UNIL | Université de Lausanne
Département d'écologie
et évolution

Providing more informative projections of climate change impact on plant distribution in a mountain environment

Thèse de doctorat ès science de la vie (PhD)

Présentée à la

Faculté de Biologie et Médecine de l'Université de Lausanne

Par

Christophe Randin

Diplômé en Biologie (Université de Lausanne)

Jury

Prof. Walter Wahli, Président

Prof. Antoine Guisan, Directeur de thèse

Prof. Jérôme Goudet, Co-directeur de thèse

Prof. Christian Körner, Expert

Dr. Wilfried Thuiller, Expert

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale

Imprimatur

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

Président	Monsieur Prof.	Walter **Wahli**
Directeur de thèse	Monsieur Prof.	Antoine **Guisan**
Co-directeur de thèse	Monsieur Prof.	Jérôme **Goudet**
Experts	Monsieur Prof.	Christian **Körner**
	Monsieur Dr	Wilfried **Thuiller**

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

Monsieur Christophe Randin

Biologiste diplômé de l'Université de Lausanne

intitulée

Providing more informative projections of climate change impact on plant distribution in a mountain environment

Lausanne, le 29 juin 2007

pour Le Doyen
de la Faculté de Biologie et de Médecine

W. Wahli

Prof. Walter Wahli

« Il existe un monde d’espace, d’eau libre, de bêtes naïves où brille encore la jeunesse de la terre. Et il dépend de nous, de nous seuls qu’ils survivent »

Samivel

« Toute vie va finir. Il y aura une chaleur croissante. Elle sera insupportable à tout ce qui vit. Il y aura une chaleur croissante et rapidement tout mourra. Et néanmoins rien encore ne se voit. »

C.F. Ramuz, 1922

« Les forêts précédent l'homme, les déserts le suivent. »

La Rochefoucauld

Table of Content

Summary

Due to their conic shape and the reduction of area with increasing elevation, mountain ecosystems were early identified as potentially very sensitive to global warming. Moreover, mountain systems may experience unprecedented rates of warming during the next century, two or three times higher than that records of the 20th century. In this context, species distribution models (SDM) have become important tools for rapid assessment of the impact of accelerated land use and climate change on the distribution plant species.

In my study, I developed and tested new predictor variables for species distribution models (SDM), specific to current and future geographic projections of plant species in a mountain system, using the Western Swiss Alps as model region. Since meso- and micro-topography are relevant to explain geographic patterns of plant species in mountain environments, I assessed the effect of scale on predictor variables and geographic projections of SDM. I also developed a methodological framework of space-for-time evaluation to test the robustness of SDM when projected in a future changing climate. Finally, I used a cellular automaton to run dynamic simulations of plant migration under climate change in a mountain landscape, including realistic distance of seed dispersal. Results of future projections for the 21st century were also discussed in perspective of vegetation changes monitored during the 20th century.

Overall, I showed in this study that, based on the most severe A1 climate change scenario and realistic dispersal simulations of plant dispersal, species extinctions in the Western Swiss Alps could affect nearly one third (28.5%) of the 284 species modeled by 2100. With the less severe B1 scenario, only 4.6% of species are predicted to become extinct. However, even with B1, 54% (153 species) may still loose more than 80% of their initial surface.

Results of monitoring of past vegetation changes suggested that plant species can react quickly to the warmer conditions as far as competition is low However, in subalpine grasslands, competition of already present species is probably important and limit establishment of newly arrived species.

Results from future simulations also showed that heavy extinctions of alpine plants may start already in 2040, but the latest in 2080.

My study also highlighted the importance of fine scale and regional assessments of climate change impact on mountain vegetation, using more direct predictor variables. Indeed, predictions at the continental scale may fail to predict local refugees or local extinctions, as well as loss of connectivity between local populations. On the other hand, migrations of low-elevation species to higher altitude may be difficult to predict at the local scale.

Résumé

La forme conique des montagnes ainsi que la diminution de surface dans les hautes altitudes sont reconnues pour exposer plus sensiblement les écosystèmes de montagne au réchauffement global. En outre, les systèmes de montagne seront sans doute soumis durant le 21ème siècle à un réchauffement deux à trois fois plus rapide que celui mesuré durant le 20ème siècle. Dans ce contexte, les modèles prédictifs de distribution géographique de la végétation se sont imposés comme des outils puissants pour de rapides évaluations de l'impact des changements climatiques et de la transformation du paysage par l'homme sur la végétation.

Dans mon étude, j'ai développé de nouvelles variables prédictives pour les modèles de distribution, spécifiques à la projection géographique présente et future des plantes dans un système de montagne, en utilisant les Préalpes vaudoises comme zone d'échantillonnage. La méso- et la microtopographie étant particulièrement adaptées pour expliquer les patrons de distribution géographique des plantes dans un environnement montagneux, j'ai testé les effets d'échelle sur les variables prédictives et sur les projections des modèles de distribution. J'ai aussi développé un cadre méthodologique pour tester la robustesse potentielle des modèles lors de projections pour le futur. Finalement, j'ai utilisé un automate cellulaire pour simuler de manière dynamique la migration future des plantes dans le paysage et dans quatre scénarios de changement climatique pour le 21ème siècle. J'ai intégré dans ces simulations des mécanismes et des distances plus réalistes de dispersion de graines.

J'ai pu montrer, avec les simulations les plus réalistes, que près du tiers des 284 espèces considérées (28.5%) pourraient être menacées d'extinction en 2100 dans le cas du plus sévère scénario de changement climatique A1. Pour le moins sévère des scénarios B1, seulement 4.6% des espèces sont menacées d'extinctions, mais 54% (153 espèces) risquent de perdre plus 80% de leur habitat initial.

Les résultats de monitoring des changements de végétation dans le passé montrent que les plantes peuvent réagir rapidement au réchauffement climatique si la compétition est faible. Dans les prairies subalpines, les espèces déjà présentes limitent certainement l'arrivée de nouvelles espèces par effet de compétition.

Les résultats de simulation pour le futur prédisent le début d'extinctions massives dans les Préalpes à partir de 2040, au plus tard en 2080.

Mon travail démontre aussi l'importance d'études régionales à échelle fine pour évaluer l'impact des changements climatiques sur la végétation, en intégrant des variables plus directes. En effet, les études à échelle continentale ne tiennent pas compte des micro-refuges, des extinctions locales ni des pertes de connectivité entre populations locales. Malgré cela, la migration des plantes de basses altitudes reste difficile à prédire à l'échelle locale sans modélisation plus globale.

Introduction

The starting point: a warming world

Drivers of global change - such as destruction and fragmentation of habitats (Pitelka et al. 1997), pollution (Emberson et al. 2001) or biological invasions (Walther et al. 2002) - have a direct and fast impact on the distribution of species and biodiversity. Climate change is one component of global change that will have a long term impact on ecosystems and the geographic distribution of species, particularly for sessile organisms like plants.

During the last three decades, strong evidence suggest that increased atmospheric concentration of greenhouse gases – and particularly carbon dioxide and methane – was most likely caused by human activities and have already begun to modify the climate (Houghton et al. 1983, Houghton and Woodwell 1989, Woodwell et al. 1998). On average, the Earth temperature already increased by 0.61 (+/-0.18K) between 1861 and 2000 (Folland et al. 2001, Root et al. 2003). This warming likely results from the parallel increase in CO_2 from 280 to 370 ppmv observed since the "Industrial Revolution", at a rate 100 times faster than over the last 20'000 years (Berger 2002).

Switzerland already experienced a temperature increase twice the global average of the northern hemisphere, with a warming of +1.35K during the 20th century at low and high elevations (Rebetez 2002, Rebetez and Reinhard in press). The warming trend became particularly important after 1975, with a warming of 0.57 °C per decade. The same change was recorded in Château-d'Oex (980 m a.s.l.; data MeteoSuisse, Fig. 1).

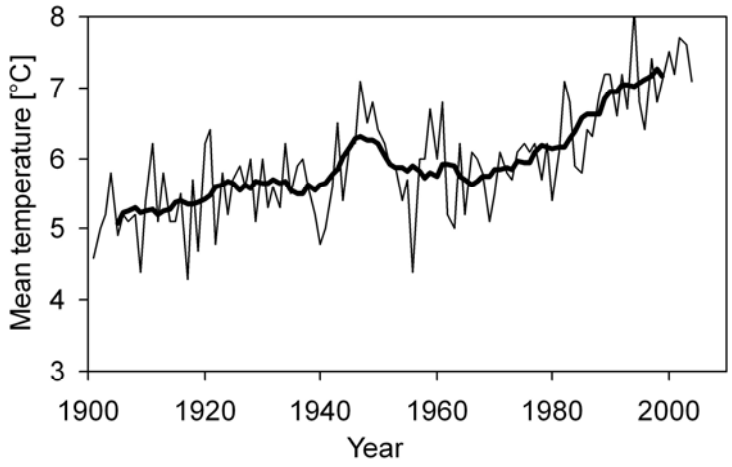


Figure 1 Temperature change during the 20th century in Château-d'Oex (980 m; data: MeteoSuisse).

During the same period, no net modification was observed for annual mean precipitations. However, some significant changes were observed with extreme values. Long periods of either strong precipitation or dry climate became increasingly common. In parallel, a decrease of snow-cover was observed at elevations below 1300 m, while nothing changed at higher elevations (Rebetez 2002).

Due to their conic shape and the reduction of area with increasing elevation, mountains ecosystems were early identified as potentially very sensitive to future global warming (Beniston 1994, Guisan et al. 1995, Beniston et al. 1996, Theurillat et al. 1998, Diaz et al. 2003). Recently, a global assessment of mountain sensitivity to climate change further indicated that these areas will experience a rate of warming two or three times greater than that recorded during the 20th century (Nogués-Bravo et al. 2006). Therefore, habitat shift and resulting species migration are susceptible to lead to important plant extinction events, especially at high-elevation, where vegetation is expected to be most affected by a long-term climate change (Grabherr et al. 1995, Walther 2003). Within the alpine and nival zones, the effect of topographic configuration (e.g. conic shape) is accentuated by a harsh climate. Because of this, at such high elevation, climate and other abiotic factors usually take much more importance than biotic interactions to explain biodiversity patterns (Grabherr et al. 1995, Körner 2003).

How can plants respond to climate change?

Plants can react in three different ways to climate change (Theurillat and Guisan 2001). It can: (i) stay (persist) in the modified climate, (ii) migrate to more suitable climates, or (iii) get extinct. To persist under climate change, plants must be able to adapt to new biotic and abiotic factors. New biotic interactions, such as competition, can occur with other plants invading by migration. In such situation, native plants have to be good competitors to resist extinction (Tilman 1994, 1997). To be able to stay and persist under new climate conditions (abiotic factors), a plant can react in five ways: (i) by changing its phenology, (ii) or physiology, (iii) by adapting genetically, (iv) through phenotypic plasticity or (v) through ecological buffering (Houghton et al. 2001, Theurillat and Guisan 2001). With regard to phenology, an extended growing season with earlier shooting and flowering in spring and later leaves colouring in autumn has already been observed in Europe (Menzel and Fabian 1999, Menzel et al. 2001, Fitter and Fitter 2002), in a wide range of situations (Chuine et al. 2000, Penuelas and Filella 2001, Minorsky 2002, Sparks and Menzel 2002, Walther et al. 2002). Concerning physiology, plants can change their growth, respiratory or photosynthetic rate with climate changes (Phillips and Gentry 1994, Cannell et al. 1998, Zhou et al. 2001, Jolly et al. 2005), or adapt genetically (Bradshaw and McNeilly 1991, Rodriguez-Trelles and Rodriguez 1998, Etterson and Shaw 2001). Alternatively, some plants may persist if climate is not the main cause of their distribution, e.g. when edaphic factors take more importance (Fig. 1: Theurillat and Guisan 2001).

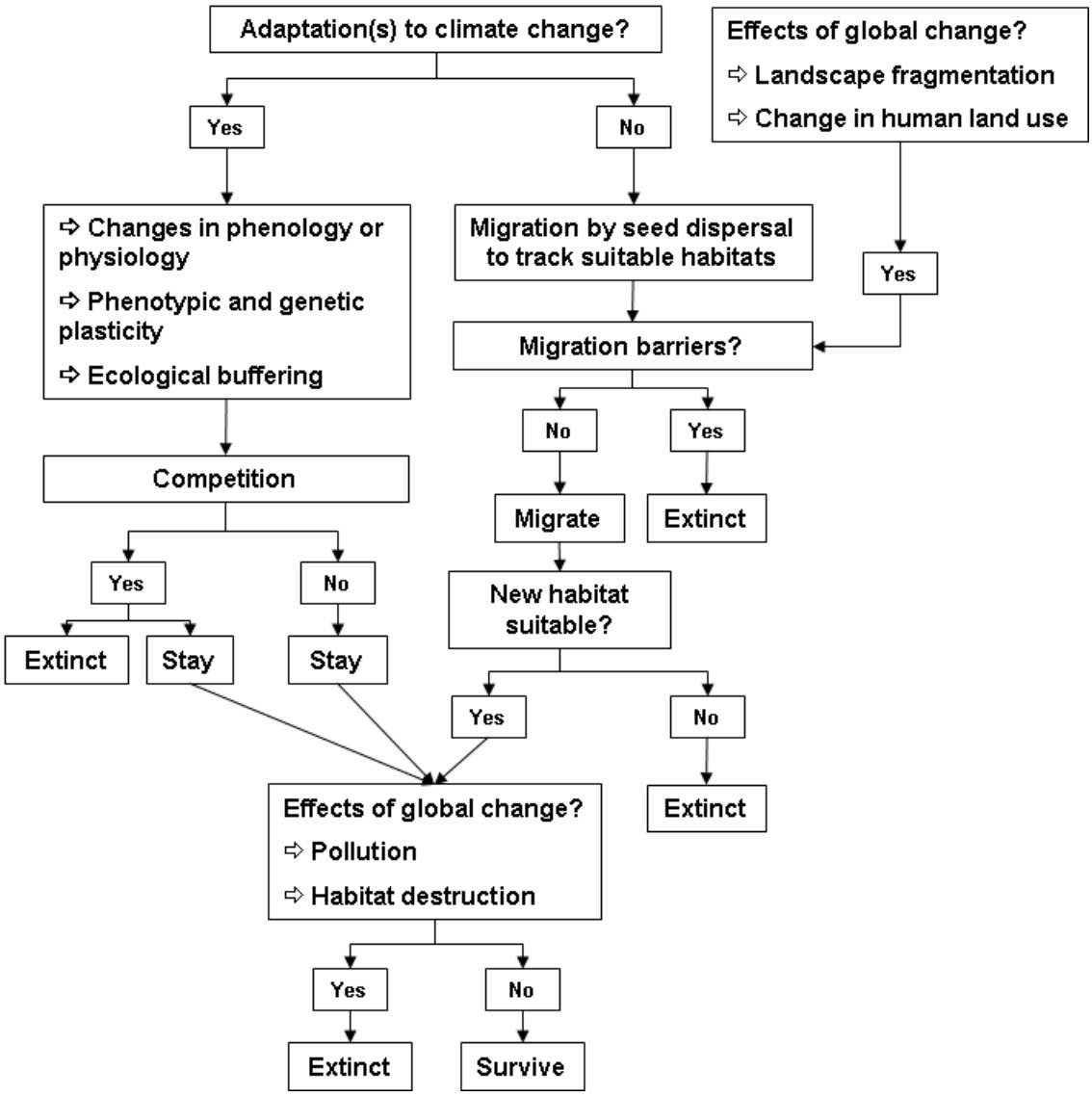


Figure 2 Potential effects of global and climate changes and potential reaction of vegetation.

If a plant cannot stand to adapt to climate changes, it has either to migrate to more suitable climates or become extinct. Many studies show that species' distributions are already expanding both poleward in latitude and upward in elevation (Grabherr et al. 1994b, Parmesan 1996, Shugart et al. 2001, Klanderud and Birks 2003, Parmesan and Yohe 2003). Successful migration can only occur if the dispersal strategy of the species allows it and if no barrier (i.e. landscape fragmentation at a broad scale or steep slope at a fine scale) impedes dispersal. Otherwise, if migration is unsuccessful, full or a quasi extinction can occur. Quasi extinction means here that a species could stay and persist in a few scattered locations in the form of relict populations, but these would be very vulnerable to any further change occurring. Extinction is particularly expected for species with long generation time, slow growth rate and/or short dispersal (Bradshaw and McNeilly 1991, Klanderud and

Birks 2003). Some extinctions have already been observed during the last 30 years (Klanderud and Birks 2003, Parmesan and Yohe 2003, Root et al. 2003).

Biological fingerprints of climate change in mountain areas

Climate warming induces plant migrations towards mountain summits or northern latitudes. Such migrations have already been documented for mid- to high-elevation summits in the Alps (e.g. Braun-Blanquet 1957, Hofer 1992, Grabherr et al. 1994a, Pauli et al. 1996, Walther et al. 2005, Vittoz et al. 2006), so as an increase in biodiversity in alpine plant communities (Braun-Blanquet 1975, Pauli et al. 1996). Similarly, an upward shift of the treeline ecotone was observed in different mountain ranges after the climate warming after following the end of the Little Ice Age (e.g. Taylor 1995, Kullman 2001, Camarero and Gutierrez 2004, e.g. Vittoz et al. in press); Gehrig-Fasel et al. in press). At lower elevations, plant species migrations were also already observed, e.g. for the mistletoe (*Viscum album*) in the Swiss Central Alps (Dobbertin et al. 2005).

Assessing the impacts of climate change on mountain plants

Three experimental ways have been used to assess the effect of climate change on plant species distribution in mountain environments: (1) *in situ* growth-chamber experiments (e.g. Theodose and Bowman 1997, De Valpine and Harte 2001, Richardson et al. 2002, e.g. Handa et al. 2006), (2) reciprocal transplantation (e.g. Clausen 1940, 1948, Núñez-Farfán 2001) or (3) laboratory experiments (e.g. Jones et al. 1998). All these approaches allow assessing physiological reactions of plants to climate change, nutrient enrichment or CO_2 increase. Experiments allow projections of plant behaviors to future by changing artificially physiologically meaningful parameters, but long term experiments are expensive and raise time-scale problems.

Monitoring networks have also been set up to identify ongoing trends. Recent projects, such as PERMANENT.PLOT.CH (Vittoz and Guisan 2003), GLORIA (Pauli et al. 2004) or MIREN (Dietz et al. 2006), are example of such initiatives. Numerous monitoring studies assessed the impact of climate change on alpine and nival vegetation (Grabherr et al. 1994a, Pauli et al. 1996, Walther et al. 2005, Vittoz et al. 2006) but only few focused on changes within grasslands at lower altitudes (Price and Waser 2000).. This is despite the subalpine zone representing, in the Alps, a much larger surface than alpine and nival zones (Theurillat and Guisan 2001). However, these monitoring studies do not allow deriving future projections, and very long-term projects (longer than human life-time) may be required to see some higher-level biological effects, e.g. on ecosystem composition, structure and functioning.

Using species distribution models to predict climate change impact on alpine plants

In the last decade, species distribution models (SDM; Guisan and Zimmermann 2000, Guisan and Thuiller 2005) have become important tools to provide rapid forecasts of the potential impact of accelerated human landuse and climate change on the distribution of organisms (Bakkenes et al. 2002, Thomas et al. 2004, Thuiller et al. 2005). SDMs attempt to provide detailed predictions of distributions by relating presence or abundance of species to environmental predictors (Elith et al. 2006). In this context, modelling the distribution of plants is particularly appropriate because valid absences can be more realistically obtained for sessile species than for mobile species, e.g. from plot surveys where exhaustive species lists are drawn (Guisan and Thuiller 2005). Furthermore, as sessile organisms, plants are also less likely to successfully track new suitable habitats under transient climate change than very mobile species such as birds.

A large number of techniques are available for modeling species' distribution (Guisan & Thuiller 2005, Elith et al. 2006). These techniques vary in how they model the distribution of the response, select relevant predictor variables, define fitted functions for each variable, weight variables contributions, allow for interactions, or predict geographic patterns of occurrence (Elith et al. 2006). Guisan & Zimmermann (2000) distinguish six main categories of modelling techniques: multiple regression and its generalized forms (GLM or GAM; e.g. Frescino et al. 2001, GLM or GAM; e.g. Bakkenes et al. 2002, Guisan et al. 2002) classification techniques (CART or RF; e.g. Franklin 1998, CART or RF; e.g. Prasad et al. 2006) environmental envelopes (Box 1981, Busby 1991), ordination techniques (Hill 1991, Gottfried et al. 1998, Guisan et al. 1999), Bayesian approaches (Aspinall 1992) and neural networks (Manel et al. 1999). Recently, new methods showed up, such as Generalised Boosted Models (GBM; Friedman et al. 2000),or Random Forest (Breiman 2001), that are combinations or improvements of the previous methods and appeared to be performing particularly well compared to previous techniques (e.g. Iverson et al. 2004, e.g. Elith et al. 2006).

So far, few studies have used SDMs to assess the possible impact of climate change on plant species distribution in mountain environments. In a recent large-scale SDM study, Thuiller *et al.* (2005) forecasted certain mountain regions of Europe (e.g. mid-elevation Alps) as being disproportionably sensitive to climate change, with up to 60% of species loss predicted by the end of the 21st century. This study was conducted using 1350 species from the Atlas Florae Europaea (AFE; Lahti and Lampinen 1999) over the entire extent of Europe at a coarse spatial resolution (model fitting at 50 km x 50 km; spatial projections at 16 km x 16 km). In a study of 85 subalpine and alpine non-woody plants of open habitats conducted at a much finer scale (20 m x 20 m) in the Austrian Alps, Dirnböck *et al.* (2003) predicted up to 40-50% potential species extinction due to climate and land-use change. The climate change scenarios used here included regionalized climatic simulations of the ECHAM4 GCM for the year 2050 and land-use changes. Guisan and Theurillat

(2000), in a sensitivity study (+1.5K,+3.0K and +4.5K) using 62 alpine or nival species, predicted strict extinctions (100% of habitat lost) only in a range 2-5%, but nearly 40% of species were predicted to loose more than 90% of their most suitable habitats.

Developing new predictor variables for SDM in mountain systems

One crucial task when building a SDM is selecting the environmental variables to be used for predictions. Various descriptors of the environment can be potentially used, reflecting topographic, climatic, edaphic, geologic or human influences. These variables can be further classified as having indirect or direct effects on the species, and constitute a resource or not for them. At fine, local scales in mountain environments, models based on indirect predictors like topographic ones (e.g. altitude or aspect; Fig.1) can still provide accurate local predictions. However, on global to meso-scales, like Switzerland and larger areas, models based on indirect predictors loose power compared to models based on more proximal predictors (Guisan and Hofer 2003). This is because at these scales, same values of indirect predictors at distant sites can translate into very different values of the more proximal predictors, due to e.g. major climatic or geologic differences taking place between entire regions within the modelled area. Hence, although models fitted with indirect predictors at the local scale can still provide accurate predictions for the same area, they will lack power to predict to other areas, i.e. they will show weak transferability (Randin et al. 2006).

In order to have accurate models of plant distribution in alpine systems that still have a good generality and transferability to other areas, we need to use preferentially predictor variables that have a direct eco-physiological or ecological impact on plants (regulators) or constitute resources for them (Austin 1980, Austin et al. 1984, Austin 1985, Austin and Heyligers 1989, Guisan and Zimmermann 2000) (Figure 1). A current limitation of theses models is the lack of spatially explicit predictors related to stresses or modifications occurring on the soil surface, such as human-related land use or geomorphologic and snow translocation processes.

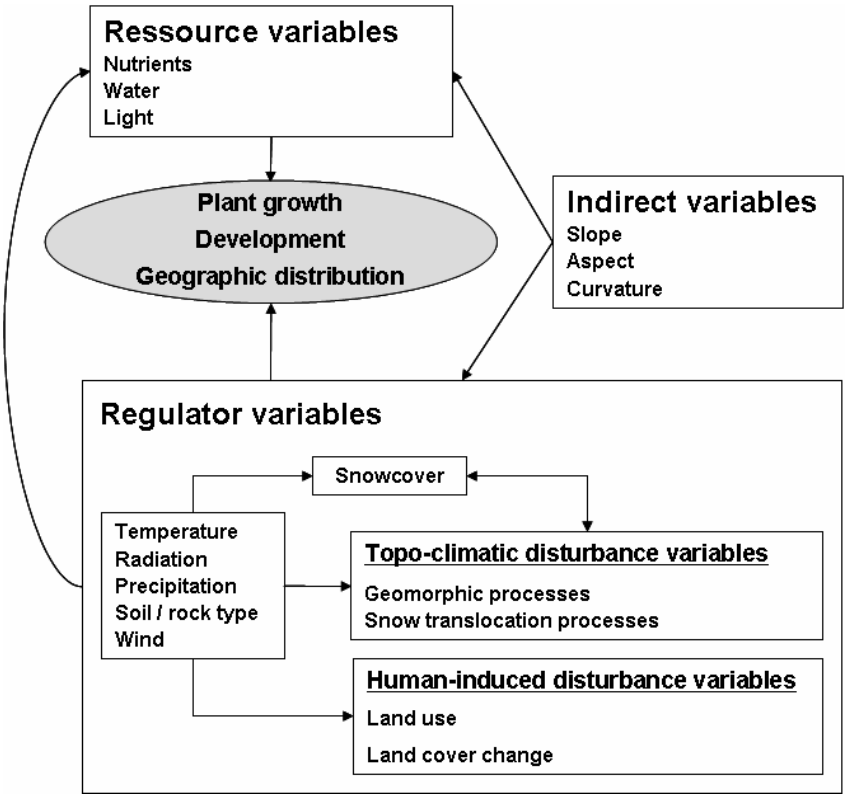


Figure 3 Hierarchical framework of resource, regulator and indirect variables used in species distribution models (adapted from Guisan & Thuiller 2005).

Another recurrent limitation of SDM is the lack of spatially explicit information on land use (Zimmermann and Kienast 1999, Guisan and Zimmermann 2000, Pearson and Dawson 2003). Although bioclimatic factors alone were shown to explain a major part of the observed patterns at large scale (Thuiller et al. 2005), agricultural land use remains a potentially useful variable for predicting plant species and communities distribution at finer scales (Fischer 1994, Dirnböck et al. 2003). At this scale, land use is partly dependent on topography and climate and should thus also be altered by changing climate (Carcaillet and Brun 2000, Rounsevell et al. 2006). It can also prove an important factor by acting as barrier or corridor for plant dispersal when predicting future species distributions using dynamic dispersal models (Carey and Brown 1994, Carey 1996, Dullinger et al. 2004). Because spatially-explicit information on land use is often difficult to obtain at fine scales, few studies using SDM have so far integrated this information, despite a large ecological knowledge accumulated on the effects of land use on plant communities.

In a mountain area like the Swiss Alps, landscape configuration in the alpine and nival vegetation belts leads to very active and dynamic hydro-mechanical processes, like geomorphologic ones. These processes affect vegetation at the local scale (Burga 1999) but also modify the species diversity at larger scale (Nichols et al. 1998). Gravity or hydrodynamic processes like rock slides, avalanches or solifluction (Johnson and Billings 1962, Erschbamer 1989, Körner 2003) are examples of

processes that have a big impact on the soil surface and strongly drive the species distribution patterns (Bretz-Guby 1994). Some adaptive traits of plant species are highly correlated with the conditions encountered in extremely disturbed situations, such as large root networks on moving rock slides (Jonasson and Callaghan 1992). In alpine landscape, the spatial and temporal persistence of the snow cover is influenced by topography and determines species composition and spatial vegetation patterns (Walsh et al. 1994, Körner 2003). Snow cover affects soil and vegetation moisture levels and offers protection from climatic stress, particularly wind desiccation (Schaefer and Messier 1995). Snowpack is also a direct reservoir of nutrients for plant growth (Bowman 1992). But snow could also be a stress factor itself through its dynamic effects like avalanche paths or wind-induced snow translocation (Erschbamer 1989, Körner 2003). Topographic position (Gottfried et al. 1998, Guisan et al. 1998, Gottfried et al. 1999, Guisan and Theurillat 2000, Dirnbock et al. 2003, Dirnböck et al. 2003), drainage surface (Leathwick et al. 1998) or distance to ridges (Moore et al. 1991, Dirnbock et al. 2002) have been used as surrogate for disturbance processes in plant distribution modelling studies. Most of these were conducted in mountain environment (Gottfried et al. 1998, Guisan et al. 1998, Gottfried et al. 1999, Guisan and Theurillat 2000, Dirnbock et al. 2003, Dirnböck et al. 2003). However, no study so far tested disturbance variables like geomorphic perturbations. I am also not aware of any study using physical and mechanistic models of snow translocation as input in SDM.

Influence of topography becomes more pronounced as elevation increases because slopes tend to be steeper, climatic vectors are stronger and because plant life becomes more dependent on decoupling from a hostile atmosphere (Körner 2003). The strong influence of local microclimate and the difficulty to produce accurate climatic maps by interpolation explain the successful use of topographic predictors in many SDM studies in mountain environment (Gottfried et al. 1999, Zimmermann and Kienast 1999, Guisan and Theurillat 2000, Dirnböck et al. 2003). Indeed, standard climate stations do not necessarily provide biologically relevant microclimatic measurements, and/or their interpolation will generate a lot of uncertainties (Guisan and Zimmermann 2000). In this case, the lack of precision can be compensated by using indirect environmental variables, i.e. by exploiting correlation between easily measurable non-causal and too-difficult-to-measure causal variables (Guisan et al. 1999). Elevation and its derivative, such as slope, aspect and curvature can be used as proxies for multiple variables such as drainage, soil depth, soil composition and all microclimatic variables co-varying with topography. Snow distribution, wind and energy allocation in structured alpine terrain are the most obvious consequences of indirect effect of topography and its interaction with climate (Billings and Bliss 1959, Körner 2003)

A remnant question is at which spatial resolution should these topographic predictors be measured to optimize the predictive power of SDM? For example, curvature at a very fine scale drives snow accumulation during winter, which will directly influence the length of the growing season or the water budget in the soil. On larger scales, the relative position of plants between ridge tops and valley

bottoms exposes them to different wind systems. Slope can be used directly to express primarily gravitation and dynamic geomorphologic processes and their influence on the composition of vegetation, e.g. through snow movement or water flow, but it can also be used indirectly to express the quantity of energy received by plants (i.e. solar radiations).

