosaic into the modelling of population dynamics. Spatially explicit models like Ramas GIS only constrain the size of populations by the carrying capacity of the subject species (see Keith et al. 2008) but not by the light availability or the vegetation structure of the area. This coupling resolves most of the criticisms formulated against the use of meta-population models at regional scale and makes the estimation of extinction risks of endangered species more realistic.

To estimate uncertainty, LAMOS provides a script interface for the design and execution of multiple simulations. Data for any or all model variables can be exported in a variety of formats. In addition, a Monte-Carlo feature provides a means for generating such scripts to explore the model's parameter space. This includes implementation of a Latin Hypercube design where prior knowledge of parameter distributions around some mean and possible correlations between parameters may be used to limit the number of simulations to a tractable number.

BioMove uses a demographic model (Fig. 3) to simulate the dynamics of the subject species within the landscape, with key demographic rates. The demographic model is a mixed age/stage population demographic model implemented in each cell. Age is used to specify the juvenile age classes. Once cohorts achieve the sub-adult age class, transitions between sub-adult and adult age classes become probabilistic (Fig. 3). Different biological traits are necessary to parameterise this demographic model, such as (st)age-specific mortality and fecundity, age at maturity and disturbance response. Similarly, the dispersal characteristics of the species are also required (mean dispersal distance, fraction of seeds long distance dispersal). Our approach differs from animal population viability analysis (PVA) in only having undirected movement between pixels (kernel seed dispersal). Models of animal populations often have directed (intelligent) movement between patches. Different seed dispersal models can be added to LAMOS by modellers to provide for particular dispersal syndromes (van der Pijl, L. 1982). This is an area of research we are currently pursuing with particular regard to the implications for modelling animal populations in BioMove.

The light availability resulting from the simulated landscape is then used inside the species demographic model to simulate the competition for light between the vegetation and the modelled species, depending on its shade tolerance at different life stages (germinant, immature and mature stage). This is a one way effect in the sense that the subject species does not influence the PFT composition in return (see below in the case of a dominant or invasive species).

We also introduce the ability to simulate the influence of environmental change on the dynamics of the vegetation structure and the subject species. What makes BioMove unique is its coupling to habitat suitability models. The habitat suitability determination for each PFT and the subject species is performed outside of BioMove for the study area. We did not include this process inside BioMove as there are many techniques with different assumptions and we considered it important to allow user choice in the method applied (Thuiller 2009, Marmion 2009). The habitat suitability (HS) model can be estimated using climate-only data (e.g., in a climate change context), soil or any resource variable and then projected using the different climate change scenarios available. The output of this process is a single 0-1 layer supplied to BioMove representing the HS for a PFT or the subject species at some point in time. Many layers may be required to represent the dynamics of land use or climate change for a given scenario. These can be read into BioMove during a model run as directed by a script file. There is no fixed time step for this process.
A set of HS layers represents a single land use/climate change scenario. A scenario can then be viewed as a parameter value and the uncertainty associated with a scenario measured in the same way as uncertainty in other BioMove parameters.

Once imported into BioMove, the HS is used to scale the demographic processes, mortality, recruitment and fecundity and has the effect of excluding recruitment of the species and PFT in unsuitable habitats (e.g. Keith et al. 2008). The scaling process and the form of the relationships between HS and the parameters are decided prior the simulations. If there is no information about the relationship a factorial design in implementing simulations is recommended to estimate the sensitivity of the species or PFT to this relationship. This approach has the potential to make the PFTs and subject species effectively variable with regard to environmental conditions. For example, morphological parameters and the timing of life stage events could be set to change with HS providing different competitive abilities between the PFTs (Fig. 2).

There are various potential applications of BioMove. At the regional scale, BioMove is useful for projecting vegetation dynamics in response to both climate and land use change. This is critical to evaluate the consequences of these dual factors to ecosystem functions and associated goods and services. Different scenarios can be envisaged according to various adaptation or mitigation plans, for example, forced growing or mowing to prevent bush encroachment (implemented as specific disturbance in BioMove). More specifically, it is highly relevant to investigate extinction risk of subject species in the face of climate change or/and land use change. Previous attempts have ignored biotic interactions and the influence of communities (e.g., Keith et al. 2008). To explore the effectiveness of proposed adaptation plans, specific corridors across protected areas can be simulated to investigate their usefulness in decreasing extinction risk.

Finally, with regard to the growing threats due to invasive species, BioMove could simulate the ecosystem modifications arising from an invasive PFT and the consequences this has for the species "at risk", native community and ecosystem functions. Given the amount of money invested each year for alien species eradication, such a tool simulating the spread of invasive species through the landscape under different eradication strategies (e.g., implemented as specific disturbance) are of crucial importance to identifying management options that maximize ecosystem integrity while minimizing eradication costs.

Three key assumptions need to be born in mind when considering the appropriateness of BioMove to a research question. These three points are in regard to coupling between various components of the model.

- Firstly, it is assumed that the habitat suitability can be established independently from BioMove simulations, that is, HS is determined by habitat suitability modelling techniques using exogenous variables such as climate, soils and topography.
- Secondly, it is assumed the population density of the subject species is sufficiently low so as not to have any competitive effect upon the PFT community. The coupling is unidirectional. This is a reasonable assumption given that one of the purposes of BioMove is to examine the viability of populations at risk rather than robust populations in high abundance, but fails when a subject species drives processes itself, as discussed below.
- Thirdly, the disturbance regime is influenced only by the PFT(s), and not by the population size of the subject species. It is the community mix of PFTs that provides fuel or food for disturbance models within LAMOS. This is also a reasonable assumption given the second point above. However,

this potentially limits BioMove for the application of modelling keystone species that may cause switches in ecosystem structure and function, such as invasive aliens. In such a case we suggest that the subject species are best modelled as a PFT in the community model rather than a low density population in the demographic model.

BioMove is written in the Delphi (v 7) programming language and incorporated within the LAMOS platform through recompilation. The software, source code and documentation are free and available from http://purl.oclc.org/NET/lamos_biomove.zip. To modify BioMove, it will be necessary to obtain and install Delphi; Abbrevia, XStringgrid (both freeware from sourceforge.net) and TeeChart (commercial). BioMove has been tested on Windows 2000 and XP. Disk space required for the download files is 60Mb. Memory requirements for a simulation with 10,000 cells and four PFTs is 23Mb and takes 1 minute to run 100 time steps (years) on a current model desktop computer. LAMOS places no constraints on spatial grain or extent, however, grid sizes greater the 10^7 pixels are a practical upper limit. Some elements of BioMove are stochastic and LAMOS provides features to manage replicate simulations. To cite BioMove or acknowledge its use, refer to this Software Note as follows, substituting the version of the application that you used for Midgley et al. 2008. BioMove – an integrated platform simulating the dynamic response of species to environmental change. Ecography 00: 000-000. (Version XX.).

Acknowledgements
BioMove has been developed with support from the California Energy Commission PIER Program (CEC-500-02-004) and is currently developed with funding from FP6 EU MACIS (No 044399 SSPI) and ECOCHANGE (066866 GOCE) projects, and the French ANR Diversitalp (ANR 07 BDIV 014) and 3Worlds (ANR-07-CIS7-001) projects. Initial development of BioMove's bioclimatic/demographic concepts was made by Reuben Roberts, Michael Rutherford and GFM of the South African National Biodiversity Institute as part of the South African Country Study on Climate Change, funded by USAID.

Albert, C. H. et al. 2008. Land-use change and subalpine tree dynamics: colonization of Larix decidua in French subalpine grasslands. — Journal of Applied Ecology 45: 659-669.
Baker, W. L. 1989. A review of models of landscape change. — Landscape Ecology 2: 111-133.
Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. — Basic and Applied Ecology 8: 387-397.
Hannah, L., Guy Midgley, Ian Davies, Frank Davis, Lydia Ries, Willfried Thuiller, James Thorne, Changwan Seo, David Stoms, Nathan Snider. 2008. BioMove — Creation of a Complex and Dynamic Model for Assessing the Imparts of Climate Change on California Vegetation. California Energy Commission, PIER Energy-Related Environmental Research Program. CEC-500-2008-060. http://www.energy.ca.gov/publications/displayOneReport.php?pubNum=CEC-500-2008-060
Heikkinen, R. K. et al. 2006. Methods and uncertainties in bioclimatic envelope modeling under climate change. — Progress in Physical Geography 30: 751-777.
Hutchinson, G. E. 1957. Concluding remarks. — Cold Spring Harbor Symposia on Quantitative Biology 22: 145-159.
Keith, D. A. et al. 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. — Biology Letters 4: 560–563.
Lacy, R. C. 1993. VORTEX: a computer simulation model for population viability analysis. — Wildlife Research 20: 45-65.
Lavorel, S. 2001. LAMOS : A LAndscape MOdelling Shell for studying the role of landscape scale processes in global change effects. — GCTE News 17: 5-6.
Marmion M., Parviainen M., Luoto M., Heikkinen R.K. & Thuiller W. (2009). Evaluation of consensus methods in predictive species distribution modelling. Diversity and Distributions, 15, 59-69.
Moore, A. D. and Noble, I. R. 1990. An individualistic model of vegetation stand dynamics. — Journal of Environmental Management 31: 61-81.
Thuiller, W. et al. 2008. Predicting global change impacts on plant species' distributions: Future challenges. — Perspectives in Plant Ecology, Evolution and Systematics 9: 137-152.
Thuiller W., Lafourcade B., Engler R. & Araujo M.B. (2009). BIOMOD – A platform for ensemble forecasting of species distributions. Ecography, 32, 369-373.
van der Pijl, L. 1982. Principles of dispersal in higher plants. Springer Verlag, Berlin, Germany
Westoby, M. 1998. A leaf-heigh-seed (LHS) plant ecology strategy scheme. — Plant and Soil 199: 213-227.

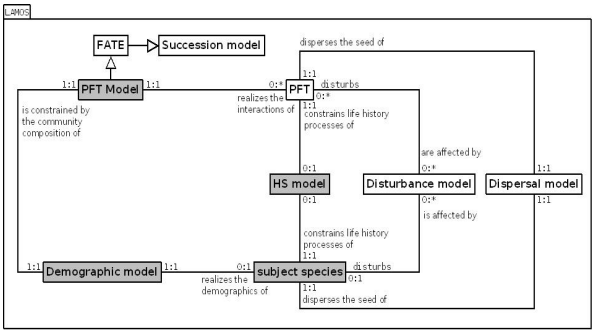


Figure 1. A conceptual UML class diagram depicting LAMOS configured with BioMove. Grey boxes are components comprising BioMove while the remainder are components already available in LAMOS. The components (boxes) their multiplicity and relationships (lines) can be read as a simple sentence. For example, "one PFT model realizes the interactions of zero or more PFT(s)." Lines with arrowheads indicate generalizations or, read the other way, specializations. For example, the PFT model is a specialization of FATE. The multiplicities indicate that BioMove can be run as a single species demographic model (zero PFTs) or a multi-PFT community model (zero subject species) or both. In addition, constraints imposed by the HS (habitat suitability model) are optional (0:1). The number of disturbance models is also optional, that is zero or more disturbances can affect zero of more PFTs. The diagram stipulates that there is one dispersal model for each PFT.

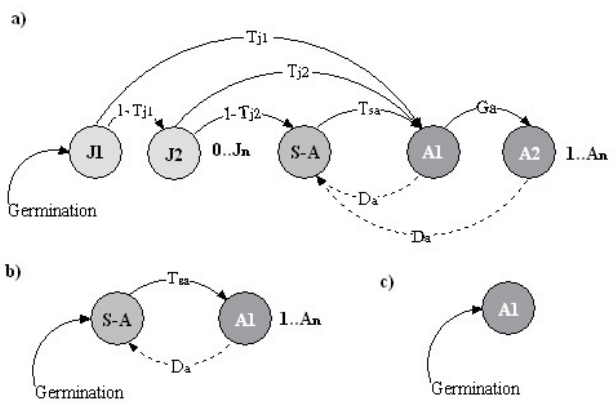
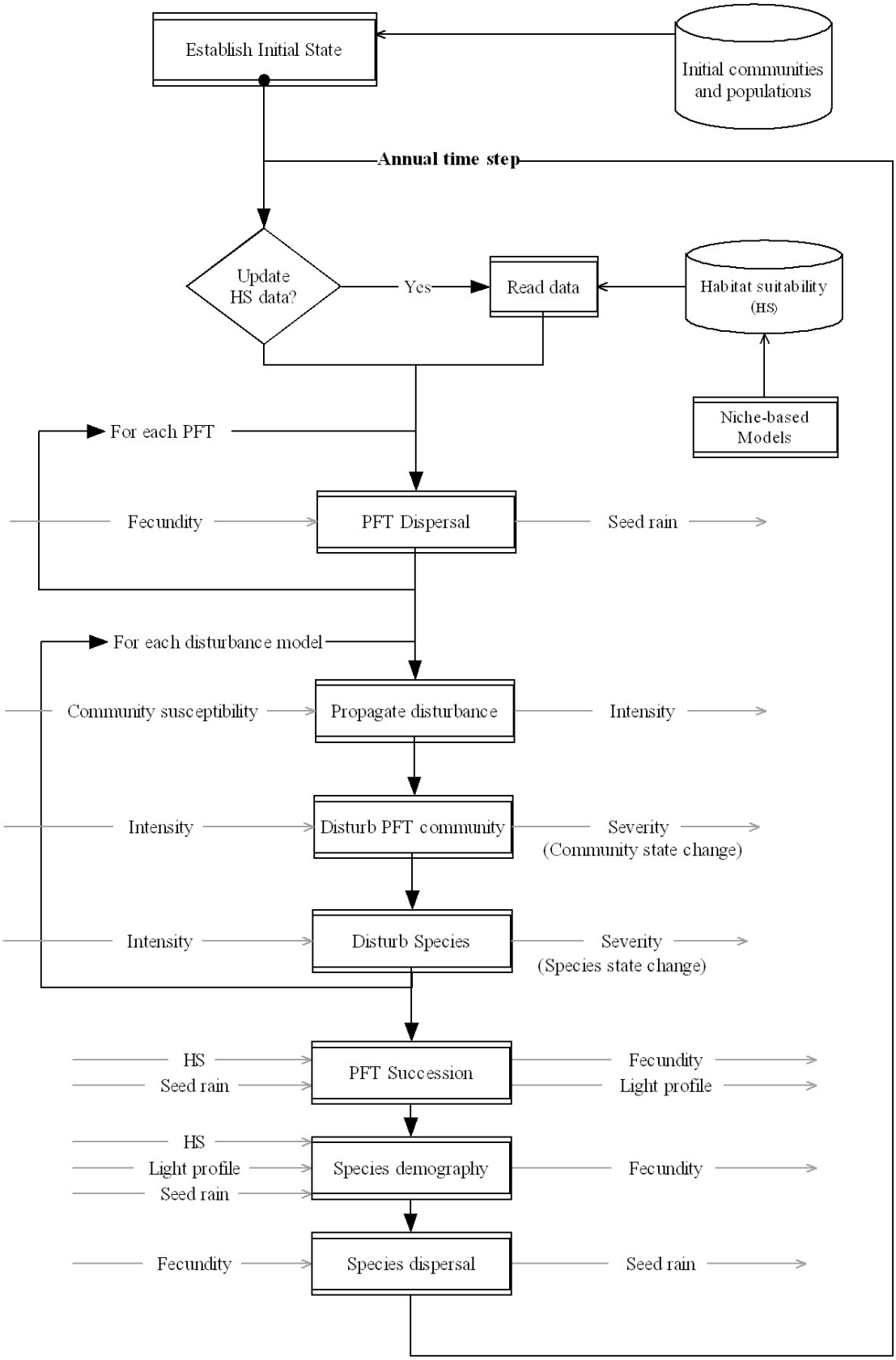


Figure 3. Transition paths for the BioMove Target Species (TS) model. The simplest arrangement a) is for germinants to recruit directly to the adult age class as is the case for annuals modelled with a yearly time step. The next level of elaboration b), is to have a single sub-adult age class to provide the opportunity to delay seed production for adults after disturbance. Beyond this, any number of Juvenile and Adult age classes can be added. Ga is the transition probability between adult age classes (Growth rate). The default value for Tj (Juvenile maturation rates) is zero which results in a time of arrival at the Sub-Adult class equal to the number of juvenile age classes. To model annuals c), the single adult age class has a mortality rate of 1 (not shown). Da is the transition from adult to sub-adult caused by disturbance (re-sprouting).

Figure 2. Flow chart of the principle sub-routines in BioMove. Solid arrow heads indicate the sequence of sub-routine calls. Horizontal arrows show the main inputs and outputs from each sub-routine.



Downloaded from rsbl.royalsocietypublishing.org on 1 October 2009

biology letters

Biol. Lett.
doi:10.1098/rsbl.2009.0669
Published online

Global change biology

Variation in habitat suitability does not always relate to variation in species' plant functional traits

Wilfried Thuiller[1,*,†], Cécile H. Albert[1,†], Anne Dubuis[2], Christophe Randin[2] and Antoine Guisan[2]

[1]*Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier 38041, Grenoble Cedex 9, France*
[2]*Department of Ecology and Evolution, Biophore 1015, Lausanne, Switzerland*
**Author for correspondence (wilfried.thuiller@ujf-grenoble.fr).*
[†]The first two authors contributed equally to the paper.

Habitat suitability models, which relate species occurrences to environmental variables, are assumed to predict suitable conditions for a given species. If these models are reliable, they should relate to change in plant growth and function. In this paper, we ask the question whether habitat suitability models are able to predict variation in plant functional traits, often assumed to be a good surrogate for a species' overall health and vigour. Using a thorough sampling design, we show a tight link between variation in plant functional traits and habitat suitability for some species, but not for others. Our contrasting results pave the way towards a better understanding of how species cope with varying habitat conditions and demonstrate that habitat suitability models can provide meaningful descriptions of the functional niche in some cases, but not in others.

Keywords: ecological niche; mixed models; information theory; intraspecific variability

1. INTRODUCTION

The concept of an ecological niche, traditionally used to describe how a species responds to the environment (Austin *et al.* 1990), has further been used in biogeography to explain why species occur where they do by relating observed presence/absence to environmental gradients, leading to the development of habitat suitability models (Guisan & Thuiller 2005). Predictions from these models are generally thought to be good indicators of habitat suitability, and thus of species' performance. The basic assumption is that, among the habitats occupied by the species, the more frequent are also the more suitable (Albert & Thuiller 2008).

To our knowledge, this assumption has been little investigated (Wright *et al.* 2006; Elmendorf & Moore 2008) and no extensive *in situ* observations have been used to assess whether this simplification of the original definition of the species niche is realistic. While it is difficult to obtain real measures of species performance along environmental gradients, it has been shown that functional traits (physiological, morphological or phenological) could be considered good surrogates (Violle *et al.* 2007). If habitat suitability models relate to change in functional traits (e.g. intraspecific variability), they give a relevant proxy for species performance. Indeed, species should have a higher performance in the core of their niche (i.e. where conditions are more suitable) than at their edges (Pulliam 2000).

In this paper, we investigate, for a set of common plant species, whether species' functional trait measured along gradients are related to predicted habitat suitability. Using 16 plant species in two study areas in the French and Swiss Alps, we demonstrate empirically that habitat suitability and trait expression closely covary for some species, whereas no relationship is detectable for other species, providing moderate support to the use of empirical habitat suitability models to deliver functionally meaningful descriptions of species' niches.

2. MATERIAL AND METHODS

We performed *in situ* trait measurements, habitat suitability modelling and confrontation of trait variation and predicted habitat suitability in two different study areas (see the electronic supplementary material, figure S1): 'Guisane' in the central French Alps (1200–2700 m) and 'Anzeindaz' in the western Swiss Alps (1100–2400 m).

(a) *Species and functional traits selection*

We selected a set of common plant species representing various alpine life forms (dwarf-woody, tall-woody, grass and herbaceous). Sixteen species were chosen in Guisane and five of these in Anzeindaz (table 1). Presence/absence data of the selected species were extracted from the database of the Conservatoire Botanique National Alpin (French site) and from the Swiss Floristic Center (CRSF) database in Geneva (see the electronic supplementary material).

We measured three functional traits strongly related to species' functional strategies (Westoby *et al.* 2002): maximum vegetative height (H_{max}, distance between the top of photosynthetic tissue and the ground) which is associated with the plant competitive vigour; leaf dry matter content (LDMC: the oven-dry mass of a leaf divided by its water-saturated fresh mass, expressed in $mg\,g^{-1}$) which is usually negatively correlated with the relative growth rate; and leaf nitrogen content (LNC: the total amount of nitrogen per unit of dry leaf mass, in $mg\,g^{-1}$) which is closely linked to the maximum photosynthetic rate (Cornelissen *et al.* 2003).

(b) *Field traits measurements*

To measure how these plant functional traits vary within species, we sampled each species in several contrasted environmental conditions. To maximize the sampled environmental heterogeneity, we stratified the study sites by two orthogonal climatic gradients for each site (Guisane: minimum temperature of the coldest month and solar radiation in August; and Anzeindaz: mean January temperature and solar radiation in August). We then combined this environmental stratification with known occurrences (see the electronic supplementary material, figure S1) to select between 6 and 18 plots for each species in each study site. In each plot, three sub-plots were randomly selected and in each sub-plot three mature and non-grazed individuals were randomly selected and measured (figure 1).

(c) *Calibration of habitat suitability models*

To calibrate habitat suitability models, we used comparable sets of non-correlated topo-climatic variables for both sites (figure 1). In Guisane we used: slope, topographic position, growing degree-days (5.56°), moisture index during the growing season, temperature of the coldest month, and annual solar radiation from the meteorological model Aurelhy (Benichou & Le Breton 1987) at 50×50 m. In Anzeindaz we used: slope, topographic position, mean annual temperature, mean annual solar radiations, and mean annual moisture index at 25×25 m (Zimmermann & Kienast 1999).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsbl.2009.0669 or via http://rsbl.royalsocietypublishing.org.

Received 14 August 2009
Accepted 3 September 2009

FirstCite e-publishing

This journal is © 2009 The Royal Society

Downloaded from rsbl.royalsocietypublishing.org on 1 October 2009

Table 1. R^2 and signs of the relationship between variation in habitat suitability and variation in species' functional traits (only R^2 higher than 5% are noted).

	Guisane/Anzeindaz		
species	LDMC	LNC	H_{max}
Carex sempervirens	69.6 [−]/—	34.1 [+]/—	—/18 [+]
Dactylis glomerata	—/—	33.7 [−]/ 48.8 [−]	—/—
Dryas octopetala	—/—	42.8 [+]/—	—/45.3 [+]
Festuca paniculata	—	—	—
Geum montanum	21.5 [+]	—	23.3 [+]
Juniperus sp.	13 [+]	—	20.9 [+]
Larix deciduas	—	45 [−]	33.8 [+]
Leucanthemum vulgare	—	64 [+]	58.5 [+]
Pinus sp.	37.4 [−]	—	31.0 [+]
Polygonum viviparum	—	62.1 [+]	—
Rhododendron ferrugineum	—	—	—
Sesleria caerulea	—/—	41.7 [+]/—	—/33.7 [−]
Salix herbacea	5.1 [+]	100 [+]	18.8 [−]
Silene nutans	—	8.5 [−]	9.7 [+]
Trifolium alpinum	—	49.5 [−]	—
Vaccinium myrtillus	—/60.2 [+]	—/25.4 [−]	63.1 [+] /51.3 [−]

We fitted models using the information-theory approach based on all possible sub-models ($2^{\text{number of candidate variables}}$) for a set of explanatory variables (see the electronic supplementary material). Inference from more than one single 'optimal' model allows the resulting habitat suitability to be the average from all possible candidate models weighted by their weights of evidence (see the electronic supplementary material).

Using this framework, we modelled species' habitat suitability with generalized additive models (see electronic supplementary material). We then extracted the predicted habitat suitability for each sampling plot where we measured the functional traits (figure 1). For each calibrated model, we calculated its predictive accuracy using the area under the curve (AUC) of a receiver operating characteristic plot (see the electronic supplementary material).

(d) *Relationship between traits and habitat suitability*

We used linear mixed models to relate measured functional trait to predicted habitat suitability (figure 1). Functional traits were defined as the response variables, habitat suitability as a fixed explanatory variable and plot and sub-plot as nested random factors (model m_1). The proportion of the plot variability explained by habitat suitability was calculated as a measure of variation explained, following Xu (2003) by

$$R^2 = 100^* \left(1 - \frac{(\sigma_{m_1})^2}{(\sigma_{m_0})^2}\right),$$

where σ_{m_0} and σ_{m_1} are the estimated error standard deviations at plot level estimated under the models with random effects only and under the model m_1, respectively.

3. RESULTS

The evaluation of the habitat suitability models revealed 'fair' predictions for six species, 'good' for height and 'excellent' for two of the 16 species modelled in Guisane, and 'good' for four and 'excellent' for one of the five species in Anzeindaz (see the electronic supplementary material, table S1).