Better validating SDM for safer future projections

A large number of studies, involving many different modelling techniques, have been conducted to predict the potential distribution of species under changing environmental conditions (Huntley et al. 1995, Sykes and Prentice 1996, Guisan and Theurillat 2000, Peterson and al. 2002, Dirnböck et al. 2003, Thuiller et al. 2005, Araújo et al. 2006). However, testing for proper geographical or temporal transferability of models and the related uncertainty has usually been neglected.

Testing model transferability should prove particularly powerful for complementing standard model evaluation procedures. The use of observations independent from the training dataset has been recommended for a proper evaluation of models (Fielding and Bell 1997, Guisan and Zimmermann 2000). If the training and test datasets are restricted to the same spatial and temporal domains, internal evaluation techniques such as data partitioning or split sample approaches are sufficient. If the predictions are to be tested for their generality and robustness (accuracy and stability of predictions for predictions in a future changing climate), Fielding & Bell (1997) recommend using a geographically (Fielding and Haworth 1995) or temporally (Boyce et al. 2002, Araújo et al. 2005) independent data set for their external evaluation.

Using the best scale for SDM projections in mountain systems

Taking the same examples as previously discussed (Guisan and Theurillat 2000, Dirnböck et al. 2003, Thuiller et al. 2005), different extinction rates are predicted, depending on the geographic extent, elevation range and resolution used to draw them. However, no comparison of predictions across scale has been attempted so far. As predictions based on SDMs are commonly used as raw material for preparing IPCC reports and eventually guiding nature managers in actions aiming to anticipate biodiversity losses in these systems (Araujo et al. 2004a, Araujo et al. 2004b), there is an urgent need to assess whether predictions at different scale are comparable.

In particular, what needs to be assessed is whether the discrepancy between predictions results from different species being considered in separate assessments, or if the discrepancy persists for the same set of species when modeled on different scales. A discrepancy between predictions may for instance result from the fact that different scales may reflect topography in different ways, then translating into divergent predictions. More generally, the same environmental variables may not bear the same meaning for a species when calculated at local or global scales (Guisan 2005).

More dynamics in species distribution models

The predicted warming will require considerable migration rates for plant species to stay under similar climatic conditions (e.g. Malcolm et al. 2002). However, so far, only few of these studies have considered dispersal (Carey and Brown 1994, Carey 1996, Iverson and Prasad 1998, Dullinger et al. 2004, Iverson et al. 2004, Broennimann et al. 2006). Instead, most of them have used the assumption of universal dispersal (i.e. a species has unlimited dispersal, its future distribution becoming the entire area projected by the habitat distribution model; Guisan and Theurillat 2000, Bakkenes et al. 2002, Dirnböck et al. 2003, i.e. a species has unlimited dispersal, its future distribution becoming the entire area projected by the habitat distribution model; Thomas et al. 2004, Thuiller et al. 2005). While this hypothesis might provide good approximations for plants that are man-dispersed or have high dispersal ability, it is likely to overestimate the future distribution of many other species, because: (i) human-driven habitat fragmentation has rendered the landscape increasingly impassable (Pitelka et al. 1997) and (ii) the speed of the expected global warming is predicted to be rapid – one or more orders of magnitude faster than past climate changes (Etterson and Shaw 2001) – and may thus, at least for certain species, require migration rates much faster than those observed during post-glacial time (Malcolm et al. 2002), meaning that not all species might be able to keep pace with climate change. Furthermore, such rapid change may not allow species to adapt (Etterson and Shaw 2001). Since universal dispersal represents a too optimistic, "best case", scenario, some authors (e.g. Thomas et al. 2004, Thuiller et al. 2004, Thuiller et al. 2005) also provide a, "worst case", no dispersal for each species to balance projections. However, the difference between these extreme projections yields large uncertainties (e.g. Thuiller 2004). Reducing such uncertainty calls for the inclusion of dispersal processes when addressing the issue of climate change impact on plant distribution.

Objectives and structure of the thesis

The main objective of my thesis was to provide more informative predictions of future impacts of climate change on alpine plant species distribution in the Western Swiss Alps. This research was initiated in 2002 within the framework of the MODIPLANT project (Guisan 2005). My thesis is divided in three temporally structured parts (Fig.4):

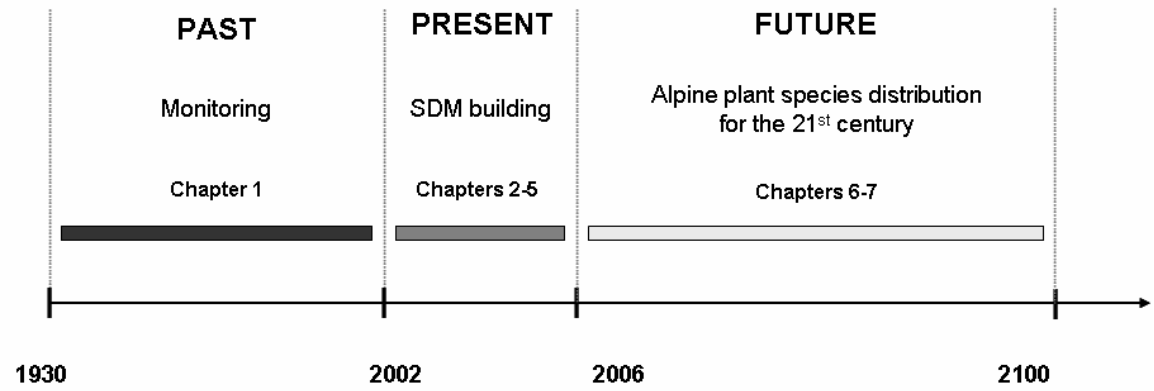


Figure 4 Structure of the thesis.

1. **Past** - Fingerprints of climate change in the Western Swiss Alps

In chapter 1, I assessed whether or not climate change affected subalpine grasslands in the Northern Swiss Alps during the 20th century.

2. **Present** - Building species distribution model (SDM)

In chapter 2 and 3, I tested the improvement of species distribution models by human- and topo-climatic-induced disturbance variables. In particular, I assessed the importance of land use in species distribution models in chapter 2 and I estimated the improvement of current topo-climatic based SDM when introducing process-based disturbance variables in chapter 3. Then I tested the effect topographic variables on the predictive power of SDM in chapter 4. In the last chapter of this part (chapter 5), I developed a modeling framework to estimate the scale for time transferability of species distribution models.

3. **Future** - Alpine plant species distribution for the 21st century

In chapter 6 I estimated the divergence between large and fine scale predictions of SDM for the end of the 21st century under four different climate change scenarios and assessed the possible causes of such divergence. In chapter 7, I performed continuous dynamic simulations of the distribution of 284 plant species under four climate change scenarios for the 21st century. For this purpose, I used cellular

automaton to compared extreme simulations without and with unlimited dispersal to simulations integrating more realistic seed dispersal distances

Annexes Methodological and theoretical contributions of the MODIPLANT project

During my PhD, I had the opportunity to set different collaborations to answer to problematic closely related to species distribution models. The MODIPLANT species dataset with its large and robust presence-absence observation was well adapted to test a new method for the evaluation of a presence-only model (chapter 8) and to evaluate a new modeling technique based on presence only (chapter 9).

Theoretical hypothesis are required when using species distribution models. In particular, the assumption that the ecological realized niche will not evolve in the future is necessary for future predictions. An appropriate scale is also necessary for a proper definition of the ecological niche. These problems are developed in chapter 10.

All the studies presented in the chapters 1 to 11 have been conducted in the same study area of the Western Alps of the Canton de Vaud or one of its sub-regions. Hence, I present here the general characteristics of the study area, the data collected there for modelling and the associated sampling strategy.

Study area

The study area includes all mountains of the Western Alps of the Canton de Vaud in Switzerland (6°60' to 7°10' E; 46°10' to 46°30' N) and covers an area of 700 km^2 (Fig. 5). Elevation ranges from 375 m in Montreux to 3210 m on the top of the Diablerets massif, where important glaciers are found (most important: Tsanfleuron with a surface of 3.8 km^2). Annual temperature and precipitation vary respectively from 8°C and 1200 mm at 600 m to -5°C and 2600 mm at 3000 m (Bouët 1985). The soil parent material is mainly calcareous. Vegetation has been and is still very influenced by human activity: the pasture is common in this region, and is observed from the bottom in the Rhone valley to the subalpine and lowers alpine zones at higher elevations. Meadows of diverse intensity of management are mainly found at lower elevations.

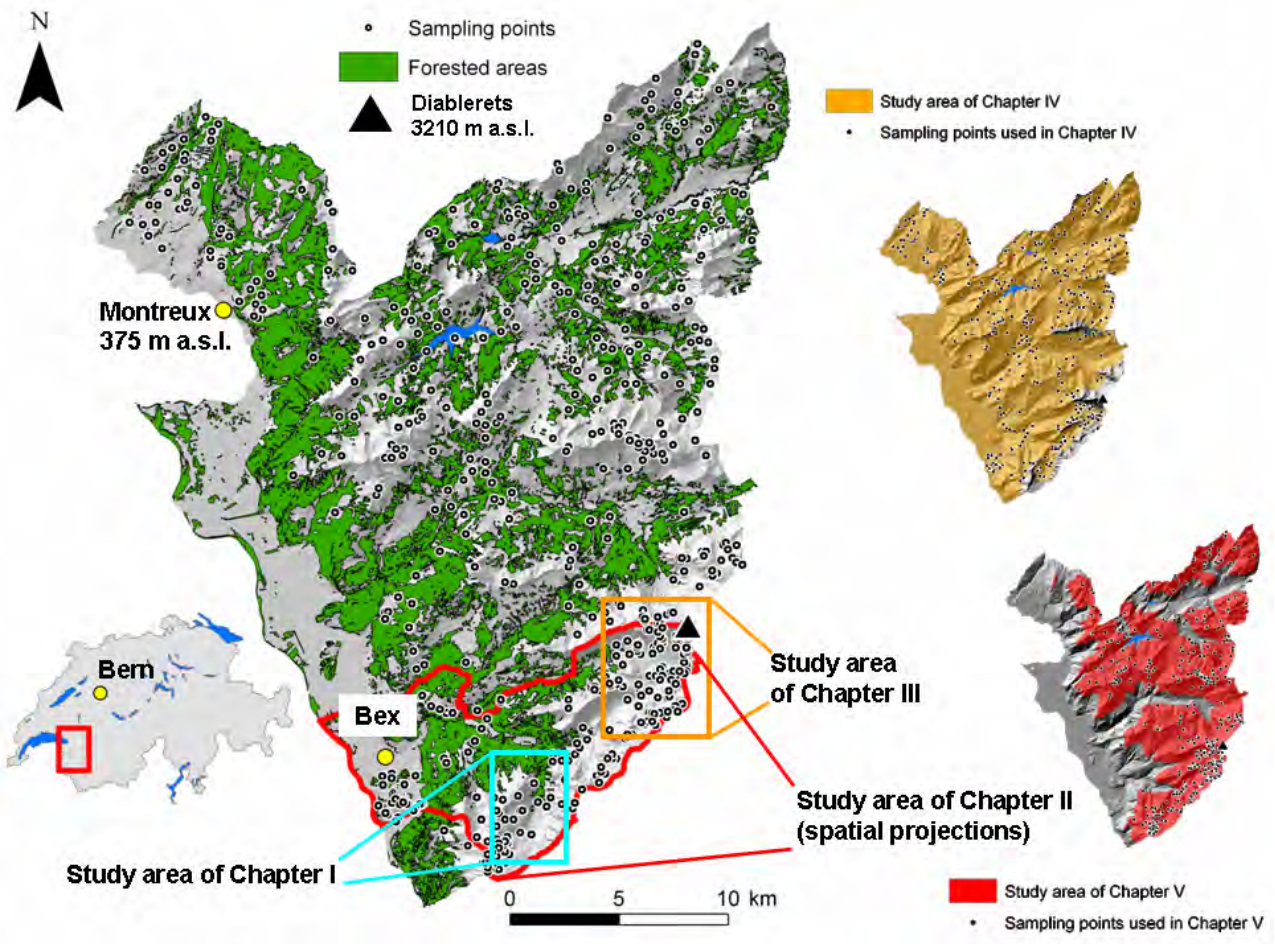


Figure 5 Main study area with locations of the sub-regions used in the difference chapters.

Sampling strategy

During summers 2002-2003, 518 vegetation plots were sampled following a random-stratified sampling strategy restricted to non-woody vegetation (grassland, rock and scree). Stratification was done by elevation, slope, aspect and simplified classes of geology representing soil permeability and $CaCO_3$ concentration. Plots were composed of a large surface of 64 m^2 and, nested within each, smaller surfaces of 16 m^2, 4 m^2 and 1 m^2 (Fig.6). The presence of all species was recorded in the four surfaces, and abundance was additionally recorded in the 4 m^2 plots following Braun-Blanquet's method (Braun-Blanquet 1964). The exact coordinates of plots were recorded on the center of the 8x8 m quadra with a Trimble Geoexplorer 3 GPS system which allows an accuracy of less than 1 m in average after differential correction. During summer 2004, a third field campaign was performed, focusing especially on geomorphologic units which were extra strata, and providing 32 additional points.

Figure 6 Design of the imbricated sampling plot with the position of the GPS on the center of the 8x8 m quadra.

A preliminary study showed that the most appropriate plot size for species distribution was the largest surface of 8x8m (Fig.7). In this exploratory study (unpublished results), model fit and predictive power were calculated for the SDMs of 230 species. Models were built by regressing the presence-absence of species to a same set of predictor variables. Model fit was measured with the adjusted D^2 (Guisan and Zimmermann 2000). Predictive power was calculated through a 10-fold cross-validation procedure using the area under the curve of a ROC plot (AUC; Fielding and Bell 1997; see chapter 2 for more details). The comparisons between surfaces were done with Wilcoxon signed-rank tests (treating samples as grouped by species). The 8x8m surface yielded significantly greater values, on average, of model fit and predictive power than those obtained from models fitted with data from smaller plot sizes (P values < 0.001).

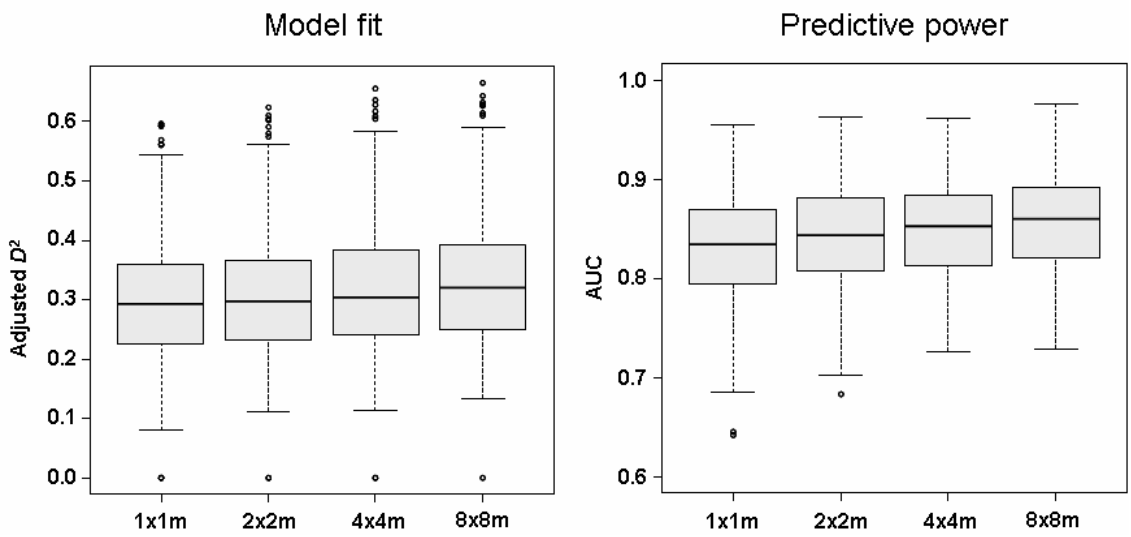


Figure 7 Model fit and predictive power of models for 230 species calibrated on the four different sampling plot size.

The 2x2m (the only one in which abundance values were additionally recorded) and 8x8m plots were used in chapter 2 whereas all plot sizes were used in chapter 4. In chapter 5, the 4x4m plots allowed comparing SDM calibrated in the Western Alps to models calibrated in the Austrian Alps. We calibrated SDMs of chapter 3, 6 and 7 based on observations of 8x8m plots.

References

Araujo, M. B., M. Cabeza, W. Thuiller, L. Hannah, and P. H. Williams. 2004a. Would climate change drive species out of reserves? an assessment of existing reserve-selection methods. Global Change Biology **10**.

Araujo, M. B., P. J. Densham, and P. H. Williams. 2004b. Representing species in reserves from patterns of assemblage diversity. Journal of Biogeography **31**:1037-1050.

Araújo, M. B., W. Thuiller, and R. G. Pearson. 2006. Climate warming and the decline of amphibians and reptiles in Europe. Journal of Biogeography **In press**.

Araújo, M. B., R. J. Whittaker, R. Ladle, and M. Erhard. 2005. Reducing uncertainty in projections of extinction risk from climate change. Global Ecology & Biogeography **In press**.

Aspinall, R. 1992. An Inductive Modeling Procedure Based on Bayes Theorem for Analysis of Pattern in Spatial Data. International Journal of Geographical Information Systems **6**:105-121.

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. Vegetatio **42**:11-21.

Austin, M. P. 1985. Continuum concept, ordination methods and niche theory. Annu. Rev. Ecol. Syst. **16**:39-61.

Austin, M. P., R. P. Cunningham, and P. M. Fleming. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio **55**:11-27.

Austin, M. P., and P. C. Heyligers. 1989. Vegetation survey design for conservation: gradsec sampling of forests in north-east New South Wales. Biol. Conserv. **50**:13-32.

Bakkenes, M., J. R. M. Alkemade, F. Ihle, R. Leemans, and J. B. Latour. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. Global Change Biology **8**:390-407.

Beniston, M. 1994. Mountain Environment in Changing Climates, Routledge edition. Routledge, London.

Beniston, M., D. G. Fox, S. Adhikary, R. Andressen, A. Guisan , J. I. Holten, J. Innes, J. Maitima, M. F. Price, and L. Tessier. 1996. Impacts of climate change on mountain regions. Chapter 5. *in* IPCC, editor. Intergovernmental panel on climate change. Second assessment report. Cambridge University Press, Cambridge.

Berger, A. 2002. Global warming 2001. Journal de Physique IV (Proceedings) **12**:19-26.

Billings, W. D., and L. C. Bliss. 1959. An Alpine Snowbank Environment and Its Effects on Vegetation, Plant Development, and Productivity. Ecology **40**:388-397.

Bouët, M. 1985. Climat et météorologie de la Suisse romande, Payot edition, Lausanne.

Bowman, W. D. 1992. Inputs and Storage of Nitrogen in Winter Snowpack in an Alpine Ecosystem. Arctic and Alpine Research **24**:211-215.

Box, E. O. 1981. Macroclimate and Plant Forms: An Introduction to Predictive Modeling in Phytogeography. Junk, The Hague.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling **157**:281-300.

Bradshaw, A. D., and T. McNeilly. 1991. Evolutionary response to Global Climatic Change. Annals of Botany **68**:5-14.

Braun-Blanquet, J. 1957. Ein Jahrhunder Florenwandel am Piz Linard (3414 m). Bull. Jard. Bot. Etat Brux. **volume jubilaire Walter Robyns**:221-232.

Braun-Blanquet, J. 1964. Pflanzensoziologie. Grundzüge der Vegetationskunde, 3 edition. Springer, Wien - New-York.

Braun-Blanquet, J. 1975. Fragmenta Phytosociologica Raetica I: Die Schneebodengesellschaften (Klasse der Salicetea herbaceae). Jahresbericht der Naturforschenden Gesesellschaft Graubünden **96**:42-71.

Breiman, L. 2001. Random Forests. Mach Learn **45**:5-32.

Bretz-Guby, N. 1994. Géomorphologie et végétation à l'étage alpin: L'exemple du Mont-Gelé. Université de Lausanne, Lausanne.

Broennimann, O., W. Thuiller, G. Hughes, G. F. Midgley, J. R. M. Alkemade, and A. Guisan. 2006. Do geographic distribution, niche properties and lifeforms explains vulnerability of plants to climate changes? . Global Change Biology **12**:1-15.

Burga, C. A. 1999. Vegetation development on the glacier forefield Morteratsch (Switzerland). Applied Vegetation Science **2**:17-24.

Busby, J. R. 1991. BIOCLIM - A bioclimate analysis and prediction system. Pages 64-68 *in* C. R. Margules and M. P. Austin, editors. Nature conservation: cost effective biological surveys and data analysis. CSIRO, Melbourne.

Camarero, J. J., and E. Gutierrez. 2004. Pace and pattern of recent treeline dynamics: Response of ecotones to climatic variability in the Spanish Pyrenees. Climatic Change **63**:181-200.

Cannell, M. G. R., J. H. M. Thornley, D. C. Mobbs, and A. D. Friend. 1998. UK conifer forests may be growing faster in response to increased N deposition, atmospheric CO2 and temperature. Forestry **71**:277-296.

Carcaillet, C., and J. J. Brun. 2000. Changes in landscape structure in the northwestern Alps over the last 7000 years: lessons from soil charcoal. Journal of Vegetation Science **11**:705-714.

Carey, P. D. 1996. DISPERSE: A cellular automaton for predicting the distribution of species in a changed climate. Global Ecology and Biogeography Letters **5**:217-226.

Carey, P. D., and N. J. Brown. 1994. The use of GIS to identify sites that will become suitable for a rare orchid, Himantoglossum hircinum L., in future changed climate. Biodiversity Lett. **2**:117-123.

Chuine, I., G. Cambon, and P. Comtois. 2000. Scaling phenology from the local to the regional level: advances from species-specific phenological models. Global Change Biology **6**:943-952.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1940. Experimental studies on the nature of species. I. Effect of varied environment on western North American plants. Washington, DC: Publication 520, Carnegie Institution of Washington.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1948. Experimental studies on the nature of species. III. Environmental responses of climatic races of Achillea. Carnegie Institute of Washington.

De Valpine, P., and J. Harte. 2001. Plant responses to experimental warming in a montane meadow. Ecology **82**:637-648.

Diaz, H. F., M. Grosjean, and L. Graumlich. 2003. Climate variability and change in high elevation regions: Past, present and future. Climatic Change **59**:1-4.

Dietz, H., C. Kueffer, and C. G. Parks. 2006. MIREN: A new research network concerned with plant invasion into mountain areas. Mountain Research and Development **26**:80-81.

Dirnbock, T., S. Dullinger, M. Gottfried, C. Ginzler, and G. Grabherr. 2003. Mapping alpine vegetation based on image analysis, topographic variables and Canonical Correspondence Analysis. Applied Vegetation Science **6**:85-96.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Dirnbock, T., R. J. Hobbs, R. J. Lambeck, and P. A. Caccetta. 2002. Vegetation distribution in relation to topographically driven processes in southwestern Australia. Applied Vegetation Science **5**:147-158.

Dobbertin, M., N. Hilker, M. Rebetez, N. E. Zimmermann, T. Wohlgemuth, and A. Rigling. 2005. The upwart shift of pine mistletoe (Viscum album ssp. austriacum) in Switzerland - The result of climate warming ? Int. J. Biometeorol. **50**:40-47.

Dullinger, S., T. Dirnbock, and G. Grabherr. 2004. Modelling climate change-driven treeline shifts: relative effects of temperature increase, dispersal and invasibility. Journal of Ecology **92**:241-252.

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

Emberson, L. D., M. R. Ashmore, F. Murray, J. C. I. Kuylenstierna, K. E. Percy, T. Izuta, Y. Zheng, H. Shimizu, B. H. Sheu, C. P. Liu, M. Agrawal, A. Wahid, N. M. Abdel-Latif, M. van Tienhoven,

L. I. de Bauer, and M. Domingos. 2001. Impacts of air pollutants on vegetation in developing countries. Water Air and Soil Pollution **130**:107-118.
Erschbamer, B. 1989. Vegetation on Avalanche Paths in the Alps. Vegetatio **80**:139-146.
Etterson, J. R., and R. G. Shaw. 2001. Constraint to Adaptive Evolution in Response to Global Warming. Science **294**:151-154.
Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence-absence models. Environmental Conservation **24**:38-49.
Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conservation Biology **9**:1466-1481.
Fischer, H. S. 1994. Simulation der räumlichen Verteilung von Pflanzengesellschaften auf der Basis von Standortskarten. Dargestellt am Beispiel des MaB-Testgebiets Davos. Pages 1 - 158 *in* Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübel, in Zürich.
Fitter, A. H., and R. S. R. Fitter. 2002. Rapid changes in flowering time in British plants. Science **296**:1689-1691.
Folland, C. K., N. A. Rayner, S. J. Brown, T. M. Smith, S. S. P. Shen, D. E. Parker, I. Macadam, P. D. Jones, R. N. Jones, N. Nicholls, and D. M. H. Sexton. 2001. Global temperature change and its uncertainties since 1861. Geophysical Research Letters **28**:2621-2624.
Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. Journal of Vegetation Science **9**:733-748.
Frescino, T. S., T. C. Edwards, and G. G. Moisen. 2001. Modeling spatially explicit forest structural attributes using Generalized Additive Models. Journal of Vegetation Science **12**:15-26.
Friedman, J. H., T. Hastie, and R. Tibshirani. 2000. Additive Logistic Regression: a Statistical View of Boosting. Annals of Statistics **28**:337-374.
Gottfried, M., H. Pauli, and G. Grabherr. 1998. Prediction of vegetation patterns at the limits of plant life: A new view of the alpine-nival ecotone. Arctic and Alpine Research **30**:207-221.
Gottfried, M., H. Pauli, K. Reitter, and G. Grabherr. 1999. A fine-scaled predictive model for changes in species distribution patterns of high mountain plants induced by climate warming. Diversity & Distributions **5**:241-251.
Grabherr, G., M. Gottfried, A. Gruber, and H. Pauli. 1995. Patterns and current changes in alpine plant diversity. Pages 167-181 *in* A. Chapin and C. Körner, editors. Arctic and alpine biodiversity: patterns, causes and ecosystem consequences. Springer, Heidelberg.
Grabherr, G., M. Gottfried, and H. Pauli. 1994a. Climate effects on mountain plants. Nature **369**:448.
Grabherr, G., M. Gottfried, and H. Pauli. 1994b. Climate Effects on Mountain Plants. Nature **369**:448-448.
Guisan, A. 2005. Niche-based models as tools to assess climate change impact on the distribution and diversity of plants in mountain reserves. *in* Proceedings of the 2nd and 3rd GLOCHAMORE workshops in L'Aquila (Italy) and Granada (Spain). UNESCO-MRI, Paris, France.
Guisan, A., T. C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling **157**:89-100.
Guisan, A., and U. Hofer. 2003. Predicting reptile distributions at the mesoscale: relation to climate and topography. Journal of Biogeography **30**:1233-1243.
Guisan, A., J. I. Holten, R. Spichiger, and L. Tessier, editors. 1995. Potential Ecological Impacts of Climate Change in the Alps and Fennoscandian Mountains. Ed. Conserv. Jard. Bot. Geneva, Geneva, Switzerland.
Guisan, A., and J.-P. Theurillat. 2000. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**:307-320.
Guisan, A., J. P. Theurillat, and F. Kienast. 1998. Predicting the potential distribution of plant species in an Alpine environment. Journal of Vegetation Science **9**:65-74.
Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters **8**:993-1009.
Guisan, A., S. B. Weiss, and A. D. Weiss. 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecology **143**:107-122.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.
Handa, I. T., C. Korner, and S. Hattenschwiler. 2006. Conifer stem growth at the altitudinal treeline in response to four years of CO2 enrichment. Global Change Biology **12**:2417-2430.
Hill, M. O. 1991. Patterns of species distribution in Britain elucidated by canonical correspondence analysis. J. Biogeogr. **18**:247-255.
Hofer, H. R. 1992. Veränderungen in der Vegetation von 14 Gipfeln der Berninagebietes zwischen 1905 und 1985. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich **58**:39-54.
Houghton, J. T., Y. Ding, and D. J. Griggs, editors. 2001. Climate change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
Houghton, R. A., J. E. Hobbie, J. M. Melillo, B. Moore, B. J. Peterson, G. R. Shaver, and G. M. Woodwell. 1983. Changes in the Carbon Content of Terrestrial Biota and Soils between 1860 and 1980 - a Net Release of Co2 to the Atmosphere. Ecological Monographs **53**:235-262.
Houghton, R. A., and G. M. Woodwell. 1989. Global Climatic-Change. Scientific American **260**:36-44.
Huntley, B., P. M. Berry, W. Cramer, and A. P. McDonald. 1995. Modelling present and potential future ranges of some European higher plants using climate response surfaces. Journal of Biogeography **22**:967-1001.
Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecological Monographs **68**:465-485.
Iverson, L. R., M. W. Schwartz, and A. M. Prasad. 2004. How fast and far might tree species migrate in the eastern United States due to climate change? Global Ecology and Biogeography **13**:209-219.
Johnson, P. L., and W. D. Billings. 1962. Alpine Vegetation of Beartooth Plateau in Relation to Cryopedogenic Processes and Patterns. Ecological Monographs **32**:105-&.
Jolly, W. M., M. Dobbertin, N. E. Zimmermann, and M. Reichstein. 2005. Divergent vegetation growth responses to the 2003 heat wave in the Swiss Alps. Geophysical Research Letters **32**.
Jonasson, S., and T. V. Callaghan. 1992. Root Mechanical-Properties Related to Disturbed and Stressed Habitats in the Arctic. New Phytologist **122**:179-186.
Jones, T. H., L. J. Thompson, J. H. Lawton, T. M. Bezemer, R. D. Bardgett, T. M. Blackburn, K. D. Bruce, P. F. Cannon, G. S. Hall, S. E. Hartley, G. Howson, C. G. Jones, C. Kampichler, E. Kandeler, and D. A. Ritchie. 1998. Impacts of rising atmospheric carbon dioxide on model terrestrial ecosystems. Science **280**:441-443.
Klanderud, K., and H. J. B. Birks. 2003. Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. The Holocene **13**:1-6.
Körner, C. 2003. Alpine plant life. 2nd Ed. Springer.
Kullman, L. 2001. Immigration of Picea abies into North-Central Sweden. New evidence of regional expansion and tree-limit evolution. Nordic Journal of Botany **21**:39-54.
Lahti, T., and R. Lampinen. 1999. From dot maps to bitmaps - Atlas Flora Europaeae goes digital. Acta Botanica Fennica **162**.
Leathwick, J. R., B. R. Burns, and B. D. Clarkson. 1998. Environmental correlates of tree alpha-diversity in New Zealand primary forests. Ecography **21**:235-246.
Malcolm, J. R., A. Markham, R. P. Neilson, and M. Garaci. 2002. Estimated migration rates under scenarios of global climate change. Journal of Biogeography **29**:835-849.
Manel, S., J. M. Dias, S. T. Buckton, and S. J. Ormerod. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. Journal of Applied Ecology **36**:734-747.
Menzel, A., N. Estrella, and P. Fabian. 2001. Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. Global Change Biology **7**:657-666.
Menzel, A., and P. Fabian. 1999. Growing season extended in Europe. Nature **397**:659-659.
Minorsky, P. V. 2002. Global warming - Effects on plants. Plant Physiology **129**:1421-1422.