The three measured traits presented strong intraspecific variability. For example, H_{max} was particularly variable, with species' coefficients of variation (CVs) between 0.19 and 0.49, while LDMC (CV in 0.08–0.25) and LNC (CV in 0.09–0.29) tended to be less variable (see the electronic supplementary material, table S2).

Overall, the link between predicted habitat suitability and functional traits was species- and trait-specific (table 1). In general, the link between habitat suitability and H_{max} was positive (e.g. *Leucanthemum vulgare*, figure 2). For some species, there were no relationships between predicted habitat suitability and any trait (e.g. *Festuca paniculata, Rhododendron ferrugineum*). Interestingly, *Vaccinium myrtillus* had a positive relationship between habitat suitability and H_{max} in the French site but a negative one in the Swiss site. The sign of the relationship between habitat suitability and LNC or LDMC was species-specific (table 1). For instance, variation of LNC of *Dactylis glomerata* was negatively correlated to variation in habitat suitability in both Swiss and French sites, while it was strongly positively correlated for *Polygonum viviparum* (figure 2). The same pattern emerges for LDMC with a strong negative correlation between habitat suitability and LDMC for *Carex sempervirens* in the French site (figure 2) and a strong positive one for *V. myrtillus* in the Swiss site (table 1).

The strength of the covariation between predicted habitat suitability and trait expression was not related to the accuracy (AUC) of the habitat suitability model ($p > 0.05$).

4. DISCUSSION

Our contrasting species-specific results corroborate previous studies investigating the relationships between habitat suitability and seed recruitment (Wright *et al.* 2006), or habitat suitability and fecundity (Elmendorf & Moore 2008). Despite an inherent individual variability within plots, models of habitat suitability predict the variation in species' functional traits reliably for some species. In general, H_{max}, an estimate of competitive ability (higher plants get more light), is positively related to habitat suitability, indicating that the more suitable a site, the greater the above-ground growth performance of individuals within a species. The relationship is more contrasted for the other two traits. Intraspecific variation in LDMC is not very well predicted by change in habitat suitability. The strong negative relationship we observe for *C. sempervirens* highlights the well-known trade-off between conservation and exploitation strategies in plant growth (Díaz *et al.* 1998). In a highly suitable habitat, plants tend to invest resources instead of conserving them, thus allocating dry matter to new tissues instead of storing it. Similar to H_{max}, intraspecific variation in LNC, an estimate of photosynthetic activity, is positively related to predicted habitat suitability. Plants exhibit a higher photosynthetic activity in more optimal conditions. Surprisingly, this positive relationship does not hold for some species, notably *D. glomerata*.

Downloaded from rsbl.royalsocietypublishing.org on 1 October 2009

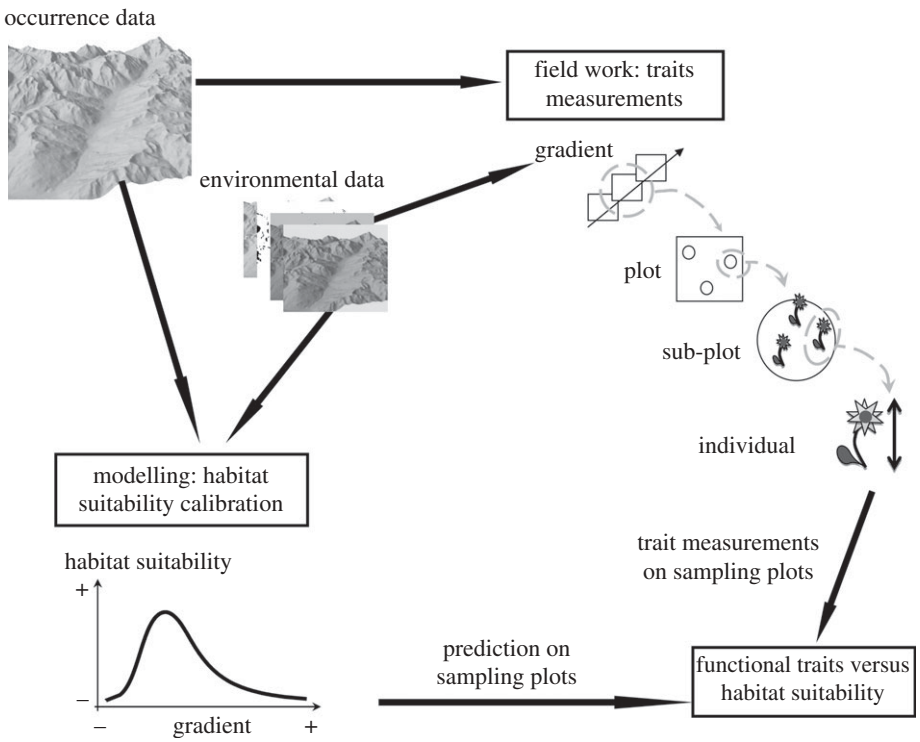


Figure 1. Flow chart of the analysis.

Although we do find a strong relationship between the variation in species' functional traits and variation in habitat suitability for some species, this is not a general pattern in our study with no relationship for half of the species. This suggests that species are affected by the environment in idiosyncratic ways (Hultine & Marshall 2000). This corroborates the results of Gerdol (2005) who showed that two species from the same life form (deciduous dwarf shrubs, *V. myrtillus* and *Vaccinium uliginosum*) can have very different growth performance and nutrient concentrations along environmental gradients. There are several non-exclusive possible explanations to this.

(i) The habitat suitability models may not account for species-specific abiotic factors. We have selected supposedly important topo-climatic variables but some are probably missing. The addition of pH and organic matter did not improve the relationships between habitat suitability and traits (results not shown), but we have not tested for the potential effects of geology, shifts in the microclimatic niches of species, land use, and geomorphic disturbances (data not available).

(ii) The habitat suitability models may not explicitly account for biotic interactions and population dynamics, which might strongly influence functional trait expression. Elmendorf & Moore (2008) showed a higher predictive performance of community-based models to predict seed recruitment than individual habitat suitability models.

(iii) There may be an inherent within-plot variability owing to uncontrolled factors (e.g. micro-habitat, phenotypic variability) which might

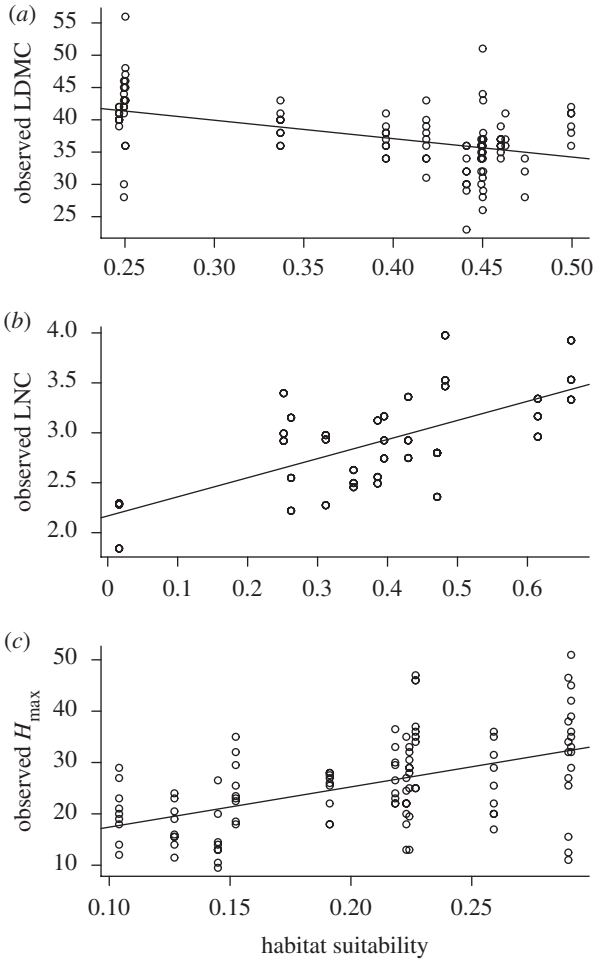


Figure 2. Variation in species' functional traits against variation in species' habitat suitability. (*a*) Maximum height of *C. sempervirens*, (*b*) leaf nitrogen content of *P. viviparum* and (*c*) leaf dry matter content of *L. vulgare*.

Downloaded from rsbl.royalsocietypublishing.org on 1 October 2009

hide strong relationships between variation in habitat suitability and variation in species' functional traits.

(iv) Intraspecific variability in functional traits may not be influenced by the environment. This is unlikely to be the case, given the strong link already highlighted between interspecific variation in traits and the environment (Westoby *et al.* 2002)

Although extremely time consuming and data demanding, the type of analysis we presented here is necessary to demonstrate whether habitat suitability models can only discriminate species presence or absence, or can additionally inform on species performance. As the formal validation of habitat suitability models remains difficult (Guisan & Thuiller 2005), confronting their predictions to measured functional traits, as we propose here, allows a complementary evaluation of their reliability and usefulness.

We thank B. Couchaud, A. Soudant, P. Saccone, F. Boucher, L. Chalmandrier, J. Icard and D. Paulin for their technical help, and D. Pio for the English editing. This research was funded by the ANR-Diversitalp (ANR-07-BDIV-014) and the EU Ecochange (066 866GOCE) projects.

Albert, C. & Thuiller, W. 2008 Favourability functions against probability of presence: advantages and misuses. *Ecography* **31**, 417–422. (doi:10.1111/j.0906-7590.2008.05221.x)

Austin, M. P., Nicholls, A. O. & Margules, C. R. 1990 Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecol. Monogr.* **60**, 161–177. (doi:10.2307/1943043)

Benichou, P. & Le Breton, O. 1987 Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie* **7**, 23–34.

Cornelissen, J. H. C. *et al.* 2003 A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Aust. J. Bot.* **51**, 335–380. (doi:10.1071/BT02124)

Díaz, S., Cabido, M. & Casanoves, F. 1998 Plant functional traits and environmental filters at a regional scale. *J. Vegetat. Sci.* **9**, 113–122. (doi:10.2307/3237229)

Elmendorf, S. C. & Moore, K. A. 2008 Use of community-composition data to predict the fecundity and abundance of species. *Conserv. Biol.* **22**, 1523–1532. (doi:10.1111/j.1523-1739.2008.01051.x)

Gerdol, R. 2005 Growth performance of two deciduous *Vaccinium* species in relation to nutrient status in a subalpine heath. *Flora* **200**, 168–174.

Guisan, A. & Thuiller, W. 2005 Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009. (doi:10.1111/j.1461-0248.2005.00792.x)

Hultine, K. R. & Marshall, J. D. 2000 Altitude trends in conifer leaf morphology and stable carbon isotope composition. *Oecologia* **123**, 32–40. (doi:10.1007/s004420050986)

Pulliam, H. R. 2000 On the relationship between niche and distribution. *Ecol. Lett.* **3**, 349–361. (doi:10.1046/j.1461-0248.2000.00143.x)

Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. 2007 Let the concept of trait be functional! *Oikos* **116**, 882–892. (doi:10.1111/j.0030-1299.2007.15559.x)

Westoby, M., Falster, D. S. & Moles, A. T. 2002 Plant ecological strategies: some leading dimensions of variation between species. *Annu. Rev. Ecol. Syst.* **33**, 125–159. (doi:10.1146/annurev.ecolsys.33.010802.150452)

Wright, J. W., Davies, K. F., Lau, J. A., McCall, A. C. & McKay, J. K. 2006 Experimental verification of ecological niche modelling in a heterogeneous environment. *Ecology* **87**, 2433–2439. (doi:10.1890/0012-9658(2006)87[2433:EVOENM]2.0.CO;2)

Xu, R. H. 2003 Measuring explained variation in linear mixed effects models. *Statist. Med.* **22**, 3527–3541. (doi:10.1002/sim.1572)

Zimmermann, N. E. & Kienast, F. 1999 Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J. Vegetat. Sci.* **10**, 469–482. (doi:10.2307/3237182)

Intraspecific functional variability: extent, structure and sources of variation

Cécile Hélène Albert[1, a], Wilfried Thuiller[1, b], Nigel Gilles Yoccoz[2, c], Alex Soudant[1, d], Florian Boucher[1, e], Patrick Saccone[1, f], Sandra Lavorel [1, 3, g]

1. Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France

2. Department of Arctic and Marine Biology, University of Tromsø, N-9037 Tromsø, Norway

3. Station Alpine Joseph Fourier, CNRS UMS 2925, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France

Key-words: alpine ecosystems, environmental gradients, intraspecific variability, leaf traits, life forms, linear mixed models, plant functional traits, temperature, radiation

Corresponding author: Cécile H. ALBERT
cecile.albert@m4x.org

Summary

1.Functional traits are increasingly used to investigate community structure, ecosystem functioning or to classify species into functional groups. These functional traits are expected to be variable between and within species. Intraspecific variability is supposed to influence and modulate species responses to environmental changes and their effects on their environment. However this hypothesis remains poorly tested and species are mostly described by mean trait values without any consideration of variability in individual trait values.

2.In this paper, we quantify the extent of intraspecific plant functional trait variability, its spatial structure and its response to environmental factors. Using a sampling design structured along two direct and orthogonal climatic gradients in an alpine valley we quantified and analysed the intraspecific variability for three functional traits (height, leaf dry matter content and leaf nitrogen content) measured on sixteen plant species with contrasting life histories.

3. Results showed a large variability of traits within species with large discrepancies between functional traits and species. This variability did not appear to be structured within populations. Between populations, the overall variability was partly explained by the selected gradients. Despite the strong effects of temperature and radiation on traits intraspecific variability, the response curves of traits along gradients seemed to be idiosyncratic.

4.'Synthesis': Giving a comprehensive quantification of intraspecific functional variability through the analysis of an original dataset, we bring new evidence that using a single trait value to describe a given species can hide large functional variation for this species along environmental gradients. These findings suggest that intraspecific functional variability should be a concern for ecologists and its recognition opens new opportunities to better understand and predict ecological patterns in a changing environment. Further analyses are however required to compare inter- and intraspecific variability.

Introduction

Plant functional traits - any measurable features at the individual level affecting directly or indirectly the overall plant fitness (Lavorel *et al.*, 1997; Violle *et al.*, 2007) - are increasingly used for instance to establish functional classifications of organisms (Lavorel *et al.*, 1997), to quantify the functional diversity of communities (de Bello *et al.*, 2009) or to parameterize plant functional types in vegetation models (Woodward & Diament, 1991). These functional traits are variable between and within species. Contrasts between species' trait values have been largely investigated in terms of response of functional traits to gradients or traits trade-offs (Diaz *et al.*, 2004; Reich *et al.*, 2003; Thuiller *et al.*, 2004; Wright *et al.*, 2005). However species are commonly described by functional trait measurements collected from a few individuals from one or few populations and averaged at the population or species level, disregarding the intraspecific functional variability. A large number of studies have been published from regional to global scale investigating the response of functional, eco-physiological or demographic traits to environmental gradients (Cordell *et al.*, 1998; Dyer *et al.*, 2001; Garnier *et al.*, 2001; Meziane & Shipley, 1999; Reich *et al.*, 1999; Ryser & Aeschlimann, 1999). However, the available literature gives an incomplete description of intraspecific variability - as studies have not been carried out with this objective in mind- and reveals a lack of knowledge about its general structure and its response to environmental gradients (McGill *et al.*, 2006, Table S1 & S2 in Supporting Information). Moreover, most studies dealing with intraspecific functional variability have focused on indirect gradients, namely altitude, latitude or longitude (Cordell *et al.*, 1998; Ryser & Aeschlimann, 1999). These indirect gradients, *sensu* Austin (1984), are unknown combinations of multiple direct environmental gradients that impact plant physiology directly (e.g., temperature, nutrient availability). Indirect gradients should not be used to describe ecological patterns as they are not comparable over time and space and do not have any ecological meaning (Körner, 2007).

Several recent studies, however, have shown that intraspecific functional variability can have significant effects on ecosystem functioning (nutrient and carbon cycles: Lecerf & Chauvet, 2008; response to herbivory: Boege & Dirzo, 2004). These studies called for a better understanding of this intraspecific functional variability, its extent, its structure and the way traits vary along the environment, in order to be fully integrated into ecological theories and studies, to further improve our understanding of biodiversity and ecosystems functioning and to predict their response to global changes (McGill *et al.*, 2006). Observed intraspecific functional variability expresses the range of possible trait values as well as the variability of what is functionally experienced by individuals of a given species living in different environmental conditions. It results from genetic diversity and phenotypic plasticity of the different genotypes, both mechanisms being potentially complementary and not exclusive (Byars *et al.*, 2007; Joshi *et al.*, 2001). Intraspecific variability indicates the potential for plant species to survive, grow and reproduce under new environmental conditions (Byars *et al.*, 2007; Joshi *et al.*, 2001). It will thus condition the response of populations and species to environmental changes and the resulting changes in communities and ecosystems. Intraspecific functional variability can occur at different levels, (1) as differences between mean traits of populations living in contrasted environmental conditions, (2) as differences between individuals of a population or (3) as differences between leaves of an individual (Shipley 1995).

We investigated the extent of intraspecific functional variability, without distinguishing genetic and plastic effects, because: (1) this phenotypic variability gives a general idea of the species' ability to cope with new environmental conditions (Milla *et al.*, 2009); (2) from a functional point of view, individual phenotypic expression and its consequences for higher levels of organisation (populations, communities, ecosystems) make the identification of underlying processes responsible for such variability a secondary priority, albeit relevant for other questions such as evolution of species niches. In this paper, we investigate the magnitude and structure of intraspecific functional variability through the analysis of an original dataset collected on sixteen contrasting species following a unique sampling design and protocol for this purpose. The study focussed on alpine ecosystems that provide over small spatial scales steep climatic gradients constraining species functional make-up (Körner, 1999). Such combinations of contrasting conditions and species allowed us to address the four following questions: (1) Magnitude: how variable are functional traits within species? (2) Structure: how is intraspecific functional variability structured spatially (between individuals, sub-populations, populations)? As neighbouring individuals are expected to be genetically related, the variability is expected to be higher between than within populations. (3) Response to gradients: how do traits respond to climatic gradients (shape and strength)? (4) Determinants: Could the functional variability within each

species (magnitude, structure and response to gradients) be related to species' characteristics such as e.g. growth form or niche breadth?

Material and methods

Study site

The study was conducted in the 25 km long Guisane valley (~260 km²; 44.9 °N, 6.6 °E) in the central French Alps (between Briançon at 1200 m and Combe Roche Noire at 3000 m a.s.l.). It is characterised by strongly contrasted climatic conditions, with monthly mean temperatures running from -6.3 °C to -1 °C for the coldest and from 10 °C to 17 °C for the warmest months. As in other valleys of the central Alps, landscape is a mosaic of coniferous and deciduous forests, shrub heaths, subalpine grasslands and alpine meadows, which were all included in the sampling strategy.

Species selection

We sampled, during summer 2007 and 2008, sixteen species (see Table 1) (i) from different life forms (graminoïd, forbs, legume, shrub, tree) to obtain the widest possible scope for the robustness and generality of detected patterns; (ii) that were sufficiently common to be found in contrasting conditions, *i.e.* widely distributed; (iii) with contrasting climatic preferences, *i.e.* species for which optimal conditions are lowland, subalpine or alpine, in order to be able to detect potential functional changes between marginal and optimal populations (Hampe & Petit, 2005). We used two sets of congeners (*Juniperus: sibirica* and *communis* and *Pinus: sylvestris* and *uncinata*) because of their apparent hybridization in the study area (P. Salomez, pers. comm.). We checked *post hoc* that these congeners were relevant functional continuums, *i.e.* had unimodal trait distributions. We calculated for each species the theta index to estimate its degree of specialisation (Fridley *et al.*, 2007). This species co-occurrence index assumes that generalist species co-occur with a larger number of species than specialists (see Appendix S1).

Sampling strategy

Based on the results of a principal component analysis (data not shown) on a set of interpolated topo-climatic variables from the meteorological model Aurelhy at a 50 m resolution (Benichou & Le Breton, 1987), we selected two uncorrelated direct gradients: mean minimal temperature in winter (Temperature) and solar radiation in August (Radiation). We hypothesized that functional traits would respond to these orthogonal gradients either linearly (in case of a strong limiting gradient) or following a bell-shaped curve (with an optimum trait expression along the gradient). They could then be expressed using the following equation with "Temp" for the Temperature and "Rad" for the Radiation:

$$Trait \approx a1 \times Temp + a2 \times Temp^2 + b1 \times Rad + b2 \times Rad^2 + ab \times Temp \times Rad + c$$

Then, following a hierarchical stratified sampling design based on a surface response methodology (Box & Draper, 1959), at least nine samples were required to parameterise the model and account for environmental heterogeneity. To select populations for each species throughout the whole valley, we applied the environmental stratification presented above to the known occurrences of this species in the study valley (dataset from the National Alpine Botanical Conservatory).

In order to determine the spatial structure of the intraspecific functional variability we further stratified the sampling design. For each species, we selected three sub-populations within each population (50 x 50 cm or 1 x 1 m for herbaceous species, and 10 x 10 m for shrubs and trees), and in each sub-population three random individuals (tussocks or ramets) were measured (Table 1). Sub-populations were scaled to the individual size in order to make them 1) contain at least three distinct individuals and 2) be sufficiently homogeneous, so that differences between sub-populations integrate environmental micro-heterogeneity. For leaf trait measurement, we collected one leaf per individual for herbaceous species and ten for shrubs and trees (see Table 1), trait values were averaged at the individual level for the analyses. Only sexually mature plants and unshaded individuals were sampled (except in woodlands). We collected non-senescent, non-grazed, non-frozen, well developed and sun-exposed leaves. We tracked the growing season according to altitude, aspect and field observations in order to sample all the individuals of a species at a similar phenological stage.

Trait selection and measurement

We measured three functional traits to characterize the ecological strategies of the studied species. (1) Maximum vegetative height (Hmax, not for trees), the distance between the top of photosynthetic tissue and the ground, is associated with the plant competition vigour and tends to be allometrically correlated with aboveground biomass (Cornelissen *et al.*, 2003). (2) Leaf dry matter content (LDMC), the dry mass of a leaf divided by its water-saturated fresh mass (Cornelissen *et al.*, 2003) expressed in mg g^{-1}, reflects plant growth rate and carbon assimilation. LDMC is considered to be a robust trait (Roche *et al.*, 2004) and is usually negatively correlated with the Relative Growth Rate (Weiher *et al.*, 1999). LDMC was measured with the partial rehydration method, which has been demonstrated to not produce significant biases as compared to the more intensive full rehydration method (Vaieretti *et al.*, 2007). (3) Leaf nitrogen content (LNC), the total nitrogen per unit of dry leaf mass (in mg g^{-1}), is closely linked to the mass-based maximum photosynthetic rate (Cornelissen *et al.*, 2003). Dried and marble-ground leaves samples of 3-5 mg were analysed with FlashEA 1112 elemental analyzer (Thermo Fisher Scientific Inc. made in Milan, Italia) at the individual or at the sub-population level for species with small leaves (Table 1).

To validate our data and compare trait values with existing literature, we extracted Hmax, LDMC and LNC from several available databases (Fig. 1: own data and see Table S1 in Appendix S1).

Statistical analysis

We used linear mixed models to represent the hierarchical sampling design (Pinheiro & Bates, 2000) and decompose the trait variation of each species. We started with intercept random models (m_0) with only random effects (population, sub-population nested into population) to decompose the variability of trait values (Hmax, LNC and LDMC) at the different levels of sampling. We then built more complex mixed models with linear and quadratic fixed effects of the gradients (m_1). We checked the normality of predicted random effects and residuals, as well as for stability of estimates obtained from different models. To determine what part of the variability at the population level was due to the fixed effects, we calculated a measure of explained variation based on the variances at the population level (σ^2) for the different models (Xu, 2003):

$R^2 = 1 - \frac{(\sigma_{m1})^2}{(\sigma_{m0})^2}$, where σ_{m0} (resp. σ_{m1}) is the estimated error standard deviation at population level estimated under model m_0 (resp. m_1). Variances were estimated by maximizing the restricted log-likelihood (REML).

We further tested if some emerging patterns in intraspecific variability could be related to species' characteristics such as growth form ("Life Form"), niche breadth ("Niche1"), optimal altitude ("Niche2"), sampling range ("Sampling") or specialisation ("Spe") (Table 1). We tested the effects of these characteristics on response variables (Y) with linear models built as follows: Y~ "Characteristics". To test the existence between species of general patterns in: (i) trait variability, we used for Y the coefficient of variation (CV) of the trait for each species (Lande, 1977); (ii) decomposition of trait variance between the different sampling levels, we used for Y the percentage of variance at the population level; (iii) the percentage of trait variation explained by gradients, we used for Y the maximal value of R^2 obtained for the species.

Statistical analyses were carried out with R 2.7.0 (R_Development_Core_Team., 2008) using the packages *nlme, stats, ade4, vioplot.*

Results

Quantification of intraspecific variability

A comparison of the altitudinal range observed for the selected species in the French Alps and our sampling showed that the sampling strategy was rather efficient and that a large part, or at least the upper part, of their range was sampled (Table 1). We found large interspecific variation across the measured traits (Fig. 1). Species mean values ranged from 2.4 cm to 44 cm for Hmax (Fig. 1a, without *Juniperus*), from 142 mg g^{-1} to 518 mg g^{-1} for LDMC (Fig. 1b) and from 9.3 mg g^{-1} to 33.1 mg g^{-1} for LNC (Fig. 1c). Each species showed large intraspecific variability. For example, the mean LDMC of *S. caerulea* was 363 mg g^{-1} but it ranged between 200 mg g^{-1} and 720 mg g^{-1}. Hmax was particularly variable with coefficients of variation (CV) for each species running from 0.19 to 0.49 while LDMC and LNC tended to be less variable (CV in 0.08-0.25 and 0.09-0.29 respectively, Fig. 1). This large variability within each species resulted from both between- and within-population levels. The case of *D. glomerata* is given as an example in the left corner of each sub-figure in Fig. 1. Study species did not have similar magnitudes of variability and no clear trend emerged across species types (Life Forms, Niche1, Niche2, Sampling, Specialization index, see Table 1). Moreover, a given species could be highly variable for one trait and not for another (*T. alpinum* had a CV of 0.46 for Hmax and of 0.09 for LNC). The comparison of the measured values with literature and existing databases (Fig. 1 and Table S1 in Appendix S1) showed that they are consistent with former measurements.

Attribution of the sources of variation

Variance decomposition

The sources of variation varied notably across traits and species (Fig. 2). For most species, these sources varied between the three traits, for example 50% of the variance in Hmax of *T. alpinum* was due to differences between populations, but only 12% for LNC and LDMC. On average, variability between populations represented a large part of the variance (with individual data, on average 47% for Hmax and 35% for LDMC). There was also a high individual variability, reaching as much as 60-80% (LDMC of *G. montanum, L. vulgare, T. alpinum*), and 53% on average for LDMC with individual data. In contrast, only a small part of the variance was due to sub-populations (with individual data, on average 10% for Hmax and 12% for LDMC).

Analysing the structure of variability within literature and databases data led to results comparable to ours, however existing data were not precise enough to achieve the same complete quantification. In particular, very few data were available for LNC and at the within-population level (Table S1 & S2 in Appendix S1).

Effects of environmental variables

All the calibrated models converged and most of them showed significant effects of the fixed factors. Most trait response shapes were quadratic with strong interactions between both temperature and radiation gradients. The models results remain, however, a complex patchwork with response curves and variation explained by models differing between species and between traits (Fig. 3 & Table S3 in Appendix S1). Overall, variations in Hmax were better explained by gradients (on average 30% of the variance at the population level explained) than variations in LNC (on average 21%), with LDMC (26%) being intermediate. Again, there were however strong differences between traits and between species (Table S3 in Appendix S1).

Discussion

A large variability and a complex hierarchical structure but no general pattern

The highly structured sampling design allowed us to answer our four key questions about the structure of intraspecific functional variability. (1) Trait values were highly variable within species but this variability depended on both traits and species, suggesting that the behaviour (response to and effect on the environment) of individuals or populations from a species is not fixed and might differ depending on the conditions. (2) A large part of the variability came from differences between populations. Moreover the variability observed within populations did not seem to be spatially structured. In selecting random individuals in small (scaled to the individual size) homogeneous sub-populations, we assumed first that the differences between sub-populations within a given population would integrate environmental micro-heterogeneity, and second that individuals within a sub-population may be genetically related, although it was beyond the scope of this study to check this assumption. We thus expected large differences between sub-populations and smaller differences within them, but we actually found the opposite with a strong individual variability within sub-populations. This indicates that the sub-populations probably did not integrate either environmental heterogeneity or genetic structure, leading to a large part of the variability remaining unexplained. As our field measurements were not designed to tease apart local adaptation (the observed local variability would be due to the presence of several genotypes within populations) and phenotypic plasticity (the observed local variability would be due to the plasticity of given genotypes), which are both integrated by measures of phenotypic variability (Byars *et al.*, 2007), more studies are needed to research causes for the idiosyncratic variability that we highlighted. Similar results have been obtained on other measured traits (e.g. specific leaf area, leaf carbon content, number of inflorescences, green biomass), however as these traits were not available for all study species, we preferred to focus this paper on three important and available traits. (3) We showed that the shapes and strengths of trait response to environmental gradients are idiosyncratic. Depending on species and traits, climatic gradients explained more or less of the variability between populations. Moreover, including more local factors (at population level) like soil (pH, organic matter content, texture), disturbance (grazing and mowing) or competition (vegetation cover) significantly increased the variability explained in models (in mean the explained part of variance increased from 31 to 59% for Hmax, from 22 to 51% for LDMC and from 17 to 33% for LNC), but did not increase generality in patterns. (4) Although we conducted our study in a very environmentally heterogeneous area and on contrasting species (different life forms, environmental requirements, and degrees of specialisation), we did not detect any pattern in terms of variability, sources of variation or response to gradients. This absence of patterns thus suggests that species are affected by the environment in idiosyncratic ways (Hultine & Marshall, 2000) and corroborates the results of Gerdol (2005), who showed that two related species from the same life form (deciduous dwarf shrubs, *Vaccinium myrtillus* and *uliginosum)* can have very different growth performance (net primary production) and nutrient concentrations (N and P) along environmental gradients.

Traits and gradients: a long story, no happy end?

Although we specifically attempted to sample trait values in order to identify their response shapes along environmental gradients, we did not obtain clearer patterns than previous studies. The apparent complexity and inconsistency of the curves obtained is then not due to a problem of gradient selection and comparability, or of sampling methodology. Nevertheless, traits are expected to respond to environmental gradients (Cornelissen *et al.*, 2003) and patterns have been found at the interspecific level at broad spatial scales along resource, disturbance or environmental gradients (Reich *et al.*, 1999; Wright *et al.*, 2004).

We propose a conceptual model (Fig. 4 & 5) to disentangle the apparent complexity of our results and of data sourced from the literature and databases. We assume that within species traits could respond in a generic way, e.g. following a bell-shaped curve along an environmental gradient, but with parameters differing between species, e.g. maximum trait value at the species' environmental optimum on the gradient (Fig. 4). Then at the intraspecific level, when studying the lower end of the relevant gradient one should find a positive relationship between the trait value and the gradient, while at the upper end a negative relationship is expected, and at an intermediate gradient position either a bell-shaped or a flat relationship are expected (Fig. 4a). Therefore, depending on the part of the gradient sampled and the environmental requirements (range and optimum) of the different studied species, the observed relationships can be very different and inconsistent along the gradient, appearing as idiosyncratic (Fig. 4b). However when the observed curves are combined in terms of species range (in the virtual example: lower part for species 3, higher part for species 1 and intermediate part for species 2), they lead to a bell-shaped curve, *i.e.* a generic

intraspecific response shape along the gradient (as in Fig. 4a). Whereas, when the observed curves are combined along the whole environmental gradient they lead to an interspecific pattern that can be very different from the generic intraspecific one (as in Fig. 4b).

Following this conceptual model, we can further interpret our results. We obtained a unimodal response for Hmax along the temperature gradient, reaching a maximum for intermediate temperature values corresponding to more suitable conditions for the species (Fig. 3 & Fig. 5-A1). For some species (*D. glomerata*, *S. caerulea*), Hmax might continue to increase with temperature because of strong light competition (Woodward, 1986). At the interspecific level, Hmax increased with temperature (R^2=17%) and there was no relationship with radiation (Fig. 5-B1). For LDMC, the conceptual reconstruction leads to an inverse bell-shaped response, showing lower LDMC indicating a higher growth rate for herbaceous species (Weiher *et al.*, 1999) in more suitable places at the core of their distribution (Fig. 3 & Fig. 5-A2). We found for example a decreasing LDMC with increasing temperature for *S. caerulea* which was measured at the upper part of its distribution. LDMC could also increase or follow a bell-shaped response to radiation, consistent with Shipley (2003). At the interspecific level, LDMC has been shown to be either decreasing (Ryser & Aeschlimann, 1999) or increasing (Roche *et al.*, 2004) with altitude; whereas we found no correlation between LDMC and the two studied gradients (Fig. 5-B2). At the intraspecific level, LNC decreased along the temperature gradient with a clearer pattern for nitrogen content per area (mmol m^{-2}) than for nitrogen content per dry mass (mg g^{-1}, Fig. 3 & Fig. 5-A3), corroborating the ecophysiological literature at the species or life form level (Cordell *et al.*, 1998; Körner, 1989; Woodward, 1986). Nitrogen has been extensively investigated along gradients because it is a key element for a variety of metabolic functions in plants, linked to their photosynthetic capacity. At the interspecific level, LNC (mg g^{-1}) decreased (R^2=10%) along the temperature gradient (Fig. 5-B3 & Fig. S1 in Appendix S1), contrary to what has been found at larger scales (Craine, 2003).

Thus, pooling species together masks more complex patterns existing at the intraspecific level and can even change the direction of the relationship, as conceptualized by Shea & Chesson (2002) for the relationship between invasion success and species richness. By using species range and environmental requirements as key parameters in trait-environmental gradient relationships, our conceptual model manages to reconcile the different results from our study and the literature. As a way (i) to limit the shortcomings resulting from gradient selection and sampling range and (ii) to reconcile the general and the species-specific aspects of trait-gradient linkages, we suggest rethinking them as trait-niche relationships; with the niche defined as a synthetic index of species environmental requirements (see Fig. 4). However, methods for evaluating this niche are beyond the scope of this study (see, Elmendorf & Moore, 2008; Thuiller *et al.*, In press; Wright *et al.*, 2006).

Consequences for the future use of functional traits

Our study offers a better knowledge of intraspecific functional variability in establishing an original framework for quantifying its extent, structure, and sources. Such a framework is a preliminary requisite to future analyses aiming to test when taking this variability into account is compulsory or not and what are its implications in diverse fields of ecological research. We first needed to advance our knowledge on the extent of intraspecific functional variability. Showing a large variability of functional traits within species, our study challenges the use of mean values of functional traits to describe species in single species studies, like the response of a species to environmental changes or to herbivory, or the effect of a species on community dynamics or on nutrient cycling. The behaviour of individuals and populations will likely depend on the location of the study along environmental gradients and on other unknown factors (the unexplained part of the variability). Species mean traits should thus be replaced by populations mean traits (in each environmental conditions) or by distribution of traits for more realism. This requires however that descriptions of traits variance are given as well as means in the trait databases (Garnier *et al.*, 2001; Weiher *et al.*, 1999). Second, concerning multi-species studies, our results suggest that further analyses are required to test the relative influence of intraspecific variability versus the interspecific one.

Acknowledgments

We would like to acknowledge the Ecrins National Park, the National Alpine Botanical Conservatory and the Joseph Fourier Alpine Station for their help and the data they provided. We also would like to thank R. Douzet, I. Boulangeat, L. Chalmandrier, J.-C. Clément, B. Couchaud, G. Girard, J. Icard, D. Lejon and F.Grassein for their help in the field and in the lab, as well as the anonymous reviewers and Scott Wilson for their helpful comments on the ms. This work was funded by ANR QDiv (ANR-05-BDIV-009) and DIVERSITALP (ANR-07-BDIV-014) projects. We received support from European Commission's FP6 MACIS (Minimization of and Adaptation to Climate change Impacts on biodiversity, N°044399) and FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, N° 066866 GOCE) projects.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1:

A- *Supplementary Materials: Specialisation index*

B- *Comparison with data from databases and literature*

Table S1: *Data from databases and literature*

Table S2: *Means and coefficients of variance comparison*

Figure S1: *Variation of LNC along the altitudinal gradient*

C- *Models results*

Table S3: *Proportion of total variability at population level explained by climatic gradients*

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

*Austin, M.P., Cunningham, R.B., & Fleming, P.M. (1984) New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio, **55**, 11-27.*

*Benichou, P. & Le Breton, O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. La Météorologie, **7**, 23-24.*

*Boege, K. & Dirzo, R. (2004) Intraspecific variation in growth, defense and herbivory in Dialium guianense (Caesalpiniaceae) mediated by edaphic heterogeneity. Plant Ecology, **175**, 59-69.*

*Box, G.E.P. & Draper, N.R. (1959) A basis for the selection of a response surface design. Journal of American Statistical Association, **54**, 622-654.*

*Byars, S.G., Papst, W., & Hoffmann, A.A. (2007) Local adaptation and cogradient selection in the alpine plant, Poa hiemata, along a narrow altitudinal gradient. Evolution, **61**, 2925-2941.*

*Cordell, S., Goldstein, G., Mueller-Dombois, D., Webb, D., & Vitousek, P.M. (1998) Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. Oecologia, **113**, 188-196.*

*Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., ter Steege, H., Morgan, H.D., van der Heijden, M.G.A., Pausas, J.G., & Poorter, H. (2003) Handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany, **51**, 335-308.*

*Craine, J.M., Lee, W.G. (2003) Covariation in leaf and root traits for native and non-native grasses along and altitudinal gradient. Oecologia, **134**, 471-478.*

*de Bello, F., Thuiller, W., Leps, J., Choler, P., Clément, J.-C., Macek, P., Sebastia, M.T., & Lavorel, S. (2009) Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. Journal of Vegetation Science, **in press**.*

*Diaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H.C., Jalili, A., Montserrat-Marti, G., Grime, J.P., Zarrinkamar, F., Asri, Y., Band, S.R., Basconcelo, S., Castro-Diez, P., Funes, G., Hamzehee, B., Khoshnevi, M., Perez-Harguindeguy, N., Perez-Rontome, M.C., Shirvany, F.A., Vendramini, F., Yazdani, S., Abbas-Azimi, R., Bogaard, A., Boustani, S., Charles, M., Dehghan, M., de Torres-Espuny, L., Falczuk, V., Guerrero-Campo, J., Hynd, A., Jones, G., Kowsary, E., Kazemi-Saeed, F., Maestro-Martinez, M., Romo-Diez, A., Shaw, S., Siavash, B., Villar-Salvador, P., & Zak, M.R. (2004) The plant traits that drive ecosystems: Evidence from three continents. Journal of Vegetation Science, **15**, 295-304.*

*Dyer, A.R., Goldberg, D.E., Turkington, R., & Sayre, C. (2001) Effects of growing conditions and source habitat on plant traits and functional group definition. Functional Ecology, **15**, 85-95.*

*Elmendorf, S.C. & Moore, K.A. (2008) Use of community-composition data to predict the fecundity and abundance of species. Conservation Biology, **22**, 1523-1532.*

*Fridley, J.D., Vandermast, D.B., Kuppinger, D.M., Manthey, M., & Peet, R.K. (2007) Co-occurrence based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. Journal of Ecology, **95**, 707-722.*

*Garnier, E., Laurent, G., Bellmann, A., Debain, S., Berthelier, P., Ducout, B., Roumet, C., & Navas, M.-L. (2001) Consistency of species ranking based on functional leaf traits. New Phytologist, **152**, 69-83.*

*Gerdol, R. (2005) Growth performance of two deciduous Vaccinium species in relation to nutrient status in a subalpine heath. Flora, **200**, 168-174.*

*Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: the rear edge matters. Ecology Letters, **8**, 461-467.*

*Hultine, K.R. & Marshall, J.D. (2000) Altitude trends in conifer leaf morphology and stable carbon isotope composition. Oecologia, **123**, 32-40.*

*Joshi, J., Schmid, B., Caldeira, M.C., Dimitrakopoulos, P.G., Good, J., Harris, R., Hector, A., Huss-Danell, K., Jumpponen, A., Minns, A., Mulder, C.P.H., Pereira, J.S., Prinz, A., Scherer-Lorenzen, M., Siamantziouras, A.S.D., Terry, A.C., Troumbis, A.Y., & Lawton, J.H. (2001) Local adaptation enhances performance of common plant species. Ecology Letters, **4**, 536-544.*

*Körner, C. (1989) The nutritional status of plants from high altitudes. A worldwide comparison. Oecologia, **81**, 379–391.*

Körner, C. (1999) Alpine Plant Life Springer-Verlag, Berlin.

*Körner, C. (2007) The use of "altitude" in ecological research. Trends in Ecology & Evolution, **22**, 570-574.*

*Lande, R. (1977) On comparing coefficients of variation. Systematic Zoology, **26**, 214-217.*

*Lavorel, S., McIntyre, S., Landsberg, J., & Forbes, T.D.A. (1997) Plant functional classifications: from general groups to specific groups based on response to disturbance. Trends in Ecology & Evolution, **12**, 474-478.*

*Lecerf, A. & Chauvet, E. (2008) Intraspecific variability in leaf traits strongly affects alder leaf decomposition in a stream. Basic and Applied Ecology, **9**, 598-605.*

*McGill, B.J., Enquist, B.J., Weiher, E., & Westoby, M. (2006) Rebuilding community ecology from functional traits. Trends in Ecology & Evolution, **21**, 178-185.*

*Meziane, D. & Shipley, B. (1999) Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. Plant Cell And Environment, **22**, 447-459.*

*Milla, R., Escudero, A., & Iriondo, J.M. (2009) Inherited variability in multiple traits determines fitness in populations of an annual legume from contrasting latitudinal origins. Annals Of Botany, **103**, 1279-1289.*

Pinheiro, J.C. & Bates, D.M. (2000) Mixed-effects models in S and S-Plus Springer, New York.

R_Development_Core_Team., ed. (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

*Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Vollin, J.C., & Bowman, W.D. (1999) Generality of leaf trait relationships: a test across six biomes. Ecology, **80**, 1955-1969.*

*Reich, P.B., Wright, I.J., Cavender-Bares, J., Craine, J.M., Oleksyn, J., Westoby, M., & Walters, M.B. (2003) The evolution of plant functional variation: Traits, spectra, and strategies. International Journal of Plant Sciences, **164**, S143-S164.*

*Roche, P., Diaz-Burlinson, N., & Gachet, S. (2004) Congruency analysis of species ranking based on leaf traits: which traits are the more reliable? Plant Ecology, **174**, 37-48.*

*Ryser, P. & Aeschlimann, U. (1999) Proportional dry-matter content as an underlying trait for the variation in the relative growth rate among 22 Eurasian populations of Dactylis glomerata. Functional Ecology, **13**, 473-482.*

*Shea, K. & Chesson, P. (2002) Community ecology theory as a framework for biological invasions. Trends in Ecology & Evolution, **17**, 170-175.*

*Shipley, B. (2003) Interspecific consistency and intraspecific variability of specific leaf area with respect to irradiance and nutrient availability. Ecoscience, **10**, 74-79.*

Thuiller, W., Albert, C.H., Dubuis, A., Randin, C., & Guisan, A. (In press (doi:10.1098/rsbl.2009.0669)) Variation in habitat suitability models does not always relate to variation in species' plant functional traits. Biology Letters.

*Thuiller, W., Lavorel, S., Midgley, G., Lavergne, S., & Rebelo, T. (2004) Relating plant traits and species distributions along bioclimatic gradients for 88 Leucadendron taxa. Ecology, **85**, 1688-1699.*

*Vaieretti, M.V., Diaz, S., Vile, D., & Garnier, E. (2007) Two measurement methods of leaf dry matter content produce similar results in a broad range of species. Annals Of Botany, **99**, 955-958.*

*Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007) Let the concept of trait be functional! Oikos, **116**, 882-892.*

*Weiher, E., van der Werf, A., Thompson, K., Roderick, M., Garnier, E., & Eriksson, O. (1999) Challenging Theophrastus: A common core list of plant traits for functional ecology. Journal of Vegetation Science, **10**, 609-620.*

*Woodward, F.I. (1986) Ecophysiological studies on the shrub Vaccinium myrtillus L. taken from a wide altitudinal range. Oecologia, **70**, 580-586.*

*Woodward, F.I. & Diament, A.D. (1991) Functional approaches to predicting ecological effects of global change. Functional Ecology, **5**, 202-212.*

*Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Warton, D.I., & Westoby, M. (2005) Modulation of leaf economic traits and trait relationships by climate. Global Ecology and Biogeography, **14**, 411-421.*

*Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. (2004) The worldwide leaf economics spectrum. Nature, **428**, 821-827.*

*Wright, J.W., Davies, K.F., Lau, J.A., McCall, A.C., & McKay, J.K. (2006) Experimental verification of ecological niche modeling in a heterogeneous environment. Ecology, **87**, 2433-2439.*

*Xu, R.H. (2003) Measuring explained variation in linear mixed effects models. Statistics in Medicine, **22**, 3527-3541.*

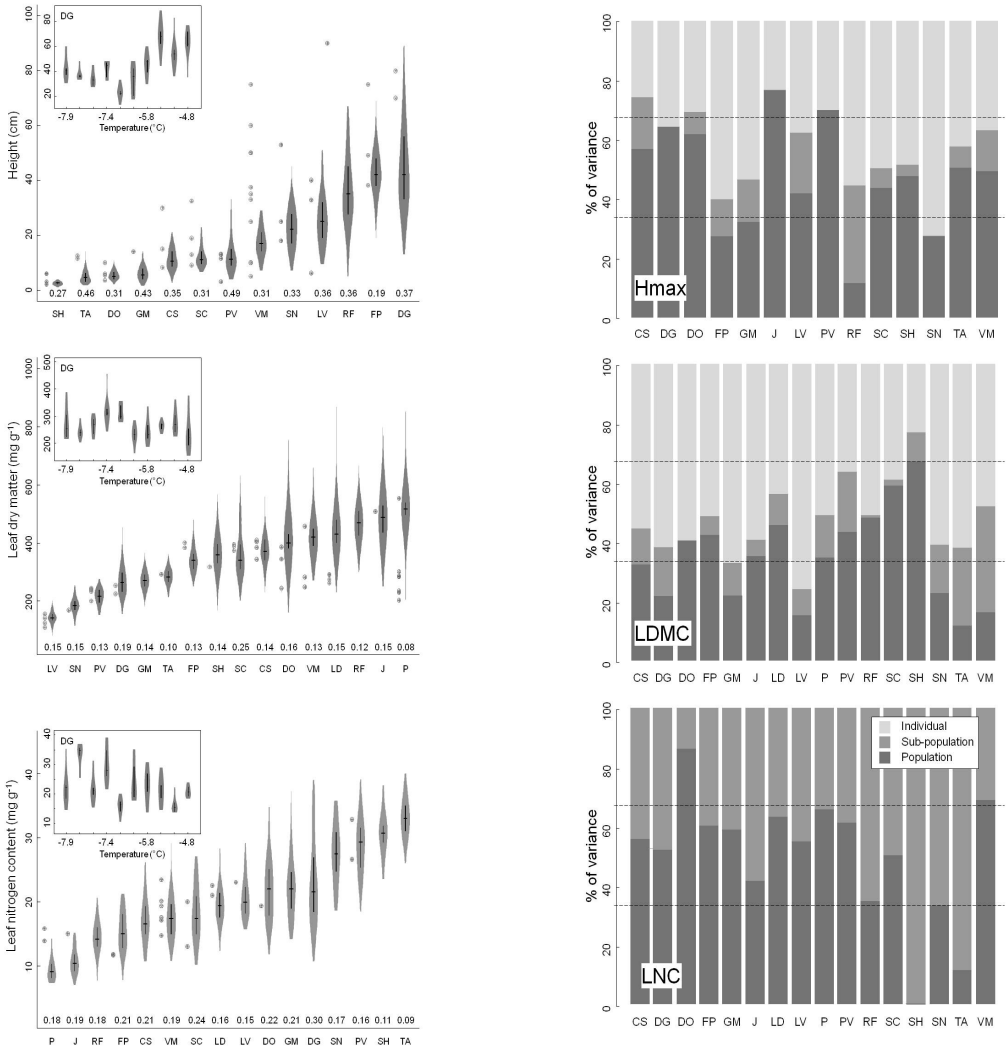


Fig. 1. Extent of the intraspecific functional variability

A violin plot is a combination of a boxplot and a kernel density plot (a non-parametric way of estimating the probability density function of a random variable). It presents (i) the density of the data estimated by kernel method (in grey) (ii) the median value (black dash) (iii) the inter-quartile range: between the first and the third ones (black segment). These figures give the extent of the intraspecific functional variability for (a) the maximum vegetative height (Hmax), (b) the leaf dry matter content (LDMC) and (c) the leaf nitrogen content (LNC), at the level of the sub-population or the individual plant. For each trait, results are presented by species ordered by their mean values (species codes come from the Table 1) and coefficients of variation are given for each species under the violin (ex: 0.15 for the LDMC of Juniperus sp.). The grey asterisks represent data from the literature and from databases (see Appendix S2 in Supporting Information) for comparison with our results. The inset within each graph presents data for D. glomerata (DG); trait values are given per population sorted by increasing temperature (minimal temperature in winter).

Fig. 2. Sources of variation within species

Relative variance decomposition at the individual, sub-population and population levels for (a) the maximal vegetative height (Hmax), (b) the leaf dry matter content (LDMC) and (c) the leaf nitrogen content (LNC, data at the sub-population level). Species are sorted by their names (species codes come from the Table 1) and the 33 and 66% thresholds are given in each case by a dashed line.