Moore, D. M., B. G. Lees, and S. M. Davey. 1991. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. Environmental Management **15**:59-71.
Nichols, W. F., K. T. Killingbeck, and P. V. August. 1998. The influence of geomorphological heterogeneity on biodiversity II. A landscape perspective. Conservation Biology **12**:371-379.
Nogués-Bravo, D., M. Araújo, M. P. Errea, and J. P. Martinez-Rica. 2006. Exposure of global mountain systems to climate warming during the 21st century. GLOBAL ENVIRONMENTAL CHANGE.
Núñez-Farfán, J. S., G. D. 2001. Evolution in chaging environments: The „synthetic work of Clausen, Keck, and Hiesey. The Quaterly review of biology **76**.
Parmesan, C. 1996. Climate and species' range. Nature **382**:765-766.
Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature **421**:37-42.
Pauli, H., M. Gottfried, and G. Grabbherr. 1996. Effects of climate change on mountain ecosystems. Upward shifting of alpine plants. World Resource Review **8**:382-390.
Pauli, H., M. Gottfried, D. Hohenwallner, K. Reiter, R. Casale, and G. Grabherr. 2004. The GLORIA field manual. Multi-summit approach. European Commission. Office for Official Publications of the European Communities, Luxembourg.
Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology and Biogeography **12**:361-371.
Penuelas, J., and I. Filella. 2001. Phenology - Responses to a warming world. Science **294**:793-+.
Peterson, A. T., and al. 2002. Future projections for Mexican faunas under global climatic change scenarios. Nature **416**:626-629.
Phillips, O. L., and A. H. Gentry. 1994. Increasing Turnover through Time in Tropical Forests. Science **263**:954-958.
Pitelka, L. F., R. H. Gardner, J. Ash, S. Berry, H. Gitay, I. R. Noble, A. Saunders, R. H. W. Bradshaw, L. Brubaker, J. S. Clark, M. B. Davis, S. Sugita, J. M. Dyer, R. Hengeveld, G. Hope, B. Huntley, G. A. King, S. Lavorel, R. N. Mack, G. P. Malanson, M. McGlone, I. C. Prentice, and M. Rejmanek. 1997. Plant migration and climate change. American Scientist **85**:464-473.
Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems **9**:181-199.
Price, M. V., and N. M. Waser. 2000. Responses of subalpine meadow vegetation to four years of experimental warming. Ecological Applications **10**:811-823.
Randin, C. F., T. Dirnbock, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography **33**:1689-1703.
Rebetez, M. 2002. La Suisse se réchauffe: effet de serre et changement climatique, Presse polytechniques et unversitaires romandes edition, Lausanne.
Rebetez, M., and M. Reinhard. in press. Monthly air temperature trends in Switzerland 1901-2000 and 1975-2004. Theoretical and Applied Climatology.
Richardson, S. J., M. C. Press, A. N. Parsons, and S. E. Hartley. 2002. How do nutrients and warming impact on plant communities and their insect herbivores? A 9-year study from a sub-Arctic heath. Journal of Ecology **90**:544-556.
Rodriguez-Trelles, F., and M. A. Rodriguez. 1998. Rapid micro-evolution and loss of chromosomal diversity in Drosophila in response to climate warming. Evolutionary Ecology **12**:829-838.
Root, T. L., J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds. 2003. Fingerprints of global warming on wild animals and plants. Nature **421**:57-60.
Rounsevell, M. D. A., P. M. Berry, and P. A. Harrison. 2006. Future environmental change impacts on rural land use and biodiversity: a synthesis of the ACCELERATES project. Environmental Science & Policy **9**:93-100.
Schaefer, J. A., and F. Messier. 1995. Scale-Dependent Correlations of Arctic Vegetation and Snow Cover. Arctic and Alpine Research **27**:38-43.

Shugart, H. H., N. H. F. French, E. S. Kasischke, J. J. Slawski, C. W. Dull, R. A. Shuchman, and J. Mwangi. 2001. Detection of vegetation change using reconnaissance imagery. Global Change Biology **7**:247-252.

Sparks, T. H., and A. Menzel. 2002. Observed changes in seasons: An overview. International Journal of Climatology **22**:1715-1725.

Sykes, M. T., and I. C. Prentice. 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.

Taylor, A. H. 1995. Forest expansion and climate change in the mountain hemlock (Tsuga mertensiana) zone, Lassen Volcanic National Park, California, USA. Artic and Alpine Research **27**:207-216.

Theodose, T. A., and W. D. Bowman. 1997. Nutrient availability, plant abundance, and species diversity in two alpine tundra communities. Ecology **78**:1861-1872.

Theurillat, J.-P., F. Felber, P. Geissler, J.-M. Gobat, M. Fierz, A. Fischlin, P. Küpfer, A. Schlüssel, C. Velluti, and G.-F. Zhao. 1998. Sensitivity of plant and soil ecosystems of the Alps to climate change. Pages 225-308 *in* P. Cebon, U. Dahinden, H. C. Davies, D. Imboden, and C. C. Jäger, editors. Views from the Alps. Regional perspectives on climate change. MIT Press, London.

Theurillat, J. P., and A. Guisan. 2001. Potential impact of climate change on vegetation in the European Alps: A review. Climatic Change **50**:77-109.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. Van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from climate change. Nature **427**:145-147.

Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology **10**:2020-2027.

Thuiller, W., M. B. Araujo, R. G. Pearson, R. J. Whittaker, L. Brotons, and S. Lavorel. 2004. Uncertainty in predictions of extinction risk. Nature **430**:34.

Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the United States of America **102**:8245-8250.

Tilman, D. 1994. Competition and Biodiversity in Spatially Structured Habitats. Ecology **75**:2-16.

Tilman, D. 1997. Community invasibility, recruitment limitation, and grassland biodiversity. Ecology **78**:81-92.

Vittoz, P., and A. Guisan. 2003. Le projet PERMANENT.PLOT.CH demande votre collaboration. Das Projekt PERMANENT.PLOT.CH bittet um Ihre Mithilfe. Bot. Helv. **113**:105-110.

Vittoz, P., S. Jutzeler, and A. Guisan. 2006. Flore alpine et réchauffement climatique: observation de trois sommets valaisans à travers le 20ème siècle. Bulletin de la Murithienne **123/2005**:49-59.

Vittoz, P., B. Rulence, T. Largey, and F. Freléchoux. in press. Effects of Climate and Land-Use Change on the Establishment and Growth of Cembran Pine (Pinus cembra L.) over the Altitudinal Treeline Ecotone in the Central Swiss Alps. Arctic Antarctic and Alpine Research.

Walsh, S. J., D. R. Butler, T. R. Allen, and G. P. Malanson. 1994. Influence of snow patterns and snow avalanches on the alpine treeline ecotone. Journal of Vegetation Science **5**:657-672.

Walther, G.-R., S. Beißner, and C. A. Burga. 2005. Trends in the upward shift of alpine plants. Journal of Vegetation Science **16**:541-548.

Walther, G. R. 2003. Plants in a warmer world. Perspectives in Plant Ecology, Evolution and Systematics **6**:169-185.

Walther, G. R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature **416**:389-395.

Woodwell, G. M., F. T. Mackenzie, R. A. Houghton, M. Apps, E. Gorham, and E. Davidson. 1998. Biotic feedbacks in the warming of the earth. Climatic Change **40**:495-518.

Zhou, L. M., C. J. Tucker, R. K. Kaufmann, D. Slayback, N. V. Shabanov, and R. B. Myneni. 2001. Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. Journal of Geophysical Research-Atmospheres **106**:20069-20083.

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.

Chapter 1

Does climate change affect subalpine grasslands in the Swiss Northern Alps?

Vittoz Pascal[1], Randin Christophe[2], Dutoit Annelise[3], Bonnet François[4] & Hegg Otto[5]

This manuscript will be submitted to Global Change Biology.

[1]University of Lausanne, Faculty of Geosciences and Environment, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[3]Ch. du Molard 3, 1805 Jongny, Switzerland

[4]La Thomasia, Jardin alpin de Pont de Nant, 1880 Les Plans-sur-Bex, Switzerland

[5]University of Berne, Geobotanical Institute, Altenbergrain 21, 3013 Bern, Switzerland

Abstract

Climate change is susceptible to have a very strong impact on mountain vegetation. Switzerland is particularly concerned with a warming twice higher than the average for Northern Hemisphere. If phenological shifts and migration of isolated species were already often observed in different ecosystems, very few studies focused on community composition in subalpine grasslands. We used permanent plots since 1935 in Schynige Platte and precisely located phytosociological relevés in 1970 in Vallon de Nant to study composition changes of subalpine grasslands in Swiss Northern Alps. In Schynige, grazing was replaced by mowing and in Nant goat pasturing was abandoned. Old and recent inventories were compared with correspondence analyses (CA) and the sociology of increasing and decreasing species was used to explain the main factor triggering the observed changes. In both regions, subalpine grasslands showed a high stability with little changesThe four inventories in Schynige showed a slow, coherent shift of plant composition with CA, mainly explained by increasing grass density. *Rhinanthus alectorolophus* and *Galium pumilum* arrived recently, probably favoured by warmer conditions as they are closer of their higher altitudinal distribution limit. *Rhinanthus* has replaced another hemiparasite, *Eurphrasia minima*. Changes were a little more important in Nant, but with a strong development of fallows and wood edges species, following pasturing abandonment. However, among them, *Brachypodium pinnatum* and *Trifolium medium* are at their higher altitudinal distribution limit and probably could increase because of warmer climate. Most of the decreasing species had an alpine distribution, but they suffered mainly of the denser and taller grass following grazing end. With a short review of already observed changes in vegetation under climate change, we concluded that plant species can react quickly to the warmer conditions as far as competition is low (e.g., colonisation of rocky summits, structure changes with treeline ecotone shift). However, in subalpine grasslands, competition of already present species is probably important and limit establishment of newly arrived species. A disturbance, like pasturing abandonment, is susceptible to accelerate the changes.

Keywords: competition, correspondence analysis, disturbance, global warming, pastures abandonment, sociology, Switzerland.

Nomenclature: Aeschimann and Heitz (1996)

Introduction

Climatic warming induces a plant migration towards the mountain summits or northern latitudes. Such migration was already often observed with an upward shift of the treeline ecotone since the end of the Little Ice Age (e.g. Taylor 1995, Kullman 2001, Camarero and Gutierrez 2004, Vittoz et al. in press) and a biodiversity increase in plant communities in alpine level (Braun-Blanquet 1975, Pauli et al. 2007) or on alpine-nival summits in the Alps (e.g. Braun-Blanquet 1957, Hofer 1992, Grabherr et al. 1994, Pauli et al. 1996, Walther et al. 2005a, Vittoz et al. 2006). Arctic ecosystems recorded similar changes with a northward shift of the treeline ecotone (Darrigo et al. 1987), invasion of shrubs (Sturm et al. 2001, Klanderud and Birks 2003). In Norway, Klanderud and Birks (2003) observed a biodiversity increase at lower altitudes and a decrease of high-altitude species at their lower-elevation sites as well as an increase of their abundance at the highest altitudes. Parmesan and Yohe (2003) calculated a shift averaging 6.1 km per decade towards the pole. In warmer conditions, migration was observed for plant species, like mistletoe (*Viscum album*) in the Swiss Central Alps (Dobbertin et al. 2005) or *Ilex aquifolium* in undergrowth of Danish and Swedish forests (Walther et al. 2005b).

Mountain ecosystems were identified as potentially very sensitive to global warming in the future (Beniston 1994, Guisan et al. 1995, Beniston et al. 1996, Theurillat et al. 1998, Diaz et al. 2003, Nogués-Bravo et al. 2006) demonstrated that mountains will experience a rate of warming two or three times higher than what was globally recorded during the 20th century. Such a trend was already observed by Rebetez and Reinhard (in press) in Switzerland which gained 1.35 °C during the 20th century at low and high elevations. The warming trend is particularly important since 1975 with a decadal rate of 0.57 °C.

In such conditions, plant distribution models predict a high risk of species extinction. For example, Guisan and Theurillat (2000) found that 2-5% of species are committed to extinction with a 1.5-4.5 °C temperature elevation (62 species considered) but nearly 40% of species were predicted to loose more than 90% of their most suitable habitats. Similarly, Dirnböck et al. (2003) predicted up to 40-50% potential species extinction due to climate and land-use change for 2050 in the Austrian Alps (85 species). Since plant communities are not likely to move as an entity under changing climatic conditions (Huntley 1991), the models may fail to predict the future potential distribution of species in a changing biotic environment (Guisan and Theurillat 2000, Pearson and Dawson 2003, Randin et al. 2006a). In fact, only few studies incorporated mechanistic rules of biotic interactions such interspecific competition or facilitation (Silvertown et al. 1992, Dullinger et al. 2005).

Different processes like facilitation, competition and disturbance could influence plant migration. Choler et al. (2001) demonstrated that facilitation in alpine plant communities appeared to allow species from lower elevation to move up the gradient, Conversely, Dullinger et al. (2005) found a negative effect of pine on the recruitment of spruce and larch at the limit of the subalpine zone, thus suggesting that plant-plant interactions could decrease the rate of migration of species under climate warming. The theory of fluctuating resource (Burke and Grime 1996) states that ecosystems are particularly vulnerable to invasions when resources become unused, following the destruction of a part of the

vegetation after a disturbance or by an exogenous income of nutrients (eutrophization). Fertilizing alpine communities is expected to have a large impact on plant communities (e.g. Chapin and Shaver 1985, Korner et al. 1997, e.g. Theodose and Bowman 1997). In addition, Alpert and Maron (2000) showed that eutrophization can facilitate settlement of species with a strong production capacity. In the context of global change, other disturbances like grazing abandonment (Hobbs and Huenneke 1992) and nitrogen input by air pollution (Galloway et al. 1995, Galloway et al. 2003) are thus susceptible to accelerate plant migration in otherwise stable communities.

If many studies measured phenology or growth of individual species in natural or experimental conditions (e.g. Harte and Shaw 1995, van Wijk et al. 2004, Hollister et al. 2005, Sebastia 2007) or monitored plant migration in open alpine-nival vegetation (e.g. Walther et al. 2005a, e.g. Pauli et al. 2007), only few studies really focused on changes in composition or dominance within grasslands in lower altitudes (Price and Waser 2000). Permanent plots might be the best tool for such studies but long-term monitoring are mostly recent (see projects PERMANENT.PLOT.CH; Vittoz and Guisan 2003, GLORIA; Pauli et al. 2004, MIREN; Dietz et al. 2006).

In order to study the influence of climate change in the second part of the 20th century on plant communities in subalpine level, we used the Schynige Platte permanent plots, among the oldest ones in Switzerland without important change since 1930 (Lüdi 1936, Hegg 1992). We completed with data from the Vallon de Nant, where precisely located inventories were realized in 1970 at the end of pasturing exploitation (Dutoit 1984). We hypothesized that in these grasslands communities at subalpine level, (i) climate change had only a restricted influence because of the competition of present species towards potential new species in subalpine grasslands and (ii) the changes were more important in Vallon de Nant after grazing abandonment than in Schynige Platte under almost constant management.

Location and methods

Study sites

Both study sites are located in the Northern Alps biogeographic Swiss region (Fig. 1; Gonseth et al. 2001). Climatic conditions are frequently wet as it is the first barrier to the lows coming from Atlantic Ocean. Annual sum of precipitations is 1600 mm in Vallon de Nant (46° 14' N, 7° 06' E; below designated by Nant) and in Schynige Platte (46° 39' N, 7° 55' E; below designated by Schynige), regularly distributed in the year (Zimmermann and Kienast 1999).

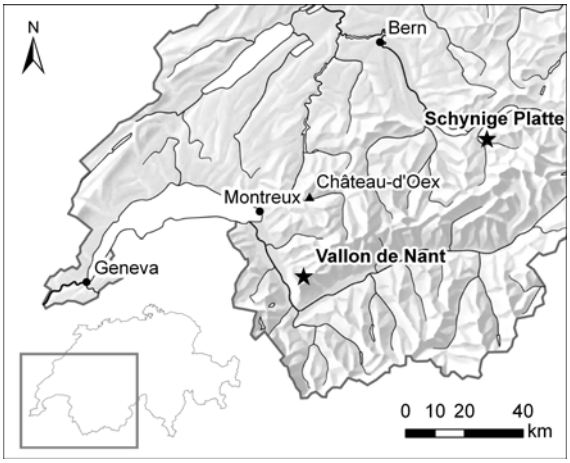


Figure 1 Location in Switzerland of the study sites (Vallon de Nant and Schynige Platte), the Château-d'Oex meteorological station and some important towns.

These sites belong to the subalpine level, naturally dominated by *Picea abies* forests, but often converted in pastures for cattle breeding since hundreds of years. The plots in Nant were situated on 2.5 km, along a valley, between 1370 and 1815 m a.s.l, what corresponds to the whole range of the subalpine level. The approximate mean annual temperature ranges between 3.5-5.3 °C (Zimmermann and Kienast 1999). In Schynige, all the plots were grouped on a restricted area (0.3 ha) at 1920 m a.sl., at the upper limit of the subalpine level. The mean annual temperature is 1.8 °C (Zimmermann and Kienast 1999).

Nant was previously grazed by cattle and goats. Goat breeding was seriously reduced in 1940 and abandoned in 1970 (Dutoit 1983). Only a small herd of cattle is still grazing the valley, mainly the flat areas. Many steep slopes are slowly colonized by *Alnus viridis*, and the other grasslands are grazed mainly by wild chamois (*Rupicapra rupicapra*). The Schynige Platte was grazed by cattle as well, but when Lüdi set the permanent plots (Lüdi 1936), he fenced the research areas. The device was given up after 1956 and cattle grazed again the area. In 1981, Hegg settled again the fences and recovered all the permanent plots. Now, the area is mown every two years (Dähler 1993).

In the region, temperature during the 20th century measured in MeteoSwiss meteorological station in Château-d'Oex (980 m a.s.l., Fig. 2) followed the same trend as recorded in

average for Switzerland (Rebetez and Reinhard in press), with a strong increase since 1975 (Fig. 2).

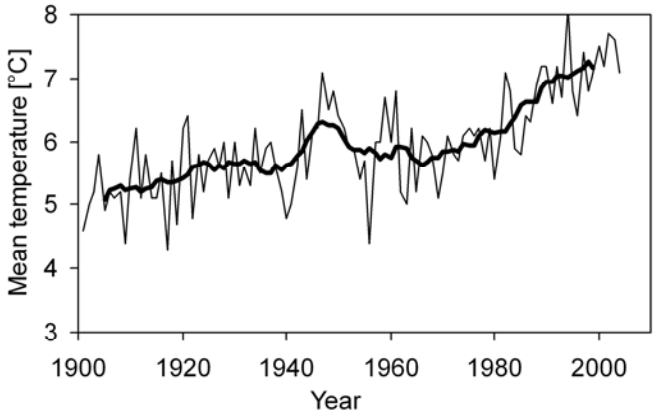


Figure 2 Temperature change during the 20th century in Château-d'Oex (980 m).

Field samplings

The vegetation of Nant was described by Dutoit (1983). The description laid on vegetation inventories following Braun-Blanquet's method (Braun-Blanquet 1964). The plots were not marked on the field but the geographic coordinates were precisely measured on a 1:10'000-map. As the valley is characterized by numerous irregularities (small ridges, gullies, very large stones, groups of trees) accurately drawn on the map, the confidence is often very high. We retained for our study the plots situated in the two dominant grassland types in the subalpine level. Both types are on calcareous soils but the *Seslerion albicantis* occupies the driest slopes (mainly on southern aspect or very steep) when the *Caricion ferrugineae* is more shady or on a more humid soil. These vegetation types share many common species, like the Poaceae *Calamagrostis varia* and *Sesleria caerulea* or the Cyperaceae *Carex sempervirens* and *C. ferruginea* which are present in almost all plots and dominate them together or alternatively. All the retained plots were inventoried between July and September in 1970, except one in 1969 and one in 1978.

In 2006, we looked for the location of the plots with a Global positioning system (GPS, Trimble GeoXT) on the basis of the geographic coordinates. Once at the right coordinates, we selected the closest site, in a radius of 30 m, with the right slope, aspect and position relative to the surface morphology. When we could be confident with the plot position, we retained it. However, when position confidence was not high enough or the place covered by shrubs (mainly *Alnus viridis*), we gave up to use it. The plots were inventoried following Braun-Blanquet (1964) on the same area as Dutoit (1983), i.e. generally 10 m^2. A similar method was already used with success in Swiss dry meadows (Fischer and Stöcklin 1997) and forests (Carraro et al. 1999). In order not to miss species which were recorded in 1970 but were outside of the new plots because of small location shifts, we completed the inventories with supplementary species visible on 100 m^2 around the plots (called extended plots). The 19 retained plots (6 in *Seslerion albicantis* and 13 in *Caricion ferrugineae*) were

distributed on both side of the valley, with a east (ENE-SE) or west (WSW-NW) aspect, and slopes ranging between 22-52° (mean 33.7 ± 7.6°).

From 1930, Lüdi started an experiment with permanent plots in Schynige to find ways to improve quality and yields of pastures dominated by *Nardus stricta*, a Poaceae refused by cattle (Lüdi 1936). He set 340 permanent plots, marked with small wooden stakes. Each was 1 m^2, and they were grouped following different fertilisation experiments (combination of NPK, manure, limestone). In each group, some plots were left without fertilisation as control plots. The sampling frequency varied through time, but we could retained 23 plots which were never fertilized and inventoried around 1935 (1932-1935), 1954 (1953-1954), in 1990 and 1999. Plant cover was estimated in percents, with 0.5 % as the smallest value. All the plots had a SSE-aspect with a slope of 20° (Dähler 1993).

Data analyses

Three types of problems are frequent in such monitoring: mistakes in species identification, overlooked species and differences in cover estimate (e.g. Leps and Hadincova 1992, Klimes et al. 2001, Scott et al. 2002, Klimes 2003, Vittoz and Guisan 2007). We limited the identification mistakes by grouping in analyses all the close species, with frequent confusions, especially in vegetative state (e.g., *Dactylorhiza maculata/Orchis mascula, Festuca rubra aggr./F. ovina aggr., Hieracium lactucella/H. pilosella/H. aurantiacum, Leontodon hispidus/L. helveticus*). This was realized on our own experience of frequent confusions and by observing the regular substitutions of species by close relatives between years. There is no way to check for overlooked species and we had to be conscious of differences in data quality. In Nant, Dutoit (1983) inventoried the plots alone in 1970, when we were 2-3 botanists in 2006. This is susceptible to spuriously increase the number of species (Vittoz and Guisan 2007). In Schynige, Lüdi had firstly an agronomic interest and it is probable that he did not spend much time to inventory rare and little covering species during the 1935-inventory. Later, the plots were inventoried with larger interests and lists are probably more complete. Finally, to be able to compare the data from both regions and to lessen inaccuracies in visual cover estimates in percent or between years cover fluctuations, we converted for the analyses all the cover estimates in the same scale (Table 1).

Table 1 Scales used for plant cover estimates in Vallon de Nant, corresponding values in percents in Schynige Platte and transformations used in analyses.

Braun-Blanquet scale in Vallon de Nant	Percents in Schynige Platte	Values used in analyses
r		0.1
+	0.5	0.5
1	1-5	1
2	6-25	2
3	26-50	3
4	51-75	4
5	76-100	5

Correspondence analysis (CA; Bénzécri 1973) was performed in R (R 2.4.1 2006) separately in both sites to put forward possible coherent shifts of the plots between the different

inventories (Berlin et al. 2000, Vittoz and Hainard 2002, Köhler et al. 2005). Species with less than 3 observations in all the plots and years were previously removed from the data table. To interpret vegetation change, in terms of changing environmental factors, weighted average ecological values were calculated for all inventories using indicator values attributed to each species (Landolt 1977). These vary between 1 and 5 for increasing light (L), temperature (T), soil humidity (F), humus amount (H), nutrient amount (N) and pH (R). The average values were correlated with the first two axes of the correspondence analysis (Pearson correlation). Similarly, in Vallon de Nant, the axes were correlated with plot altitude, slope and calculated radiation (total amount of energy received in June-August related to the slope and aspect; Zimmerann and Kienast 1999). Finally, to quantify the difference between two relevés, we used the Whittaker percentage similarity, as employed in Legendre and Legendre (1979) and averaged the results for each year.

Increasing species were selected following three different criterions: (N) for the new species, absent in the first inventory and present in at least 3 plots in one of the following inventories; (F) for species with an increasing frequency of 3 plots or more between the first and last inventory; (C) for species with an increasing mean cover of 0.3 or more between the first and last inventory and present in at least 3 plots at each inventory. In Schynige, strongly fluctuating species between inventories (succession of increases and decreases) were not retained. Similarly, the decreasing species were selected following three criterions: (d) for the disappeared species, present in at least 3 plots in one of the inventories and absent in the last one; (f) for species with a decreasing frequency of 3 plots or more between the first and last inventory; (c) for species with a decreasing mean cover of 0.3 or more between the first and last inventory and present in at least 3 plots at each inventory. In Nant, the d and f criterions were established by comparing the original data with the species list in the extended plots and in Schynige, strongly fluctuating species between inventories (succession of increases and decreases) were not retained. The sociologic affinity of decreasing and increasing species was compared with the complete list of species in each site following Ellenberg et al. (1991). Finally, we calculated an elevation index with the altitudinal range indicated by Lauber and Wagner (1991). This reference takes into account the species distribution in the Northern Swiss Alps and gives for each one its optimum and possible distribution in altitudinal levels. We attributed a 1 for the collinean level, 2 for the mountain, 3 for subalpine and 4 for the alpine level. The weighted mean was calculated for each species with a 2 coefficient for the optimum and a 1 coefficient for a possible distribution. These means were compared with a Wilcoxon test.

Results

The number of inventoried species increased in Nant between 1970 and 2006 of 6.8 ± 4.2 species in *Seslerion albicantis* and of 6.0 ± 6.7 in *Caricion ferrugineae* (Fig. 3). The fluctuations were more irregular in Schynige (Fig. 3) with firstly an strong increase from 1935 to 1954 (10.7 ± 4.3) followed by a decrease (8.0 ± 6.5 from 1954 to 1999).

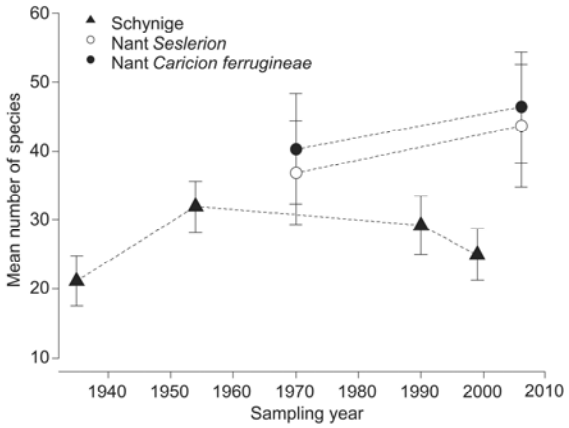


Figure 3 Mean number of species (with standard deviation) observed in the different samplings in the Vallon de Nant (1970 and 2006) and the Schynige Platte (1935, 1954, 1990 and 1999). The Vallon de Nant data are separated in the two vegetation types.

In both sites, CA showed a general trend in the vegetation. In Nant, most of the plots shifted between 1970 and 2006 in the direction of the inferior right corner (Fig. 4). The shift was approximately similar in both vegetation types (*Seslerion albicantis* and *Caricion ferrugineae*).

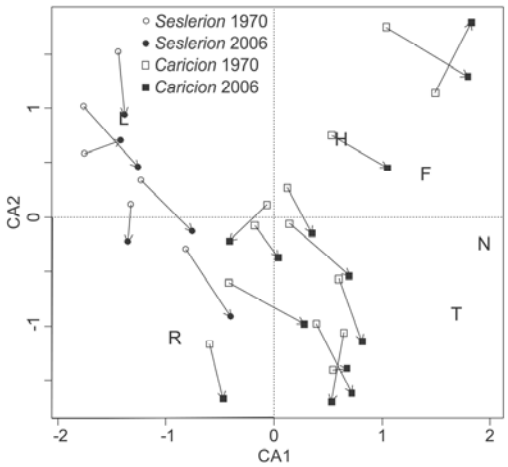


Figure 4 Correspondence analysis with the relevés of Vallon de Nant. Symbols separate sampling years and vegetation types (*Seslerion albicantis* and *Caricion ferrugineae*). Arrows are indicating the shift of each plot. The first axis represents 11.4% of the total variance and the second 7.4%. The ecological indicator values (Landolt 1977) are projected by using the Pearson correlation of the weighted average of each inventory with its position along the two axes (correlation is multiplied by 2).