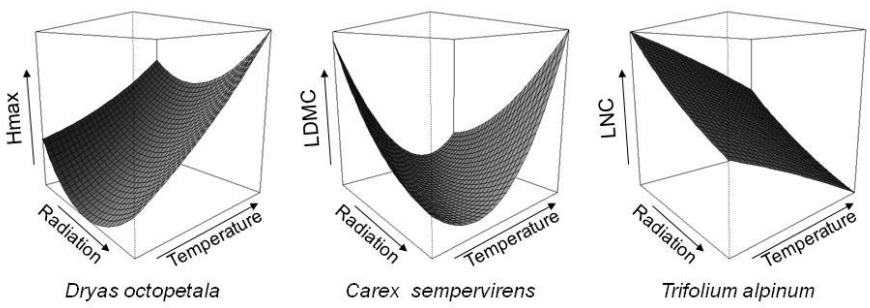


Fig. 3. Traits response shapes along gradients

These examples show the diverse shapes (significant relationships, see Table S3) resulting from models with the trait as the response and linear, quadratic and interaction between the gradients as the explanatory variables. T. alpinum has been measured in the middle of its range and D. octopetala and C. sempervirens have been measured at the upper end of their ranges.

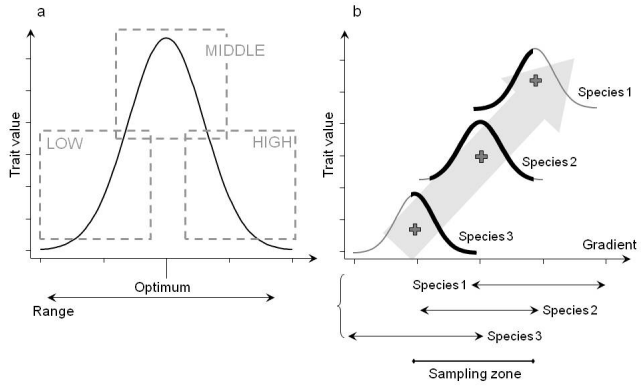


Fig. 4. Trait response along the gradients at the inter- and intraspecific levels: the conceptual model

(a) Assumed generic intraspecific trait response shape along species range: the trait is following a bell-shaped curve, reaching the maximum trait value at the species' environmental optimum on the gradient and decreasing towards the ends of the species range. The trait can be measured in the field at the low, middle or upper part of the species range, leading to different observed relationships between trait and gradient. (b) Theoretical multi-species analysis. Three virtual species are sampled following the gradient. One species is sampled at the lower part of its range, one species at the middle and one species at the upper part: the observed relationships then appear as idiosyncratic. The interspecific pattern is very different from the generic intraspecific one.

Fig. 5. Trait response along the gradients at the inter- and intraspecific levels: application of the model to the three measured traits

(A) Intraspecific variation of LNC, Hmax and LDMC (grey line) with examples for two measured species (dark arrows) along their ranges (with species range limits at low and high temperature and optimum at intermediate temperature). (B) Interspecific variation of LNC, Hmax and LDMC (grey arrow) expressed as the overlay of intraspecific responses for three hypothetical species (dark arrows) along a gradient (decreasing temperature). Pooling intraspecific patterns along the gradient may lead to interspecific patterns that are different from intraspecific ones.

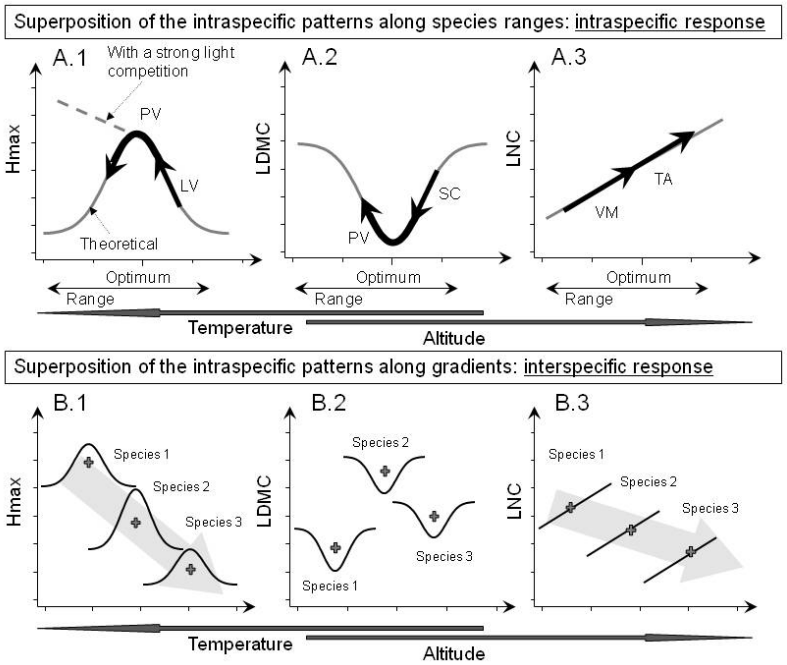


Table 1. Characteristics of the sixteen species sampled

For all sampled species are given: their names (species were sampled in 2007, except species with an asterisk which were sampled in 2008), their labels, their families, their life forms (G: grass, H: herb, S: shrub, T: tree), the number of sites where they were sampled, the number of measured individuals, the number of measured leaves per individual plant, the way the Leaf Nitrogen Content (LNC) was sampled (SP: samples pooled by sub-populations, I: measured individually). 'Niche1' represents species' altitudinal range (considered as the inter 10-90% quantiles length from the presence points observed by the Conservatoire National Botanique des Alpes: CBNA), 'Niche2' represents the optimal altitude of the species (mode of the distribution curve of presence along altitude, data from CBNA), 'Sampling' represents the altitudinal range sampled in the study and 'Spe' represents species' Fridley's specialisation index (see Appendix S1 in Supporting Information). The 'environmental requirements and sampling' describes the altitudinal range of the species between 0 and 4000 m.a.s.l.: (1) white rectangle give the full altitudinal range from the presence points observed by the CBNA; (2) hatched rectangle give the inter-quantiles (10-90%) of the altitudinal range and (3) dark rectangle give the altitudinal range we sampled.

Species	Label	Family	Life Form	Sites	Individuals	Leaves	LNC	Niche 1	Niche 2	Sampling	Spe	Environmental requirements and Sampling
Carex sempervirens Vill.	CS	*cyperaceae*	G	12	108	1	SP	897	1990	899	0.54	
Dactylis glomerata L.	DG	*poacea*	G	10	90	1	I	1448	1073	624	0.62	
Dryas octopetala L.	DO	*rosaceae*	S	15	135	10	I	771	2040	878	0.52	
Festuca paniculata (L) Schinz & Thell	FP	*poacea*	G	16	144	1	I	605	2040	831	0.48	
Geum montanum L.	GM	*rosaceae*	H	17	153	1	I	781	2308	760	0.46	
Juniperus communis L. / sibirica Lodd. Ex Burgsd	J	*cupressaceae*	S	14	126	10	SP	1597	1504	1156	0.60	
Larix decidua Mill.	LD	*pineaceae*	T	18	162	10	I	870	1883	747	0.62	
*Leucanthemum vulgare Lam. **	LV	*asteraceae*	H	12	108	1	SP	1385	1278	920	0.53	
Pinus sylvestris L. / uncinata Ramond ex.DC	P	*pineaceae*	T	9	81	10	SP	1329	1221	720	0.59	
*Polygonum viviparum L.**	PV	*polygonaceae*	H	12	108	1	SP	770	2247	525	0.57	
Rhododendron ferrugineum L.	RF	*ericaceae*	S	7	63	10	I	737	1981	626	0.55	
Salix herbacea L.	SH	*salicaceae*	S	7	63	10	SP	492	2471	329	0.45	
Sesleria caerulea (L.) Ard	SC	*poacea*	G	11	99	1	SP	1456	1748	916	0.61	
*Silene nutans L. **	SN	*caryophyllaceae*	H	12	108	1	SP	1386	1621	895	0.58	
Trifolium alpinum L.	TA	*fabaceae*	H	13	117	1	SP	585	2279	871	0.43	
Vaccicium myrtillus L.	VM	*ericaceae*	S	14	126	10	I	1004	1683	719	0.60	
Total	**16**	**11**	**4**	**199**	**1791**	**8595**	**1179**	**-**	**-**	**-**	**-**	**0 m.a.s.l. 2000 4000**

Appendix S1:

A- Supplementary Materials

Specialisation index:
The degree of specialisation-generalisation of each species was determined using the recently proposed Fridley *theta* index (Fridley *et al.*, 2007). This measure is calculated from the beta diversity among the plots with a given species. It is particularly interesting because it makes no hypothesis about the environmental factors controlling species' distribution niches. For each species, a fixed number of 10 plots were randomly sampled among all plots in which the species was present, and all co-occurring species were added up. This sum was corrected by subtracting the alpha diversity of the sampled plots. This procedure, repeated 50 times, gave a mean value that was the quantitative index of species' specialization (theta).

B- Comparison with data from databases and literature

To validate our data and compare trait values and variability with existing literature, we extracted LNC, LDMC and Hmax from several available databases. ANDROSACE combines several other databases (VISTA, Garnier *et al.*, 2007; LEDA, Knevel *et al.*, 2003) . "Botanists" for the origin of some Hmax values means data come from internet sites of amateur botanists. We also used data directly from literature (Gerdol, 2005; He *et al.*, 2006; Peratoner *et al.*, 2007; Reich *et al.*, 1999; Ryser & Aeschlimann, 1999; Thomas *et al.*, 2007; Welker *et al.*, 1997). The collected data are summarised in the following table (Table S1). More data can be found on Hmax and LDMC than on LNC. More data can be found on populations' mean traits than on populations' trait variability.

From the data in Table S1, one can extract coefficients of variation (CV) between and within plots (populations). For the between plots level (CVb), mean and variance can be calculated from the collected data, CV is then the ratio between the root mean square of the variance and the mean. For the within plot level (CVi), one can calculate a CV for each plot where SD is available and then calculate a mean resulting CV. More information is available in databases and literature on the inter-plot than on the intra-plot variability. More information is available on the variability of Hmax and LDMC than on the one of LNC.

Table S1. Data from databases and literature

Trait value found in databases and literature for the three traits studied (Hmax, Leaf dry matter content and Leaf nitrogen content) for the studied species. Mean and SD (standard deviation) are given.

	Hmax			LDMC			LNC		
Libelle	Mean	SD	Ref	Mean	SD	Ref	Mean	SD	Ref
Carex sempervirens	15		Androsace	40.98	0.82	Androsace	16.94		Androsace
	8.15	3.9	Androsace	40.46	1.4	Androsace	15.06		Androsace
	30	10	Botanists	38.5	1.9	Androsace	12.7		Androsace
				34.4	2.3	Androsace	14.8		Androsace
Dactylis glomerata	80	40	Androsace	25.3	1.6	Ryser 1999			
	70	50	Botanists						
Dryas octopetala	3.5	1.5	Androsace	34.5		Androsace	1.93		Androsace
	6	4	Botanists	24.4	1.2	Androsace			
	10	5	Botanists	38.7		Androsace			
	5.4	1.4	Welker 1997						
Festuca panuiculata	38.18	8	Androsace	40.2	2.4	Androsace	1.16		Androsace
	49.1	9.6	Androsace	38.5	5.2	Androsace	1.18		Androsace
	75	25	Botanists						
GM	14		Androsace						
Juniperus communis	100		Androsace	50.9		Androsace	1.5		Thomas 2007
& sibirica	150		Androsace						
	160		Androsace						
	50		Androsace						
	30		Androsace						
	120		Androsace						
	100		Androsace						
	3.5	2.5	Botanists						
Larix decidua				29.2	3.9	Androsace	2.25		Androsace
				26.2	1.25	Androsace	2.1		Reich 1999
				29	0.8	Androsace			
				27.5	0.6	Androsace			
Leucanthemum vulgare	32.8		Androsace	12.4		Androsace	2.3		Androsace
	40		Androsace	14.1		Androsace			
	6.1		Androsace	15.5		Androsace			
				11		Androsace			
Pinus sylvestris				28.6		Androsace	1.58		Androsace
& uncinata				28.4		Androsace	1.39		Reich 1999
				55.5		Androsace			
				22.9		Androsace			
				30.1		Androsace			
				23.5		Androsace			
				20.3		Androsace			
Polygonum viviparum	13		Androsace	24.5		Androsace	2.66		Androsace
	3		Androsace	23.9		Androsace	3.28		He 2006
	11.5	2.5	Androsace	23.3		Androsace			
	13.05	3.7	Androsace	20	2	Androsace			
				24	2	Androsace			
Rhododendron ferrugineum	90		Androsace						
Sesleria caerulea	9		Androsace	37.4	2.7	Androsace	1.3		Androsace
	12.9	2.55	Androsace	39.6	1.1	Androsace	2		Androsace
	18.9		Androsace	38.8	2.9	Androsace			
	32.5	17.5	Botanists						
Salix herbacea	6		Androsace	31.8	2.9	Androsace			
	2		Androsace						
	3		Androsace						
	6		Androsace						
Silene nutans	53		Androsace	16.8		Androsace			
	25		Androsace						
	18		Androsace						
Trifolium alpinum	11.45		Androsace	29.3		Androsace	2.3	0.65	Peratoner 2007
	12.5	7.5	Botanists						
Vaccinium myrtillus	25		Androsace	45.7		Androsace	1.47		Androsace
	35		Androsace	28.2		Androsace	1.71		Androsace
	60		Androsace	24.8		Androsace	1.93	0.02	Gerdol 2005
	50		Androsace				2.01	0.02	Gerdol 2005
	5		Androsace				1.75	0.02	Gerdol 2005
	33		Androsace				2.34		Reich 1999
	10		Androsace						
	75		Androsace						
	10		Androsace						
	10		Androsace						
	120		Androsace						
	50		Androsace						

Species	Hmax				LDMC				LNC			
	Inter-plot		Intra-plot		Inter-plot		Intra-plot		Inter-plot		Intra-plot	
	Own	Lit	Own	Lit	Own	Lit	Own	Lit	Own	Lit	Own	Lit
Carex sempervirens	11/0.28	17/0.63	0.22	0.4	37/0.09	38/0.08	0.09	0.04	17/0.18	14/0.12		
Dactylis glomerata			0.22	0.6			0.16	0.06				
Dactylis (Ryser 1999)					26/0.06	25/0.11						
Dryas octopetala	5.5/0.26	6/0.43	0.18	0.46	41/0.09	32/0.22	0.1	0.05				
Festuca paniculata	41/0.13	54/0.35	0.16	0.24								
Larix decidua					44/0.08	29/0.05	0.12	0.06				
Leucanthemum vulgare	26/0.23	26/0.67			14/0.08	13/0.15						
Pinus					52/0.03	30/0.39						
Polygonum viviparum	12/0.43	10/0.47	0.26	0.25	22/0.09	23/0.07	0.09	0.09				
Sesleria caerulea	12/0.23	18/0.56	0.21	0.36	36/0.2	39/0.02	0.14	0.05				
Salix herbacea	2.6/0.22	4.2/0.48										
Silène nutans	22/0.2	32/0.58										
Vaccinium myrtillus	17/0.24	40/0.76	0.2	0.6					1.74/0.14	1.88/0.29		
Vaccinium (Gerdol 2005)									1.74/0.14	1.92/0.09		

Table S2. Means and coefficients of variance comparison between our data and data from literature and databases

For each of the three traits (Hmax, LDMC and LNC), mean/CVb are given at the inter-plot level estimated from our own data (own) and from the data from the Table S1 (Lit). At the intra-plot level the CVi are given, estimated from our own data (own) and from the data from the Table S1 (Lit).

The variability found in our study for Hmax was lower at inter- and intra-plot levels than what can be observed in former data. For LDMC, the variability that we observed at the inter- and intra-plot levels was either similar to what can be found in former data, or we obtained a lower inter-plot variability and a higher intra-plot variability. For LNC, we observed an inter-plot variability which was consistent with what can be found in former data, but the intra-plot variability was never described in those data.

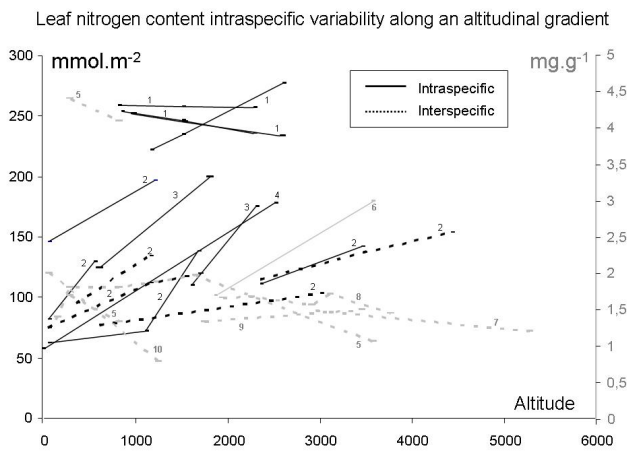


Figure S1. Variation of LNC in response to altitude

Data from the literature on the variability of leaf nitrogen content are presented as means along the altitudinal gradient. Depending on the source, data are in mmol.m^-2 or in mg.g^-1, they are at the intra or the interspecific level: 1(Hultine & Marshall, 2000), 2(Körner, 1989), 3(Sparks & Ehleringer, 1997), 4(Cordell et al., 1998), 5(Reich et al., 1999), 6(Bowman, 1995), 7(He et al., 2006), 8(Luo et al., 2005), 9(Wegner et al., 2003), 10(Craine, 2003).

C- Models results

Species/Trait	Hmax	LDMC	LNC
CS	10	79	28
DG	78	11	4
DO	72	11	44
FP	0	0	25
GM	65	0	0
J	79	0	23
LD		3	36
LV	13	36	0
P		50	15
PV	0	0	22
RF	0	48	0
SC	49	0	0
SH	20	96	0
SN	40	8	0
TA	0	41	100
VM	49	29	35

Table S3. Proportion of total variability at plot level explained by climatic gradients

References

Bowman, W.D. (1995). Altitudinal variation in leaf N and P concentrations in frasera speciosa (green gentian). Open Space and Moutain Parks, Boulder.

*Cordell, S., Goldstein, G., Mueller-Dombois, D., Webb, D., & Vitousek, P.M. (1998) Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. Oecologia, **113**, 188-196.*

*Craine, J.M., Lee, W.G. (2003) Covariation in leaf and root traits for native and non-native grasses along and altitudinal gradient. Oecologia, **134**, 471-478.*

*Fridley, J.D., Vandermast, D.B., Kuppinger, D.M., Manthey, M., & Peet, R.K. (2007) Co-occurrence based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. Journal Of Ecology, **95**, 707-722.*

*Garnier, E., Lavorel, S., Ansquer, P., Castro, H., Cruz, P., Dolezal, J., Eriksson, O., Fortunel, C., Freitas, H., Golodets, C., Grigulis, K., Jouany, C., Kazakou, E., Kigel, J., Kleyer, M., Lehsten, V., Leps, J., Meier, T., Pakeman, R., Papadimitriou, M., Papanastasis, V.P., Quested, H., Quetier, F., Robson, M., Roumet, C., Rusch, G., Skarpe, C., Sternberg, M., Theau, J.P., Thebault, A., Vile, D., & Zarovali, M.P. (2007) Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: A standardized methodology and lessons from an application to 11 European sites. Annals Of Botany, **99**, 967-985.*

*Gerdol, R. (2005) Growth performance of two deciduous Vaccinium species in relation to nutrient status in a subalpine heath. Flora, **200**, 168-174.*

*He, J.S., Wang, Z., Wang, X., Schmid, B., Zuo, W., Zhou, M., Zheng, C., Wang, M., & Fang, J. (2006) A test of the generality of leaf trait relationships on the Tibetan Plateau. New Phytologist, **170**, 835-848.*

*Hultine, K.R. & Marshall, J.D. (2000) Altitude trends in conifer leaf morphology and stable carbon isotope composition. Oecologia, **123**, 32-40.*

*Knevel, I.C., Bekker, R.M., Bakker, J.P., & Kleyer, M. (2003) Life-history traits of the Northwest European flora: The LEDA database. Journal Of Vegetation Science, **14**, 611-614.*

*Körner, C. (1989) The nutritional status of plants from high altitudes. A worldwide comparison. Oecologia, **81**, 379–391.*

*Luo, T., Luo, J., & Pan, Y. (2005) Leaf traits and associated ecosystem characteristics across subtropical and timberline forests in the Gongga Moutains, Eastern Tibetan Plateau. Oecologia, **142**, 261-273.*

*Peratoner, G., Joergensen, R.G., & Spatz, G. (2007) Growth of Trifolium alpinum: Effects of soil properties, symbionts and pathogens. Ecological engineering, **30**, 349-355.*

*Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Vollin, J.C., & Bowman, W.D. (1999) Generality of leaf trait relationships: a test across six biomes. Ecology, **80**, 1955-1969.*

*Ryser, P. & Aeschlimann, U. (1999) Proportional dry-matter content as an underlying trait for the variation in the relative growth rate among 22 Eurasian populations of Dactylis glomerata. Functional Ecology, **13**, 473-482.*

*Sparks, J.P. & Ehleringer, J.R. (1997) Leaf carbon isotope discrimination and nitrogen content for riparian trees along elevational transects. Oecologia, **109**, 362-367.*

*Thomas, P.A., El-Barghathi, M., & Polwart, A. (2007) Biological Flora of the British Isles: Juniperus communis L. Journal Of Ecology, **95**, 1404-1440.*

*Wegner, C., Wunderlich, M., & Kessler, M. (2003) Foliar C:N ratio of ferns along an Andean Elevational Gradient. Biotropica, **35**, 486-490.*

*Welker, J.M., Molau, U., Parsons, A.N., Robinson, C.H., & Wookey, P.A. (1997) Responses of Dryas octopetala to ITEX environmental manipulations: a synthesis with circumpolar comparisons. Global Change Biology, **3**, 61-73.*

Intraspecific functional variability: on the relative importance and structure of intra *versus* interspecific variability

Cécile Hélène Albert[1, a], Wilfried Thuiller[1, b], Nigel Gilles Yoccoz[3, c], Anne Dubuis[4, d], Antoine Guisan[4, e], Rolland Douzet[2, f], Serge Aubert[2, g], Sandra Lavorel [1, h]

1. Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France
2. Station Alpine Joseph Fourier, CNRS UMS 2925, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France
3. Department of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics, University of Tromsø, N-9037 Tromsø, Norway
4. Department of Ecology & Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

Corresponding author: Cécile H. ALBERT
cecile.albert@m4x.org

Abstract

Defining ecological strategies by identifying leading axes of variation among species has been central in Ecology to simplify the complexity of the plant world. Positioning species along such axes requires the selection of key functional traits to describe, group and rank species according to their functions. Candidate traits should be robust, *i.e.* vary more between than within species. This crucial hypothesis has rarely been empirically tested whereas it seems essential to disentangle the relative importance of the intra- and interspecific components of the overall trait variability. Using an environmentally stratified sampling design in two alpine regions we tested the relative importance and structure of intra versus interspecific functional variability for thirteen plant species with contrasted life histories and traits. Using robust and innovative single- and multi-trait statistical approaches, we showed a non-negligible contribution of intraspecific variability (~30%) to the overall functional traits variability. Intraspecific variability did not affect the definition of functional strategies, although the strong intraspecific variability led to a strong overlap between species trait distributions within the trait space, resulting in a functional continuum rather than clear-cut classification. in conclusion, we bring here new evidence that intraspecific functional variability should not be considered as negligible, supporting individual rather than population-level comparisons. We further suggest that species should be represented by trait distributions instead of simple non-informative averages.

Key-words: Plant functional traits, leaf traits, interspecific and intraspecific variability, alpine ecosystems, environmental gradients, linear mixed models, functional strategies, species ranking

Introduction

The long-lasting quest for general principles simplifying the diversity and complexity of the plant world (van Humboldt and Bonpland 1807) has led to the definition of plant functional strategies describing and predicting how plants will respond to different environmental conditions. The definition of such ecological strategies - e.g. Grime's C.S.R. (competitive, stress-tolerant, ruderal) triangle (Grime 1977)- is based on the determination of leading axes of variation (Westoby et al. 2002) and specialization among species, like the major resource economics trade-off (Grime et al. 1997, Diaz et al. 2004). The empirical application of such a framework requires the selection of a set of key functional traits - any measurable features at the individual level affecting its fitness directly or not (Violle et al. 2007) – that can be used to describe and group species according to their functions, either in terms of response to or effects on their biotic and abiotic environments (Westoby 1998, Weiher et al. 1999). The selected traits should express the species positions along leading axes of variation and should be more variable between than within species to allow consistent and meaningful species rankings (Garnier et al. 2001, Roche et al. 2004, Al Haj Khaled et al. 2005).Functional traits have been used extensively during the last decade for several empirical and theoretical purposes: to establish functional classifications of organisms (Weiher et al. 1999); to quantify the functional diversity of communities (de Bello et al. 2009); to identify relationships between ecosystem functioning and its associated goods and services (Diaz et al. 2007); and to parameterize plant functional types in landscape and global vegetation models (Albert et al. 2008) or plant individuals in trait-based models (Norberg et al. 2001, Savage et al. 2007).

in spite of the increasing use of functional traits in various ecological fields, the essential assumption of robustness has often remained implicit or expert-based while there have been very few empirical investigations comparing intra- and interspecific trait variability (but see, Wilson et al. 1999, Garnier et al. 2001, Roche et al. 2004). Macroecological studies, which have focused on the way species mean traits are driven by general trade-offs (Wright et al. 2005) or environmental gradients (Thuiller et al. 2004), have generally implicitly ignored intraspecific variability. in contrast, studies which have focused on functional, genetic, or demographic intraspecific variability in response to environmental gradients, have generally been conducted for a single species (Cordell et al. 1998, Ryser and Aeschlimann 1999) or considering mean traits per treatment or population (Meziane and Shipley 1999, Shipley 2000).

There are now increasing evidence that intraspecific functional variability does matter, with significant effects on community dynamics and ecosystem functioning (Boege and Dirzo 2004, Lecerf and Chauvet 2008). It is barely known, however, (1) whether this affects general trade-offs among species traits, usually established on mean values, and (2) how individual phenotypes are distributed around these means (but see, Reich et al. 1999, but see, Fonseca et al. 2000, Westoby et al. 2002).

Here we quantified the relative importance and structure of the intra- and interspecific functional trait variability and the impact of using individual measures on the definition of functional strategies within alpine ecosystems. These ecosystems are characterized by steep climatic gradients over small spatial scales that constrain species functional make-up (Körner 1999). We specifically designed a hierarchical sampling strategy on thirteen common species over steep alpine climatic gradients in two study sites to address the following questions: (i) How is functional trait variability structured at the intra- and interspecific levels? (ii) How robust are the well-established interspecific functional trade-offs when considering intraspecific variability? (iii) Do the same trade-offs apply within species? (iv) Can individuals be classified by their traits when intraspecific functional variability is accounted for?

Material & methods

Field sites

Two valleys were selected for this study. The first one called GUISANE (6°21' to 6°40' E; 44°5' to 45°4' N) hereafter is located in the central French Alps (1200 - 2600 m). The second one called ANZEINDAZ (6°60' to 7°10' E; 46°10' to 46°30' N) hereafter is located in the western Swiss Alps (1076 - 2313 m), north west of the Lake Geneva near Montreux. GUISANE (resp. ANZEINDAZ) is characterized by mean annual temperatures running from 0°C to 6.3°C (resp. -3.5°C to 8°C) and annual precipitations from 600 mm to 1200 mm (resp. 1300 mm to 2300 mm), on calcareous or schist (resp. calcareous) substrates mainly. Both sites encompass a heterogeneous range of representative mountain vegetation types (e.g. coniferous forests, shrub heaths, subalpine grasslands and alpine meadows).

Species selection

We selected thirteen common shrub, grass and forb species that could be encountered over contrasted conditions along each of

the two valleys. These species cover a large set of plant families which prevent phylogenetic influences (11 families) and were assumed to represent a large proportion of the vegetation type diversity of the selected alpine valleys. To represent herbaceous species we selected four graminoïds (FP: *Festuca paniculata* Schinz & Thell., DG: *Dactylis glomerata* L. *, SC: *Sesleria caerulea* (L.) Ard *, CS: *Carex sempervirens* Vill.*) known to exemplify contrasted ecological strategies (Gross *et al.* 2007, Quétier *et al.* 2007), four forbs (PV: *Polygonum viviparum* L., GM: *Geum montanum* L., SN: *Silene nutans* L., LV: *Leucanthemum vulgare* Lam.) and one legume species (TA: *Trifolium alpinum* L.). Shrub species were represented by two dwarf shrubs from high altitude (DO: *Dryas octopetala* L.*, SH: *Salix herbacea* L.), one evergreen subalpine shrub (RF: *Rhododendron ferrugineum* L.) and one deciduous shrub (VM: *Vaccinium myrtillus* L.*). All species were sampled in GUISANE and only species marked with an asterisk were sampled in ANZEINDAZ.

Sampling strategy

We first stratified the two sites along two orthogonal interpolated topo-climatic variables from meteorological models: data from the meteorological model AURELHY at 50x50 m (Benichou and Le Breton 1987) in GUISANE (winter temperature and solar radiation in August) and data from WSL at 25x25 m (Zimmermann and Kienast 1999) in ANZEINDAZ (mean January temperature and solar radiation in August). We then followed a hierarchical stratified sampling combined with known occurrences of each species (GUISANE: phytosociological dataset from the National Alpine Botanical Conservatory; ANZEINDAZ: Swiss Floristic Center database, http://www.crsf.ch) to select sampling plots covering at best environmental heterogeneity (from 7 to 17 plots per species).

Within each plot we measured nine randomly selected sexually mature and unshaded individuals (tussocks or ramets). Non senescent, non grazed, non frost-damaged leaves were collected for trait measurements, one leaf per individual for herbaceous species and ten per individuals for shrub species (means per individual were then used). We tracked the growing season according to altitude, aspect and field observations in order to sample all individuals of a given species at a similar phenological stage.

Trait selection and measurements

Five functional traits reflecting the ecological strategies of the studied species were measured. (1) Maximum vegetative height (Hmax) is the distance between the top of photosynthetic tissue and the ground, which is associated with plant competitive vigour and tends to be allometrically correlated with aboveground biomass (Cornelissen et al. 2003). (2) Specific leaf area (SLA) is the one-sided area of a fresh leaf divided by its oven-dry mass; SLA is usually well correlated with relative growth rate (Cornelissen et al. 2003). (3) Leaf dry matter content (LDMC) is the oven-dried mass of a leaf divided by its water-saturated fresh mass (Cornelissen et al. 2003), here expressed in $mg.g^{-1}$; it was measured using the partial rehydration method, which has been proven to give similar results that the full rehydration method (Vaieretti et al. 2007); LDMC can be used as a proxy for plant growth rate and carbon assimilation. (4) Leaf nitrogen content (LNC) is the total amount of nitrogen per unit of dry leaf mass (in $mg.g^{-1}$), which is closely linked to the mass-based maximum photosynthetic rate (Cornelissen et al. 2003). (5) Leaf carbon content (LCC) is the total amount of carbon per unit of leaf dry mass (in $mg.g^{-1}$). For both LNC and LCC, dried and marble-ground leaves samples of 3-5 mg were analyzed with FlashEA 1112 elemental analyzer (Thermo Fisher Scientific inc. made in Milan, Italia) at the individual level or with three individuals pooled together for species with too small leaves (*C. sempervirens*, *L. vulgare*, *P. viviparum*, *S. herbacea*, *S. caerulea*, *S. nutans*, *T. alpinum).*

Statistical analysis

Hierarchy in the variance sources for each measured trait

in order to decompose the variability of each trait according to different levels of variation (interspecific, intraspecific between plots and intraspecific within plots), we used linear mixed models (Bolker et al. 2009), which are appropriate for representing hierarchical data. Models included either no fixed effects (written as fixed~1, m0) or a species fixed effect (fixed~ Species, m1) and a random intercept plot effect (random ~ 1| Plot). Estimated standard deviations at plot level (σ) were used to determine the percentage of variability attributed to each level (interspecific, intraspecific between plots and intraspecific within plots). As plots were represented as a random effect, we determined the between-plot part of variance with σ^2 divided by the overall variability. As species were represented as a fixed effect, we calculated a measure of explained variation based on the variances at plot-level for the different models following (Xu 2003):

$R^2 = 1 - \frac{(\sigma_{m1})^2}{(\sigma_{m0})^2}$, where σ_{m0} and σ_{m1} are the estimated error standard deviation at plot level estimated under models m1 and m0 respectively. Variances were estimated by maximizing the restricted log-likelihood (REML).

in order to compare the two study sites and assess whether the variance structures were robust independently of the life form considered, we performed this variance decomposition for each study site and also for data reduced to herbaceous species.

Intra- and interspecific functional structures

To decompose the overall variance across the five traits and determine the main axes discriminating species in the trait space and the main axes discriminating individuals within each species in the trait space, we used between and within species principal component analyses (PCA, Dolédec and Chessel 1991). Between PCA assesses the correlation structure based on groups means and find linear combinations of variables maximizing the between groups variance. Within PCA decomposes the structure within groups, by considering the data centered on group means. The inertia calculated in each of these analyses represents the part of the variance due to between and within group levels respectively (Dolédec and Chessel 1991).

As different numbers of plots – and thus individuals - had been sampled for each species, we used a re-sampling procedure to balance the datasets (equal number of individuals). First we determined the smallest number (SN) of individuals measured within species (SN=63). Then we sub-sampled the original dataset by drawing randomly one thousand balanced data subsets with SN individuals per species, and ran the within and between analyses over these data subsets. The resulting distributions of between- and within parts of the variance were tight enough to assume that the re-sampling procedure did not affect variance decomposition. Finally within each species, the variance decomposition within and between plots was assessed through between PCAs.

In addition we run PCA on data at the individual level, data aggregated by plots and data aggregated by species to observe whether the structures were maintained at these different levels.

Illustrations of overlaps among species in the multi-trait space

a. Attributing a given leaf to the correct species from its trait values

We determined the probability of correctly attributing a given leaf to the correct species. To do so, we performed a linear discriminant analysis (LDA), a statistical method used to find the linear combination of continuous explanatory variables best separating two or more classes of a categorical variable (Venables and Ripley 2002). It is close to between PCA but maximizes the ratio of the between groups variance to the total, and can be projected onto new observations.

For each leaf of herbaceous and shrub species measured in GUISANE (1374 leaves), we calibrated a LDA with the five leaf functional traits (LDMC, SLA, LNC, LCC) on a dataset containing all the other measured leaves. The calibrated LDA was then used to predict to which species a given leaf belongs to. Finally we calculated the probability P_{ij} to attribute a leaf from the species *i* to the species *j*. The same analysis was performed on herbaceous and shrub species common to both sites, the LDA

being calibrated on the GUISANE dataset and projected on the ANZEINDAZ dataset.

b. Functional distance between individuals

To determine whether a given individual was more distant to individuals from other species than from its own within the trait space, we calculated the Mahalanobis distance (Mahalanobis 1936) between all pairs of individuals (herbaceous species) with the five traits collected, thus accounting for correlations between traits.

All statistical analyses were computed with R 2.9.1 (R Development Core Team 2008) libraries: nlme, stats, ade4, MASS)

Results

Hierarchy in the variance sources for each measured trait

Our field data covered a large range of values for each trait (e.g. 90% of LDMC values between 150 and 470 $mg.g^{-1}$) due to large interspecific differences. They also revealed a large intraspecific variability depending on both traits and species. This intraspecific variability led to a strong overlap between species trait distributions resulting in a continuum of trait values over the full data set (Fig. 1).

Variance decompositions were similar between the whole datasets and the restricted herbaceous datasets (Fig. 2), with a partition around 70/30% for inter/intraspecific variation. These decompositions were relatively similar across the two sites, even if the interspecific part was always slightly lower (particularly for LNC) in ANZEINDAZ due to a smaller set of less contrasted species in comparison to GUISANE. in particular in ANZEINDAZ, two species among the five measured shared similar trait values (*C. sempervirens* and *S. caerulea*). In contrast these variance decompositions differed across traits (around 80 / 20 % for Hmax, LDMC and SLA, around 65 / 35 % for LNC). The intraspecific part of the variance was always above 20 % and reached up to 70 % for LNC in ANZEINDAZ (Fig. 2).

When variance decompositions between and within-species were recalculated from the mean traits by plot, instead of individual values, the between species part of the variance increased by around 13 % (not shown), leading approximately to a 85 / 15 % partition (between/within).

Intra- and interspecific functional structures

Quantification: between and within variance

When the five traits were considered together as a global functional strategy, the repeated variance decompositions using within and between PCA analyses led to very narrow distributions of intra- and interspecific contributions (interspecific: mean=0.70 and sd=0.007, intraspecific : mean=0.30 and sd=0.007). This result showed that individuals have equivalent contributions to the variance, and implied that on average 70% (resp. 60%) of the inertia was due to interspecific differences (always significant *p-value* <0.001) and 30% (resp. 40%) was due to intraspecific variability in the GUISANE dataset (resp. ANZEINDAZ). The intraspecific part of the variance was divided into between-plots (on average 47%) and within plots variance (on average 53%, Fig. 3).

The same analysis performed over the data aggregated at plot level indicated a higher part of the variance due to interspecific differences (86%, not shown).

Main axes of variation between species

The variance decompositions through intra- and interspecific analyses allowed the identification of the main axes structuring the dataset. The interspecific structure consisted in a strong first axis (53% of the variance) mainly explained by LDMC (positively correlated) and SLA (negatively correlated), and a second axis (29% of the variance) explained by Hmax (positively correlated) (Fig. 4a). This general interspecific structure was stable when the analysis was restricted to herbaceous and/or shrub species alone. Moreover, simple PCAs performed at the individual, plot and species levels presented similar main axes of variations.

The location of species centres of gravity within the trait space differed between grass, forb and shrub species and also between fast-growing exploitative species (*D. glomerata*, *L. vulgare*) and slow-growing conservative species (*F. paniculata*, *S. caerulea*, *C. sempervirens*, Fig. 5b). After the addition of individual data, the strategies observed at the species level were conserved, while overlaps between the ellipses encompassing individuals from each species (using 65% of each species individuals) increased due to the noise added by individual variability.

Main axes of variation within species

The within species structure presented no clear general pattern (Fig. 4b & c, axis 1: 31% of the variance, axis 2: 27%, axis 3: 20%). The PCA conducted on each species separately showed that the functional structures were not fully consistent across species: the first axis (from 33 to 42% of the variance) was mainly rougly driven by SLA and LDMC for all species, but the second axis was driven by LNC and LCC and/or Hmax and the correlations between these traits and SLA and LDMC were not consistent (Fig. 4b & c and 5a & b).

Illustrations of overlaps among species in the multi-trait space

Attributing a given leaf to the correct species from its trait values

After calibration and projection of the LDA over the GUISANE dataset, the probabilities to attribute given leaves to the correct species were high overall (above 70% for 10 species; min=36%), but there were also important misclassifications (Fig. 6). For example *T. alpinum* which is the only legume, characterized by high and slightly variable LNC, was well predicted (96%), while *S. caerulea* which is strongly variable for LDMC and LNC was predicted accurately at only 36%. Adding Hmax in the LDA only marginally improved the prediction of *S. caerulea* (45%).

When the LDA was calibrated with the GUISANE dataset and projected over the ANZEINDAZ one, similar results were also obtained: 100% of leaves correctly attributed for *D. octopetala* and only 14% for *S. caerulea.*

Functional distances between individuals

The Mahalanobis distances between pairs of individuals in the trait space can be represented as a density of Mahalanobis distances between individuals from a given species and individuals from other species. Individuals with extreme values within a given species could be closer to individuals from other species than to individuals from their own species (Fig. 7). Two contrasted cases were observed: either the species was very different from all other species (e.g. *F. paniculata*) and only few marginal individuals were similar to marginal individuals from other species; or a major number of individuals were functionally very close to those of another species (e.g. *C. sempervirens* close to *S. caerulea* and *G. montanum*). For instance 95% of the *C. sempervirens* individuals were below a 7.7 distance threshold to other *C. sempervirens* individuals. In contrast, 71% of *S. caerulea* individuals were also below the 7.7 distance threshold to *C. sempervirens* individuals, as were 62% of *G. montanum* individuals, but only 25% of the *F. paniculata* individuals and 16% of the *L. vulgare*.

Discussion

Robustness of key assumptions in functional trait analyses

Relative importance of inter- versus intraspecific on the overall traits variability

A key assumption behind the use of functional traits to rank species according to particular functions is that intraspecific variability in trait values is negligible in comparison to interspecific variability (Garnier et al. 2001, Roche et al. 2004, Al Haj Khaled et al. 2005). Our results challenge this assumption. intraspecific variability remained generally lower than interspecific variability but not negligible, as both single- and multi-trait analyses identified a relatively strong component of intraspecific variability (around 30% of the total variability). The relative importance of inter- and intraspecific variability obviously strongly depends on the species considered, however we specifically focused on contrasted species within the alpine flora in order to maximize the interspecific variability. Deciding whether intraspecific variability is negligible or not will therefore depend on the studied system and the selected traits and species. intraspecific functional variability is however clearly under-estimated when traits are averaged at the population level rather than used at the individual level.

Interspecific functional strategies are robust to intraspecific variability…

At the interspecific level our results corroborate trade-offs in resource economics - expressed by the negative correlation between SLA and LDMC- described in the literature (Grime et al. 1997, Diaz et al. 2004, Gross et al. 2007). The second main axis of variation was given by plant height, consistently with former results (Diaz et al. 2004, Gross et al. 2007). This recurring pattern was obtained for both study sites and for the different life forms. Therefore, functional trade-offs and strategies are robust to individual-level measurements: the observed general patterns in the trait space are similar at the species, population and individual level. However, intraspecific functional strategies and trade-offs were less pronounced than between species. Consequently individual measurements add an unexplained noise to the observed patterns leading to overlaps between species' positions in the trait space.

…, but ordination and thus ranking might not be

In our analysis plots spread in the trait space around the species means (Fig. 5a) and individuals spread around the plot means, leading to overlaps between the scatters of points defined by individuals of each species. The mechanisms underlying the spread of individuals around plot means cannot be answered in this study as this would require genetic analyses and/or transplant experiments to estimate the respective roles of plasticity and genetic adaptation. The way plots are spread around the species means is not structured by the environmental gradients directly (Fig. 5a) but probably more by species environmental preferences (Thuiller et al. in press). Consequently a study in a given set of environmental conditions (e.g. cold conditions) may not lead exactly to the same ordination of species in the trait space than a study in other environmental conditions (e.g. hot conditions). These potential changes in species ordination in the trait space could lead to changes in species ranking with respect to a single trait. These observations challenge former studies on species rankings with respect to a single trait (Garnier et al. 2001, Roche et al. 2004), and suggest that species ordination should be established for specific environmental conditions and not using species trait means measured on populations in significantly different conditions.

Functional continuum and consequences for empirical and theoretical functional ecology

Attributing species to functional strategies according to the similarity of their trait values and therefore functions requires that a given species belongs to a single strategy. However, measurements of traits at the individual level lead to a continuous distribution of traits rather than to a discrete one. In particular, a large number of individuals from a given species can be functionally closer to individuals from other species than from its own species and the scatter of points defined by individuals from each species are largely overlapping in the trait space. Moreover individuals are not very well attributed to their species from their leaf traits. On this basis it appears hazardous (1) to describe a species by a unique fixed mean trait value and (2) to classify species into functional groups described by unique fixed mean trait values (mean of the species means within the functional group or mean values for a characteristic species). Our results thus suggest replacing the unique fixed mean trait values by the use of trait value distributions -as a more realistic depiction of nature- as for example through the use of trait-environment response curves or random distributions around the means, as has already been done for quantifying functional diversity (e.g. Mouillot et al. 2005, Leps et al. 2006). This is of primary importance in vegetation modelling (Norberg et al. 2001, Kleidon et al. 2007), as progress on this issue is essential to link the dynamics of individuals and plant communities to ecosystem functioning (Diaz et al. 2007, Suding et al. 2008).

Conclusion

All together, our results demonstrate a non-negligible contribution of intraspecific variability to the overall functional traits variability. Intraspecific variability did not affect the definition of functional strategies, although the strong intraspecific variability led to a strong overlap between species trait distributions within the trait space, resulting in a functional continuum rather than clear-cut classification.

We can thus distinguish two different methods for different research questions concerning the importance of intraspecific functional variability:

1. To describe functional strategies when there are broad contrasts across broad environmental gradients and/or across species, e.g. from different life forms, it is not necessary to include intraspecific variability, as suggested by previous literature. 2. In contrast, considering close species or populations along functional axes of variation at the local scale can lead to overlaps among species (individuals are not necessarily functionally distinct between close species). In such cases, studies must use trait values measured in the relevant environmental conditions and not in contrasted or unknown conditions. in addition, analyses should preferentially use individual measurements and trait syndromes at individual rather than, as commonly done, at population or species level.

Acknowledgments

We would like to acknowledge the Ecrins National Park, the National Alpine Botanical Conservatory and the Joseph Fourier Alpine Station for their help and the data they provided. We also would like to thank I. Boulangeat, L. Chalmandrier, B. Couchaud, G. Girard, J. Icard, D. Lejon, A. Soudan, P. Saccone, F. Boucher and D. Paulin for their help in the field and in the lab and L. Gallien for helpful comments on the ms. This work was funded by ANR QDiv (ANR-05-BDIV-009) and DIVERSITALP (ANR-07-BDIV-014) projects. We also received support from European Commission's FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, N° 066866 GOCE) projects.

References

*Al Haj Khaled, R., M. Duru, J. P. Theau, S. Plantureux, and P. Cruz. 2005. Variation in leaf traits through seasons and N-availability levels and its consequences for ranking grassland species. Journal of Vegetation Science **16**:391-398.*

Albert, C. H., W. Thuiller, S. Lavorel, I. D. Davies, and E. Garbolino. 2008. Land-use change and subalpine tree dynamics: colonization of Larix decidua in French subalpine grasslands. Journal of Applied Ecology:659-669.

*Benichou, P., and O. Le Breton. 1987. Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. La Météorologie **7**:23-24.*

*Boege, K., and R. Dirzo. 2004. Intraspecific variation in growth, defense and herbivory in Dialium guianense (Caesalpiniaceae) mediated by edaphic heterogeneity. Plant Ecology **175**:59-69.*

*Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J. S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution **24**:127-135.*

*Cordell, S., G. Goldstein, D. Mueller-Dombois, D. Webb, and P. M. Vitousek. 1998. Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. Oecologia **113**:188-196.*

*Cornelissen, J. H. C., S. Lavorel, E. Garnier, S. Díaz, N. Buchmann, D. E. Gurvich, P. B. Reich, H. ter Steege, H. D. Morgan, M. G. A. van der Heijden, J. G. Pausas, and H. Poorter. 2003. Handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany **51**:335-308.*

*de Bello, F., W. Thuiller, J. Leps, P. Choler, J. C. Clement, P. Macek, M. T. Sebastia, and S. Lavorel. 2009. Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. Journal of Vegetation Science **20**:475-486.*

*Diaz, S., J. G. Hodgson, K. Thompson, M. Cabido, J. H. C. Cornelissen, A. Jalili, G. Montserrat-Marti, J. P. Grime, F. Zarrinkamar, Y. Asri, S. R. Band, S. Basconcelo, P. Castro-Diez, G. Funes, B. Hamzehee, M. Khoshnevi, N. Perez-Harguindeguy, M. C. Perez-Rontome, F. A. Shirvany, F. Vendramini, S. Yazdani, R. Abbas-Azimi, A. Bogaard, S. Boustani, M. Charles, M. Dehghan, L. de Torres-Espuny, V. Falczuk, J. Guerrero-Campo, A. Hynd, G. Jones, E. Kowsary, F. Kazemi-Saeed, M. Maestro-Martinez, A. Romo-Diez, S. Shaw, B. Siavash, P. Villar-Salvador, and M. R. Zak. 2004. The plant traits that drive ecosystems: Evidence from three continents. Journal of Vegetation Science **15**:295-304.*

*Diaz, S., S. Lavorel, F. de Bello, F. Quetier, K. Grigulis, and M. Robson. 2007. incorporating plant functional diversity effects in ecosystem service assessments. Proceedings of The National Academy of Sciences of The United States of America **104**:20684-20689.*

*Dolédec, S., and D. Chessel. 1991. Recent developments in linear ordination methods for environmental sciences. Advances in Ecology, India **1**:133-155.*

*Fonseca, C. R., J. M. Overton, B. Collins, and M. Westoby. 2000. Shifts in trait-combinations along rainfall and phosphorus gradients. Journal of Ecology **88**:964-977.*

*Garnier, E., G. Laurent, A. Bellmann, S. Debain, P. Berthelier, B. Ducout, C. Roumet, and M.-L. Navas. 2001. Consistency of species ranking based on functional leaf traits. New Phytologist **152**:69-83.*

*Grime, J. P. 1977. Evidence for existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. American Naturalist **111**:1169-1194.*

*Grime, J. P., K. Thompson, R. Hunt, J. G. Hodgson, J. H. C. Cornelissen, I. H. Rorison, G. A. F. Hendry, T. W. Ashenden, A. P. Askew, S. R. Band, R. E. Booth, C. C. Bossard, B. D. Campbell, J. E. L. Cooper, A. W. Davison, P. L. Gupta, W. Hall, D. W. Hand, M. A. Hannah, S. H. Hillier, D. J. Hodkinson, A. Jalili, Z. Liu, J. M. L. Mackey, N. Matthews, M. A. Mowforth, A. M. Neal, R. J. Reader, K. Reiling, W. RossFraser, R. E. Spencer, F. Sutton, D. E. Tasker, P. C. Thorpe, and J. Whitehouse. 1997. Integrated screening validates primary axes of specialisation in plants. Oikos **79**:259-281.*

*Gross, N., K. N. Suding, S. Lavorel, and C. Roumet. 2007. Complementarity as a mechanism of coexistence between functional groups of grasses. Journal of Ecology **95**:1296-1305.*

*Kleidon, A., K. Fraedrich, and C. Low. 2007. Multiple steady-states in the terrestrial atmosphere-biosphere system: a result of a discrete vegetation classification? Biogeosciences **4**:707-714.*

Körner, C. 1999. Alpine Plant Life. Springer-Verlag, Berlin.