The first CA axis was correlated with increasing light, temperature, soil humidity and nutriment amount and with a decreasing pH (Table 2). The second axis was correlated with increasing light and humus and with decreasing temperature and pH (Table 2). Moreover, there was a negative correlation of the first axis with altitude (r=-0.502, p=0.001) and slope (r=-0.428, p=0.007) and a positive correlation of the second axis with altitude (r=0.393, p=0.015).

Table 2 Pearson correlation of mean ecological values (Landolt 1977) with the first two axes of the CA for Vallon de Nant and Schynige Platte. NS not significant, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

	Vallon de Nant		Schynige Platte	
	Axis 1	Axis 2	Axis 1	Axis 2
L	-0.720****	0.460**	0.297**	0.559****
T	0.818****	-0.440**	-0.374***	-0.326**
F	0.676****	0.203^{NS}	0.077^{NS}	0.422****
H	0.280^{NS}	0.364*	-0.539****	-0.142^{NS}
N	0.940****	-0.116^{NS}	0.491****	0.323**
R	-0.489**	-0.550***	0.822****	-0.042^{NS}

In Schynige, the plots shifts between inventories were irregular (not shown), with high interannual fluctuations, but the plots showed together a constant shift between 1935 and 1999 (Fig. 5). The first axis of the CA was correlated with increasing light, nutrient and pH and with decreasing temperature and humus. The second axis was correlated with light, humidity and nutrient and with decreasing temperature (Table 2). However, most of these correlations are rather low.

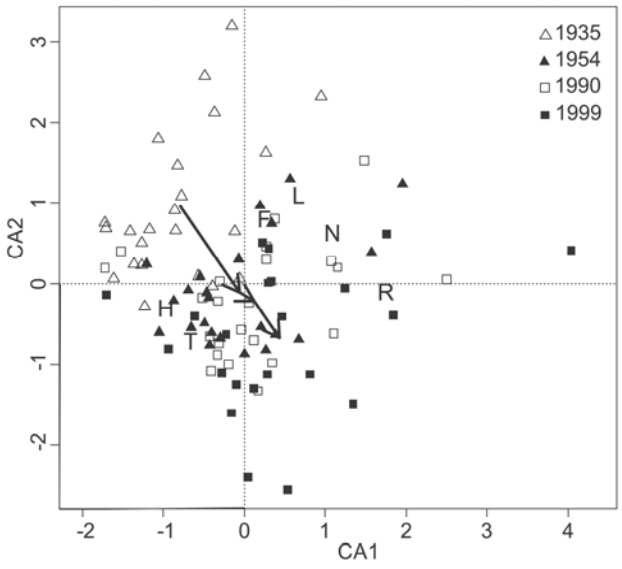


Figure 5 Correspondence analysis with the relevés of Schynige Platte. Symboles separate sampling years. Arrows are indicating the shift of the average position of each year. The first axis represents 8.0% of the total variance and the second 7.1%. The ecological indicator values (Landolt 1977) are projected by using the Pearson correlation of the weighted average of each inventory with its position along the two axes (correlation is multiplied by 2).

The Whittaker percent similarity showed that the mean changes were lower in Schynige than in Nant. Mean values between 1970 and 2006 were 0.77 ± 0.11 for *Seslerion albicantis* and 0.85 ± 0.09 for *Caricion ferrugineae* in Nant. In Schynige, the mean values decreased through the monitored period: 0.75 ± 0.16 in the period 1935-1954, 0.65 ± 0.10 in the period 1954-1990 and 0.56 ± 0.11 in the period 1990-1999. It was 0.82 ± 0.15 if the complete period 1935-1999 was considered.

36 species were considered as increasing and 12 were as decreasing in Nant (Table 3 and 4). *Centaurea montana* was the only new species, *Heracleum sphondylium, Galium anisophyllon, Carex ornithopoda, Rubus saxatilis* the highest frequency increases and *Trifolium medium, Brachypodium pinnatum, Calamagrosits varia* the highest cover increases. Only *Trifolium badium* disappeared, no species decreased in frequency and *Carex sempervirens, Alchemilla conjuncta aggr., Rhinanthus alectorolophus* showed the highest cover decreases. These species showed mainly an increase of species associated to fallows, wood edges or with a large ecology and a strong decrease of species associated to alpine grasslands (Fig. 6). Altogether, increasing species have a significantly lower elevation index than decreasing species (Fig. 7).

Table 3 Increasing species in Vallon de Nant. Change is the type of increase: **N** new species, **F** frequency increase and **C** cover increase.

	Frequency			Mean cover			Change
Species	1970	2006	Increase	1970	2006	Increase	
Centaurea montana	0	3	3		0.10		N
Festuca rubra aggr.	8	12	4	0.94	1.26	0.32	FC
Brachypodium pinnatum	3	6	3	0.67	1.42	0.75	FC
Aposeris foetida	3	7	4	0.50	1.01	0.51	FC
Chaerophyllum villarsii	3	6	3	0.50	0.95	0.45	FC
Heracleum sphondylium s.l.	3	10	7	0.50	0.43	-0.07	F
Galium anisophyllon	9	15	6	0.50	0.23	-0.27	F
Carex ornithopoda	2	7	5	0.50	0.27	-0.23	F
Rubus saxatilis	2	7	5	0.75	0.54	-0.21	F
Petasites paradoxus	2	6	4	0.30	0.68	0.38	F
Campanula scheuchzeri	2	6	4	0.50	0.37	-0.13	F
Carduus defloratus s.str.	8	12	4	0.51	0.30	-0.21	F
Daphne mezereum	1	5	4	0.10	0.34	0.24	F
Hieracium murorum aggr.	10	14	4	0.55	0.44	-0.11	F
Laburnum alpinum	1	5	4	0.50	0.18	-0.32	F
Ranunculus montanus	11	15	4	0.64	0.48	-0.16	F
Selaginella selaginoides	2	6	4	1.25	0.30	-0.95	F
Viola biflora	1	4	3	0.50	0.90	0.40	F
Agrostis stolonifera	4	7	3	1.25	1.01	-0.24	F
Alchemilla vulgaris aggr.	1	4	3	1.00	0.80	-0.20	F
Astrantia major	12	15	3	0.88	0.87	-0.01	F
Briza media	5	8	3	0.50	0.58	0.08	F
Crepis pyrenaica	10	13	3	0.66	0.70	0.04	F
Gentiana lutea	4	7	3	0.40	0.40	0.00	F
Knautia dipsacifolia s.str.	9	12	3	0.78	0.76	-0.02	F
Laserpitium latifolium	10	13	3	0.90	1.06	0.16	F
Leucanthemum vulgare aggr.	16	19	3	0.59	0.51	-0.08	F
Polygonum viviparum	7	10	3	0.50	0.30	-0.20	F
Potentilla erecta	9	12	3	0.83	0.61	-0.23	F
Salix appendiculata	1	4	3	0.10	0.30	0.20	F
Soldanella alpina	2	5	3	1.50	0.50	-1.00	F
Vincetoxicum hirundinaria	5	8	3	0.50	0.63	0.13	F
Trifolium medium	4	5	1	0.50	1.42	0.92	C
Calamagrostis varia	16	17	1	2.31	2.94	0.63	C
Carex flacca	5	7	2	0.60	1.00	0.40	C
Larix decidua	3	3	0	0.37	0.67	0.30	C

Table 4 Decreasing species in Vallon de Nant. The disappearance of species is evaluated between the original relevé and the extended area of the new relevé. The Change is the type of decrease: d disappeared species and c cover decrease.

Species	Frequency (extended plot)			Frequency			Mean cover			Change
	1970	2006	Decrease	1970	2006	Decrease	1970	2006	Decrease	
Trifolium badium	3	0	-3	3	0	-3	0.50			d
Carex sempervirens	15	18	3	15	16	1	3.67	2.06	-1.60	c
Alchemilla conjuncta aggr.	13	18	5	13	13	0	1.27	0.46	-0.81	c
Hedysarum hedysaroides	3	4	1	3	4	1	1.33	0.63	-0.71	c
Rhinanthus alectorolophus	9	12	3	9	9	0	1.06	0.42	-0.63	c
Anthyllis vulneraria subsp. alpestris	8	10	2	8	8	0	1.00	0.40	-0.60	c
Erica carnea	6	7	1	6	4	-2	1.58	1.00	-0.58	c
Trifolium pratense s.str.	9	11	2	9	8	-1	0.94	0.40	-0.54	c
Dryas octopetala	4	4	0	4	3	-1	1.50	1.03	-0.47	c
Linum alpinum	6	12	6	6	6	0	0.83	0.47	-0.37	c
Polygala chamaebuxus	10	12	2	10	10	0	0.90	0.57	-0.33	c
Helianthemum nummularium s.l.	16	18	2	16	17	1	1.60	1.29	-0.31	c
Euphrasia hirtella	5	7	2	5	4	-1	0.50	0.20	-0.30	c

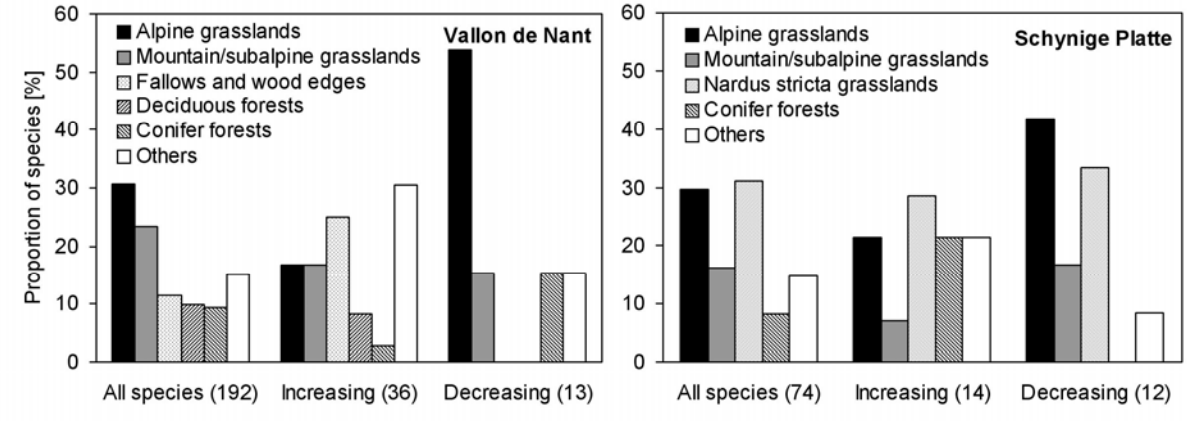


Figure 6 Sociologic affinity (Ellenberg et al. 1991) of the species in the Vallon de Nant (left) and Schynige Platte (right): for the complete list of observed species (all species), the increasing and the decreasing species. The number of species considered in each category is between brackets.

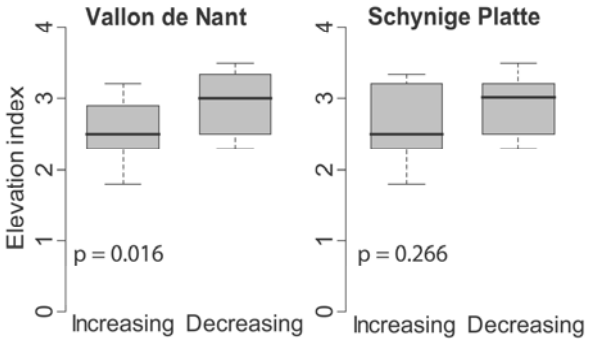


Figure 7 Elevation index of the increasing and decreasing species in Vallon de Nant (left) and Schynige Platte (right). The central lines are the mean, the boxes the second and third quartiles and horizontal lines indicate the outliers. The p-values are from a Wilcoxon test.

13 species were considered as increasing and 12 as decreasing in Schynige (Table 5 and 6). 8 species were new, but most of them appeared with a high frequency in 1954, what can be associated with incomplete inventories in 1935. Only *Rhinanthus alectorolophus* and *Galium pumilum, Ajuga reptans* seemed to be really new. Similarly, all the increasing species in frequency had their change before 1954 and all the cover increase were low. *Trifolium repens, Euphrasia minima* were the disappeared species with the highest original frequency. *Crepis aurea, Plantago alpina, Carex pallescens* showed the most important frequency decrease. The sociologic affinity of this small number of species (Ellenberg *et al.* 1991) was unclear compared to the complete list of species (Fig. 6). Species with large ecology and from conifer forests gained the most (only 3 species each) and alpine grasslands lost the most (5 species). There was no significant difference of elevation index between increasing and decreasing species (Fig. 7).

Table 5 Increasing species in Schynige Platte. Change is the type of increase: **N** new species, **F** frequency increase and **C** cover increase.

	Frequency					Mean cover					Change
Species	1935	1954	1990	1999	Increase	1935	1954	1990	1999	Increase	
Rhinanthus alectorolophus	0	0	1	15	15			0.5	0.6	0.1	NF
Galium pumilum	0	1	5	6	6		0.5	0.5	0.7	0.2	NF
Homogyne alpina	0	13	17	15	15		0.6	0.8	1.0	0.4	NC
Carlina acaulis s.l.	0	3	4	4	4		0.7	1.0	1.0	0.3	NC
Polygala alpestris	0	12	7	8	8		0.5	0.5	0.6	0.1	N
Stellaria graminea	0	5	4	6	6		0.5	0.5	0.7	0.2	N
Ajuga reptans	0	1	3	3	3		0.5	0.7	0.7	0.2	N
Polygala chamaebuxus	4	19	20	20	16	0.9	0.8	0.9	0.9	0.1	F
Campanula scheuchzeri	2	16	15	15	13	0.5	0.5	0.6	0.7	0.2	F
Bartsia alpina	1	14	14	13	12	0.5	0.6	0.6	0.9	0.4	F
Trifolium badium	1	7	5	5	4	1.0	0.6	0.6	0.7	-0.3	F
Crepis conyzifolia	16	22	23	21	5	0.9	1.1	1.1	1.2	0.3	C
Vaccinium myrtillus	17	22	22	20	3	0.9	1.1	0.8	1.2	0.3	C
Gentiana acaulis	18	19	21	20	2	0.7	0.7	0.8	1.0	0.3	C

Table 6 Decreasing species in Schynige Platte. Change is the type of decrease: **d** disappeared species, **f** frequency decrease and **c** cover decrease.

	Frequency					Mean cover					Change
Species	1935	1954	1990	1999	Decrease	1935	1954	1990	1999	Decrease	
Trifolium repens s.str.	7	0	1	0	-7	0.9		0.5		-0.4	d
Euphrasia minima	5	6	16	0	-5	0.5	0.6	0.5		0.0	d
Poa alpina	3	1	0	0	-3	0.7	0.5			-0.2	d
Ligusticum mutellina	2	3	2	0	-2	1.0	0.5	0.8		-0.2	d
Crepis aurea	20	15	11	3	-17	1.1	0.7	0.6	0.7	-0.4	fc
Calluna vulgaris	16	20	7	8	-8	1.8	1.2	1.1	1.0	-0.8	fc
Plantago alpina	23	23	18	9	-14	1.1	1.0	0.8	0.9	-0.1	f
Carex pallescens	19	12	17	8	-11	0.8	0.6	0.6	0.8	0.0	f
Festuca rubra aggr.	16	20	13	7	-9	0.8	1.1	0.8	1.1	0.3	f
Phleum alpinum	11	15	14	4	-7	0.7	0.7	0.6	0.8	0.0	f
Alchemilla vulgaris aggr.	5	6	4	2	-3	0.7	0.6	0.6	1.0	0.3	f
Arnica montana	23	23	23	21	-2	1.9	2.5	1.2	1.4	-0.4	c

Discussion

Vegetation changes

Altogether, observed changes in both sites were rather limited. Structuring and dominant species stayed identical in both sites and changed only in cover. In Schynige, new species represented about 10% on a 65-years survey (Table 5), whereas in the same time frame biodiversity increase higher than 100% were frequently recorded on alpine-nival summits in the Alps (e.g. Grabherr et al. 2001, Walther et al. 2005a, Vittoz et al. 2006). The proportion of new species was even lower in Nant (0.5%; Table 3) but 16% showed an increasing cover or frequency. In approximately the same period, alpine-nival summit biodiversity increased by 26% in Grisons (Walther et al. 2005a).

The coherent change observed in the plots of Nant (Fig. 4) was mainly related to pasturing abandonment in the valley. Fallows and wood edges species (e.g., *Centaurea montana, Chaerophyllum villarsii, Trifolium medium*) represented almost the majority of the increasing group, only exceeded by species with large ecology. But among them, many are common in such conditions as well *(Astrantia major, Carex flacca, Gentiana lutea, Rubus saxatilis;* Oberdorfer 1990*)*. The developments of *Brachypodium pinnatum* and *Trifolium medium* were particularly relevant as they colonize efficiently abandoned grasslands (Köhler et al. 2005). However, these species developed at the alpine grasslands' detriment (Fig. 6), like *Trifolium badium, Carex sempervirens*, *Alchemilla conjuncta* or *Anthyllis vulneraria subsp. alpestris*. These alpine species probably suffered from the competition with taller and denser plants which are no more hampered by grazing (Poyry et al. 2006), but an influence of climate warming may not be excluded. Decreasing species had a higher elevation index than increasing species (Fig. 7) and some of the increasing are at their higher altitudinal distribution limit in the region, like *Brachypodium pinnatum, Trifolium medium* or *Vincetoxicum hirundinaria* (Lauber and Wagner 1991, data from MODIPLANT project; Guisan 2005, Randin et al. 2006b). All these changes were translated on the CA plot with a general shift correlated with decreasing light conditions (more species of shadowy situations) and increasing temperature following the correlations with ecological indicator values (Fig. 4).

In Schynige, changes were unidirectional as well (Fig. 5) but weak, especially after 1954 and with a low variance explained by axes. No clear trend with ecological indicator values was visible with CA plot, with often contradictory changes along both axes. Most of the new species or with an increasing frequency were between 1935 and 1954 (Table 5). This important change has to be considered with cautious because of perhaps incomplete inventories in 1935. Only three species (*Rhinanthus alectorolophus, Galium pumilum, Ajuga reptans*) showed clear, recent increases, especially important for the first one between 1990 and 1999. Future inventories will show whether it is a real trend or temporary fluctuations, but these species are at their higher altitudinal distribution limit (Lauber and Wagner 1991). The high climate warming rate of the last decades (Fig. 2; Rebetez and Reinhard in press) may explain their recent development. Decreasing species belonged to different groups, with a small majority of alpine grasslands, but on a too limited number of species for a solid interpretation (Fig. 6). Some are low growing species (*Euphrasia minima, Crepis aurea,*

Plantago alpina, Arnica montana) which may have suffered of the change from an annual grazing to biannual mowing, corresponding with an increasing density of grasses (result not shown). This same management change may explain the decrease of *Calluna vulgaris*, which suffers from mowing when avoided by cattle, and the increase of *Vaccinium myrtillus*, with opposite characteristics (Welch 1998).

Climate change and mountain vegetation

Influence of climate change on vegetation was already demonstrated with numerous examples (Walther et al. 2002, Parmesan and Yohe 2003). In mountains, upward shifts were often recorded but mostly for species in situation of low competition.

Inventories on alpine-nival summits in the Alps showed since a long time (Braun-Blanquet 1957) the increasing plant diversity during the 20th century (Hofer 1992, Grabherr et al. 1994, 2001, Walther et al. 2005a, Vittoz et al. 2006). Climate is the major factor limiting their altitudinal distribution and species colonize areas dominated by rocks, with a low plant cover and thus a low competition. Similarly, the observed upward shift of the altitudinal treeline ecotone in many mountain regions (Taylor 1995, Kullman 2001, Camarero and Gutierrez 2004, Vittoz et al. in press) was possible because trees, by their taller size on alpine grasslands, did not enter in competition with already present species.

Changes at the community level were rarely recorded under climate warming for vegetation. Braun-Blanquet (1975) observed during 26 years a snowbed community in the Alps and recorded a strong increase of cover (12-80%) and flowering plant diversity (3-10 species) between 1921 and 1947. But, like for forest altitudinal shift, there was a structure change from mosses dominated vegetation in the middle of rocks to taller flowering plants. Such a structure change was recorded as well in a summer warming and winter snow manipulation experiment in Alaskan arctic tundra. Shrub cover and size increased to the prejudice of shadow sensitive lichens and bryophytes (Wahren et al. 2005).

At lower altitudes, a 200 m altitudinal shift was recorded for the mistletoe (*Viscum album* subsp. *austriacum*) in Swiss Central Alps during the 20th century (Dobbertin et al. 2005). This parasite of *Pinus sylvestris* could quickly react to climate warming as its host was present almost 1000 m higher than *Viscum* original distribution and there is no other species in this niche. Similarly, *Ilex aquifolium*, whose distribution is mainly limited by winter temperature, extended recently its distribution to the north (Walther et al. 2005b). This migration was facilitated by its regular presence in gardens, but the low competition in forest undergrowth for shade tolerant species like *Ilex aquifolium* was certainly determinant as well.

Vegetation changes in community with dominance shift was observed in South Switzerland forests where evergreen broad-leaved species became more competitive than deciduous species following a lengthening of the growing season and a strong decrease in frost days (Walther 2002, Walther et al. 2002). Introduced evergreen species, like *Cinnamomum glanduliferum* or *Trachycarpus fortunei*, are now dominating the forest undergrowth and locally replace the native deciduous *Castanea sativa* at the tree layer (Walther 2002).

Conversely to all these studies recording major changes, our survey showed that changes under climate warming were up to now very limited in semi-natural subalpine grasslands. Newly incoming species, close to their physiological limit for temperature, have probably to compete with already growing ones. Moreover, interannual fluctuations of non-temperature factors may dampen the progression of the species (Hollister et al. 2005). Furthermore, the high longevity of mountain plants may slow species replacement (Theurillat and Guisan 2001). Choler et al. (2001) observed that facilitation might help species to move up along the altitudinal gradient, but this advantage could be limited to small species with an isolated growing form, like *Rhinanthus alectorolophus, Galium pumilum, Ajuga reptans* in Schynige. These can grow in structuring grasses environment, without suffering from a too high competition. *Rhinanthus alectorolophus*, as a hemiparasite with preference for Poaceae (Hartl and Wagenitz 1975), avoids even their competition and has a facilitated migration because of its wide host range (Phoenix and Press 2005).

Absence of changes was found in a four years warming experiments in subalpine grassland in West-Central Colarado (Price and Waser 2000). The authors explained this result, compared to similar experiments in Arctic tundra with high changes, by summer dry conditions which cancelled out effects of earlier snowmelt. But drought is only rarely a problem in the Northern Alps and is certainly not the cause of the low observed changes. Another hypothesis is insufficient dispersal of low subalpine or montane plants to colonize the studied sites (Malcolm et al. 2002). However, this could result from higher changes in Schynige than in Nant because of the isolated situation of this valley separated from lower altitudinal pastures by a large forested area and a complex topography.

More important changes were observed in Nant than in Schynige. Whittaker percent similarity showed the same rate of changes for 65 years in Schynige as for 36 years in Nant. In addition, the number of increasing species in Nant was more important than in Schynige, especially without considering the dubious new species between 1935 and 1954. Those differences are probably to relate to the pasturing abandonment in Nant, acting as a disturbance on the vegetation (Hobbs and Huenneke 1992). It was observed that small-scale disturbances are the driving force for species changes and fluctuations in communities (van der Maarel 1996, Gigon 1997). The end of grazing changes competition balance between species: unpalatable or little palatable species which are advantaged under grazing, like *Carex sempervirens* or *Alchemilla conjuncta*, are replaced by palatable ones, generally limited by grazing, like *Trifolium medium* (Troxler pers. com.). Falling rocks or cryoturbation are other possible disturbances in the steep slopes of Nant, generally surmounted by cliffs. Whatever the disturbance, some of thermophilous species took this opportunity to move upslope under better climatic conditions.

In both sites, but with a stronger trend in Nant, decreasing species were mainly alpine plants. Alpine and arctic plants react more with a development of reproductive organs than with a growth increase under warmer conditions (Hollister et al. 2005). Moreover, they are generally low and thus little competitive in front of tall grasses and herbs which were no more grazed in Nant. This corresponds with previous hypothesis of alpine species exclusion by light competition with taller subalpine plants (Theurillat and Guisan 2001) and

observation of competition exclusion in the lowest part of the species distribution (Choler et al. 2001). The disappearance of *Euphrasia minima* in Schynige seems particularly interesting as this species seems to have directly suffered of the arrival of *Rhinanthus alectorolophus*. Both are unspecialized hemi-parasites (Hartl and Wagenitz 1975) but *Euphrasia* was at its lower altitudinal limit and was replaced by *Rhinanthus* during the warmest decade in the 20th century (Houghton et al. 2001).

The high stability of subalpine grasslands to climate warming could lead to a vegetation composition more and more unbalanced with climatic conditions. Experiments and data are insufficient to guess precisely what will happen in the future. However, it is conceivable that major and quick changes will follow the dieback of one structuring species after exceptional climatic conditions (e.g., severe drought), pathogens outbreak suddenly able to develop in higher altitudes or a combination of these factors, like commonly observed with trees (e.g. Rebetez and Dobbertin 2004, Suarez et al. 2004).

Conclusion

Plant migration under climate warming may be quick when the species are not submitted to competition with already present plants. This can occure when the new species occupy a free niche in the plant community, or when there is a clear structural change, like invading trees in alpine grasslands. In such situation, plants can keep in pace with the climate change. However, in the case of subalpine grasslands, in balance with other ecological factors, changes seem to be very low, like observed in Schynige Platte. The already established species probably competed with newly arrived ones and they restricted establishment to species growing isolated in the middle of the structuring plants. However, changes can be triggered following a disturbance. In Vallon de Nant, the grazing abandonment played this role. Changes, although still limited, showed a trend towards fallow or shrub communities, but with some thermophilous species colonizing. Low growing alpine species decreased but certainly suffered more from the higher, no more grazed plants, than from the warmer conditions. This corresponds to the hypothesis that vegetation changes in mountains will be more driven by anthropogenic managements than by climate change (Körner 2005). However, these results are limited to two sites in the subalpine level of the Alps and supplementary similar data are necessary to confirm these trends.

Acknowledgments

This research has been supported by the Federal Office for the Environment FOEN (Switzerland), Centre de conservation de la faune et de la nature (Canton de Vaud) and Fondation Audemars Piguet.

References

Aeschimann, D., and C. Heitz. 1996. Index synonymique de la flore de Suisse et territoires limitrophes. CRSF, Genève.

Alpert, P., and J. L. Maron. 2000. Carbon addition as a countermeasure against biological invasion by plants. Biological Invasions **2**:33-40.

Beniston, M. 1994. Mountain Environment in Changing Climates, Routledge edition. Routledge, London.

Beniston, M., D. G. Fox, S. Adhikary, R. Andressen, A. Guisan , J. I. Holten, J. Innes, J. Maitima, M. F. Price, and L. Tessier. 1996. Impacts of climate change on mountain regions. Chapter 5. *in* IPCC, editor. Intergovernmental panel on climate change. Second assessment report. Cambridge University Press, Cambridge.

Bénzécri, J.-P. 1973. L'analyse des correspondances. Dunod, Paris.

Berlin, G. A. I., A. C. Linusson, and E. G. A. Olsson. 2000. Vegetation changes in semi-natural meadows with unchanged management in southern Sweden, 1965-1990. Acta Oecologica-International Journal of Ecology **21**:125-138.

Braun-Blanquet, J. 1957. Ein Jahrhunder Florenwandel am Piz Linard (3414 m). Bull. Jard. Bot. Etat Brux. **volume jubilaire Walter Robyns**:221-232.

Braun-Blanquet, J. 1964. Pflanzensoziologie. Grundzüge der Vegetationskunde, 3 edition. Springer, Wien - New-York.

Braun-Blanquet, J. 1975. Fragmenta Phytosociologica Raetica I: Die Schneebodengesellschaften (Klasse der Salicetea herbaceae). Jahresbericht der Naturforschenden Gesesellschaft Graubünden **96**:42-71.

Burke, M. J. W., and J. P. Grime. 1996. An experimental study of plant community invasibility. Ecology **77**:776-790.

Camarero, J. J., and E. Gutierrez. 2004. Pace and pattern of recent treeline dynamics: Response of ecotones to climatic variability in the Spanish Pyrenees. Climatic Change **63**:181-200.

Carraro, G., P. Gianoni, R. Mossi, F. Klötzli, and G.-R. Walther. 1999. Observed changes in vegetation in relation to climate warning. Final report NRP 31, Zürich.

Chapin, F. S., and G. R. Shaver. 1985. Individualistic Growth-Response of Tundra Plant-Species to Environmental Manipulations in the Field. Ecology **66**:564-576.

Choler, P., R. Michalet, and R. M. Callaway. 2001. Faciliation and competition on gradients in alpine plant communities. Ecology **82**:3295-3308.

Dähler, W. 1993. Langfristige Auswirkungen menschlicher Eingriffe in alpine Vegetation. Geobotanica Helvetica **69**:1-139.

Darrigo, R., G. C. Jacoby, and I. Y. Fung. 1987. Boreal Forests and Atmosphere Biosphere Exchange of Carbon-Dioxide. Nature **329**:321-323.

Diaz, H. F., M. Grosjean, and L. Graumlich. 2003. Climate variability and change in high elevation regions: Past, present and future. Climatic Change **59**:1-4.

Dietz, H., C. Kueffer, and C. G. Parks. 2006. MIREN: A new research network concerned with plant invasion into mountain areas. Mountain Research and Development **26**:80-81.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Dobbertin, M., N. Hilker, M. Rebetez, N. E. Zimmermann, T. Wohlgemuth, and A. Rigling. 2005. The upwart shift of pine mistletoe (Viscum album ssp. austriacum) in Switzerland - The result of climate warming ? Int. J. Biometeorol. **50**:40-47.

Dullinger, S., T. Dirnbock, R. Kock, E. Hochbichler, T. Englisch, N. Sauberer, and G. Grabherr. 2005. Interactions among tree-line conifers: differential effects of pine on spruce and larch. Journal of Ecology **93**:948-957.