*Lecerf, A., and E. Chauvet. 2008. Intraspecific variability in leaf traits strongly affects alder leaf decomposition in a stream. Basic and Applied Ecology **9**:598-605.*

*Leps, J., F. de Bello, S. Lavorel, and S. Berman. 2006. Quantifying and interpreting functional diversity of natural communities: practical considerations matter. Preslia **78**:481-501.*

*Mahalanobis, P. C. 1936. On the generalised distance in statistics. Proceedings of the National institute of Sciences of India **2**:49-55.*

*Meziane, D., and B. Shipley. 1999. interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. Plant Cell and Environment **22**:447-459.*

*Mouillot, D., W. Stubbs, M. Faure, O. Dumay, J. A. Tomasini, J. B. Wilson, and T. D. Chi. 2005. Niche overlap estimates based on quantitative functional traits: a new family of non-parametric indices. Oecologia **145**:345-353.*

*Norberg, J., D. P. Swaney, J. Dushoff, J. Lin, R. Casagrandi, and S. A. Levin. 2001. Phenotypic diversity and ecosystem functioning in changing environments: A theoretical framework. Proceedings of The National Academy of Sciences of The United States of America **98**:11376-11381.*

R.Development.Core.Team., editor. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

*Reich, P. B., D. S. Ellsworth, M. B. Walters, J. M. Vose, C. Gresham, J. C. Vollin, and W. D. Bowman. 1999. Generality of leaf trait relationships: a test across six biomes. Ecology **80**:1955-1969.*

*Roche, P., N. Diaz-Burlinson, and S. Gachet. 2004. Congruency analysis of species ranking based on leaf traits: which traits are the more reliable? Plant Ecology **174**:37-48.*

*Ryser, P., and U. Aeschlimann. 1999. Proportional dry-matter content as an underlying trait for the variation in the relative growth rate among 22 Eurasian populations of Dactylis glomerata. Functional Ecology **13**:473-482.*

*Savage, V. M., C. T. Webb, and J. Norberg. 2007. A general multi-trait-based framework for studying the effects of biodiversity on ecosystem functioning. Journal of Theoretical Biology **247**:213-229.*

*Shipley, B. 2000. Plasticity in relative growth rate and its components following a change in irradiance. Plant Cell and Environment **23**:1207-1216.*

*Suding, K. N., S. Lavorel, F. S. Chapin, J. H. C. Cornelissen, S. Diaz, E. Garnier, D. Goldberg, D. U. Hooper, S. T. Jackson, and M. L. Navas. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. Global Change Biology **14**:1125-1140.*

Thuiller, W., C. H. Albert, A. Dubuis, C. Randin, and A. Guisan. in press. Variation in habitat suitability models do not always relate to variation in species' plant functional traits. biology letters.

*Thuiller, W., S. Lavorel, G. Midgley, S. Lavergne, and T. Rebelo. 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 Leucadendron taxa. Ecology **85**:1688-1699.*

*Vaieretti, M. V., S. Diaz, D. Vile, and E. Garnier. 2007. Two measurement methods of leaf dry matter content produce similar results in a broad range of species. Annals of Botany **99**:955-958.*

van Humboldt, A., and A. Bonpland. 1807. Essai sur la géographie des plantes, Paris.

Venables, W. N., and B. D. Ripley. 2002. Modern Applied Statistics with S, Fourth edition edition.

*Violle, C., M. L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier. 2007. Let the concept of trait be functional! Oikos **116**:882-892.*

*Weiher, E., A. van der Werf, K. Thompson, M. Roderick, E. Garnier, and O. Eriksson. 1999. Challenging Theophrastus: A common core list of plant traits for functional ecology. Journal of Vegetation Science **10**:609-620.*

*Westoby, M. 1998. A leaf-heigh-seed (LHS) plant ecology strategy scheme. Plant and Soil **199**:213-227.*

*Westoby, M., D. S. Falster, A. T. Moles, P. A. Vesk, and I. J. Wright. 2002. Plant ecological strategies: Some leading dimensions of variation between species. Annual Review of Ecology and Systematics **33**:125-159.*

*Wilson, P. J., K. Thompson, and J. G. Hodgson. 1999. Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. New Phytologist **143**:155-162.*

*Wright, I. J., P. B. Reich, J. H. C. Cornelissen, D. S. Falster, E. Garnier, K. Hikosaka, B. B. Lamont, W. Lee, J. Oleksyn, N. Osada, H. Poorter, R. Villar, D. I. Warton, and M. Westoby. 2005. Assessing the generality of global leaf trait relationships. New Phytologist **166**:485-496.*

*Xu, R. H. 2003. Measuring explained variation in linear mixed effects models. Statistics in Medicine **22**:3527-3541.*

*Zimmermann, N. E., and F. Kienast. 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. Journal of Vegetation Science **10**:469-482.*

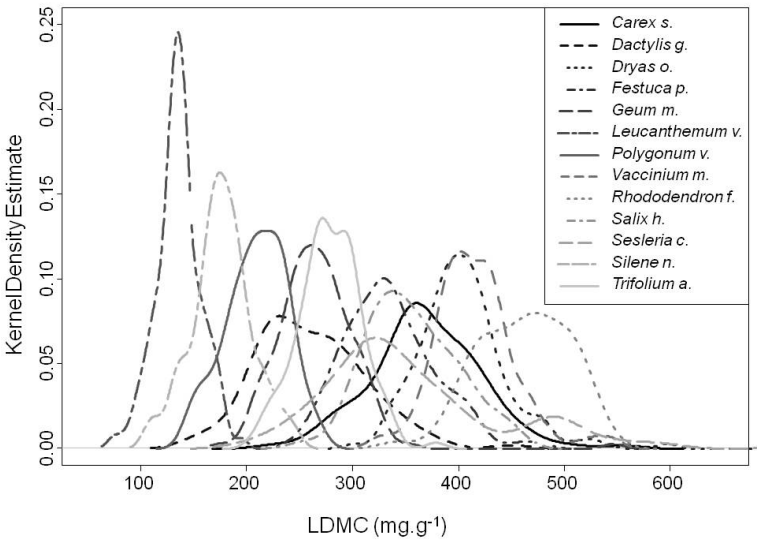


Fig. 1 Distributions of the leaf dry matter content among species

Distribution of leaf dry matter content for the thirteen studied species using a kernel density estimation. It is a non-parametric way of estimating the probability density function of a random variable.

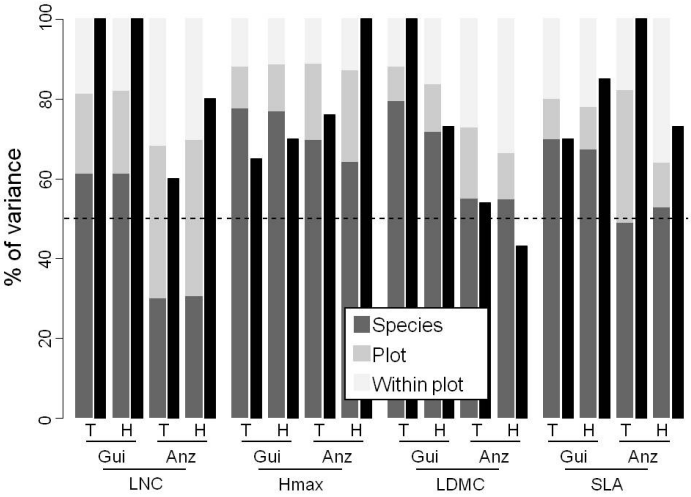
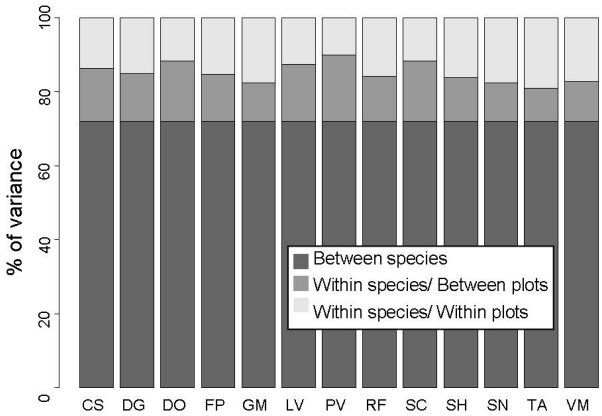


Fig. 2 Variance decomposition in inter- and intraspecific contributions for each trait

Variance decomposition for the four selected traits given as percentages for the different levels: species, plots and within plots (increasingly dark shades of grey). The decomposition is given for both sites (GUISANE: Gui and ANZEINDAZ: Anz) and with all the species (T) or only herbaceous species (H). The black bars are the observed variance given as % of the highest variance observed per trait.

Fig. 3 Variance decomposition in inter- and intraspecific contributions for all the traits together

Variance decomposition resulting from between PCA analyses on the whole GUISANE dataset, taking into account the five traits together. The decomposition gives the percentages of variability due to the three levels of the analysis: between species, within species-between plots and within species-within plots (grey levels). The decomposition into between and within species variability was estimated using a single between PCA on the whole dataset. Within species variability was decomposed into between plots and within plots using between PCA on each species data.



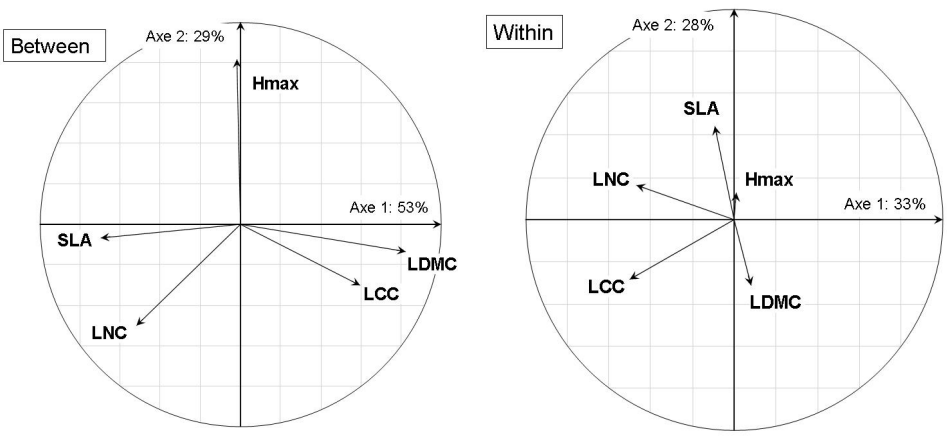
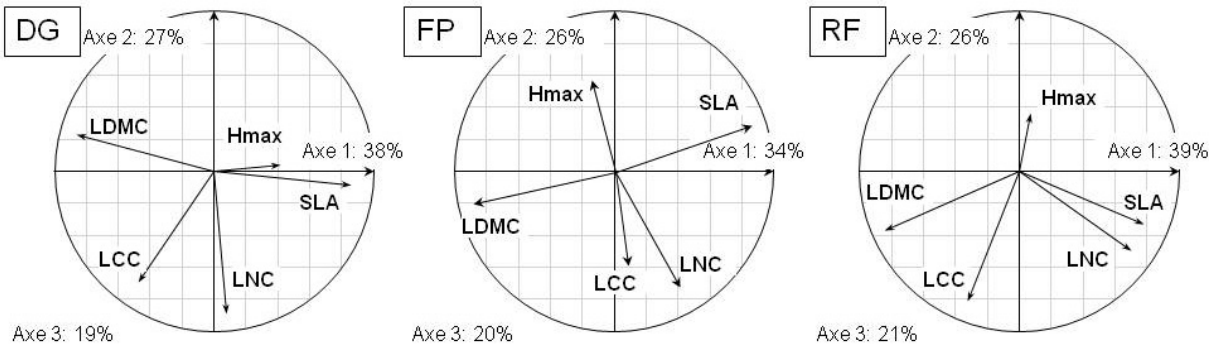


Fig. 4 Multidimensional structure within the trait space: intra- and interspecific trade-offs

a. Inter-specific: correlation circles between the five measured traits and the two first PCA axes, obtained from the between and within PCA. As our dataset contained different numbers of individuals per species, we used a multi re-sampling procedure to perform analyses on balanced datasets (equal number of lines in each group). The correlation circles are the average correlations obtained after a thousand simulations.

b. intra-specific: correlations circles between the five measured traits within three species (DG= D. glomerata, FP= F. paniculata, RF= R. ferrugineum) and the two first PCA axes. PCAs were performed on the data representing each species. For each axis the percentage of inertia is presented.



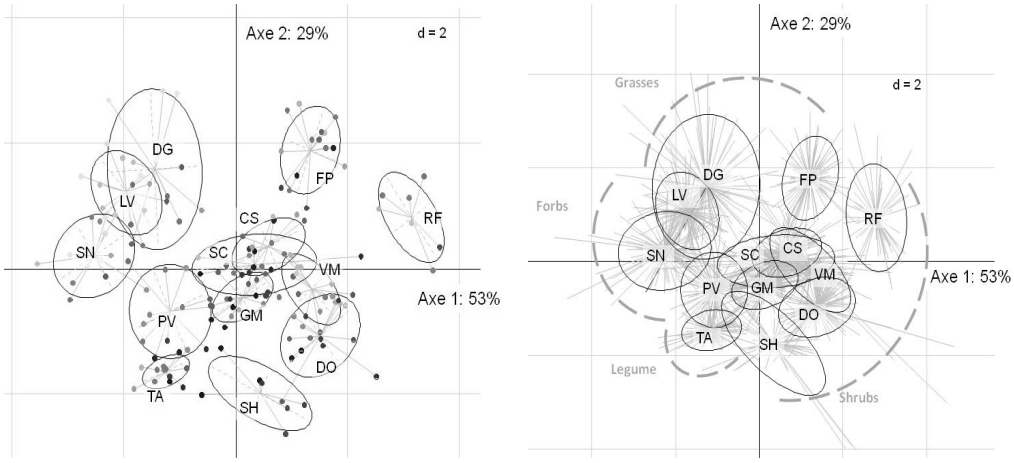


Fig. 5 Dispersion of species, plots and individuals in the trait space

Dispersion of the different levels (species, plots and individuals) of the study in the trait space. The ellipsoids of inertia are encompassing 65% of the individuals of each species.

a. Projection of the species centers of gravity and the different plots in the trait space (first and second main axes) defined by the between PCA. Stars are centered on the species means. Mean traits of plots are represented by dots, whose color depends on the temperature (Tminh) of the plot. The darker dots represent the coldest plots and the lighter dots the hottest plots.

b. Projection of species centers of gravity and all individuals in the trait space (first and second main axes) defined through the between PCA. The grey half circles define the different life forms and their place in the trait space.

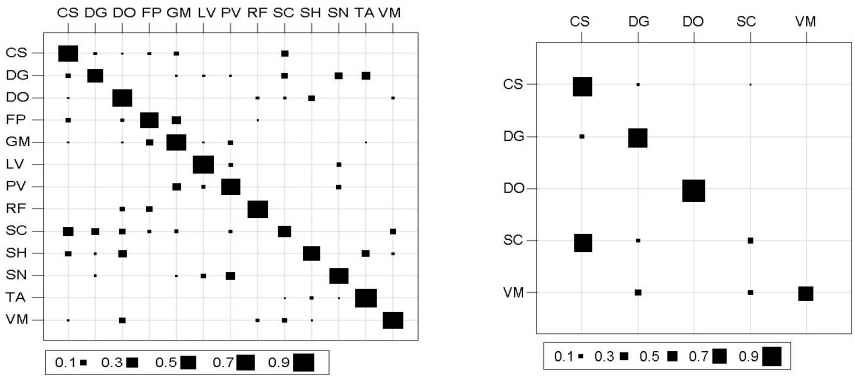


Fig. 6 Probabilities to attribute a given leaf to the different species from its traits

The values of the probabilities to attribute a given leaf to a species from its functional traits are given. Predicted species are presented in columns and observed species in lines. Results are given for the thirteen species. The black squares are proportional to the probabilities. All squares that are outside the diagonal are errors of prediction.

a. LDA calibrated and projected on the whole GUISANE dataset.

b. LDA calibrated on the common species of the GUISANE dataset and projected on the ANZEINDAZ dataset.

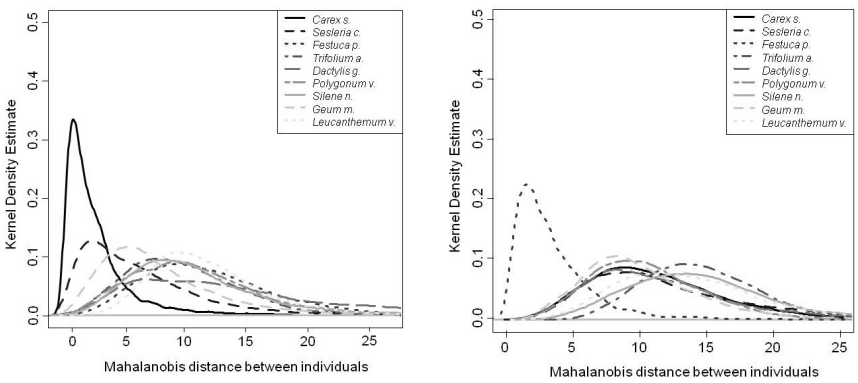


Fig. 7 Functional distances between individuals

Distribution of the Mahalanobis functional distance between measured individuals using a kernel density function. It is a non-parametric way of estimating the probability density function of a random variable. a. Distribution of the distance between the individuals of Carex sempervivens and individuals from the other species. b. Distribution of the distance between the individuals of Festuca paniculata and individuals from the other species.

Sampling in ecology and evolution– Avoiding pitfalls by integrating questions and designs

Cécile H. Albert[1a], Nigel G. Yoccoz[2b], Thomas C. Edwards, Jr.[3c], Catherine H. Graham[4d], Niklaus E. Zimmermann[5e], Wilfried Thuiller[1f]

1. Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France

2. Department of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics, University of Tromsø, N-9037 Tromsø, Norway

3. U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah 84322-5290 USA

4. Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794 USA

5. Land Use Dynamics, Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland

Corresponding author: Cécile H. ALBERT cecile.albert@m4x.org

Summary

Sampling is a key issue when dealing with most of ecological and evolutionary questions. The importance of a sound sampling design tailored to the raised question has already been discussed field by field in the ecological literature and the sampling literature has provided useful tools and recommendations to sample and analyse ecological data. However given the increasing trend toward analysing large and complex datasets, it remains unclear how much data quantity is worth compared to data quality. Our objective is thus to draw attention on the main shortcomings resulting from inappropriate samplings and to highlight the importance of sampling designs for understanding patterns of biological properties and processes generating these patterns. Using a semi-virtual experiment as an illustration we discuss major sampling issues that have been neglected so far. We reveal how the selection of the space (e.g. geographic, climatic), in which the sampling is designed, influences the patterns that can be ultimately detected. We also demonstrate the inefficiency of common sampling designs to reveal response curves between ecological variables and climatic gradients. Consequently these results show that the sampling theory mainly focusing on determining unbiased and precise estimates of means and variances might be inappropriate to provide good estimates of response curves. All ecological questions dealing with response of organisms (e.g. demography, functional traits, and abundance) to environmental heterogeneity and environmental changes are however focused on such curves or model estimations. We finally discuss the use of virtual experiments and model-based designs to help defining appropriate sampling designs so that they can effectively fulfil requirements of both sampling theory and field work logistics.

***Keywords**: semi-virtual simulations, model-based sampling, sampling space, sampling frame, sample size, sampling design, inference, response curve*

Introduction –principles and definitions of sampling

Knowledge of how the environment influences living organisms, their distribution, their abundance, their interactions, and how the underlying evolutionary processes affect these patterns is at the core of ecological and evolutionary sciences. Most research in this domain can be reduced to four principal questions – ***Why*** is a particular response or pattern observed? ***Where*** in geographical, temporal or evolutionary space does the response or pattern occur? ***How much*** of a response or element does exist? ***How is*** the response or pattern ***changing*** over time or in space? Answers to these questions are obtained by measuring relevant biological variables on items of ecological or evolutionary interest in the system under study (Fig. 1). It is, however, rarely possible to measure these variables on all possible items – called sampling units – relevant to the questions being posed. Instead, we employ sampling designs to select a subset of the sampling units – the sample – from which we obtain the measurements necessary to answer any of the four general questions (Fig. 2 & Appendix 1). The sample itself is drawn from the set of known and measurable sampling units - the sampling frame - using procedures called sampling designs (Fig. 2 & Appendix 1).

There are two main types of sampling designs - probabilistic and non-probabilistic - that strongly differ in terms of estimation methodology and reliability (Lemeshow and Levy 1999). In a probabilistic design, every sampling unit from the frame has a known non-zero probability of selection, resulting in unbiased estimates of the mean and the variance for the variables of interest. In contrast, some sampling units in a non-probabilistic design have no chance of being selected from the frame, or their selection probability cannot be accurately determined and is therefore ignored at the estimation stage. This can lead to biased estimates and may result in erroneous conclusions (Lemeshow and Levy 1999). Statistical theory emphasizes that the choice of a sampling method -which requires a definition of the units and frame, sampling design, and sampling effort- influences what patterns and processes are revealed in subsequent analyses, the confidence in the resulting estimates, and the validity of extrapolation to other systems or time periods (Lemeshow and Levy 1999, Olsen et al. 1999, Skalski et al. 2005). It is not only the quantity, but also the quality of the data that matters (Björklund and Bergek 2009, Kelling et al. 2009).

The importance of a sound sampling design tailored to the question is well known in ecological studies. For example, the field of species distribution modelling, a ***where*** question (Fig. 1), has clearly benefited from groundwork laid by Austin and others (Austin 1987, Austin et al. 2006) demonstrating the importance of sampling across environmental gradients. More recently, Hirzel and Guisan (2002) discussed different sampling designs for species distribution modelling. Edwards *et al.* (2006) showed the peril of ignoring sampling issues; radically different model results and spatial patterns were obtained depending on sampling design. In other ecological fields, the importance of sound sampling designs has also emerged when dealing with sparsely distributed organisms of conservation concern (e.g., lichens, Edwards et al. 2004; butterflies, Haddad et al. 2008), or more common organisms of commercial and ecological concern (e.g., forests, Schreuder et al. 1999; fishes, Courbois et al. 2008, Hughes and Peck 2008), or of mixed ecological and public health concern (e.g., radioactivity monitoring, Scott et al. 2008).

There is no doubt that ecologists and evolutionary biologists face considerable sampling problems, but applying appropriate sampling theories to field research is essential, even if often difficult to implement. Unfortunately, it appears that many scientists are unaware of or deliberately ignore the diversity of statistical sampling theories available for resolving technical and logistic dilemmas common to field studies. This has often led to the implementation of so-called convenience sampling designs (Schreuder et al. 1999, Rosenstock et al. 2002), such as following transects or roads (e.g., US Breeding Bird survey, Peterjohn 1994, Bart et al. 1995), which, like all non-probabilistic designs, suffer from unknown biases. In contrast, sampling theory generally focuses on the unbiased estimation of simple parameters such means or variances that may be inappropriate to fully address any of the four key questions described above. For example, ***how*** and ***how much is it changing*** questions rely on estimating response curves or the magnitude of changes using sophisticated statistical modelling (Guisan and Thuiller 2005). Extending the use of classical sampling principles on the estimation of means and variances to the estimation of such response curves is not straightforward.

Sampling issues are and have always been an important aspect of ecological study. However, there is no universal sampling design for all research circumstances, and an appropriate sampling design for one question may not necessarily be appropriate for other questions. For these reasons, and in order to reconcile sampling theory and ecological field sampling, we suggest a comprehensive framework for applied sampling in ecological sciences. We motivate this discussion by:

(1) Providing a synthetic review of the different sampling shortcomings emerging from the ecological literature. We focus this review on two types of shortcomings, study extent and sample representativeness, and the crucial choices associated with delineation of the sampling frame, and ecological space and gradients.

(2) Identifying important links among different sampling shortcomings and the selection of the sampling space. We also exemplify how the sampling design may influence our ability to effectively address the hypotheses (i.e., patterns) generated from these data, or test hypotheses about causal factors generating these patterns. We end with a discussion of how sampling decisions also affect our ability to interpret species response curves.

(3) Discussing the use of virtual experiments and model-based designs to define appropriate sampling designs.

Pitfalls in sampling designs

We do not intend to present an exhaustive list of problems but rather to depict the key shortcomings resulting from inappropriate sampling designs. We illustrate this with few striking examples from the literature. The examples come from studies that generally led to interesting new results, but remained difficult to generalize because of debatable sampling choices. In some cases, comparably little effort would be needed to clearly improve sampling choices and resulting analyses and inferences.

1. Study extent and representativeness

There is a strong extent-intensity trade-off ranging from intensive surveys of one particular site or biological population to sparsely sampled broad overviews of an entire region or continent (Olsen et al. 1999). Selecting the extent of the study and intensity of sampling is crucial as they both influence the representativeness of the target population by the selected frame.

Lack of generalization: The intensive small-extent surveys with few or no repetitions often lead to accurate local results that are difficult to generalize to broader extents. It is indeed not statistically correct to draw any generalization from few intensive local surveys, like for example, extrapolating the consequences of environmental change on species using demographic studies of one or two contrasting populations (Charmantier et al. 2008, Grosbois et al. 2008). Although such studies based on long-term series are particularly strong to uncover unknown and complex mechanisms, their specific context (e.g. evolutionary history, past population demography, land use, disturbance regimes) tends to prevent any extrapolation to the whole species or similar group of species. As important these studies are, they clearly need to be supplemented by simpler but larger scale studies based on relevant designs to validate the proposed revealed mechanisms (Bråthen et al. 2007). In such large scale studies, there should be more populations surveyed and populations should be selected with a proper sampling design among all the possible populations of the species.

Incomplete or biased information: In contrast to small-extent surveys, large extent surveys often rely on rough *a priori* selected gradients (e.g. latitude or temperature in Garcia et al. 2000, Lecerf and Chauvet 2008) often poorly informative at local scales, since they jointly subsume many underlying factors that are locally or physiologically important (e.g., soil, pollution, drought). Making inference from a given gradient however assumes that all other involved gradients vary independently of the one of interest (Freedman 2006), which is unlikely in studies dealing with few sites spread at continental or global scales. Moreover, irrespective of the scale, environmental gradients should be sampled and thus represented in a balanced manner. For instance, when studying the response of a species along a temperature gradient, high temperatures should not be over-represented compared to low temperatures (Wright et al. 2005, Cingolani et al. 2007). A representative and balanced design ensures the quality and robustness of response curves obtained in a given analyses, whereas an unbalanced design (e.g. a large amount of points in one part of the curve) may seriously hamper detection of the true response or at least reduce the precision of estimates. Sampling several gradients at a time increases the need for careful consideration of design since some combinations of gradients may be very rare and need to be targeted carefully when sampling.

Truncated gradient: Both the definitions of the study extent and of the sampling intensity are crucial and should aim at encompassing the whole extent of the response shape of a given biological variable along the gradients of interest. For instance, estimating accurately the shape of a biological pattern – such as the response of a functional trait to an environmental gradient, which in many cases will be non-linear – requires sampling across the entire range of variation of both the environmental and biological variables of interest (Mohler 1983, Austin 1987, Thuiller et al. in press). Truncated gradients can lead to erroneous statistical inference, which in turn can affect ecological interpretation. Works by Van Horne (2002) and Thuiller *et al.* (2004) illustrate how incorrect interpretations could be drawn if only a part of an important environmental gradient is sampled to estimate the niche of a species. This may occur, e.g., when using administrative instead of natural boundaries for delineating the study extent. The resulting response curves of a species might appear truncated – possibly expressing a negative (e.g., on the colder part of the temperature gradient), a positive (e.g., on the warmer part of the temperature gradient) or no relationship (e.g., on the intermediate part of the temperature gradient) – where the full response actually is unimodal. Such truncated sampling obviously leads to spurious predictions if the response is used to extrapolate a species distribution in space or time (Thuiller et al. 2004).

2. Frame, spaces and gradients: a crucial choice.

There exists no universal sampling design for all biological studies. The selection of the sampling frame and of the environmental gradients implied in the sampling process should be made with regard to the biological patterns of interest. Obviously, it would not be efficient to sample based on soil stratification if the aim was to determine a species response to a temperature gradient. This example seems trivial on the surface, but selecting the appropriate frame, space and gradients is not always so evident. While the sampling is done in one space (e.g. geographical), the pattern of interest might occur in a different space (e.g. climatic, Figure a. in Box 1). In such a case, the gradient related to the response is not sampled in a balanced way (see *Incomplete or biased information)*. As an example, climatic influence on species distributions, abundances or demography may result in patterns in the geographical space induced by spatial autocorrelation of the climate or dispersal limitation (Araujo et al. 2009).

Indirect gradients: Latitudinal and altitudinal gradients are extensively used in ecological studies (Cordell et al. 1998, Wegner et al. 2003, Luo et al. 2005). They are, however, indirect gradients (*sensu* Austin 1987) of unknown or complex combinations of multiple environmental gradients that directly impact organisms (e.g. temperature, nutrient and resource availability, photoperiod). Consequently, indirect gradients are hard to interpret and are usually not transferable in time and space. For example, an altitudinal increase of 1000 m does not have the same impact on environmental conditions in the Rocky Mountains and on Mount Kenya (Körner 2007, McCain 2009). Indirect gradients may nonetheless be pertinent for descriptive studies (Rahbek 1995, Rickart 2001), but not as surrogates of climatic gradients to derive reliable statistical inferences (Körner 2007).

A posteriori selection of additional gradients: The selection of gradients of interest should be done before the sampling, so that it is part of the design and therefore allows for an effective, unbiased sampling. Gradients that have not been included in the sampling design could be used as co-variables, or sometimes to post-stratify the sample, but *a posteriori* changes in frame and gradients being studied can lead to *incomplete or biased information*, to *truncated gradients*, and are unlikely to result in efficient designs (Santamaria 2003, Wang and Gao 2004, Cingolani et al. 2007).

Avoiding pitfall by integrating questions and designs

1. A story of sampling spaces

The virtual sampling experiment presented in Box 1 illustrates the importance of the selection of the appropriate sampling space. The sampling frame, here defined as the set of all geographical pixels in the study site, can be represented in different spaces (Fig. 1 in Hutchinson 1957): geographic space (e.g. latitude, longitude), topographic space (e.g. elevation), climatic space (e.g. temperature, precipitations), and biological space (e.g. habitat suitability). In our example the sampling frame was defined in the geographic space. From this frame, one can generate a set of sampling units that are not redundant in this space because one geographical location corresponds to only one possible unit. However, when projected into the climatic space, the frame looks less representative than in the geographic space, with a large part of the sampling units representing hot and dry conditions and only few of them representing other climatic conditions. Clearly, switching from one space to another can have consequences on the sampling (Box 1, Fig. a). For instance, sampling along roads leads to a complete coverage of the latitudinal gradient, but to an incomplete coverage of the altitudinal and temperature gradients, with the full omission of the cold mountainous areas that are far from roads. The transfer of a set of points from one space to another is predetermined by the choice of sampling frame and units. If done inappropriately, such choices may result in data that suffer from *truncated gradients problems*. Such big discrepancies resulting from the transfers from one space to another also explain why analyses based on an *a posteriori selection of additional gradients* or on *indirect gradients* may not be powerful for drawing strong inferences about a given biological variable.

2. Influence of the sampling design: a matter of quality or quantity

How the sampling design may influence the patterns revealed from the sampled data is rarely discussed in ecological studies where the focus is generally made on the estimation of simple parameters such as population averages, size or measures of diversity (e.g., Yoccoz et al. 2001, Edwards et al. 2006). But as our virtual experiment illustrates, the method used to implement the sampling can influence the estimates by affecting the selection probabilities of the sampling units (Appendix 2). Sampling probabilities are determined analytically in the case of probabilistic designs, and evaluated numerically in the case of non-probabilistic designs, and they can be both displayed in the geographical space. The three example designs (following the roads, stratified along altitude, and model-based design, Box 1, Fig. a), all lead to very low but non-null probabilities and to very distinct probability patterns. This result illustrates that among the three designs, the sampling units differ in selection probabilities, which has severe consequences for the final parameter estimates: the same sample resulting from two different sampling designs will not result in identical parameter estimates (Lemeshow and Levy 1999, Appendix 2). .

Second, our virtual experiment strongly illustrates that the efficiency of the sampling design depends on the question and the parameters to be estimated (e.g. mean *vs* optimum, Box 1 & Appendix 2). To be reliable, a sampling design should lead to parameter estimates that are both unbiased and precise. For the estimations of both the mean and the Euclidean distance between the 'true optimum' and the optimum of the estimated response, the different designs led to very different results in terms of bias and precision (Box 1 & Appendix 2). In both cases sample size had a strong (and expected) effect on the precision of estimates; the biggest samples resulted in more precise estimates. The estimate of the mean was closer to truth in our virtual experiment with data drawn from the simple random and the systematic climatic designs than from any other sampling design. The estimated mean exhibited a much higher variability when using an altitudinal stratified random and model-based design; it was underestimated in the design following roads because cold areas (higher values) were less frequently sampled (see *truncated gradients*). In summary, if the basic aim is to estimate the mean value of a biological variable over a study area, the simple random sampling design, one of the simplest probability-based designs, gave results that were closest to truth. The estimate of the variable's optimum was closer to the truth when calculated from the model-based design than from any other sampling design. Three designs, systematic climatic, simple random and altitudinal stratified random, resulted in less precise estimates of the optimum. The design following roads resulted in a very poor estimate of the response shape and its optimum (see *truncated gradients* & a *posteriori selection of additional gradient*). It is not surprising that the model-based design performed better than the other methods, as it represents an iterative learning method. The so-called surface response methodology, which is derived from model-based sampling, was developed to estimate the kind of response shapes used in our virtual example (Box and Hunter 1957, Wu and Hamada 2000).

In our virtual experiment, sample size was not the most influential factor in our ability to draw biological inference; indeed, sampling design had a stronger influence on the precision of parameter estimation than had sample size. This means that an appropriate, probability-based sampling design could reduce the sampling effort and costs associated with sampling, an important attribute given that resources for ecological investigation are generally limited.

3. Estimation of response curves

Several ecological questions, including those related to ***how much*** and ***how is changing***, rely on accurate estimation of response curves and surfaces (Austin 1985). Making unbiased and reliable estimations of model parameters and efficiently selecting the most appropriate models has not been sufficiently covered in the sampling literature. In fact, we think there is a need to find robust approaches and evaluation methods that will achieve goals similar to those of classical sampling – unbiased and precise parameter estimates - but in terms of statistical models. Such models, which are approximations to the truth, are of unknown form (e.g. multivariate, nonlinear and include interaction terms) before a given study is started. They depend on optimization criteria such as AIC or cross-validated measures of predictive ability to be selected appropriately (Hastie et al. 2001, Burnham and Anderson 2004). Moreover estimating relationships between the biological or ecological variable of interest (e.g. growth, functional traits, abundance) and geographical or climatic gradients or time requires that the sampling frame encompasses the entire variable response curve along a given gradient, from its optimum to its edges (see *truncated gradients*). For example, to study the influence of temperature on a bird species' demography, it is necessary that the sampling frame includes populations from both the core and edge of its distribution along the temperature gradient. Our virtual experiment indicates that a sampling procedure that effectively captures variation along a gradient can be created using the surface response methodology (Box and Hunter 1957, Wu and Hamada 2000) from which we derived our model-based design. However, this sampling method has rarely been used in ecology (Inouye 2001), probably because, as Bayesian approaches in statistics, this methodology relies on *a priori* knowledge or assumptions on the patterns of interest.

Conclusion

Having *a priori* knowledge and a detailed hypothesis of how the main variables of the system will impact the variables of interest is key for developing sampling designs that lead to reliable, representative estimation of parameters of interest, not just falsifiable predictions as is often emphasized. The importance of

efficient sampling designs has been emphasized before, but we suggest that we lack a strategy linking sampling principles (particularly probability sampling and stratification) and the crucial role played by modeling in ecology and evolution. We note that *a priori* knowledge of the system is a challenging task (Clark 2005), but in many cases it will lead to the selection of an appropriate sampling design. There are potentially three ways to generate an efficient sampling design prior to start field work: (1) use expert knowledge, (2) conduct a preliminary descriptive study, and (3) design semi-virtual experiments. Researchers generally rely on expert knowledge, literature and past analyses when developing a sampling design. Preliminary descriptive analyses can also be used to develop model-based design. In this case, a very simple sampling scheme (e.g. random design) is conducted on the variable of interest and then the response between the environmental and biological variable is evaluated. Based on this preliminary knowledge, a comprehensive sampling design can be developed (Brito et al. 1999, Edwards et al. 2005, Guisan et al. 2006, Conroy et al. 2008, Williams et al. 2009). While the former methods have been used extensively in ecology, semi-virtual experiments seem to be rarely used for sampling even though they have already shown to be important tools to investigate ecological processes (Austin et al. 2006, Kraft et al. 2007, Zurell et al. 2009). In this study, the semi-virtual experiment allowed us to use real data to build different spaces in which the sampling was then performed. Different sampling designs can then be simulated based on initial hypotheses and including the costs of sampling, the difficulty to reach the sampling sites, and so on. While a semi-virtual experiment should be built with some expert knowledge, it provides an additional tool for developing sampling designs; one that we think has not been sufficiently exploited in ecology and evolution.

Acknowledgments

The idea for this manuscript arose at a workshop on the species' ecological niche, its stability and evolution at Riederalp, Switzerland (August 2008). CA and WT acknowledge support from the ANR Diversitalp (ANR-07-BDIV-014) and SCION (ANR-08-PXXX-003) projects. WT, NGY, TCE and NEZ received support from European Commission's FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, N° 066866 GOCE) project.

Appendices

They are available at http://www.oikos.ekol.lu.se/app.html

APPENDIX 1: Main definitions in sampling theory

Table S1. Sampling designs: probabilistic vs non-probabilistic

APPENDIX 2: A semi-virtual case study

Fig. S1. Study site and the virtual variable of interest
Fig. S2. Optimum of the variable, optimal conditions and distance between the true and the estimated optima

References

Araujo, M. B., Thuiller, W. and Yoccoz, N. G. 2009. Reopening the climate envelope reveals macroscale associations with climate in European birds. - Proceedings of The National Academy Of Sciences Of The United States Of America 106: E45-E46.

Austin, M. P. 1985. Continuum concept, ordination methods, and niche theory. - Ann. Rev. Ecol. Syst. 16: 39-61.

Austin, M. P. 1987. Models for the analysis of species' response to environment gradients. - Vegetatio 69: 35-45.

Austin, M. P., Belbin, L., Meyers, J. A., Doherty, M. D. and Luoto, M. 2006. Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. - Ecological Modelling 199: 197-216.

Bart, J., Hofschen, M. and Peterjohn, B. G. 1995. Reliability of the breeding bird survey: Effects of restricting surveys to roads. - Auk 112: 758-761.

Björklund, M. and Bergek, S. 2009. On the relationship between population differentiation and sampling effort: is more always better? - Oikos 118: 1127-1129.

Box, G. E. P. and Hunter, J. S. 1957. Multi-Factor Experimental-Designs For Exploring Response Surfaces. - Annals Of Mathematical Statistics 28: 195-241.

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.

Brito, J. C., Crespo, E. G. and Paulo, O. S. 1999. Modelling wildlife distributions: logistic multiple regression vs overlap analysis. - Ecography 22: 251-260.

Burnham, K. P. and Anderson, D. R. 2004. Multimodel inference - understanding AIC and BIC in model selection. - Sociological Methods & Research 33: 261-304.

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B. and Sheldon, B. C. 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. - Science 320: 800-803.

Cingolani, A. M., Cabido, M., Gurvich, D. E., Renison, D. and Diaz, S. 2007. Filtering processes in the assembly of plant communities: Are species presence and abundance driven by the same traits? - Journal of Vegetation Science 18: 911-920.

Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. - Ecology Letters 8: 2-14.

Conroy, M. J., Runge, J. P., Barker, R. J., Schofield, M. R. and Fonnesbeck, C. J. 2008. Efficient Estimation Of Abundance For Patchily Distributed Populations Via Two-Phase, Adaptive Sampling. - Ecology 89: 3362-3370.

Cordell, S., Goldstein, G., Mueller-Dombois, D., Webb, D. and Vitousek, P. M. 1998. Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. - Oecologia 113: 188-196.

Courbois, J.-Y., Katz, S. L., Isaak, D. J., Steel, E. A., Thurow, R. F., Wargo Rub, A. M., Olsen, T. and Jordan, C. E. 2008. Evaluating probability sampling strategies for estimating redd counts: an example with Chinook salmon (Oncorhynchus tshawytscha). - Canadian Journal of Fisheries and Aquatic Sciences 65: 1814-1830.

Edwards, T. C., Cutler, D. R., Geiser, L., Alegria, J. and McKenzie, D. 2004. Assessing rarity of species with low detectability: Lichens in Pacific Northwest forests. - Ecological Applications 14: 414-424.

Edwards, T. C., Cutler, D. R., Zimmermann, N. E., Geiser, L. and Alegria, J. 2005. Model-based stratifications for enhancing the detection of rare ecoligical events. - Ecology 86: 1081-1090.

Edwards, T. C., Cutler, D. R., Zimmermann, N. E., Geiser, L. and Moisen, G. G. 2006. Effects of underlying sample survey designs on the utility of classification tree models in ecology. - Ecological Modelling 199: 132-141.

Freedman, D. A. 2006. Statistical models for causation - What inferential leverage do they provide? - Evaluation review 30: 691-713.

Garcia, D., Zamora, R., Gomez, J. M., Jordano, P. and Hodar, J. A. 2000. Geographical variation in seed production, predation and abortion in Juniperus communis throughout its range in Europe. - Journal Of Ecology 88: 436-446.

Grosbois, V., Gimenez, O., Gaillard, J.-M., Pradel, R., Barbraud, C., Clobert, J., Møller, A. P. and Weimerskirch, H. 2008. Assessing the impact of climate variation on survival in vertebrate populations. - Biological Review 83: 357-399.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N. G., Lehmann, A. and Zimmermann, N. E. 2006. Using niche-based models to improve the sampling of rare species. - Conservation Biology 20: 501-511.

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

Haddad, N. M., Hudgens, B., Damiani, C., Gross, K., Kuefler, D. and Pollock, K. 2008. Determining optimal population monitoring for rare butterflies. - Conservation Biology 22: 929-940.

Hastie, T., Tibshirani, R. and Friedman, J. 2001. The elements of statistical learning. Data mining, inference and prediction. - Springer-Verlag.

Hirzel, A. H. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling. - Ecological Modelling 157: 331-341.

Hughes, R. M. and Peck, D. V. 2008. Acquiring data for large aquatic resource surveys: the art of compromise among species, logistics, and reality. - Journal of the North American Benthological Society 27: 837-859.

Hutchinson, G. E. 1957. Concluding remarks. - Cold Spring Harbor Symposia on Quantitative Biology 22: 145-159.

Inouye 2001. response surface experimental designs for investigating interspecific competition. - Ecology 82: 2696-2706.

Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G. and Hooker, G. 2009. Data-intensive Science: A New Paradigm for Biodiversity Studies. - Biosciences 59: 613-620.

Körner, C. 2007. The use of "altitude" in ecological research. - Trends in Ecology & Evolution 22: 570-574.

Kraft, N. J. B., Cornwell, W. K., Webb, C. O. and Ackerly, D. D. 2007. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. - American Naturalist 170: 271-283.

Lecerf, A. and Chauvet, E. 2008. Intraspecific variability in leaf traits strongly affects alder leaf decomposition in a stream. - Basic and Applied Ecology 9: 598-605.

Lemeshow, S. and Levy, P. S. 1999. Sampling of Populations: Methods and Applications. - Wiley-Interscience.

Luo, T., Luo, J. and Pan, Y. 2005. Leaf traits and associated ecosystem characteristics across subtropical and timberline forests in the Gongga Moutains, Eastern Tibetan Plateau. - Oecologia 142: 261-273.

McCain, C. M. 2009. Vertebrate range sizes indicate that mountains may be 'higher' in the tropics. - Ecology Letters 12: 550-560.

Mohler, C. L. 1983. Effect of sampling pattern on estimation of species distributions along gradients. - Vegetatio 54: 97-102.

Olsen, A. R., Sedransk, J., Edwards, D., Gotway, C. A., Liggett, W., Rathbun, S., Reckhow, K. H. and Young, L. J. 1999. Statistical issues for monitoring ecological and natural resources in the United States. - Environmental Monitoring and Assessment 54: 1-45.

Peterjohn, B. G. 1994. The North American breeding bird survey. - Journal of American Birding Association 26: 386-398.

Rahbek, C. 1995. The Elevational Gradient Of Species Richness - A Uniform Pattern. - Ecography 18: 200-205.

Rickart, E. A. 2001. Elevational diversity gradients, biogeography and the structure of montane mammal communities in the intermountain region of North America. - Global Ecology and Biogeography 10: 77-100.

Rosenstock, S. S., Andersonb, D. R., Giesenc, K. M., Leukeringd, T. and Carterd, M. F. 2002. Landbird counting techniques: Current practices and an alternative. - Auk 119: 46-53.

Santamaria, L. 2003. Plant performance across latitude: the role of plasticity and local adaptation in an aquatic plant. - Ecology 84.

Schreuder, H. T., Hansen, M. and Kohl, M. 1999. Relative costs and benefits of a continuous and a periodic forest inventory in Minesota. - Environmental Monitoring and Assessment 59: 135-144.

Scott, E. M., Dixon, P., Voigt, G. and Whicker, W. 2008. Choice and criteria for selection of sampling strategies in environmental radioactivity monitoring. - Applied Radiation and Isotopes 66: 1575-1581.

Skalski, J. R., Ryding, K. E. and Millspaugh, J. J. 2005. Wildlife demography: analysis of sex, age and count data.

Thuiller, W., Albert, C. H., Dubuis, A., Randin, C. and Guisan, A. in press. Variation in habitat suitability models does not always relate to variation in species' plant functional traits. - Biology Letters doi:10.1098/rsbl.2009.0669.

Thuiller, W., Brotons, L., Araujo, M. B. and Lavorel, S. 2004. Effects of restricting environmental range of data to project current and future species distributions. - Ecography 27: 165-172.

Van Horn, B. 2002. Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. - In: Scott, J. M., Heglund, P. J., Morrison, M. L., Haufler, J. B., Raphael, M. G., Wall, W. A. and Samson, F. B. (eds.), Predicting species occurrences: issues of accuracy and scale. Island Press.

Wang, R. and Gao, Q. 2004. Morphological responses of Leymus chinensis (Poaceae) to the large-scale climatic gradient along the North-east China Transect (NECT). - Diversity and Distributions 10: 65-73.

Wegner, C., Wunderlich, M. and Kessler, M. 2003. Foliar C:N ratio of ferns along an Andean Elevational Gradient. - Biotropica 35: 486-490.

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M. and Schwartz, M. W. 2009. Using species distribution models to predict new occurences for rare plants. - Diversity and Distributions 15: 565-576.

Wright, I. J., Reich, P. B., Cornelissen, J. H. C., Falster, D. S., Groom, P. K., Hikosaka, K., Lee, W., Lusk, C. H., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Warton, D. I. and Westoby, M. 2005. Modulation of leaf economic traits and trait relationships by climate. - Global Ecology And Biogeography 14: 411-421.

Wu, C. F. J. and Hamada, M. 2000. Experiments: Planning, Analysis, and Parameter Design Optimization.

Yoccoz, N. G., Nichols, J. D. and Boulinier, T. 2001. Monitoring of biological diversity in space and time. - Trends in Ecology & Evolution 16: 446-453.

Zurell, D., Jeltsch, F., Dormann, C. F. and Schroder, B. 2009. Static species distribution models in dynamically changing systems: how good can predictions really be? - Ecography 32: 733-744.

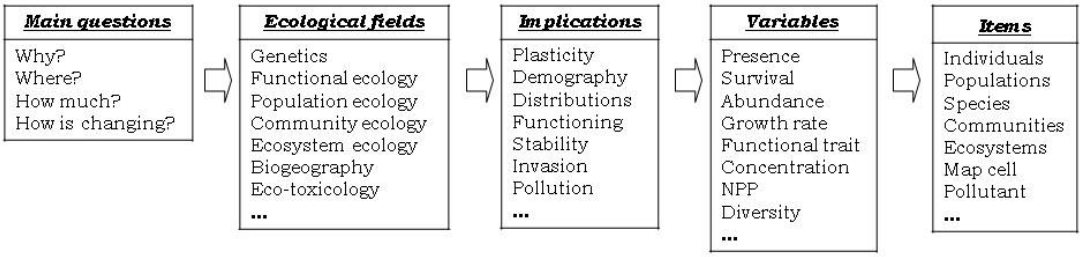


Fig. 1. Decomposing ecological questions There are four basic ecological questions that can be addressed in different ecological fields with different implications. These questions concern different kind of variables that are measurable on various possible items

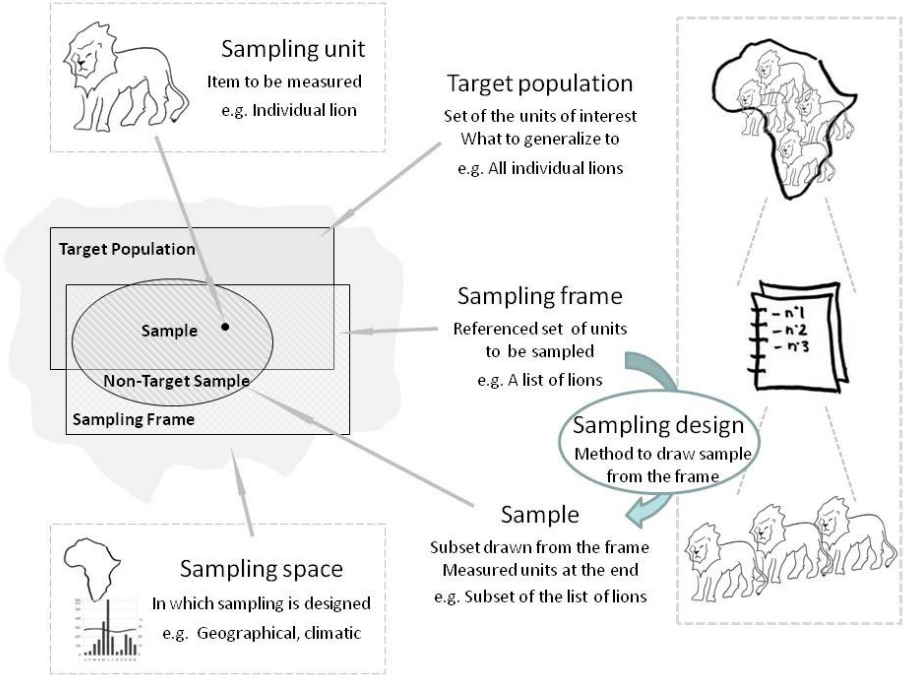


Fig.2. Sampling concepts and vocabulary

Box 1. A semi-virtual case study

1. Definition of the case study

Study site: 30,000 km^2 including the French Alps and the Rhône Valley.