Dutoit, A. 1983. végétation de l'étage subalpin du vallon de Nant. Lausanne.

Dutoit, A. 1984. La végétation de l'étage subalpin du vallon de Nant. University of Lausanne, Lausanne.

Ellenberg, H., H. E. Weber, R. Düll, V. Wirtz, W. Werner, and D. Paulissen. 1991. Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobotanica **18**.

Fischer, M., and J. Stöcklin. 1997. Local extinctions of plants in remnants of extensively used calcareous grasslands 1950-1985. Conservation Biology **11**:727-737.

Galloway, J. N., J. D. Aber, J. W. Erisman, S. P. Seitzinger, R. W. Howarth, E. B. Cowling, and B. J. Cosby. 2003. The nitrogen cascade. Bioscience **53**:341-356.
Galloway, J. N., W. H. Schlesinger, H. Levy, A. Michaels, and J. L. Schnoor. 1995. Nitrogen-Fixation - Anthropogenic Enhancement-Environmental Response. Global Biogeochemical Cycles **9**:235-252.
Gigon, A. 1997. Fluktuationen des Deckungsgrades und die Koexistenz von Pflanzenarten in Trespen-Halbtrockenrasen (Mesobromion). Phytocoenologia **27**:275-287.
Gonseth, Y., T. Wohlgememuth, B. Sansonnens, and A. Buttler. 2001. Les régions biogéographiques de la Suisse. OFEFP, Bern.
Grabherr, G., M. Gottfried, and H. Pauli. 1994. Climate Effects on Mountain Plants. Nature **369**:448-448.
Grabherr, G., M. Gottfried, and H. Pauli. 2001. Long-term monitoring of mountain peaks in the Alps. Pages 153-177 *in* C. A. Burga and A. Kratochwil, editors. Biomonitoring: general and applied aspects on regional and global scales. Kluwer, Dordrecht.
Guisan, A. 2005. Niche-based models as tools to assess climate change impact on the distribution and diversity of plants in mountain reserves. *in* Proceedings of the 2nd and 3rd GLOCHAMORE workshops in L'Aquila (Italy) and Granada (Spain). UNESCO-MRI, Paris, France.
Guisan, A., J. I. Holten, R. Spichiger, and L. Tessier, editors. 1995. Potential Ecological Impacts of Climate Change in the Alps and Fennoscandian Mountains. Ed. Conserv. Jard. Bot. Geneva, Geneva, Switzerland.
Guisan, A., and J.-P. Theurillat. 2000. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**:307-320.
Harte, J., and R. Shaw. 1995. Shifting Dominance within a Montane Vegetation Community - Results of a Climate-Warming Experiment. Science **267**:876-880.
Hartl, D., and G. Wagenitz. 1975. Dycotyledones, 4. Teil, 2 edition. Paul Parey, Berlin.
Hegg, O. 1992. Long term influence of fertilization in a Nardetum. The experimental field fo Dr. Werner Lüdi on the Schynige Platte. VEGETATIO **103**:133-134.
Hobbs, R. J., and L. F. Huenneke. 1992. Disturbance, diversity, and invasion: implications for conservation. Conservation Biology **6**:324-337.
Hofer, H. R. 1992. Veränderungen in der Vegetation von 14 Gipfeln der Berninagebietes zwischen 1905 und 1985. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich **58**:39-54.
Hollister, R. D., P. J. Webber, and C. Bay. 2005. Plant response to temperature in Northern Alaska: Implications for predicting vegetation change. Ecology **86**:1562-1570.
Houghton, J. T., Y. Ding, and D. J. Griggs, editors. 2001. Climate change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
Huntley, B. 1991. How plant respond to climate change: migration rates, individualism and the consequences for plant communities. Ann. Bot. **67**:15-22.
Klanderud, K., and H. J. B. Birks. 2003. Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. The Holocene **13**:1-6.
Klimes, L. 2003. Scale-dependent variation in visual estimates of grassland plant cover. J. Veg. Sci. **14**:815-821.
Klimes, L., M. Dancak, M. Hajek, I. Jongepirova, and T. Kucera. 2001. Scale-dependent biases of species counts in a grassland. J. Veg. Sci. **12**:699-704.
Köhler, B., A. Gigon, P. J. Edwards, B. Krüsi, R. Langenauer, A. Lüscher, and P. Ryser. 2005. Changes in the species composition and conservation value of limestone grasslands in Northern Switzerland after 22 years of contrasting managements. Perspectives in Plant Ecology Evolution and Systematics **7**:51-67.
Körner, C. 2005. The green cover of mountains in a changing environment. Pages 367-375 *in* U. M. Huber, H. K. M. Bugmann, and M. A. Reasoner, editors. Global change and mountain regions. Springer, Dordrecht.
Korner, C., M. Diemer, B. Schappi, P. Niklaus, and J. Arnone. 1997. The responses of alpine grassland to four seasons of CO2 enrichment: a synthesis. Acta Oecologica-International Journal of Ecology **18**:165-175.
Kullman, L. 2001. Immigration of Picea abies into the North-Central Sweden. New evidence of regional expansion and tree-limit evolution. Nordic Journal of Botany **21**:39-54.
Landolt, E. 1977. Ökologische Zeigerwerte zur Schweizer Flora. Veröff. Geobot. Inst. ETH Stiftung Rübel Zürich **64**:1-208.

Lauber, K., and G. Wagner. 1991. Flora des Kantons Bern, Haupt edition, Bern.
Legendre, L., and P. Legendre. 1979. Ecologie numérique. Tome 2: la structure des données écologiques. Masson, Paris.
Leps, J., and V. Hadincova. 1992. How reliable are our vegetation analyses ? J. Veg. Sci. **3**:119-224.
Lüdi, W. 1936. Experimentelle Untersuchungen an alpiner Vegetation. Ber. Schweiz. Bot. Gesell. **46**:632-381.
Malcolm, J. R., A. Markham, R. P. Neilson, and M. Garaci. 2002. Estimated migration rates under scenarios of global climate change. Journal of Biogeography **29**:835-849.
Nogués-Bravo, D., M. Araújo, M. P. Errea, and J. P. Martinez-Rica. 2006. Exposure of global mountain systems to climate warming during the 21st century. GLOBAL ENVIRONMENTAL CHANGE.
Oberdorfer, E. 1990. Pflanzensoziologische Exkursionsflora, 6 edition. Ulmer, Stuttgart.
Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature **421**:37-42.
Pauli, H., M. Gottfried, and G. Grabbherr. 1996. Effects of climate change on mountain ecosystems. Upward shifting of alpine plants. World Resource Review **8**:382-390.
Pauli, H., M. Gottfried, D. Hohenwallner, K. Reiter, R. Casale, and G. Grabherr. 2004. The GLORIA field manual. Multi-summit approach. European Commission. Office for Official Publications of the European Communities, Luxembourg.
Pauli, H., M. Gottfried, K. Reiter, C. Klettner, and G. Grabherr. 2007. Signals of range expansions and contractions of vascular plants in the high Alps: observations (1994-2004) at the GLROIA master site Schrankogel, Tyrol, Austria. Global Change Biology **13**:147-156.
Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology and Biogeography **12**:361-371.
Phoenix, G. K., and M. C. Press. 2005. Effects of climate change on parasitic plants: the root hemiparasitic Orobanchaceae. Folia Geobotanica **40**:205-216.
Poyry, J., M. Luoto, J. Paukkunen, J. Pykala, K. Raatikainen, and M. Kuussaari. 2006. Different responses of plants and herbivore insects to a gradient of vegetation height: an indicator of the vertebrate grazing intensity and successional age. Oikos **115**:401-412.
Price, M. V., and N. M. Waser. 2000. Responses of subalpine meadow vegetation to four years of experimental warming. Ecological Applications **10**:811-823.
R. 2006. R 2.4.1 A language and environment. *in*.
Randin, C. F., T. Dirnbock, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan. 2006a. Are niche-based species distribution models transferable in space? Journal of Biogeography **33**:1689-1703.
Randin, C. F., T. Dirnbock, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan. 2006b. Are niche-based species distribution models transferable in space? Journal of Biogeography **33**:1689-1703.
Rebetez, M., and M. Dobbertin. 2004. Climate change may already threaten Scots pine stands in the Swiss Alps. Theor. Appl. Climatol. **79**:1-9.
Rebetez, M., and M. Reinhard. in press. Monthly air temperature trends in Switzerland 1901-2000 and 1975-2004. Theoretical and Applied Climatology.
Scott, D., N. G. Bayfield, A. Cernusca, and D. A. Elston. 2002. Use of a weighing lysimeter system to assess the effects of trampling on evapotranspiration of montane plant communities. Canadian Journal of Botany **80**:675-683.
Sebastia, M. T. 2007. Plant guilds drive biomass response to global warming and water availability in subalpine grassland. Journal of Applied Ecology **44**:158-167.
Silvertown, J., S. Holtier, J. Johnson, and P. Dale. 1992. Cellular Automaton Models of Interspecific Competition for Space - the Effect of Pattern on Process. Journal of Ecology **80**:527-534.
Sturm, M., C. Racine, and K. Tape. 2001. Climate change - Increasing shrub abundance in the Arctic. Nature **411**:546-547.
Suarez, M. L., L. Ghermandi, and T. Kitzberger. 2004. Factors predisposing episodic drought-induced tree mortality in Nothofagus-site, climatic sensitivity and growth trends. Journal of Ecology **92**:954-966.
Taylor, A. H. 1995. Forest expansion and climate change in the mountain hemlock (Tsuga mertensiana) zone, Lassen Volcanic National Park, California, USA. Artic and Alpine Research **27**:207-216.
Theodose, T. A., and W. D. Bowman. 1997. Nutrient availability, plant abundance, and species diversity in two alpine tundra communities. Ecology **78**:1861-1872.
Theurillat, J.-P., F. Felber, P. Geissler, J.-M. Gobat, M. Fierz, A. Fischlin, P. Küpfer, A. Schlüssel, C. Velluti, and G.-F. Zhao. 1998. Sensitivity of plant and soil ecosystems of the Alps to climate change. Pages 225-

308 *in* P. Cebon, U. Dahinden, H. C. Davies, D. Imboden, and C. C. Jäger, editors. Views from the Alps. Regional perspectives on climate change. MIT Press, London.
Theurillat, J. P., and A. Guisan. 2001. Potential impact of climate change on vegetation in the European Alps: A review. Climatic Change **50**:77-109.
van der Maarel, E. 1996. Pattern and process in the plant community: Fifty years after A.S. Watt. Journal of Vegetation Science **7**:19-28.
van Wijk, M. T., K. E. Clemmensen, G. R. Shaver, M. Williams, T. V. Callaghan, F. S. Chapin, J. H. C. Cornelissen, L. Gough, S. E. Hobbie, S. Jonasson, J. A. Lee, A. Michelsen, M. C. Press, S. J. Richardson, and H. Rueth. 2004. Long-term ecosystem level experiments at Toolik Lake, Alaska, and at Abisko, Northern Sweden: generalizations and differences in ecosystem and plant type responses to global change. Global Change Biology **10**:105-123.
Vittoz, P., and A. Guisan. 2003. Le projet PERMANENT.PLOT.CH demande votre collaboration. Das Projekt PERMANENT.PLOT.CH bittet um Ihre Mithilfe. Bot. Helv. **113**:105-110.
Vittoz, P., and A. Guisan. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. Journal of Vegetation Sciences **18**.
Vittoz, P., and P. Hainard. 2002. Impact of free-range pigs on mountain pastures in the Swiss Jura. Applied Vegetation Science **5**:247-254.
Vittoz, P., S. Jutzeler, and A. Guisan. 2006. Flore alpine et réchauffement climatique: observation de trois sommets valaisans à travers le 20ème siècle. Bulletin de la Murithienne **123/2005**:49-59.
Vittoz, P., B. Rulence, T. Largey, and F. Freléchoux. in press. Effects of Climate and Land-Use Change on the Establishment and Growth of Cembran Pine (Pinus cembra L.) over the Altitudinal Treeline Ecotone in the Central Swiss Alps. Arctic Antarctic and Alpine Research.
Wahren, C.-H. A., M. D. Walker, and M. S. Bret-Harte. 2005. Vegetation responses in Alaskan arctic tundra after 8 years of a summer warming and winter snow manipulation experiment. Global Change Biology **11**:537-552.
Walther, G.-R. 2002. Weakening of climatic constraints with global warming and its consequences for evergreen broad-leaved species. Folia Geobotanica **37**:129-139.
Walther, G.-R., S. Beißner, and C. A. Burga. 2005a. Trends in the upward shift of alpine plants. Journal of Vegetation Science **16**:541-548.
Walther, G. R., S. Berger, and M. T. Sykes. 2005b. An ecological 'footprint' of climate change. Proceedings of the Royal Society B-Biological Sciences **272**:1427-1432.
Walther, G. R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature **416**:389-395.
Welch, D. 1998. Yellow Marsh Saxifrage (Saxifraga hirculus).
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.

Chapter 2

Importance of "land use" in predictive habitat modelling of plant species distribution in the Swiss Alps

Randin Christophe[1], Jaccard Hélène[1], Vittoz[2] Pascal and Guisan Antoine[1]

This manuscript will be submitted to Journal of Vegetation Science.

[1]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]University of Lausanne, Faculty of Geosciences and Environment, Bâtiment Biophore, 1015 Lausanne, Switzerland

Abstract

Predictive modelling of the geographic distribution of habitat for plants is a growing discipline in ecology. Several models were developed for predicting the distribution of alpine plant species, using climatic and topographical variables. However, lack of information on human land use is a major and recurrent limitation in these models. Few studies have taken land use into account, although grazing and fertilisation effects on plant species richness have long been recognised. At large spatial scales, land use or land cover variables were shown not to improve the predictive accuracy of models based on bioclimatic variable only, whereas land use was shown as an important explanatory variable to predict the distribution of plant communities at finer scales.

Here, we assessed if a land use variable and a index of grazing can improve the fit and the predictive power of species distribution models based on topo-climatic variables and yield better spatial predictions at fine scale in an alpine landscape.

By regressing presence/absence data for 270 species on climatic and topographic variables, on average, the land use (LU) variable and the grazing index (IG) improved significantly in average the fit of models, but not their predictive power. In contrast, the predictive power of abundance models was significantly improved when adding both LU and IG. ANOVA showed that the reduction of variance by LU and IG variables was significant for 30% and 26% of the species respectively. On the other hand, partial regression analysis showed that individual contribution of LU and IG to the deviance explained by models was very weak on average (1.1% and 0.2% for LU and IG).

Changes in projected surface were important for some species when including LU. This is an important issue for climate change projections because such differences in an assessment over a large number of species may change significantly trends of extinction.

Keywords: generalized linear models (GLM), habitat distribution, vegetation, land use, model evaluation, partial regression analysis, Western Swiss Alps

Nomenclature: Aeschimann and Heitz (1996)

Introduction

A recurrent limitation of predictive species distribution models (Guisan and Zimmermann 2000a, Guisan and Thuiller 2005) is the lack of spatially explicit information on land use (Zimmermann and Kienast 1999, Guisan and Theurillat 2000a, Pearson and Dawson 2003). Although bioclimatic factors were shown to explain a major part of the observed patterns at large scale (Thuiller et al. 2005), agricultural land use remains a potentially useful variable for predicting plant species and communities distribution at finer scales (Fischer 1994, Dirnböck et al. 2003). At this scale, land use is partly dependent on topography and climate and should thus also be altered by changing climate (Carcaillet and Brun 2000, Rounsevell et al. 2006). It can also prove an important factor by acting as barrier or corridor for plant dispersal when predicting future species distributions using dynamic dispersal models (Carey 1996, Dullinger et al. 2004). Because spatially-explicit information on land use is often difficult to obtain at fine scales, few models have so far integrated this information, despite a large ecological knowledge accumulated on the effects of land use on plant communities.

Grasslands management has a strong impact on the floristic composition (Milchunas and Lauenroth 1993). A main factor is nutrient concentration, with heavily fertilized grasslands being dominated by a limited number of fast-growing tall plants (Eler et al. 2005). Conversely, unfertilized grasslands in low swards, where luminous conditions prevail on the soil allowing rosette plants like orchids to develop, have greater plant diversity (e.g. Stevens et al. 2004, Maurer et al. 2006, Spiegelberger et al. 2006). In meadows, mowing frequency strongly influences floristic composition, as plants need time to complete their reproduction cycle. In intensive meadows, rotation time is too fast for many plants to complete their cycle, favouring species with high vegetative reproduction abilities (Ellenberg 1986). Cattle activity often increases plant diversity in pastured grasslands (Olff and Ritchie 1998). The spatial distribution of grazing intensity is irregular, with more luminous conditions at soil level in strongly grazed sites. Trampling selects tolerant species and creates gaps propitious to pioneer species (Kohler et al. 2006). Dung fertilizes the soil locally and often stimulates growth the following year(Schreiber 1969, Kohler et al. 2004). Moreover, as cattle selects the most edible plants, toxic or spiny plants tend to invade pastures, or to create favourable sites for grazing sensitive plants or tree seedlings (Smit et al. 2005). These different processes are strongly dependent on the size of the herd.

The present study aimed to investigate the importance of land use for presence-absence and abundance models of plant species distribution in a mountainous landscape. We quantified the additive variance of land use compared to pure topo-climatic models of presence-absence and we compared the fit and the predictive powers of presence-absence versus abundance models. Predictive maps of presence-absence models were also prepared for a limited area to test the effect of land use spatially.

Methods

Study area

The study area includes all mountains of the Western Alps of the Canton de Vaud in Switzerland (6°60′ to 7°10' E; 46°10′ to 46°30′ N) and covers an area of 700 km^2 (Fig. 1). Elevation ranges from 375 m in Montreux to 3210 m on the top of the Diablerets massif, where important glaciers are found. Annual temperature and precipitation vary respectively from 8°C and 1200 mm at 600 m to -5°C and 2600 mm at 3000 m (Bouët 1985). The soil parent material is mainly calcareous. Vegetation has been and is still very influenced by human activity: the pasture is common in this region, and is observed from the bottom in the Rhone valley to the subalpine and lowers alpine zones at higher elevations. Meadows of diverse intensity of management are mainly found at lower elevations.

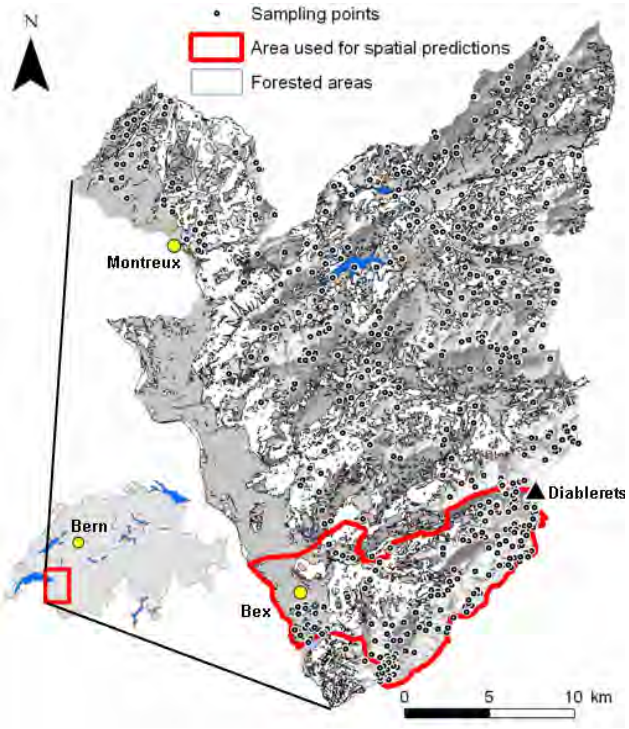


Figure 1 Geographical location of the study area (Western Alps of the Canton de Vaud) in Switzerland.

Vegetation data

During summers 2002-2004, 483 vegetation plots were sampled following a random-stratified sampling strategy restricted to non-woody vegetation (grassland, rock and scree). Stratification was done by elevation, slope, aspect and simplified classes of geology. Plots were composed of a large surface of 64 m^2 and, nested within each, a smaller surface of 4 m^2 (Lassueur et al. 2006, Randin et al. 2006). The

presence of all species was recorded in both surfaces, and abundance was additionally recorded in the 4 m^2 plots using a 4-classes semi-quantitative cover-abundance scale (modified from Guisan et al. 1998): 0: absent, 1: ≤1%, 2: 2–5%, 3: 5–12%, 4: 13–100%).

Climate and topographical variables

Climate data were derived from the national meteorological stations at different altitudes. Long-term monthly means for average temperature (°C) and sum of precipitation (mm) for the period 1961-1990 were used. A digital elevation model (DEM) was used to spatially interpolate the climatic data. Method of computation and description of the variable are developed in Zimmermann and Kienast (1999). In this study, we used three climatic and two topographic variables at a resolution of 25 m (Table 1). These variables are expected to be of greatest ecophysiological significance (Guisan and Zimmermann 2000a, Pearson et al. 2002, Körner 2003, Guisan and Thuiller 2005). Degree-days of the growing season were derived from interpolated daily temperatures using a threshold value of zero degree. Moisture index is calculated as the difference between precipitation and evapotranspiration, and thus expresses the amount of soil water potentially available at a site. The sum of mean daily values for the months of June, July and August were used. These represent the warmest months of the year, and thus are expected to be of greatest ecophysiological significance (Zimmermann and Kienast 1999). The potential global solar radiations were calculated over the year. Topographic position and slope were derived from the DEM. Positive values of topographic position express relative ridges, tops and exposed sites, whereas negative values indicate sinks, valleys or toe slopes.

Table 1 Topographic and climatic variables used in this study to model the distribution of species.

Variables	Units	Details	Method	References
Temperature degree days	°C * day * $year^{-1}$	Sum of days multiplied by temperature > 0 C	ARCInfo AML	Zimmermann & Kienast (1999)
Moisture index (average of monthly values June–August)	mm * day^{-1}	Monthly average of daily atmospheric H_2O balance	ARCInfo AML	Zimmermann & Kienast (1999)
Global solar radiation (sum of monthly values)	kJ * m^{-2} * day^{-1}	Monthly average of daily global solar radiation	ARCInfo AML	Zimmermann & Kienast (1999)
Slope	degrees	Slope inclination	DEM, ARCInfo, GRID routine	ESRI (2005)
Topographic position	uniteless	Integration of topographic features (ridge, slope, toe slope) at various spatial scales	ARCInfo AML	Zimmermann (1996)

Land use data

Land use maps did not exist for the whole study area at the start of the study. Land use data were thus sampled for each plot by interviewing farmers and local agricultural officers. The resulting land use variable (LU) is qualitative. A description of each category is given in table 2.

Table 2 Description of the different categories of land use present in the study area.

Category	Description
Exploited meadow	Meadow as a principal use for hay production generally mown 2-3 times a year, fertilized, sometimes grazed in autumn
Middle intensity meadow (SCE)*	Meadow lightly fertilized (manure or compost), mown once a year
Unfertilized meadow (SCE)*	Meadow not fertilized, mown once a year in July.
Pasture	Pasture grazed by cattle of different kind, usually fertilized, sometimes mown, varying intensity
Unfertilized pasture (SCE)*	Pasture not fertilized at low and middle elevation
High mountain pasture	Pastured grazed only during summer, not fertilized, limited quantity of cattle by surface
Unexploited area	Not exploited (rocks, rock fall)

*SCE: ecological compensation surface, see text for more explanation

In Switzerland, payments to farmers have been introduced to maintain and enhance species richness in grasslands. This resulted in an expansion of ecological compensation surfaces (SCE) where restrictions on fertilizers and mowing period have been imposed. Accurate information on land use is thus available for these SCE, but for other surfaces, the farmer is relatively free to exploit as he wishes.

Mostly cows and sheep are found in the study area, but no information exists on the spatial distribution of each cattle type. A grazing index (IG) was developed, which takes a value of 1 when the plot is grazed (pastures and mountain pasture) and 0 else (unexploited or mown).

To test the effect of land use on spatial predictions, we developed a spatially-explicit land use map for the Commune of Bex (Fig.1). Although restricted to this subarea, it remains suitable because: (i) it encompasses the entire altitudinal gradient (390 - 3180 m); (ii) it includes areas of high species richness; (iii) all categories of land use were represented in sufficient number. This map was created by linking the land use information, obtained from persons in charge of agricultures and inspectors of cattle, to the cadastral map (Office de l'information sur le territoire, http://www.geoplanet.vd.ch) and other geographic information data on landcover (Vector25, Swisstopo, http://www.swisstopo.ch). The vector file obtained was converted into a raster grid at a 5-m resolution to keep a good agreement with the boundaries of the cadastral parcels. The layer was only then reclassified in seven binary raster for each of the category of land use presented in table 2 at the resolution of 25 m. All geographic analyses were conducted in ArcGIS 9.1 (ESRI 2005).

Statistical analyses

All statistical analyses were performed in the R software (R 2.4.12006). Analyses were run on all species with more than 20 occurrences (270 plant species from lowland to high alpine areas, Appendix A).

Presence/absence models (GLM)

For each species, a GLM (McCullagh and Nelder 1989) with a binomial variance and a logistic link function was fitted using presence-absence as the response variable and climate and topographic variables only as predictors (TC model). An Akaike information criterion (AIC)-based stepwise procedure in both directions was used to select the most significant predictors in the TC model (Akaike 1973). Up to second-order polynomials (linear and quadratic terms) were allowed for each predictor, with the linear term being forced in the model each time the quadratic term was retained. In a second step, the land use variable (LU model) and the grazing index (IG model) were then added independently to the TC model.

Model fit and model predictive power were both estimated with the Somer's concordance D_{xy} index (Harrell 2001) based on the Wilcoxon-Mann-Whitney two-samples rank test. D_{xy} takes values between 0 and 1, depicting totally random vs. perfectly discriminating models, respectively. Negative values would signify predictions worse than random. Model fit was calculated on the calibration datatset, whereas the predictive power of the TC, LU and IG models was evaluated by running a 10-fold cross-validation (see van Houwelingen and Le Cessie 1990, Randin et al. 2006 for the exact implementation) on the full dataset to obtain a pseudo-independent evaluation dataset. During the cross-validation procedure, the original prevalence of the species presences and absences in the data set was maintained in each fold.

Abundance models (LRM)

We derived ordinal logistic regression models (proportional odds models) using the cover-abundance values of plant species as the response and environmental variables as the predictors. Proportional odds models are based on the cumulative probabilities of the classes, stemming from successive logistic regression models (McCullagh 1980, McCullagh and Nelder 1989, Guisan and Harrell 2000, Harrell 2001). Models were fitted using the Design Library (Azola and Harrell 2001) and up to second-order polynomials (linear and quadratic terms) were allowed for each predictor. Predictors were selected by stepwise backward selection. As for the presence-absence models, a TC model was fitted with all topo-climatic variables and land use (LU model) and the grazing index (IG model) were then added independently to the TC model. The additional deviance explained by LU and IG was tested with an ANOVA. Ordinal models were only fitted to those species showing sufficient variation in abundance. This resulted in 34 LU models and 40 IG models.

Model fit and predictive power were also estimated with the Somer's concordance index D_{xy} (Harrell 2001), which values can be interpreted as for presence-absence models Predictive power was calculated using a 10-folds cross-validation, using the Design library.

Quantifying the importance of land use

Three different methods were used to quantify the contribution of land use and the grazing index in the GLM: difference in model fit and predictive power between the models, ANOVA and variance partitioning. Variance partitioning allows the comparisons of two groups of predictors, whereas the ANOVA quantifies the additional variance explained by the LU and IG variables in the TC model. Implementations of these three different methods are presented below.

(1) Difference in model fit and predictive power of the GLM

A general estimation of model fit with and without LU or IG variables is obtained by calculating the mean *Dxy*. This was done both on the training and validation (i.e. obtained by cross-validation) datasets. The contribution is simply measured as the difference in *Dxy* between LU (or IG respectively) and TC models.

(2) ANOVA

The reduction in deviance caused by the addition of land use or the grazing index to the topo-climatic model was assessed by an ANOVA model comparison test.

(3) Variance partitioning

To quantify the variance added by the LU and IG variables, a variance partitioning approach based on partial regression analyses was used. This approach allows partitioning of the variance into four identifiable fractions of explained variance (Borcard et al. 1992, Legendre 1993; Figure 2): (a) pure topo-climatic, (b) shared topo-climatic and land use, (c) pure land use and (d) unexplained variation.

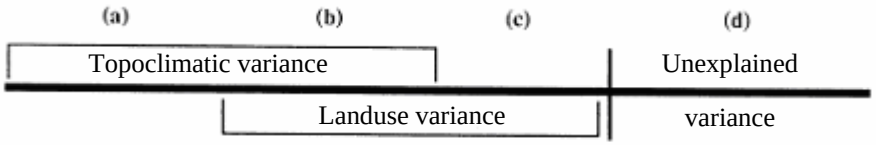


Figure 2 Variance partitioning among the target variable. The fraction of variance due to the topo-climatic variable, of the interaction between land use and topo-climatic variables, of land use and the unexplained variation are represented by letter a, b, c and d respectively.

To assess the different fraction, an iterative method was used, and the Nagelkerke R^2 (designated as D^2; Guisan and Zimmermann 2000b) was used as an estimator of the explained variance (Table 3). Regressions of the response variable are computed either on topo-climatic variables (model 1), on land use (model 2), or on land use and topo-climatic variables (model 3). Then a GLM on residuals of model 1 was run against land use (model 4) and reciprocally with residuals of model 2 against topo-climatic variables (model 5). The same was done with the grazing index instead of land use.

Table 3 Iterative method to perform partial regression analysis and rescaled D^2. This method allows estimating the fraction of variance of the three groups of variables: topo-climatic variables (TC), land use (LU) and the grazing index (IG).