Variable of interest: a virtual biological variable (e.g. growth rate, functional trait, allele frequency, abundance), constrained by temperature and precipitations, that varies between 0 and 100 and can be measured over the study area (Fig. a). It follows a Gaussian shape along both selected climatic gradients plus a random noise following a normal distribution (mean= 0, standard deviation= 4).

2. Virtual samplings

Sampling unit: pixel of the geographical map (1 x 1 km)

Sampling frame: the geographical map of the study site

Sampling designs: five different designs of increasing complexity reflecting common sampling designs in ecology (Appendix 1, Table S1): (i) simple random, (ii) following the main roads, (iii) stratified random along altitude, (iv) systematic in the climatic space (grid-based), and (v) a model-based design .

Sample size: 50, 200, and 1000 sampling units were drawn for each design.

Simulations: each combination of design x sample size was run a thousand times. For all designs except the last, the probability p_i to sample the unit i was determined analytically. We ran the model-based design 10 000 times to numerically determine the probability p_i of the unit i of being sampled.

3. Evaluation methods of the different samplings

Two biological parameters were estimated and compared to the 'true' values.

Mean: weighted mean variable (Horwitz-Thompson estimator, see Appendix 1); e.g. mean genetic diversity or mean functional trait in a landscape.

Optimum: optimal conditions in the environmental space for the biological variable (e.g. niche optimum). Assumed that the variable can be estimated using a polynomial model with linear and quadratic terms for temperature and precipitations.

$$Variable \approx temperature + temperature^2 + precipitations + precipitations^2 .$$

The Euclidean distance between the "true" and the estimated optimum in the environmental space was then calculated (see Appendix 2).

4. Results

Probability to sample a sampling unit: the probability patterns are very different between the different sampling designs leading to different probabilities for a given unit to be included in the sample (Fig. a) and then to different calculation for means and optimum (Fig. b).

Sample size: the larger the sample, the higher the precision in the estimations (Fig. b).

Sample designs: bias and precision of the estimates strongly depend on the sampling design. Selecting an appropriate design reduces the need of a larger sample. The best sampling design is not the same for both mean and optimum estimation, it as to be chosen in regards with the questions of the study (Fig. b).

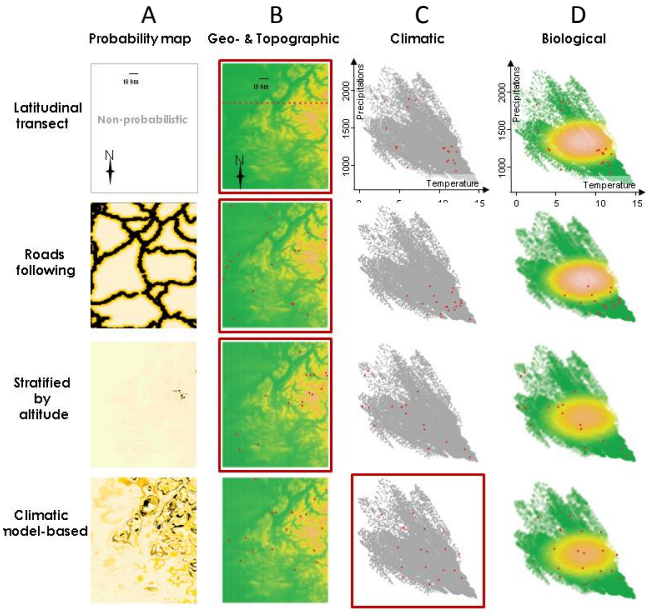


Fig. a. Importance of the sampling space onto which the sampling design is set up.

Different sampling designs: the red points represent the samples drawn from the frame through different sampling designs. From top to bottom, a latitudinal transect sampling (drawn from the frame within the geographic space), a distance to road sampling (drawn from the frame within the geographic space), an altitudinal equally stratified (drawn from the frame within the topographic space), a surface response sampling (drawn from the frame within the climatic space). The red boxes indicate in which space the sampling has been done. Axes have not been repeated on all figures, they are the same for all rows. ***Sampling probability (A):*** *patterns resulting from the three different sampling designs, i.e. for each unit of the frame (i.e. in this case each pixel from the map) its probability to be sampled for four contrasting sampling designs (the darker points represent the higher probabilities). These maps of probability are differing and will thus result in different weighting in follow-up calculations.* ***Different spaces (B, C, D):*** *from left to right the topographic space represented within the geographical map (green for low altitudes and brown for high altitudes), the climatic space represented by temperatures and precipitations and the biological space represented by the variable of interest mapped into the climatic space (green for low values and brown for high values).*

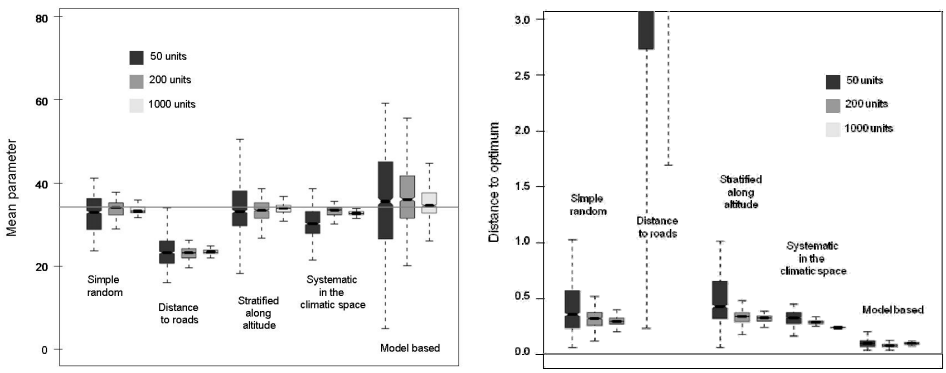


Fig. b. Effects of sampling designs and sample size on the precision of parameter estimations

Illustration of the sampling design effect on the precision and bias of the estimation of two biological variables: (a) the mean variable and (b) distance between the "true" optimum corresponding to environmental optimal conditions for the observed variable and the estimated one. Results are given for the five sampling designs: (i) simple random, (ii) following the main roads, (iii) stratified random along altitude, (iv) systematic in climatic space, and (v) a model-based design. The boxplots of the estimated values give the median, and the first and third quartiles of the estimates (mean and distance to optimum). Results are given for 50, 200 and 1000 sampling units. (A) Distribution of the estimations of the parameter mean. The grey line is the true value on the studied area. (B)Distribution of the Euclidian distances between the true optimum of the parameter and the one estimated from virtual measurements. The grey line corresponds to 0, the best value that can be obtained.

Appendix S1:

Main definitions in sampling theory

The basic unit to be measured is called a **sampling unit**. Units should be distinct and easy to define (e.g. individuals of the species *Panthera leo,* L.); however, there are many situations with sampling units not being readily identified. This is the case when studying species that do not have clear individuals (e.g. plants with vegetative growth) or when units are defined in terms of location or area (e.g. grid-cells of a map). In the later case, there is an infinite number of units that may or may not overlap.

The set containing all sampling units of interest is called the **target population** of the study; it is the set from which one wants to derive generalized patterns. This statistical definition of a population (e.g. all African lions) should not be confounded with the biological definition of population. A statistical target population can be infinite in size or at least uncountable and often only a small fraction of the target population is accessible for measurements.

Given that it is not possible to measure and even to know or list the entire population, it is generally necessary to define what is called a **sampling frame**, which is a finite list of all sampling units that could be measured. The sampling frame ideally coincides with the target population, but may differ for various reasons such as accessibility or size. The **sample**, a set of units to be measured, is drawn from the sampling frame.

There exist two main types of **sampling designs**: probabilistic and non-probabilistic (Table S1). In a probabilistic design, every sampling unit from the frame is given a non-null probability of being sampled, and this probability is known or defined. In a non-probabilistic design, some sampling units from the frame have no chance of being selected or their selection probability cannot be accurately determined. These two designs strongly differ in terms of estimation and reliability. Probability-based sampling leads to unbiased estimates of the mean of the variables of interest and of their variability. In contrast, non-probabilistic sampling is sensitive to unknown biases and may result in erroneous conclusions (Lemeshow and Levy 1999).

For a probabilistic design, the non-null probability to sample the ith sampling unit can be called p_i. The unbiased estimator of the mean variable for the target population is then: $\sum_i \frac{1}{p_i} \times Variable_i$, where $Variable_i$ is the variable measured on the ith sampling unit. Sampled sampling units that are less likely to be sampled are given a higher weight, as they are under-represented in the observed sample.

Sampling design	**Probability design**	**Definition**	**Examples / refs**
Simple Random	Yes	All units in the frame have the same probability of being selected	Biological invasions (Stohlgren et al. 2001)
Stratified Random	Yes	The frame is first divided in strata, which are then sampled using random sampling, often with strata-specific probabilities (more variable strata are sampled more intensively)	Land use (Austrheim et al. 2008) Fisheries management (Courbois et al. 2008)
Grid-Based	Yes (Random start) /No (non random start)	Usually in geographic space, each sample is selected based on regular coordinates. The start can be chosen at random or not.	Species distribution modelling (Teixeira et al. 2001, Barbosa et al. 2003)
Multi-stage - Hierarchical	Yes/No (depending on the use of probability sampling at different stages)	Different frames are selected at different stages (e.g. populations at the first stage, individuals at the second) or at different scales (e.g. selecting first large units, then smaller ones).	Functional traits (Wilson et al. 1999, Thuiller et al. in press) Forest resources (Opsomer et al. 2007).
Multi- phase – Adaptive/Learning	Yes	A first selection of units is expanded based on the presence of e.g. species./ New points are defined from the prior knowledge obtained from few first estimations/ The sample is expanded until a given precicion is reached.	Species distribution modelling (Brito et al. 1999) Search for rare species occurrences (Guisan et al. 2006, Williams et al. 2009) Estimation of abundance (Conroy et al. 2008).
Purposive/Targeted	No	Sampling units are selected to maximize contrasts …	Biological invasions (Howard et al. 2004) Biological conservation (Franklin et al. 1996).
Convenience	No	Sampling units are those available (eg literature or museum) or easily accessible	Bird survey along roads (Peterjohn 1994, Bart et al. 1995)

Table S1. Sampling designs: probabilistic vs non-probabilistic

Austrheim, G., Mysterud, A., Pedersen, B., Halvorsen, R., Hassel, K. and Evju, M. 2008. Large scale experimental effects of three levels of sheep densities on an alpine ecosystem. - Oikos 117: 837-846.

Barbosa, A. M., Real, R., Olivero, J. and Vargas, J. M. 2003. Otter (Lutra lutra) distribution modelling at two resolution scales suited to conservation planning in the Iberian Peninsula. - Biological Conservation 114: 377-387.

Bart, J., Hofschen, M. and Peterjohn, B. G. 1995. Reliability of the breeding bird survey: Effects of restricting surveys to roads. - Auk 112: 758-761.

Brito, J. C., Crespo, E. G. and Paulo, O. S. 1999. Modelling wildlife distributions: logistic multiple regression vs overlap analysis. - Ecography 22: 251-260.

Conroy, M. J., Runge, J. P., Barker, R. J., Schofield, M. R. and Fonnesbeck, C. J. 2008. Efficient Estimation Of Abundance For Patchily Distributed Populations Via Two-Phase, Adaptive Sampling. - Ecology 89: 3362-3370.

Courbois, J.-Y., Katz, S. L., Isaak, D. J., Steel, E. A., Thurow, R. F., Wargo Rub, A. M., Olsen, T. and Jordan, C. E. 2008. Evaluating probability sampling strategies for estimating redd counts: an example with Chinook salmon (Oncorhynchus tshawytscha). - Canadian Journal of Fisheries and Aquatic Sciences 65: 1814-1830.

Franklin, A. B., Anderson, D. R., Forsman, E. D., Burnham, K. P. and Wagner, F. W. 1996. Methods for collecting and analysing demographic data on the northern spotted owl. - Studies in Avian Biology 17: 12-20.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N. G., Lehmann, A. and Zimmermann, N. E. 2006. Using niche-based models to improve the sampling of rare species. - Conservation Biology 20: 501-511.

Howard, T. G., Gurevitch, J., Hyatt, L., Carreiro, M. and Lerdau, M. 2004. Forest invasibility in communities in southeastern New York. - Biological Invasions 6: 393-410.

Lemeshow, S. and Levy, P. S. 1999. Sampling of Populations: Methods and Applications. - Wiley-Interscience.

Opsomer, J. D., Breidt, F. J., Moisen, G. G. and Kauermann, G. 2007. Model-assisted estimation of forest resources with generalized additive models. - Journal Of The American Statistical Association 102: 400-409.

Peterjohn, B. G. 1994. The North American breeding bird survey. - Journal of American Birding Association 26: 386-398.

Stohlgren, T. J., Otsuki, Y., Villa, C. A., Lee, M. and Belnap, J. 2001. Patterns of plant invasions: a case example in native species hotspots and rare habitats. - Biological Invasions 3: 37-50.

Teixeira, J., Ferrand, N. and Arntzen, J. W. 2001. Biogeography of the golden-striped salamander Chioglossa lusitanica: a field survey and spatial modelling approach. - Ecography 24: 618-624.

Thuiller, W., Albert, C. H., Dubuis, A., Randin, C. and Guisan, A. in press. Variation in habitat suitability models does not always relate to variation in species' plant functional traits. - Biology Letters doi:10.1098/rsbl.2009.0669.

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M. and Schwartz, M. W. 2009. Using species distribution models to predict new occurences for rare plants. - Diversity and Distributions 15: 565-576.

Wilson, P. J., Thompson, K. and Hodgson, J. G. 1999. Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. - New Phytologist 143: 155-162.

Appendix S2:

A semi-virtual case study

To illustrate the consequences of sampling decisions and to highlight the fact that careful consideration improves inference, we developed a comprehensive semi-virtual example. Semi-virtual studies are useful because they depict the "true" pattern of interest and therefore can be used to test and compare sampling designs (Hirzel and Guisan 2002, Austin et al. 2006, Meynard and Quinn 2007). Our case study does not pretend to cover all possibilities but provides an example of how sampling design can influence inference. We selected a real study area with strong climate gradients that is not correlated with topography. This choice of a real landscape allows us to maintain realism in our simulations. The study area encompasses 30,000 km² located in south-eastern France. It includes part of the French Alps and the Rhône Valley (Fig. S1).

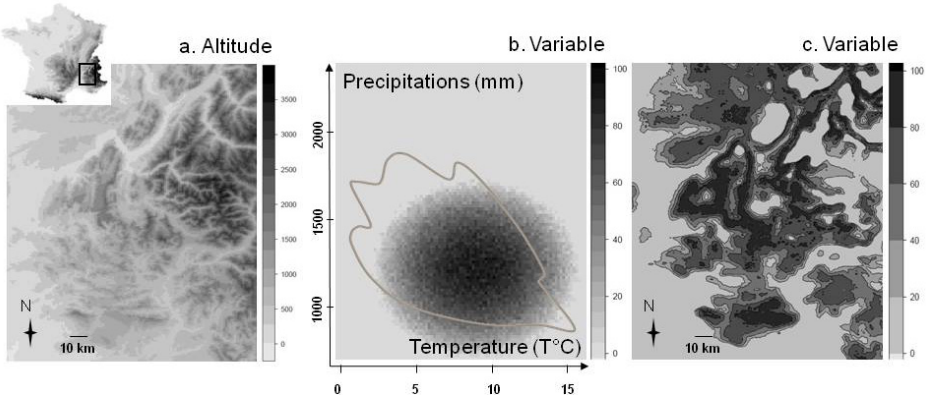


Fig. S1. Study site and the virtual variable of interest.

The study site is located in the south-east France

a. Altitudinal map of the study site including a part of the French Alps and of the Rhône valley.

b. Construction of the virtual biological variable as a Gaussian curve along both temperature and precipitations gradients plus a noise. The grey line represents the environmental conditions existing in the study area.

c. Map of the virtual ecological variable through the study area. The virtual variable ranges from 0 to 100 (e.g. habitat suitability, relative biomass, relative fecundity)

Sampling units were defined as pixels of the map (at 1×1 km) identified by their geographic coordinates (latitude and longitude). The sampling frame was the list of all those pixels. The area was further described by elevation data derived from the shuttle radar topography mission (SRTM3, http://www2.jpl.nasa.gov/srtm/) and climate. For simplicity, we selected two uncorrelated climatic gradients from the French National Climatic Network (*Aurhely*, (Benichou and Le Breton 1987): mean annual temperature and total annual precipitation.

We defined a virtual climatically-constrained biological variable that varies between 0 and 100 and can be measured over the study area. This variable could be growth rate, habitat suitability, abundance, survival, radioactivity concentration, net primary productivity, allele frequency, a functional trait, etc (Fig. S1.c). As many biological variables show an optimum at some intermediate environmental conditions and decline gradually towards the boundaries (Hirzel and Guisan 2002, Gaston 2003), we formulated a variable that followed a Gaussian shape along both selected climatic gradients. To add some realism, and because two gradients may not be sufficient to describe the environmental conditions that influence the distribution of a biological variable, we added a random noise following a normal distribution (mean=0, standard deviation=4) to simulate the other factors influencing the organism of interest (Mohler 1983, Hirzel and Guisan 2002). To avoid the problems of sampling-induced bias due to the detectability of sampling units – an important topic not evaluated in our simulation - we assumed that the variable of interest occurs everywhere and is stable over time (Wintle et al. 2005, Mackenzie 2006).

We tested the effects of pre-sampling choices on this virtual biological variable. Samples were drawn from the sampling frame following five different designs of increasing complexity and that reflect common sampling designs in ecology (Table S1): (i) simple random, (ii) following the main roads, (iii) stratified random along altitude, (iv) systematic in the climatic space (grid-based), and (v) a model-based design. In the fifth design, a sub-sample was drawn first by using a uniform sampling design in the climatic space (i.e. design iv). Here, the response to the two climatic gradients was first estimated using a polynomial model with linear and quadratic terms for temperature and precipitations on a first part of the sample:

$$Variable \approx temperature + temperature^2 + precipitations + precipitations^2$$

This prior, model-based knowledge of how the response generally varies along environmental gradients was used to define the final sampling (second part), which followed a central composite factorial design centered on the estimated optimal conditions (OC) and surface response methodology (Box and Hunter 1957, Wu and Hamada 2000). The two parts of the overall sample were chosen equal, but the optimal allocation is likely to depend on the problem.

To evaluate the effect of sample size relative to each design, we collected 50, 200, and 1000 sampling units per design. Each combination of design and sample size was run a thousand times. For all designs except the last, the probability p_i to sample the unit i was determined analytically. We ran the model-based design 10 000 times to numerically determine the probability p_i of the unit i of being sampled.

From the virtual samples, we estimated two biological parameters that we compared to the 'true' values: mean and optimum of the biological variable (see below). These are only two examples among all parameters that could be estimated for the variable (e.g. sum, variance, range, response to gradients).

(1) We estimated the mean (e.g. mean genetic diversity or mean functional trait in a landscape) as the mean variable weighted by $1/p_i$ (see Appendix 1); it gives the order of magnitude of the variable of interest.

(2) We also estimated the optimal conditions (optimum) in environmental space for the biological variable (e.g. niche optimum or trait optimum) from a linear model as explained above. Finally, we calculated the Euclidean distance between the "true" and the estimated optimum in environmental space; this distance gives an idea of how well the optimum was estimated.

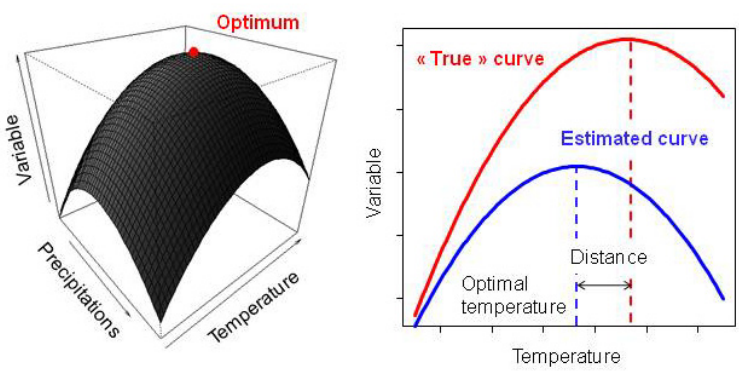


Fig. S2. Optimum of the variable, optimal conditions and distance between the true and the estimated optima
A. The virtual ecological variable (e.g., growth rate, habitat suitability, abundance) with an optimum at intermediate environmental conditions and a gradual decline towards the boundaries. It is a Gaussian curve along both selected climatic gradients: temperature and precipitations.
B. Schematic representation of the distance between the true optimal conditions and the one estimated from the sampling (e.g. following temperature).

Austin, M. P., Belbin, L., Meyers, J. A., Doherty, M. D. and Luoto, M. 2006. Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. - Ecological Modelling 199: 197-216.

Benichou, P. and Le Breton, O. 1987. Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. - La Météorologie 7: 23-24.

Box, G. E. P. and Hunter, J. S. 1957. Multi-Factor Experimental-Designs For Exploring Response Surfaces. - Annals Of Mathematical Statistics 28: 195-241.

Gaston, K. J. 2003. The structure and dynamics of geographic ranges.

Hirzel, A. H. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling. - Ecological Modelling 157: 331-341.

Mackenzie, D. I. 2006. Modeling the Probability of Resource Use: The Effect of, and Dealing with, Detecting a Species Imperfectly. - Journal of Wildlife Management 70: 368-374.

Meynard, C. N. and Quinn, J. F. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. - Journal of Biogeography 34: 1455-1469.

Mohler, C. L. 1983. Effect of sampling pattern on estimation of species distributions along gradients. - Vegetatio 54: 97-102.

Wintle, B. A., Kavanagh, R. P., McCarthy, M. A. and Burgman, M. A. 2005. Estimating and dealing with detectability in occupancy surveys for forest owls and arboreal marsupials. - Journal of Wildlife Management 69: 905-917.

Wu, C. F. J. and Hamada, M. 2000. Experiments: Planning, Analysis, and Parameter Design Optimization.

Summary

Intraspecific functional variability being poorly known, this work aimed at quantifying this variability, understanding the role of environmental gradients on this variability, disentangling its relative importance compared to interspecific variability and testing its impact on the calculation of functional diversity indices and a model of vegetation dynamics. Following a stratified and hierarchical sampling strategy, we collected traits on sixteen contrasted common species in a French alpine valley. We showed a strong intraspecific variability of traits. This variability was lower than interspecific one, but not negligible, it results partly from environmental effects and could be explained by species habitat models; it also strongly altered the calculation of functional diversity indices. In conclusion, intraspecific functional variability should be more often considered.

Résumé

Face à une méconnaissance de la variabilité intraspécifique des traits fonctionnels végétaux, ce travail visait à la quantifier, comprendre le rôle des gradients environnementaux sur celle-ci, quantifier son importance relative vis-à-vis de la variabilité interspécifique et tester son impact sur le calcul d'indices de diversité fonctionnelle et un modèle de dynamique de la végétation. Suivant une stratégie d'échantillonnage stratifiée et hiérarchique, nous avons collecté des traits sur seize espèces communes d'une vallée alpine française. Nous avons montré qu'il existait une forte variabilité intraspécifique des traits. Celle-ci était plus faible que la variabilité interspécifique, mais non négligeable, elle résultait en partie d'effets environnementaux, pouvait être expliquée par les modèles d'habitat de chaque espèce et modifiait fortement le calcul des indices de diversité fonctionnelle. En conclusion la variabilité fonctionnelle intraspécifique devrait être plus souvent considérée.

Mots-clés

Variabilité intraspécifique, Traits fonctionnels végétaux, Quantifier, Gradients environnementaux, Diversité fonctionnelle, Dynamique de la végétation, Stratégie d'échantillonnage, Ecologie alpine française