Model	Formula	D^2 of model	Rescaled D^2
1	p/a ~ TC	$D^2 1$	$D^2 1$
2	p/a ~LU or IG	$D^2 2$	$D^2 2$
3	p/a ~ TC + LU / IG	$D^2 3$	$D^2 3$
4	Residuals(model 1) ~ LU / IG	$D^2 4$	$D^2 4' = (1- D^2 1) * D^2 4$
5	Residuals (model 2) ~ TC	$D^2 5$	$D^2 5' = (1- D^2 2) * D^2 5$

The residuals are of type "response" and the absolute value is taken, thus a binomial family can be use. As the scale of variance of model 4 and 5 is not the same as the one in model 1 to 3, it needs to be rescaled. In fact, when running the model on the residuals the null variance is reset to 100% whereas it is, in reality, equal to the unexplained variance of the model on which the residuals were issued. Therefore D^2 of model 4 and 5 is multiplied by the unexplained variance of model 1 and 2 respectively. The different fractions of variation estimation are presented in table 4. The extended model contains the land use variable in addition to the variables selected in the topo-climatic model (model 1).

Table 4 Estimators of the different fraction of variance.

Fraction of variance	Estimator
TC	$D^2 4'$
LU	$D^2 5'$
TC + LU / TC + IG	$D^2 3 - (D^2 4' + D^2 5')$
Unexplained	$1 - D^2 3$

Potential habitat maps

Three maps were prepared for some exemplar species, based on: (1) TC model, (2) LU model, (3) their difference. Maps 1 and 2 were reclassified into presence-absence using an AUC-optimized threshold and a mask based on forests, roads, urbanized areas and rivers was further applied to avoid predicting the species in impossible situations. Map 3 was prepared by subtracting maps 1 and 2 (i.e. TC – LU). These latter maps were reclassified into three categories: (i) no or few differences between

TC and LU model (values between -0.3 and 0.3), (ii) LU probabilities of occurrence higher than TC (from -1 to -0.3) and, reciprocally, (iii) LU probabilities lower than TC (from 0.3 to 1). All maps were calculated from GLM outputs in the ArcGIS 9.1 software (ESRI 2005).

Results

Presence-absence models (GLM)

The three predictors most often selected in the topo-climatic (TC) model,were growing degree days, moisture index and slope. The average values of Somer's D_{xy} for the 270 species on the calibration and evaluation datasets for the TC, land use (LU) and grazing index (IG) models all yielded values above 0.7 (up to 0.8), signifying good models on average (Table 5). Model fits were significantly higher on average for LU and IG models compared to pure TC models (+3.5% for LU and +1.0% for IG; Figure 3). By contrast, predictive power of TC models remained significantly higher than that of LU models, but not significantly different from IG models (-1.5% for LU and -6.2% for IG; Fig.3). Significant differences were, in general, very small.

Table 5 Average values of the *Dxy* on the calibration and evaluation dataset for the TC, LU and IG models for the 270 species.

Model	*Dxy* cal (Model fit)	*Dxy* eval (Predictive power)
TC	0.76	0.72
TC + LU	0.79	0.71
TC + IG	0.77	0.72

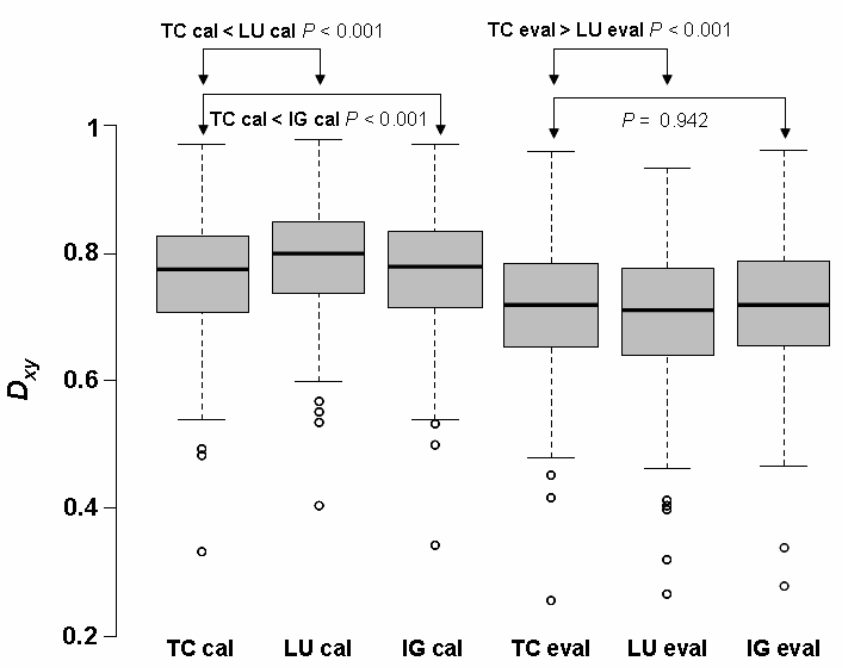


Figure 3 Boxplots of model fit (cal = *Dxy* on the calibration dataset) and predictive power (eval = *Dxy* on the evaluation dataset) of the three models (TC = topoclimatic, LU = with landuse, IG = with grazing index). The *P* values indicate the significance of the Wilcoxon signed rank tests between models. The > or < symbols indicate the direction of the test.

Difference in explanatory power and predictive ability

The difference in model fit (*Δ Dxy cal)* was higher on average for LU models (*Δ Dxy cal* = 0.027) than for IG models (mean *Δ Dxy cal* = 0.008, Table 6). The *Δ Dxy cal* was the highest for *Pimpinella major* (*Δ Dxy cal* = 0.136) between TC and LU models and for *Carex nigra* (*Δ Dxy cal* = 0.065) between TC and IG models. The weakest difference between TC and LU models was observed for *Viola hirta* (*Δ Dxy cal* = - 0.015) and between TC and IG models for *Viola biflora* (*Δ Dxy cal* = - 0.013).

On the contrary, the difference in predictive power (*Δ Dxy eval)* was on average lower for LU models (*Δ Dxy eval* = -0.011) than for IG models (*Δ Dxy eval* = 0.000, no difference on average, Table 6). The highest *Δ Dxy eval* between TC and LU models was observed for *Pimpinella major* (*Δ Dxy eval* = 0.119), and between TC and IG models for *Valeriana montana* (*Δ Dxy eval* = 0.072). The weakest difference between TC and LU models was observed for *Tussilago farfara* (*Δ Dxy eval* = - 0.16) and between TC and IG models for *Urtica dioica* (*Δ Dxy eval* = - 0.079).

Table 6 Differences in model fit (*Δ Dxy cal*) and predictive power (*Δ Dxy eval*) between pure topoclimatic models (TC) and expanded models (LU and IG).

	Δ Dxy cal		**Δ Dxy eval**	
	TC - LU	**TC - IG**	**TC - LU**	**TC - IG**
mean	0.027	0.008	-0.011	0.000
min	-0.015	-0.013	-0.016	-0.079
max	0.136	0.065	0.119	0.072

ANOVA

Land use contributed to a significant variance reduction in 30% of models (*P* value at 0.05 level) and grazing index in 26% of models. The reduction of variance was more often significant when land use rather than grazing index was considered as additional variable (81 times for LU against 71 for IG).

Partial regression analysis

The contribution of land use to explain additional variance in models was on average rather weak (+1.1%, Table 7), reaching +7.3% in the best model for *Pimpinella major*. The contribution was not greater than +0.5% and +1% for 35.6% and 59.2% of the species' models respectively.

Table 7 Percentage of mean, minimum and maximum of the explained variance for each identifiable fraction when considering land use with the topo-climatic variables.

	Landuse			Grazing Index		
	mean	min	max	mean	min	max
TC	2.2	<0.0001	16.7	8.3	<0.0001	56.3
LU	1.1	<0.0001	7.4	0.2	<0.0001	1.5
TC + LU	33.5	5.8	70.4	26.2	4.5	64.5
unexplained	61.2	29.3	90.4	65.4	30.2	92.5

The contribution of the grazing index was also on average weak (+0.2%), with the highest contribution observed for the model of *Festuca rubra* (+1.4%). The contribution was not greater than +0.5% and +1% for 90.7% and 97.4% of the models respectively.

The five species having the greatest D^2 values for land use (Figure 4) were *Pimpinella major* (74%)*, Achillea millefolium* (44%)*, Veronica persica* (42%)*, Geum montanum* (41%) and *Centaurea jacea* (39%). Topoclimatic variables alone could explain on average more than 2.2 % of the variance and at best 16.7 % of the variance (in some models; Table 7). Land use and topoclimatic variable explained usually a great proportion of variance in common (shared variance of 33.5% on average, with a maximum value of 70.4%).

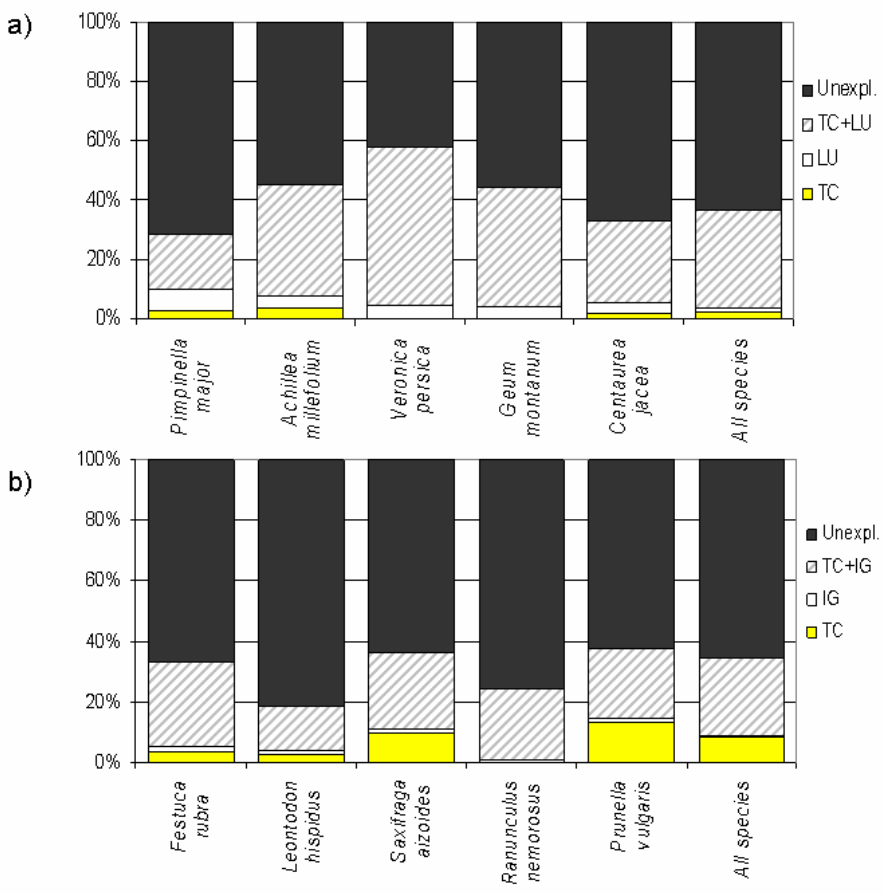


Figure 4 Fraction of explained variance among the target variables for the five species having the highest D^2 for land use (a) and growing index (b).

The five species having the highest D^2 value for the grazing index (Figure 4) were *Festuca rubra* (15%), *Leontodon hispidus* 14%), *Saxifraga aizoides* (13%), *Ranunculus nemorosus* (13%) and *Prunella vulgaris* (12%). As for land use, grazing index and topoclimatic variable had often a great proportion of shared variance (26.2% on average, with a maximum value at 64.5%).

Abundance models (LRM)

As for presence-absence logistic models, the explanatory power of models was on average significantly higher for models including LU and IG variables, compared to pure TC models (see boxplot in Figure 5). Contrary to GLM presence-absence models, the values of predictive power of LU and IG models were significantly higher (LU: -0.8% for GLM and +84.9% for LRM; IG: -0.7% for GLM and +88.6% for LRM) than those of TC models in the case of LRM abundance models, with important differences observed overall. The three species which models showed the greatest increase in predictive power were *Daucus carotta*, *Vaccinium myrtillus* and *Rhinanthus alectorolophus*.

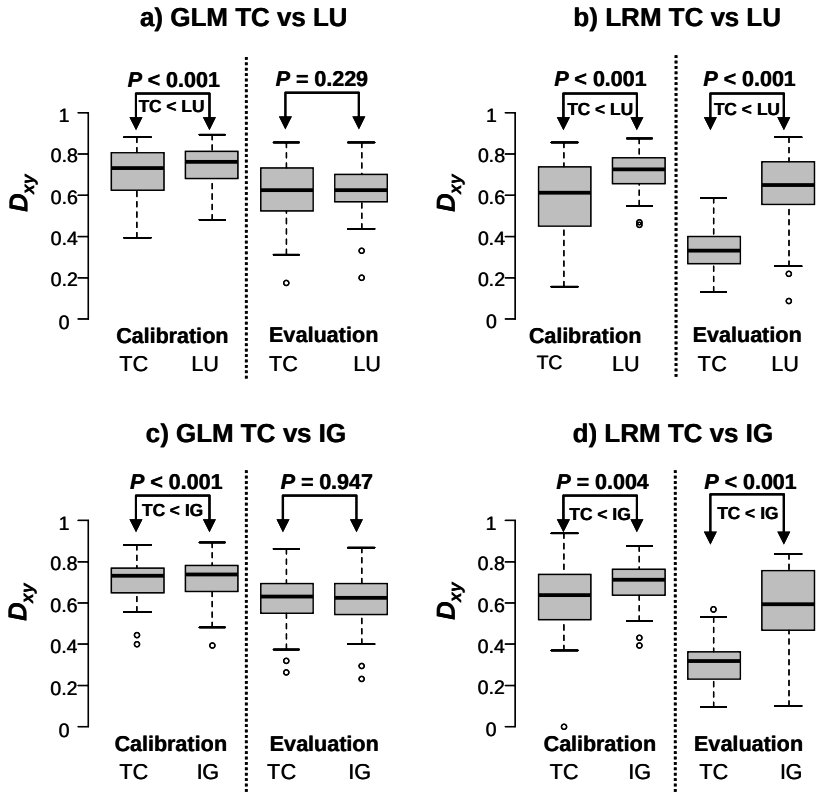


Figure 5 Boxplots of model fit (calibration dataset) and predictive power (validation dataset) of TC, LU and IG models (TC = topo-climatic, LU = with landuse, IG = with grazing index) for GLM and LRM. The pvalues indicate the significance of the Wilcoxon signed rank tests between models. The indications below give the direction of the test. GLM and LRM models were calibrated on a subset of 34 and 40 species for the LU and IG models respectively.

Potential distribution maps for p/a models

Four species were selected to illustrate spatial projections from p/a models, based on model performances (Table 8). *Pimpinella major* and *Geum montanum* showed the best improvement by the LU variable on the fit and predictive power of GLMs and from the perspective of partial regression analysis. *Festuca rubra* also showed good improvement of both fit and predictive power and is a dominant species in nutrient-rich grasslands. For these three species, *P* values of ANOVA were significant for the LU variable. *Saxifraga oppositifolia* was chosen because it is one example of species for which the LU and IG variables did not improve the model fit and the predictive power of GLM. Moreover, the LU variable of this species had a low contribution in the partial regression analysis. Ecological relevance and coherence of selected species distribution among land use categories was finally verified (Figure 6). A map of the land use in the area used for spatial predictions (Figure 7) was produced to help interpreting the distribution maps of the four selected species.

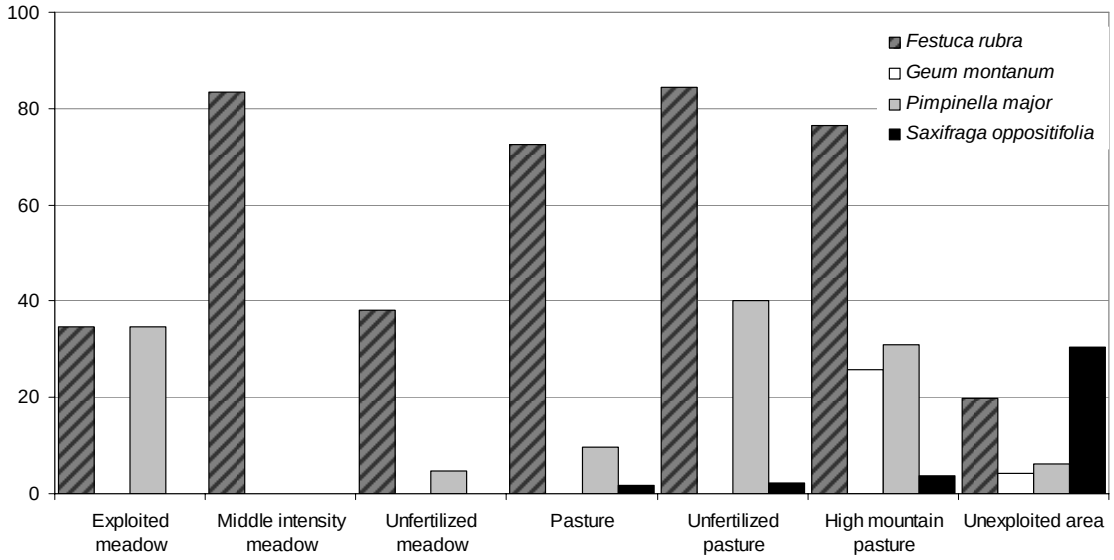


Figure 6 Percentage of occurrence of *Festuca rubra*, *Geum montanum*, *Pimpinella major* and *Saxifraga oppositifolia* among each categories of land use.

Table 8 Performances of p/a models (GLM) for the selected species used for spatial projections.

	p/a models (GLM)										
	Model fit			**Predictive power**			**ANOVA**	**Partial regression**			
	Dxy cal TC	*Dxy cal* LU	Δ *Dxy cal* (TC-LU)	*Dxy eval* TC	*Dxy eval* LU	Δ *Dxy eval* (TC-LU)	*P* value of LU variable	%TC	%LU	%TC+LU	%unexplained
Pimpinella major	0.585	0.721	0.136	0.532	0.651	0.119	<0.001	2.68	7.36	18.56	71.40
Geum montanum	0.790	0.872	0.082	0.723	0.816	0.093	<0.001	0.00	4.06	40.24	55.70
Festuca rubra aggr.	0.690	0.737	0.047	0.675	0.699	0.024	<0.001	0.00	1.45	51.55	47.00
Saxifraga oppositifolia	0.829	0.829	0.000	0.780	0.740	-0.040	0.990	0.00	0.06	40.44	59.50

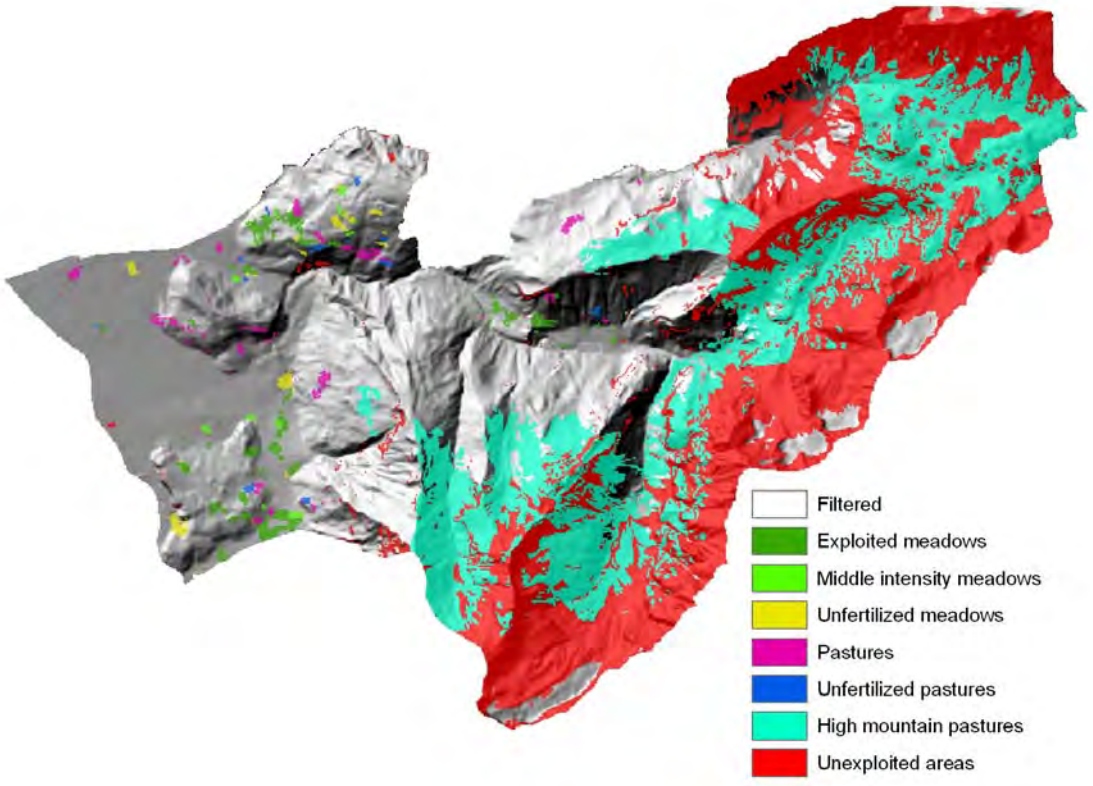


Figure 7 Land use in the area of spatial projection.

Visual observation of the difference between TC and LU maps showed, for instance for *Festuca*, *Pimpinella* and *Geum*, that LU models predict a smaller extent of suitable habitats in unexploited areas and larger extent of suitable habitats in high mountain pastures (Figure 8 and Table 9). These differences are spatially important, reaching an extent of 4.5 km^2 for *Festuca rubra*. The differences between TC and LU models are less important for *Saxifraga*. Only a few suitable surfaces in high mountain pastures are less important for the LU model. *Pimpinella* is more predicted with LU model in exploited meadows (+126.5%) and *Festuca* in middle intensity meadows (+109.5%).

Table 9 Difference in spatial projections when adding LU variable in TC models.

	Pimpinella major			*Geum montanum*			*Festuca rubra*			*Saxifraga oppositifolia*		
	Surface unchanged	% predicted with LU		Surface unchanged	% predicted with LU		Surface unchanged	% predicted with LU		Surface unchanged	% predicted with LU	
Land use categories	km^2	+	-	km^2	+	-	km^2	+	-	km^2	+	-
Exploited meadows	0.40	126.5	0.0	0.90	0.0	0.0	0.86	0.0	4.4	0.90	0.0	0.0
Middle intensity meadows	0.12	0.0	16.4	0.14	0.0	0.0	0.07	109.5	0.0	0.14	0.0	0.0
Unfertilized meadows	0.23	0.0	5.7	0.24	0.0	0.0	0.20	0.0	22.5	0.24	0.0	0.0
Pastures	0.38	0.0	31.4	0.49	0.0	0.0	0.48	1.2	2.6	0.49	0.0	0.0
Unfertilized pastures	0.10	42.3	0.0	0.14	0.0	0.0	0.11	28.3	0.0	0.14	0.0	0.0
High mountain pastures	16.68	18.8	0.0	16.48	20.3	0.0	15.17	30.6	0.0	18.82	0.0	5.3
Unexploited areas	23.57	0.0	14.2	22.93	0.1	17.3	22.21	0.0	21.2	26.17	0.0	2.8

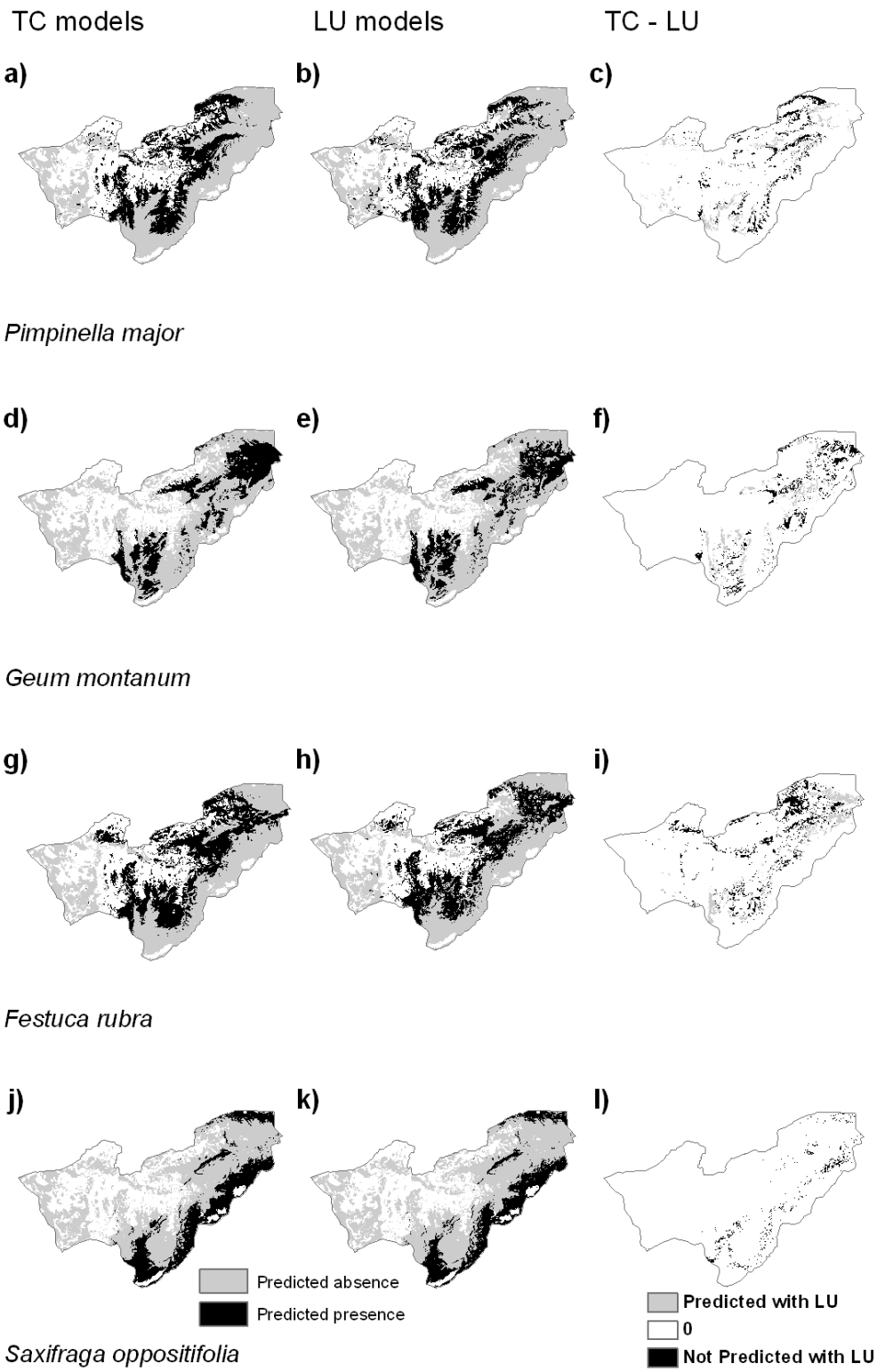


Figure 8 Potential filtered habitat maps for the four exemplar species: (a-c) *Festuca rubra*, (d-f) *Geum montanum*, (g-i) *Pimpinella major* and (j-l) *Saxifraga oppositifolia*). The three maps for each species correspond (by column) to: TC models (a, d, g, j), LU models (b, e, h, k) and the difference between TC and LU models (c, f, I, l).

Discussion

This study aimed at quantifying the importance of land use in predictive habitat distribution models of plants in a mountainous landscape. Four different approaches were used to quantify the contribution of the land use variable (LU) and the grazing index (IG) in presence-absence models (p/a models): model fit, predictive power, additive deviance explained, and partial regressions. A general result is that, on average, LU and IG improved significantly in average the fit of models, but not their predictive power. In contrast, the predictive power of abundance models (LRM) was significantly improved when adding LU or IG variables. ANOVA showed that the reduction of variance by LU and IG variables was significant for 30% and 26% of the species respectively. On the other hand, partial regression analysis showed that individual contribution of LU and IG to the deviance explained by models was very weak on average (1.1% and 0.2% for LU and IG).

Comparison of the different approaches

In this study, one classical (GLM stepwise using ANOVA for model comparison) and partial analysis have been used. The simple GLM comparison approach was not sufficient to precisely assess the contribution of land use in topo-climatic models, particularly when a shared variance is expected. Partial regression analysis is probably the more useful method to assess this information because it allows the partitioning of variance across four identifiable fractions. For land use, values of contribution were very weak, but it highlighted that shared variance is very important between land use and topo-climatic variables. This is coherent with the observation of Thuiller et al. (2004) showing an important relation between bioclimatic and land cover variables and could explain why so few models are significantly improved by adding land use.

At the European scale, land cover did not improve significantly pure bioclimatic models, with.herbs representing the group for which the least improvement was observed (Thuiller et al. 2004). At the regional scale of our study, incorporation of LU and IG in topo-climatic models significantly improved the deviance explained for one third species. However caution is needed when comparing results from these two studies, as several differences exist between them. First, categories of agriculture in Thuiller et al. (2004) derive from a land cover – rather than land use – map (i.e. arable lands, grasslands and permanent crops) and thus are broader than ours. Second, at such coarse resolution, many different land use categories can indeed occur within a single land cover mapping unit, thus making more difficult to identify simple, unambiguous relationships between species' distributions and land cover. Third, according to hierarchical theory, on such continental scale, the distribution of herbs is expected to respond primarily to climatic drivers (Guisan and Thuiller 2005), whereas land use (e.g. agricultural practices) is likely to gain importance while moving toward finer scales (Pearson and Dawson 2004), as observed by Fisher (1994) or Dirnböck et. al (2003).

Land use and species ecology

The analytical framework used in this study – based on multiple approaches – clearly shows that land use can improve both the fit and the predictive power of topo-climatic models, but this is mainly observed when abundances are used as the response variable rather than presence-absences. Furthermore, not all species responded in a same way to the different approaches used to quantify land use contribution. As our results showed overall a lower contribution of grazing index compared to land use, hereafter we only discuss results for the latter.

For the presence-absence models, *Pimpinella major* was the species for which the inclusion of land use increased the model fit and predictive power the most. It is also the species with highest LU contribution from the partial regression analysis. This is not surprising, as *Pimpinella major* is a frequent species in fertilized mountain meadows (*Polygono-Trisetion*; Oberdorfer 1990). On the contrary, *Geum montanum*, the species showing the second best contribution of LU, is mostly found in oligotrophic meadows and in pastures, and is thus affected by fertilization (Oberdorfer 1990). Hence, land use improved the amount of explained deviance for both species fleeing some types of agricultural practices (*Geum montanum*) or being favored by these practices (i.e. benefiting from higher nutrient loading; *Pimpinella major)*. Consequently a very important contribution of land use is to avoid predicting *Geum montanum* or other oligotrophic species (e.g. *Coeloglossum viride*) in fertilized grasslands. Finally, models for some species were neither improved by LU nor by IG variables. *Saxifraga oppositifolia* is an example of such species. It is an alpine herb belonging to the *Thlaspion rotundifolii* and *Drabion hoppeanae* communities (Oberdorfer 1990; Delarze et *al*., 1998). These types of vegetation are usually not grazed due to their weak productivity and very low grass cover.

Land use and changing climate

Despite the overall small amount of deviance explained by land use, it might nevertheless remain an important factor to include in models used to derive projections of climate change impact on plant species and diversity (e.g. Guisan & Theurillat 2000). We see two reasons for this. First, land cover variables (which including some land use types) were shown to be closely related to climate (Thuiller et al. 2004). Thus a change in climate may also lead to a change in land use, but not necessarily in the same direction. Scenarios of land use changes driven by changes in climate and socio-economic variables exist (Gellrich & Zimmermann in press), which can serve as support to derive climate and land use change scenarios. Second, agricultural practices have changed during the last century. At low elevations, intensification of agriculture has occurred which led to widespread changes in the distribution and abundance of many different taxa (Benton et *al.*, 2000). At higher elevations, pasture abandonment is a general trend in Europe because of the lack of competitiveness of mountain agriculture in a globalized world. Such land use changes are likely to modify the sole influence of climate change (by either enforcing or counteracting its effect, depending on the species).

In the Austrian Alps, interaction of these changing scenarios could potentially lead to 25-48% loss of non-forest habitats, depending on the climate change scenario assumed and provided that current pasture pressure remains (Dirnböck et al. 2003). If abandonment of pastures continues, a 42-64% loss of non-forest habitat is expected.

Alpine species have been recognized as being at great risk of local range reduction in case of climate change (Guisan and Theurillat 2000b), but are unlikely to be threaten by changing land use practices. On the other hand, subalpine species may be be less threatened by a changing climate than by heavy land use changes, thus leading to complex scenarios for mountain areas in general (Körner 2005). Further investigations are needed to better understand these interactions between land use and climate, which will need to take pasture abandonment and agricultural intensification into account.

Our results additionally showed that considering, or not, land use can yield important changes in term of surface area predicted by these models. This is an important issue for climate change projections, because the species' range size maintained in the future will determine the species sensitivity to changing conditions but also the future degree of threat of the species (e.g. according to IUCN criteria; Thuiller et al. 2005). Many such differences occurring in an assessment of a large number of species can change significantly the predicted trends of extinction.

Limitations and perspectives

The main limitation of this type of study is the lack of spatially-explicit land use information. In our case, it was obtained for a limited study area only by compiling different sources, and therefore is not available for most alpine areas with sufficient spatial precision. As numerization of agricultural data is currently going on in many European countries, improved data might become available in a close future, which might allow generalizing our exercise to other alpine study areas. Detection of land use from satellite imagery (remote sensing data) is another perspective for obtaining spatially-explicit land use information. The latter approach has already been used successfully for predicting rare plant species richness in agricultural landscapes (Luoto et *al.* 2002). As a future development, we might use our land use data to perform supervised classifications of satellite images and obtain a generalized land use information over the whole study area. However, it seems unlikely that discrimination will be possible for all land use categories with this approach.

The intra- and inter-annual variability of land use practices is also an important limitation. At low elevations, meadows are regularly grazed in early spring and in autumn, and exploitation type is strongly dependant on the farmer's economic situation, which can lead to change in the type of exploitation from one year to the next. Mountain agriculture has also deeply evolved over the last two decades with sometime heavy adaptation in farm management practices taking place. Still,

change in vegetation needs many years to become apparent, and thus the composition of species observed at one time reflects the combination of possibly different agricultural practices over the previous years. A solution to account for inter-annual variability is thus to include historical data, but accounting for intra-annual changes in land use is more difficult, as accurate reporting of agricultural practices across a single year is rarely achieved.

The highest improvement of the fit and the predictive power of models after adding LU was achieved with abundance data. Our results are thus consistent with those of Dirnböck et al. (2003), where land use accounted for more than 50% of explained deviance. Indeed species distribution and abundance react gradually to agricultural treatment and nutrient availability. Historical factors, which locally increased or decreased the abundance, can be expressed through the land use variable. Source-sink dynamic can also explain patterns in abundance (Pulliam 2000). Some species, may occur in some land use category (e.g. due to random dispersion), but would never be dominant there. As a result, change in relative abundance may be noticeable between land use types, whereas change in species composition (presence-absence) may not. Further studies should test, over a larger number of species, whether variation in abundance is more sensitive to change in land use than presence-absence.

Grazing, fertilization and mowing were also shown to have an effect on species richness (Han and Mark, 1998; Jeangros and Bertola, 2002). Hence, it would be interesting, in further studies, to estimates the contribution of land use in species richness and community models. This has not been done in this study because the aim was to assess a general estimation of the contribution in distribution models for a wide range of species.

Conclusion

In summary, this study showed that:

1. Land use is an important variable to take in account when predicting species' distributions, but its contribution is highly dependant on the species considered; Species of oligotrophic habitats do not react in a same way to land use than fertilization-dependent species.
2. The contribution of land use is mostly significant when abundance, rather than simple presence-absence, is used to fit the model. With presence-absence data, land use is confounded with topographic and climatic variables. More investigations is required to elucidate this relationship.
3. Reliable, spatially-explicit land use information is still missing for many mountain areas. One promising approach is to extrapolate land use information from satellite imagery, in order to cover wider areas with complete information.
4. Historical data should be gathered to account for inter-annual variability in land use such as agricultural intensification and pasture abandonment, because these are known to take much importance in alpine habitats.,
5. This study has assessed the importance of land use to explain and predict single species' distributions. Further studies should also use these models to assess land use contribution to explain community composition and diversity.

Acknowledgements

This study was supported by the Centre de Conservation de la Nature (SFFN) of the Canton de Vaud, by the Swiss National Science Foundation (SNF Grant No 3100A0-110000) and the European Commission (FP6 projects ECOCHANGE and MACIS). We thank all the people who helped with the field work in the Swiss Western Alps, particularly Stéfanie Maire, Dario Martinoni, Séverine Wydler, Chantal Peverelli, Valentine Hof, Grégoire Vuissoz, Lorenzo De Stefani, Roxane Milleret, Pierre Bermane Favrod-Coune, Daniel Cherix, Anny and Willy Berra. C.R. and A.G. greatly benefited from interactions with colleagues within the Swiss National Centre for Competence in Research (NCCR) "Plant survival in natural and agricultural ecosystems" (http:// www.unine.ch/nccr).

References

Aeschimann, D., and C. Heitz. 1996. Index synonymique de la flore de Suisse et territoires limitrophes. CRSF, Genève.

Akaike, H. 1973. Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models. Biometrika **60**:255-265.

Azola, C., and F. E. Harrell. 2001. An introduction to S-Plus
and the Hmisc and Design libraries. *in*. University of Virginia
School of Medicine.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology **73**:1045-1055.

Bouët, M. 1985. Climat et météorologie de la Suisse romande, Payot edition, Lausanne.

Carcaillet, C., and J. J. Brun. 2000. Changes in landscape structure in the northwestern Alps over the last 7000 years: lessons from soil charcoal. Journal of Vegetation Science **11**:705-714.

Carey, P. D. 1996. DISPERSE: A cellular automaton for predicting the distribution of species in a changed climate. Global Ecology and Biogeography Letters **5**:217-226.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401 - 417.

Dullinger, S., T. Dirnbock, and G. Grabherr. 2004. Modelling climate change-driven treeline shifts: relative effects of temperature increase, dispersal and invasibility. Journal of Ecology **92**:241-252.

Eler, K., M. Vidrih, and F. Batic. 2005. Vegetation characteristics in relation to different management regimes of calcareous grassland: A functional analysis using plant traits. Phyton-Annales Rei Botanicae **45**:417-426.

Ellenberg, H. 1986. Vegetation Mitteleuropas mit dem Alpen in ökologischer Sicht, 4. Ed edition. Ulmer, Stuttgart.

ESRI. 2005. ARCInfo Version 9.1. *in*. Environmental Systems
Research Institute, Inc., Redlands, CA.

Fischer, H. S. 1994. Simulation der räumlichen Verteilung von Pflanzengesellschaften auf der Basis von Standortskarten. Dargestellt am Beispiel des MaB-Testgebiets Davos. Pages 1 - 158 *in* Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübel, in Zürich.

Guisan, A., and F. E. Harrell. 2000. Ordinal response regression models in ecology. Journal of Vegetation Science **11**:617-626.

Guisan, A., and J.-P. Theurillat. 2000a. Assessing alpine plant vulnerability to climate change: a modeling perspective. Integrated Assessment **1**:307 - 320.

Guisan, A., and J.-P. Theurillat. 2000b. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**:307-320.

Guisan, A., J.-P. Theurillat, and F. Kienast. 1998. Predicting the potential distribution of pland species in an alpine environment. Journal of Vegetation Science **9**:65 - 74.

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

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

Guisan, A., and N. E. Zimmermann. 2000b. Predictive habitat distribution models in ecology. Ecological Modeling **135**:147-186.

Harrell, F. E. 2001. Regression Modeling Strategies. Springer-Verlag, New York. USA.

Kohler, F., F. Gillet, J.-M. Gobat, and A. Buttler. 2004. Seasonal vegetation changes in mountain pastures due to simulated effects of cattle grazing. Journal of Vegetation Science **15**:143-150.

Kohler, F., F. Gillet, J. M. Gobat, and A. Buttler. 2006. Effect of cattle activities on gap colonization in mountain pastures. Folia Geobotanica **41**:289-304.

Körner, C. 2003. Alpine plant life. 2nd Ed. Springer.

Lassueur, T., S. Joost, and C. F. Randin. 2006. Very high resolution digital elevation models: Do they improve models of plant species distribution? Ecological Modelling **198**:139-153.
Legendre, P. 1993. Spatial Autocorrelation: Trouble or New Paradigm? Ecology **74**:1659-1673.
Maurer, K., A. Weyand, M. Fischer, and J. Stocklin. 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.
McCullagh, P. 1980. Regression models for ordinal data. J. R. Statist. Soc. B **42**:109-142.
McCullagh, P., and J. A. Nelder. 1989. Generalized Linear Models. 2nd edition. Chapman and Hall, London.
Milchunas, D. G., and W. K. Lauenroth. 1993. Quantitative Effects of Grazing on Vegetation and Soils Over a Global Randge of Environments. Ecological Monographs **63**:327 - 366.
Olff, H., and M. E. Ritchie. 1998. Effects of herbivores on grassland plant diversity. Trends in Ecology & Evolution **13**:261-265.
Pearson, D., and T. P. Dawson. 2004. Modelling species distribution in Britain: a hierarchical integration of climate and land-cover data. Ecography **27**:285-298.
Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology and Biogeography **12**:361-371.
Pearson, R. G., T. P. Dawson, P. M. Berry, and P. A. Harrison. 2002. SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. Ecological Modelling **154**:289-300.
R. 2006. R 2.4.1 A language and environment. *in*.
Randin, C. F., T. Dirnbock, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography **33**:1689-1703.
Rounsevell, M. D. A., P. M. Berry, and P. A. Harrison. 2006. Future environmental change impacts on rural land use and biodiversity: a synthesis of the ACCELERATES project. Environmental Science & Policy **9**:93-100.
Schreiber, K. F. 1969. Beobachtungen über die Entstehung von "Buckelweiden" auf den Hochflächen des Schweizer Jura. Erdkunde **23**.
Smit, C., D. Béguin, A. Buttler, and H. Müller-Schärer. 2005. Safe sites for tree regeneration in wooded pastures: A case of associational resistance? Journal of Vegetation Science **16**:209-214.
Spiegelberger, T., O. Hegg, D. Matthies, K. Hedlund, and U. Schaffner. 2006. Long-term effects of short-term perturbation in a subalpine grassland. Ecology **87**:1939-1944.
Stevens, C. J., N. B. Dise, J. O. Mountford, and D. J. Gowing. 2004. Impact of nitrogen deposition on the species richness of grasslands. Science **303**:1876 - 1879.
Thuiller, W., M. B. Araujo, and S. Lavorel. 2004. Do we need land-cover data to model species distributions in Europe? Journal of Biogeography **31**:353-361.
Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the United States of America **102**:8245-8250.
van Houwelingen, J. C., and S. Le Cessie. 1990. Predictive value of statistical models. Statistics in Medicine **9**:1303-1325.
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.

Appendix A

Species used with GLM

Acer pseudoplatanus	*Crepis pyrenaica*	*Ligusticum mutellina*
Achillea atrata	*Crocus albiflorus*	*Linaria alpina s.str.*
Achillea millefolium	*Cruciata laevipes*	*Linum alpinum*
Acinos alpinus	*Cynosurus cristatus*	*Linum catharticum*
Adenostyles glabra	*Dactylis glomerata*	*Lolium perenne*
Agrostis alpina	*Dactylorhiza maculata*	*Lotus alpinus*
Agrostis capillaris	*Daphne mezereum*	*Lotus corniculatus*
Agrostis rupestris	*Daucus carota*	*Luzula alpinopilosa*
Agrostis schraderiana	*Deschampsia cespitosa*	*Luzula campestris*
Agrostis stolonifera	*Doronicum grandiflorum*	*Luzula multiflora*
Ajuga reptans	*Draba aizoides*	*Luzula sylvatica*
Alchemilla conjuncta aggr.	*Dryas octopetala*	*Medicago lupulina*
Alchemilla coriacea aggr.	*Elyna myosuroides*	*Myosotis alpestris*
Alchemilla glabra aggr.	*Epilobium anagallidifolium*	*Myosotis arvensis*
Alchemilla xanthochlora aggr.	*Erigeron uniflorus*	*Myosotis decumbens*
Androsace chamaejasme	*Euphorbia cyparissias*	*Nardus stricta*
Anemone narcissiflora	*Euphrasia hirtella*	*Nigritella rhellicani*
Anthoxanthum odoratum aggr.	*Euphrasia minima*	*Parnassia palustris*
Anthriscus sylvestris	*Euphrasia rostkoviana s.str.*	*Pedicularis foliosa*
Anthyllis vulneraria s.l.	*Euphrasia salisburgensis*	*Pedicularis verticillata*
Aposeris foetida	*Festuca arundinacea s.l.*	*Phleum alpinum aggr.*
Arabis alpina s.str.	*Festuca ovina aggr.*	*Phleum hirsutum*
Arabis caerulea	*Festuca pratensis s.l.*	*Phleum pratense*
Arnica montana	*Festuca quadriflora*	*Phleum rhaeticum*
Arrhenatherum elatius	*Festuca rubra aggr.*	*Phyteuma orbiculare*
Asplenium viride	*Festuca violacea aggr.*	*Phyteuma spicatum*
Aster bellidiastrum	*Fragaria vesca*	*Picea abies*
Astrantia major	*Fraxinus excelsior*	*Pimpinella major*
Athamanta cretensis	*Galeopsis tetrahit*	*Pimpinella saxifraga aggr.*
Bartsia alpina	*Galium album*	*Plantago alpina*
Bellis perennis	*Galium anisophyllon*	*Plantago atrata s.str.*
Biscutella laevigata	*Galium megalospermum*	*Plantago lanceolata*
Botrychium lunaria	*Galium pumilum*	*Plantago major s.str.*
Brachypodium pinnatum	*Gentiana acaulis*	*Plantago media*
Briza media	*Gentiana bavarica*	*Poa alpina*
Bromus erectus s.str.	*Gentiana campestris s.str.*	*Poa minor*
Bromus hordeaceus	*Gentiana clusii*	*Poa pratensis*
Calamagrostis varia	*Gentiana lutea*	*Poa supina*
Campanula barbata	*Gentiana purpurea*	*Poa trivialis s.str.*
Campanula cochleariifolia	*Gentiana verna*	*Polygala alpestris*
Campanula rhomboidalis	*Geranium sylvaticum*	*Polygala chamaebuxus*
Campanula rotundifolia	*Geum montanum*	*Polygala vulgaris s.str.*
Campanula scheuchzeri	*Geum rivale*	*Polygonum bistorta*
Cardamine pratensis	*Geum urbanum*	*Polygonum viviparum*
Carduus defloratus s.str.	*Glechoma hederacea s.str.*	*Polystichum lonchitis*
Carex caryophyllea	*Globularia cordifolia*	*Potentilla aurea*
Carex ferruginea	*Gymnadenia conopsea*	*Potentilla crantzii*
Carex flacca	*Gypsophila repens*	*Potentilla erecta*
Carex montana	*Hedysarum hedysaroides*	*Potentilla sterilis*
Carex nigra	*Helianthemum nummularium s.l.*	*Primula auricula*
Carex ornithopoda	*Helictotrichon pubescens*	*Primula elatior s.str.*
Carex pallescens	*Helictotrichon versicolor*	*Primula veris s.l.*
Carex panicea	*Heracleum sphondylium s.l.*	*Pritzelago alpina s.str.*
Carex sempervirens	*Hieracium bifidum aggr.*	*Prunella grandiflora*
Carex sylvatica	*Hieracium lactucella*	*Prunella vulgaris*
Carlina acaulis subsp. caulescens	*Hieracium murorum aggr.*	*Pulsatilla alpina s.str.*
Carum carvi	*Hieracium villosum aggr.*	*Ranunculus aconitifolius*
Centaurea jacea s.str.	*Hippocrepis comosa*	*Ranunculus acris s.l.*
Centaurea montana	*Holcus lanatus*	*Ranunculus alpestris*
Centaurea scabiosa s.l.	*Homogyne alpina*	*Ranunculus bulbosus*
Cerastium fontanum subsp. vulgare	*Hypericum maculatum s.str.*	*Ranunculus montanus aggr.*
Cerastium latifolium	*Hypericum perforatum s.str.*	*Ranunculus nemorosus aggr.*
Chaerophyllum aureum	*Hypochaeris radicata*	*Ranunculus repens*
Chaerophyllum hirsutum aggr.	*Juncus effusus*	*Rhinanthus alectorolophus*
Cirsium acaule	*Juniperus communis subsp. nana*	*Rhododendron ferrugineum*
Cirsium palustre	*Knautia arvensis*	*Rumex acetosa*
Cirsium spinosissimum	*Knautia dipsacifolia s.str.*	*Rumex alpestris*
Clinopodium vulgare	*Laserpitium latifolium*	*Rumex alpinus*
Coeloglossum viride	*Lathyrus pratensis*	*Rumex crispus*
Colchicum autumnale	*Leontodon autumnalis*	*Sagina saginoides*
Crepis aurea	*Leontodon hispidus s.l.*	*Salix reticulata*
Crepis biennis	*Leucanthemum vulgare aggr.*	*Salix retusa*

Species used with GLM	Species used with GLM/LRM and LU variable	Species used with GLM/LRM and IG variable
Salvia pratensis	*Agrostis capillaris*	*Achillea millefolium*
Sanguisorba minor s.str.	*Ajuga reptans*	*Agrostis capillaris*
Saxifraga aizoides	*Alchemilla glabra aggr.*	*Ajuga reptans*
Saxifraga moschata s.l.	*Bellis perennis*	*Alchemilla glabra aggr.*
Saxifraga oppositifolia	*Brachypodium pinnatum*	*Anthriscus sylvestris*
Saxifraga paniculata	*Centaurea jacea s.str.*	*Astrantia major*
Scabiosa lucida	*Cirsium acaule*	*Campanula cochleariifolia*
Sedum atratum	*Dactylis glomerata*	*Carex flacca*
Selaginella selaginoides	*Daucus carota*	*Clinopodium vulgare*
Senecio doronicum	*Festuca rubra aggr.*	*Dactylis glomerata*
Sesleria caerulea	*Galium album*	*Daucus carota*
Silene acaulis	*Galium anisophyllon*	*Festuca rubra aggr.*
Silene dioica	*Gentiana verna*	*Galium album*
Silene vulgaris s.l.	*Helictotrichon pubescens*	*Galium anisophyllon*
Soldanella alpina	*Heracleum sphondylium s.l.*	*Gentiana lutea*
Solidago virgaurea s.l.	*Hieracium villosum aggr.*	*Gentiana verna*
Stellaria graminea	*Hypericum maculatum s.str.*	*Geranium sylvaticum*
Taraxacum alpinum aggr.	*Leontodon hispidus s.l.*	*Helictotrichon pubescens*
Taraxacum officinale aggr.	*Lotus alpinus*	*Hieracium villosum aggr.*
Thesium alpinum	*Pimpinella major*	*Hypericum maculatum s.str.*
Thlaspi repens	*Plantago atrata s.str.*	*Knautia dipsacifolia s.str.*
Thymus praecox subsp. polytrichus	*Plantago lanceolata*	*Leontodon hispidus s.l.*
Thymus pulegioides s.str.	*Polygonum viviparum*	*Leucanthemum vulgare aggr.*
Tofieldia calyculata	*Potentilla crantzii*	*Lotus alpinus*
Tragopogon pratensis subsp. orientalis	*Prunella grandiflora*	*Luzula multiflora*
Traunsteinera globosa	*Ranunculus alpestris*	*Plantago atrata s.str.*
Trifolium badium	*Rhinanthus alectorolophus*	*Plantago lanceolata*
Trifolium medium	*Sanguisorba minor s.str.*	*Plantago major s.str.*
Trifolium pratense s.str.	*Scabiosa lucida*	*Poa pratensis*
Trifolium repens s.str.	*Silene vulgaris s.l.*	*Polygonum viviparum*
Trifolium thalii	*Taraxacum officinale aggr.*	*Potentilla crantzii*
Trisetum flavescens	*Trisetum flavescens*	*Potentilla erecta*
Trollius europaeus	*Vaccinium myrtillus*	*Ranunculus nemorosus aggr.*
Tussilago farfara	*Vicia sepium*	*Rhinanthus alectorolophus*
Urtica dioica		*Silene vulgaris s.l.*
Vaccinium gaultherioides		*Solidago virgaurea s.l.*
Vaccinium myrtillus		*Taraxacum officinale aggr.*
Vaccinium vitis-idaea		*Trifolium thalii*
Valeriana montana		*Trisetum flavescens*
Valeriana tripteris		*Veronica serpyllifolia s.l.*
Veratrum album subsp. lobelianum		
Veronica alpina		
Veronica aphylla		
Veronica arvensis		
Veronica chamaedrys		
Veronica officinalis		
Veronica persica		
Veronica serpyllifolia s.l.		
Vicia cracca s.str.		
Vicia sativa s.l.		
Vicia sepium		
Viola biflora		
Viola calcarata		
Viola hirta		

Chapter 3

Introducing process-based geomorphic and climatic variables into predictive models of alpine plant distribution in the Western Swiss Alps

Randin Christophe[1], Vuissoz Grégoire[1], Liston Glen[3], Vittoz[2] Pascal and Guisan Antoine[1]

This manuscript will be submitted to Arctic, Antarctic and Alpine Research.

[1]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]University of Lausanne, Faculty of Geosciences and Environment, Bâtiment Biophore, 1015 Lausanne, Switzerland

[3]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado 80523, U.S.

Abstract

Topographic predicting variables have been successfully used as surrogate for disturbance processes in plant distribution modelling studies in mountain environment. However, no study so far tested disturbance variables like geomorphic perturbations or snow translocation processes.

In this study, we integrated two process-based disturbance variables into presence-absence species distribution models based on commonly used topo-climatic variables. We developed a geomorphic perturbation index and an index of snow distribution by wind and we assessed how they improved the predictive power and the model fit of 91 sub-alpine and alpine plant species in the Western Swiss Alps. We also quantified the additive variance of process-based variables compared to pure topo-climatic models of presence-absence. Finally, predictive maps of presence-absence models were prepared for a limited set of species to test the effect of process-based variables spatially.

A general result was that, on average, the geomorphic (GI) and snow distribution (SDI) index improved significantly the fit of models, but their predictive power as evaluated by cross-validation was not better or lower than that of topoclimatic models. The reduction of variance by GI and SDI variables was significant for 13.2% and 9.9% of the species, respectively. In addition, partial regression analyses showed that individual contribution of GI and SDI to the deviance explained by models was very weak on average (2.4% for GI and 1.7% for SDI). In contrast, the maximum contribution of GI and SDI to the predictive power of models was important, reaching 11.9% and 35.6% respectively. Finally, spatial predictions of models suggest that the inclusion of GI into TC models led to a doubling of predicted surfaces for some categories, whereas significant divergence with TC models were observed with SDI models for the two species considered.

The contributions of GI and SDI were highly dependant on the species considered and their ecological characteristics. These geomorphic disturbance variables may thus be regarded as key variables for some species in mountain systems, especially in a context of climate change. Considering them or not can yield important changes in term of surface area predicted by SDMs. These changes in projected surfaces were interpreted as ecologically relevant when considering the two process-based disturbance variables. This is an important issue when preparing climate change projections, because the species' range size maintained in the future will determine the species sensitivity to changing conditions.

Keywords: generalized linear models (GLM), habitat distribution, vegetation, geomorphology, snow translocation, partial regression analysis, Western Swiss Alps, independent evaluation.

Nomenclature: Aeschimann and Heitz (1996)

Introduction

Alpine environments with their great landscape variability represent a challenge for understanding and modelling plant species distribution (Guisan et al. 1998). Large altitudinal gradients that generate hydro-mechanical disturbances and contrasted energy fluxes between exposed and shaded faces can explain the great observed diversity of species distribution patterns.

One crucial task when building a species distribution models (SDM; Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005) is selecting the environmental variables to be used for predictions. Various descriptors of the environment can be potentially used, reflecting topographic, climatic, edaphic, geologic or human influences. These variables can be further classified as having indirect or direct effects on the species, and constitute a resource or not for them. At fine, local scales in mountain environments, models based on indirect predictors like topographic ones (e.g. altitude or aspect) can still provide accurate local predictions. However, on global to meso-scales, like Switzerland and larger areas, models based on indirect predictors loose power compared to models based on more proximal predictors (Guisan and Hofer, 2003). This is because at these scales, same values of indirect predictors at distant sites can translate into very different values of the more proximal predictors, due to e.g. major climatic or geologic differences taking place between entire regions within the modelled area. Hence, although models fitted with indirect predictors at the local scale can still provide accurate predictions for the same area, they might lack power to predict to other areas, i.e. they can show weak transferability (Randin et al., 2006).

In order to have accurate models of plant distribution in alpine systems that still have a good generality and transferability to other areas, we need to develop and use preferentially predictor variables that have, as much as possible, a direct eco-physiological or ecological impact on plants (regulators) or constitute resources for them (Austin, 1980; Austin et al., 1984; Austin, 1985; Austin and Heyligers, 1989; Guisan and Zimmermann, 2000) (Fig. 1). A current limitation of many of theses models is the lack of spatially explicit predictors expressing dynamic processes occurring at – and possibly modifying - ground surface, such as geomorphic and snow translocation processes.

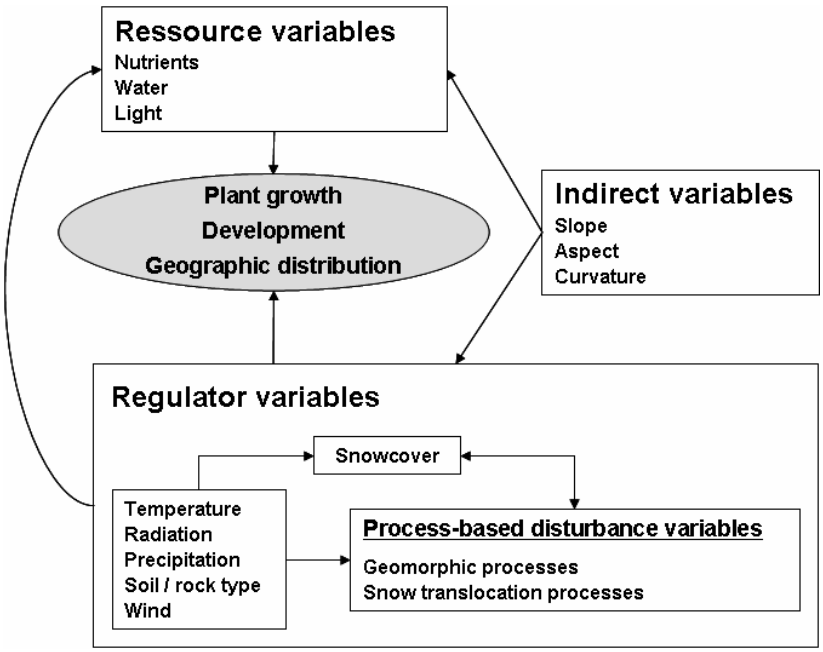


Figure 1 Hierarchical modeling framework of resource, regulator and indirect variables used in species distribution models (adapted from Guisan & Thuiller 2005).

At high elevation in mountain areas, like those found in the alpine and nival vegetation belts of the Swiss Alps, the very rugged landscape configuration leads to very active and dynamic hydro-mechanical processes, like geomorphic ones. These processes affect vegetation at the local scale (Burga, 1999) but also modify the species diversity at larger scale (Nichols et al., 1998). Gravity or hydrodynamic processes like rock slides, avalanches or solifluction (Johnson and Billings, 1962; Erschbamer, 1989; Körner, 2003) are examples of processes that have a big impact on the soil surface and strongly drive patterns of species distribution (Bretz-Guby, 1994). Some adaptive traits of plant species are highly correlated with the conditions encountered in extremely disturbed situations, such as large root networks on moving rock slides (Jonasson and Callaghan, 1992).

In alpine landscape, the spatial and temporal distribution (depth, persistence) of snow cover is influenced by topography but also by wind (Greene et al., 1999; Liston et al., 2000; Tappeiner et al., 2001). The snow cover itself determines species composition and spatial patterns of vegetation (Walsh et al., 1994; Körner, 2003). Snow cover offers protection against climatic stress, particularly wind desiccation (Schaefer and Messier, 1995),and constitutes a direct reservoir of nutrients for plant growth (Bowman, 1992). But snow also affects soil and vegetation moisture levels and can represent a severe stress factor through its dynamic effects like avalanche paths or wind-induced snow translocation (Erschbamer, 1989; Körner, 2003). As a result, its overall effect if to decrease plant productivity (Billings and Bliss, 1959) and photosynthetic activity (Körner, 2003).

Topographic position (Gottfried et al., 1998; Guisan et al., 1998; Gottfried et al., 1999; Guisan and Theurillat, 2000; Dirnböck et al., 2003), drainage surface (Leathwick et al., 1998) or distance to ridges (Moore et al., 1991; Dirnbock et al., 2002) have been successfully used as surrogate for disturbance processes in plant distribution modelling studies. Most of these were conducted in mountain environment (Gottfried et al., 1998; Guisan et al., 1998; Gottfried et al., 1999; Guisan and Theurillat, 2000; Dirnböck et al., 2003). However, no study so far tested disturbance variables like geomorphic perturbations. We are also not aware of any study using physical and mechanistic models of snow distribution as input in plant potential distribution (SDM).

In this study, we integrated two process-based disturbance variables into presence-absence species distribution models based on commonly used topo-climatic variables. We developed a geomorphic perturbation index and an index of snow distribution by wind and we assessed how they improved the predictive power and the model fit of 91 sub-alpine and alpine plant species in the Western Swiss Alps. We also quantified the additive variance of process-based variables compared to pure topo-climatic models of presence-absence. Finally, predictive maps of presence-absence models were prepared for a limited set of species to test the effect of process-based variables spatially.

For this, we transformed the vegetation relevant information contained in maps of geomorphology into a perturbation index and we calculated an index of snow distribution by wind with a dynamic physical model.

Methods

Study area

The study area of Anzeindaz (7°7′ to 7°11' E; 46°15′ to 46°18′ N) is a west-east oriented plateau of 25 km^2 located in the western Alps of Switzerland (Fig.2). This area is bordered on the North by the Diablerets massif and to the South by the Muveran massif. Elevation ranges from 1900 to 3210 m at the top of the Diableret peak. Annual temperature and precipitation vary respectively from 1°C and 1800 mm at 1900 m to -5°C and 2600 mm at 3000 m along the elevation gradient (Bouët, 1985). The soil parent material is mainly calcareous.

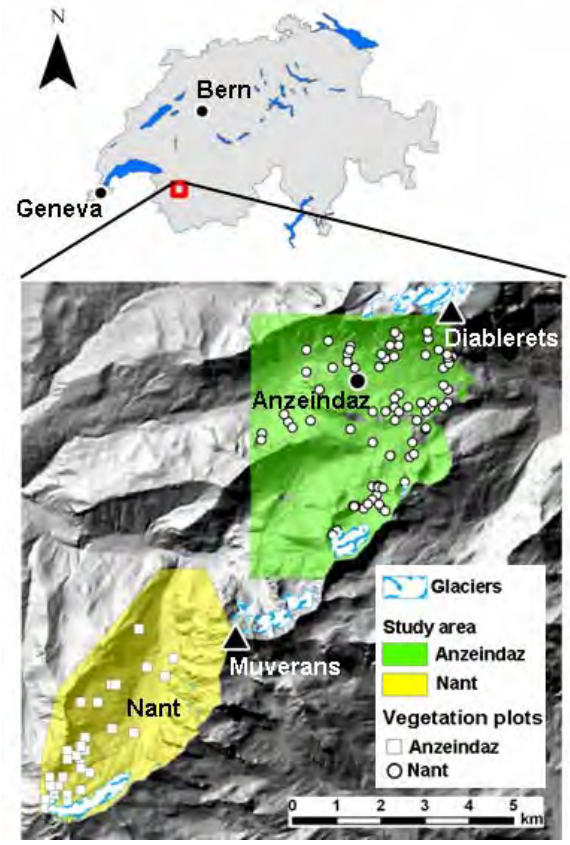


Figure 2 Geographic location of the study area with Anzeindaz and Nant regions. Nant is used for an independent evaluation of models.

The lowest part of the study area of Anzeindaz is just below the natural treeline ecotone. However, centuries of wood exploitation and cattle grazing has since long lowered this ecocline (Villaret, 1973; Gehrig-Fasel et al., 2007). Trees are now absent of the area as subalpine and alpine grasslands are still grazed every summer by cattle. This probably had little effect on the composition of alpine meadows (Randin et al., in prep; chapter 2). The study area of Anzeindaz holds 300 plant

species, which corresponds to one tenth of the Swiss flora (Villaret, 1973; Aeschimann and Burdet, 1994).

This area is also well known from geographers because of to the wide range of geomorphic processes occurring and several studies provide very detailed geomorphic maps of the area (Loye and Pahud, 1996; Rappaz and Hunziker, 1996).

A small valley in the vicinity, the Vallon de Nant was used as independent evaluation area for the validation of predictive models (Fig.2). This area is similar to the study area of Anzeindaz as it well represents the alpine environment. However the Vallon de Nant has geographic peculiarities and presents less variations of slope aspect than the calibration area, due to its unique south-north orientation. Nevertheless, this area was chosen because of the availability of geomorphic maps (Phillips, 1993).

Species data

During summers 2002-2003, 49 vegetation plots were sampled (as part of the sampling of a larger area; see Randin et al. 2006) following a random-stratified sampling strategy restricted to non-woody vegetation (grassland, rock and scree). Stratification was done by elevation, slope, aspect and simplified classes of geology. Plots were composed of a surface of 64 m^2 (8 x 8 m) and presence of all species was recorded in this surface. We registered the exact coordinates of the centre of each plot using a Geoexplorer 3 Trimble® GPS system, with an accuracy of less than 1 m after differential post-correction. A second random-stratified sampling was performed during summer 2004, focusing especially on geomorphic units. In addition, geomorphic strata were used, providing 32 additional plots. 91 species occurring in $\geq$ 10 plots were used in analyses (Appendix A). The Nant dataset was collected during summers 2002-2003 using the same random-stratified sampling strategy. It is composed of 31 vegetation plots of 64-m^2 with presence-absence records for the same 91 species.

Modelling framework

All steps of the analysis are summarized in Fig. 3. Two process-based disturbance variables (PBD) were added to topo-climatic (TC) species distribution models (SDM) and their additional contribution to predictive power was tested. These were geomorphic classes (ordinal) and an index of snow distribution by wind.(continuous). Hereafter, we refer to TC models for SDMs including only topo-climatic predictors, and to PDB models for those including topo-climatic and process-based variables (geomorphic index – GI and snow distribution index - SDI).

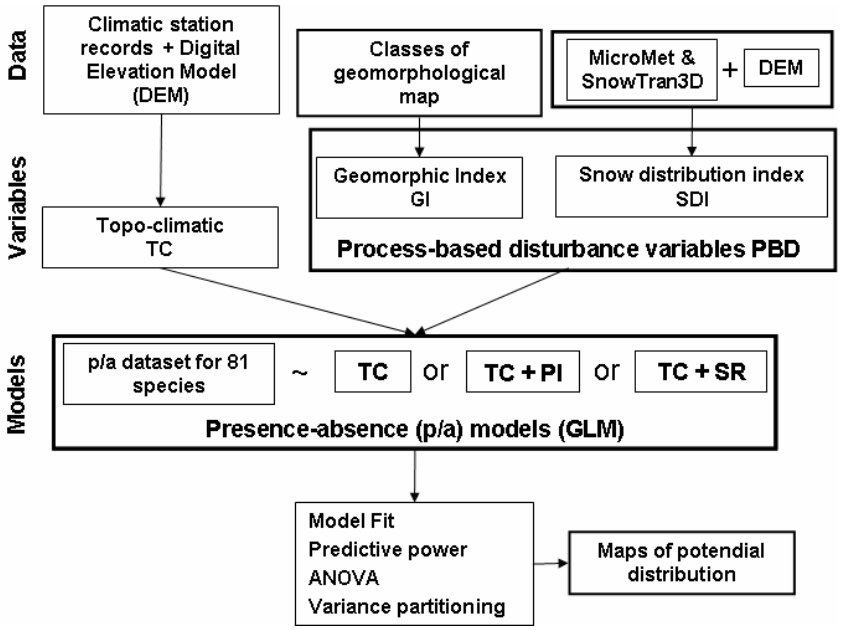


Figure 3 Experimental design of the study.

TC and PBD models were calibrated in Anzeindaz and model fit was evaluated for both models. Predictive power of TC and PBD models was evaluated with a cross-validation procedure on the training data (Anzeindaz) and with independent data (Nant). ANOVA was used to test the significance of variance reduction of PBD variables when added to TC models. We used variance partitioning to quantify the independent contribution of TC and PBD variables in SDMs. TC and PBD models were then implemented into a GIS to produce distribution maps. Difference in spatial predictions between TC and PBD models were compared visually.

Process-based disturbance variables

Published geomorphic maps of Anzeindaz and Nant (Phillips, 1993; Rappaz and Hunziker, 1996) were first digitalized in the ArcGIS 9.1 software (ESRI, 2005). Geomorphic units were then merged into three classes characterized by increasing intensity of perturbation (Table 1) and of potential impact on the vegetation.

Table 1 Classification of geomorphologic units by their potential levels of perturbation on the vegetation.

Geomorphologic perturbations			
No perturbation	**Low**	**Medium**	**High**
0	**1**	**2**	**3**
Soil on stabilized rock	Stabilized glacial deposit	Rock fall	Avalanche paths
	Stabilized dejection cone	Gravity accumulation	Rock scree
	Lapiaz	Solifluxion	Active rock glacier
		Active glacial deposit	

The calculation of the snow distribution index (SDI) was made with the physically-based numerical SnowTran-3D snow transport model (SnowTran-3D; Liston and Sturm, 1998), using the 25 m digital elevation model (DEM) and data on average dominant winds. Local wind direction and intensity was calculated in a quasi-physically-based meteorological model (MicroMet; Liston and Elder, 2006). A 1 m uniform snow layer was then applied to the whole area as the initial state and SnowTran-3D was run to simulate where snow will accumulate and where it will disappear. Snow was blown during 5 days with an average dominant wind of 280° at 15m/s (Meteoswiss). The resulting map represents an index of snow redistribution by wind. It ranges between 0 and positive infinite values, with positive values where snow accumulates and null values where snow completely disappears.

Climatic and topographic variables

Climate data were derived from the national meteorological stations at different altitudes. Long-term monthly means for average temperature (°C) and sum of precipitation (mm) for the period 1961-1990 were used. The 25 m DEM was used to spatially interpolate the climatic data, and thus all derived data were also at this resolution of 25 m. Method of computation and description of the variable are given in Zimmermann and Kienast (1999). In this study, we used three climatic and two topographic variables (Table 2) expected to be of greatest ecophysiological significance for plants (Guisan and Zimmermann, 2000; Pearson et al., 2002; Körner, 2003; Guisan and Thuiller, 2005). Degree-days above 0°C were derived from interpolated daily temperatures and summed up for the growing season. Moisture index was calculated as the difference between precipitation and evapotranspiration, and thus expressed the amount of soil water potentially available at a site. The sum of mean daily values for the months of June, July and August and the annual potential global solar radiations were were additionally calculated. Topographic position and slope were derived from the DEM. Positive values of topographic position express relative ridges, tops and exposed sites, whereas negative values indicate sinks, valleys or toe slopes.

Table 2 Topographic and climatic variables used in the present study to model the distribution of species.

Variables	Units	Details	Method	References
Temperature degree days	°C * day * year^{-1}	Sum of days multiplied by temperature > 0 C	ARCInfo AML	Zimmermann & Kienast (1999)
Moisture index (average of monthly values June–August)	mm * day^{-1}	Monthly average of daily atmospheric H_2O balance	ARCInfo AML	Zimmermann & Kienast (1999)
Global solar radiation (sum of monthly values)	kJ * m^{-2} * day^{-1}	Monthly average of daily global solar radiation	ARCInfo AML	Zimmermann & Kienast (1999)
Slope	degrees	Slope inclination	DEM, ARCInfo, GRID routine	ESRI (2005)
Topographic position	uniteless	Integration of topographic features (ridge, slope, toe slope) at various spatial scales	ARCInfo AML	Zimmermann (1996)

Statistical analyses

All statistical analyses were performed in the R software (R 2.4.1 A language and environment, 2006). Analyses were run on all species with more than 10 occurrences in the training dataset (91 plant species, Appendix).

For each species, a generalized linear model (GLM; (McCullagh and Nelder, 1989) with a binomial variance and a logistic link function was fitted using presence-absence as the response variable and climate and topographic variables as predictors (TC model). An Akaike information criterion (AIC)-based stepwise procedure in both directions was used to select the most significant predictors (Akaike, 1973). Up to second-order polynomials (linear and quadratic terms) were allowed for each predictor, with the linear term being forced in the model each time the quadratic term was retained. In a second step, the geomorphic index (GI model) and the index of snow distribution by wind (SDI model) were then added independently to the TC model.

Model fit and predictive power of the TC, GI and SDI models were both estimated with the Somer's concordance D_{xy} index (Harrell, 2001), by applying the models to the training data and to pseudo-independent data obtained by 10-fold cross-validation (see van Houwelingen and Le Cessie, 1990; see Randin et al., 2006 for the exact implementation) respectively. During the cross-validation, the original prevalence of the species was maintained in each fold. D_{xy} is based on the Wilcoxon-Mann-Whitney two-samples rank test and takes values between 0 and 1, ranging from totally random to perfectly discriminating models. Negative values signify predictions worse than random..

A fully external evaluation was made by predicting all models on the Nant dataset and comparing predictions and observations again with Somer's concordance D_{xy}. This represents a fully independent evaluation as recommended by Fielding and Bell (1997). As the independent evaluation area (Nant) does not provide a sufficient number of sampling plots (31 plots) for a robust evaluation, it can only be used as an indicative and explorative test set for evaluating model predictions.

Three different methods were used to quantify the contribution of process-based variables in the GLM: (1) difference in model fit and predictive power between the models, (2) ANOVA and (3) variance partitioning. Variance partitioning allows the comparisons of two groups of predictors, whereas the ANOVA quantifies the additional variance explained by GI and SDI variables in the TC model. Implementations of these three different methods are presented below.

A general estimation of model fit and predictive power with and without GI or SDI variables is obtained by calculating the mean D_{xy}. This was done on the training and validation datasets. The contribution is simply measured as the difference in D_{xy} between GI or SDI models and TC models. The reduction in deviance caused by the addition of GI or SDI to the TC model was assessed by an ANOVA model

comparison test. To quantify the variance added by the GI and SDI variables, a variance partitioning approach based on partial regression analyses was used. This approach allows partitioning the variance into four identifiable fractions of explained variance (Borcard et al., 1992; Legendre, 1993): (i) pure TC variables, (ii) shared TC and GI (or SDI) , (iii) pure GI (or SDI) and (iv) unexplained variation.

Potential habitat maps

Three maps were prepared for some exemplar species, based on: (1) TC model, (2) GI (or SDI) model, (3) their difference. Maps 1 and 2 were reclassified into presence-absence using an AUC-optimized threshold and a mask based on forests, roads, urbanized areas and rivers was further applied to avoid predicting the species in impossible situations. Map 3 was prepared by subtracting maps 1 and 2 (i.e. TC – GI (or SDI)). These latter maps were reclassified into three categories: (i) no or few differences between TC and GI (or SDI) model (values between -0.3 and 0.3), (ii) GI (or SDI) probabilities of occurrence higher than TC (from -1 to -0.3) and, reciprocally, (iii) GI (or SDI) probabilities lower than TC (from 0.3 to 1). All maps were calculated from GLM outputs in the ArcGIS 9.1 software (ESRI, 2005).

Results

Model improvement

The average values of Somer's D_{xy} for the 91 species on the calibration (model fit) and evaluation (predictive power) datasets for models with topo-climatic variables (TC), perturbation index (GI) and snow distribution index (SDI) yielded mean values around 0.6 (close to 0.8 for the calibration), signifying that few models failed on average (Table 3). GI contributed to a significant variance reduction for 13.2% of the species and SDI for 9.9% of species (ANOVA, *P* value at 0.05 level).

The boxplot (Fig. 4) showed that the model fits were on average significantly higher for models including GI and SDI variables, compared to the pure TC models. By contrast, the predictive power of TC models evaluated by cross-validation remained significantly higher than that of GI and SDI models. However, there were no significant differences in predictive power between TC and GI and between TC and SDI models (Fig.4) when models were evaluated in the independent study area (Nant). In addition, D_{xy} were very low for all models in Nant and significant differences remained, overall, very small.

Table 3 Model evaluation and number of species with a significant reduction of variance for GI and SDI in the TC models.

	Evaluation of models (GLM) with D_{xy}						**ANOVA TC model + GI (or SDI)**
	Model fit		**Predictive power**				
	Calibration		**Cross-validation**		**Nant dataset**		**Number of species *P* < 0.05**
	TC	**GI**	**TC**	**GI**	**TC**	**GI**	
Min	-0.12	0.11	-0.12	0.00	-0.46	-0.46	12
Mean	0.79	0.82	0.63	0.58	0.30	0.31	**% of species**
Max	1.00	1.00	0.83	0.84	0.97	0.97	13.2
	TC	**SDI**	**TC**	**SDI**	**TC**	**SDI**	**Number of species *P* < 0.05**
Min	-0.12	0.08	-0.12	-0.19	-0.46	-0.46	9
Mean	0.79	0.82	0.63	0.59	0.30	0.28	**% of species**
Max	1.00	1.00	0.88	0.92	0.97	0.93	9.9

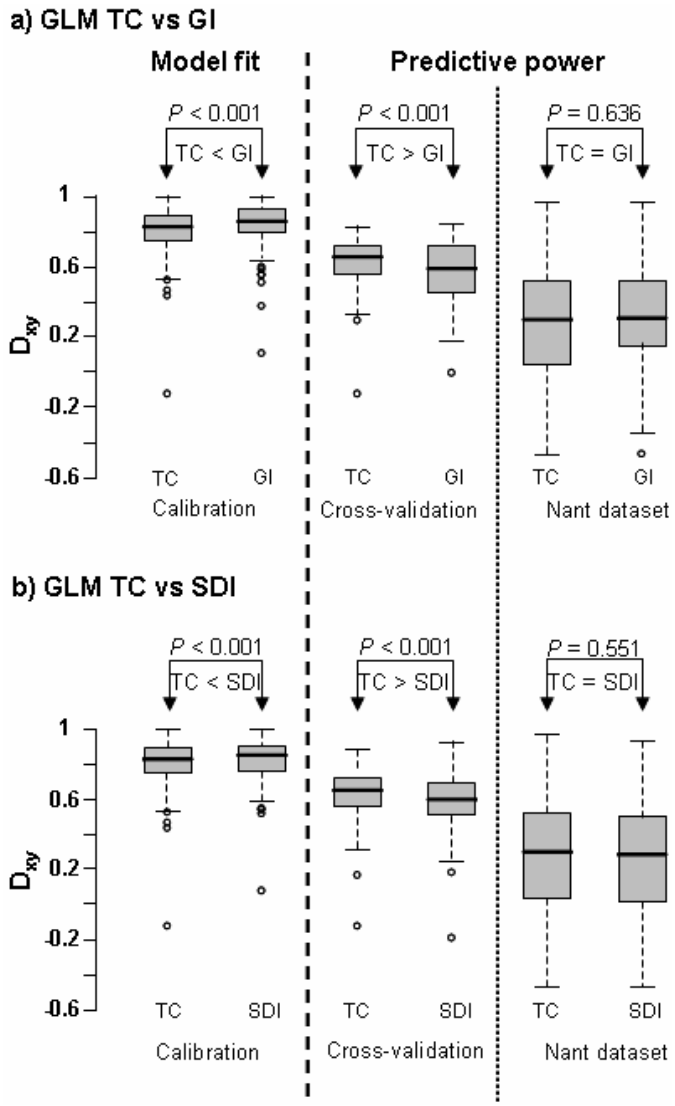


Figure 4 Boxplots of (a) model fit (*Dxy* on the calibration dataset) and (b) predictive power (*Dxy* on the dataset from cross-validation and from Nant) of the three models (TC = topoclimatic, GI = with geomorphic index, SDI = with snow distribution index). The *P* values indicate the significance of the Wilcoxon signed rank tests between models. The > or < symbols indicate the direction of the test.

Difference in model fit and predictive power

GI and SDI both improved the model fit of TC models but these differences were low on average. The difference in model fit (*Δ Dxy cal)* between GI and TC (*Δ Dxy cal* = 0.035) and between SDI and TC (*Δ Dxy cal* = 0.031) were similar (Table 4). The *Δ Dxy cal* was the highest for *Botrychium lunaria* (*Δ Dxy cal* = 0.228) between TC and GI models and between TC and SDI models (*Δ Dxy cal* = 0.198). The weakest difference between TC and GI models was observed for *Hieracium murorum* (*Δ Dxy cal* = - 0.574) and between TC and SDI models for *Poa alpina* (*Δ Dxy cal* = - 0.012).

On the contrary, the difference in predictive power (*Δ Dxy eval)* was on average lower than TC models for both GI (*Δ Dxy eval* = -0.050 for cross-validation and *Δ Dxy eval* = -0.317 in Nant) and SDI models (*Δ Dxy eval* = -0.330 and (*Δ Dxy eval* = -0.050 for cross-validation and *Δ Dxy eval* = -0.050 in Nant; see Table 4) when evaluated with cross-validations or in Nant. The highest *Δ Dxy eval* between TC and GI models for the evaluation with cross-validation was observed for *Botrychium lunaria* (*Δ Dxy eval* = 0.119), and between TC and SDI models for *Sedum atratum* (*Δ Dxy eval* = 0.356). The highest differences are more important for evaluation in Nant (*Deschampsia cespitosa* with *Δ Dxy eval* = 0.374 for GI and *Alchemilla glabra* with *Δ Dxy eval* = 0.410 for SDI). The weakest difference between TC and GI when evaluated by cross-validation was observed for *Hieracium murorum* (*Δ Dxy eval* = -0.219) and for *Festuca quadriflora* (*Δ Dxy eval* = -1.050) when evaluated with independent data in Nant. Similarly, The weakest difference between TC and SDI models was observed for *Luzula multiflora* (*Δ Dxy eval* = -0.351) when evaluated by cross-validation and for *Androsace chamaejasme* (*Δ Dxy eval* = -1.047) when evaluated with independent data in Nant.

Table 4 Differences in model fit (*Δ D_{xy} cal*) and predictive power (*Δ D_{xy} eval*) between pure topoclimatic models (TC) and expanded models (GI and SDI). Predictive power was evaluated with an internal validation (cross-validation: CV) and with an external validation (on the Nant dataset).

	Δ *Dxy cal*		**Δ *Dxy eval***			
			GI - TC		**SDI - TC**	
	GI - TC	**SDI - TC**	**CV**	**Nant**	**CV**	**Nant**
Mean	0.035	0.031	-0.050	-0.317	-0.039	-0.330
Min	-0.574	-0.012	-0.219	-1.050	-0.351	-1.047
Max	0.228	0.198	0.119	0.374	0.356	0.410

Partial regression analysis

The contribution of GI to explain additional variance in TC models was on average rather weak (2.4%, Table 5), reaching 20.7% in the best model for *Pulsatilla alpina* (Figure 5). The contribution of SDI was also on average weak (1.7%), with the highest contribution observed for the model of *Poa alpina* (8.5%) (Table 5, Fig.5).

Table 5 Percentage of mean, minimum and maximum of the explained variance for each identifiable fraction when considering process-based disturbance (PBD) with topo-climatic variables (TC).

	Geomorphic index (GI)			Snow distribution index (SDI)		
	Mean	Min	Max	Mean	Min	Max
TC	2.0	0.0	24.8	9.1	0.0	46.1
PB	2.4	0.0	20.7	1.7	0.0	8.5
TC + PBD	44.4	0.0	100.0	37.8	0.4	100.0
Unexplained	51.1	0.0	100.0	51.4	0.0	93.9

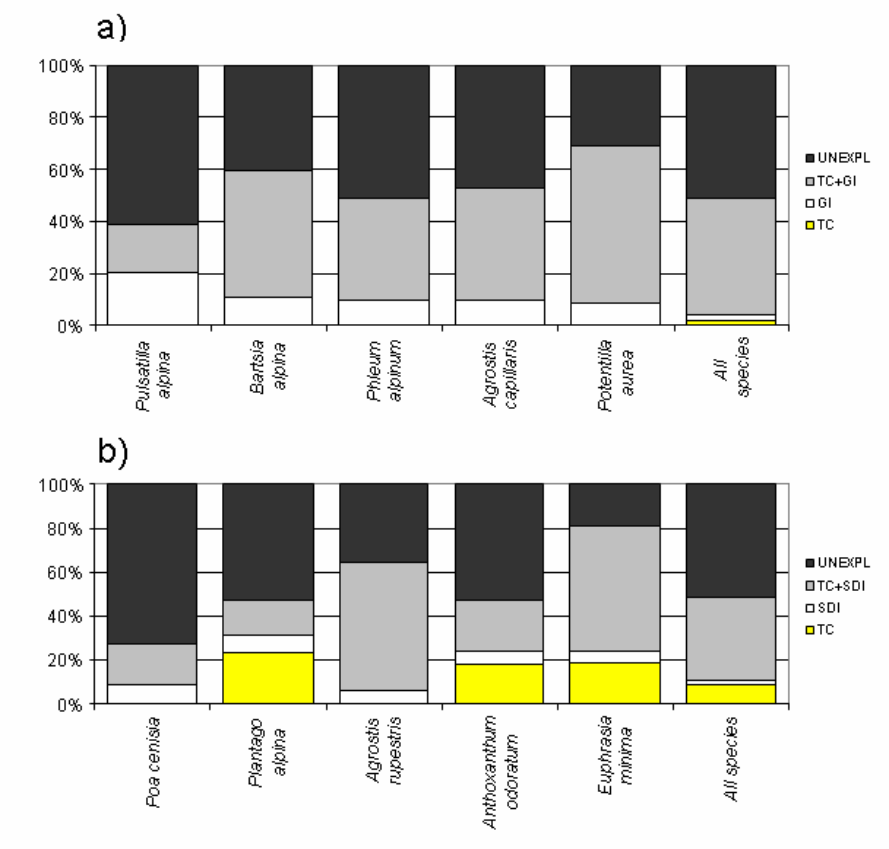


Figure 5 Fraction of explained variance among the target variables for the five species having the highest D^2 for land use (a) and growing index (b).

Potential distribution maps for TC, GI and SDI models

Four species were selected to illustrate spatial projections from presence-absence models, based on model performances (Table 6). *Agrostis capillaris* and *Galium megalospermum* were selected to show the improvement by GI variable whereas *Sedum atratum* and *Silene acaulis* were selected for improvement by SDI variable. A map of the GI and SDI variables in the study area (Fig.6) was produced to help interpreting the spatial predictions for these selected species.

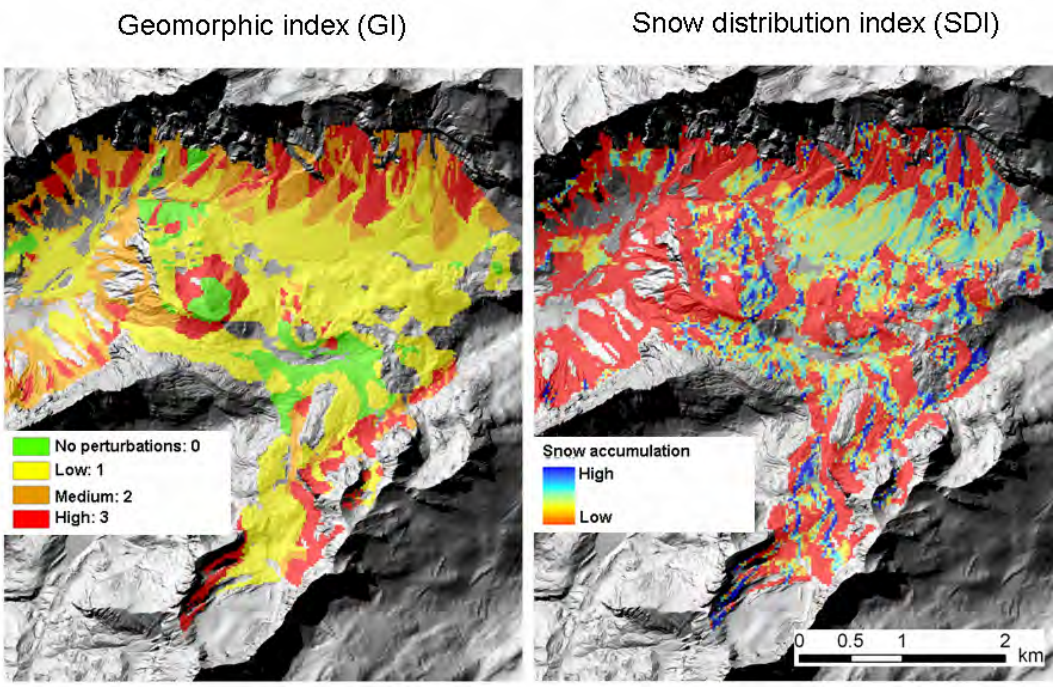


Figure 6 Maps of the geomorphic index (GI) and the index of snow distribution (SDI) used as predicing variables in distribution models.

Comparisons of spatial projections of TC and GI models (Fig.7; Table 7) showed that *A. capillaris* was more often predicted on surfaces without or with low geomorphic perturbations. On the contrary, *G. megalospermum* was more often predicted on surfaces with medium to high perturbations. When comparing TC and SDI models, boxplots (Fig.8) showed that occurrences of *S. atratum* were more often predicted on areas with high accumulation of snow whereas *S. acaulis* was more often predicted on surfaces with low snow accumulation when including SDI into models. These differences were all significant (Wilcoxon; $P < 0.001$).

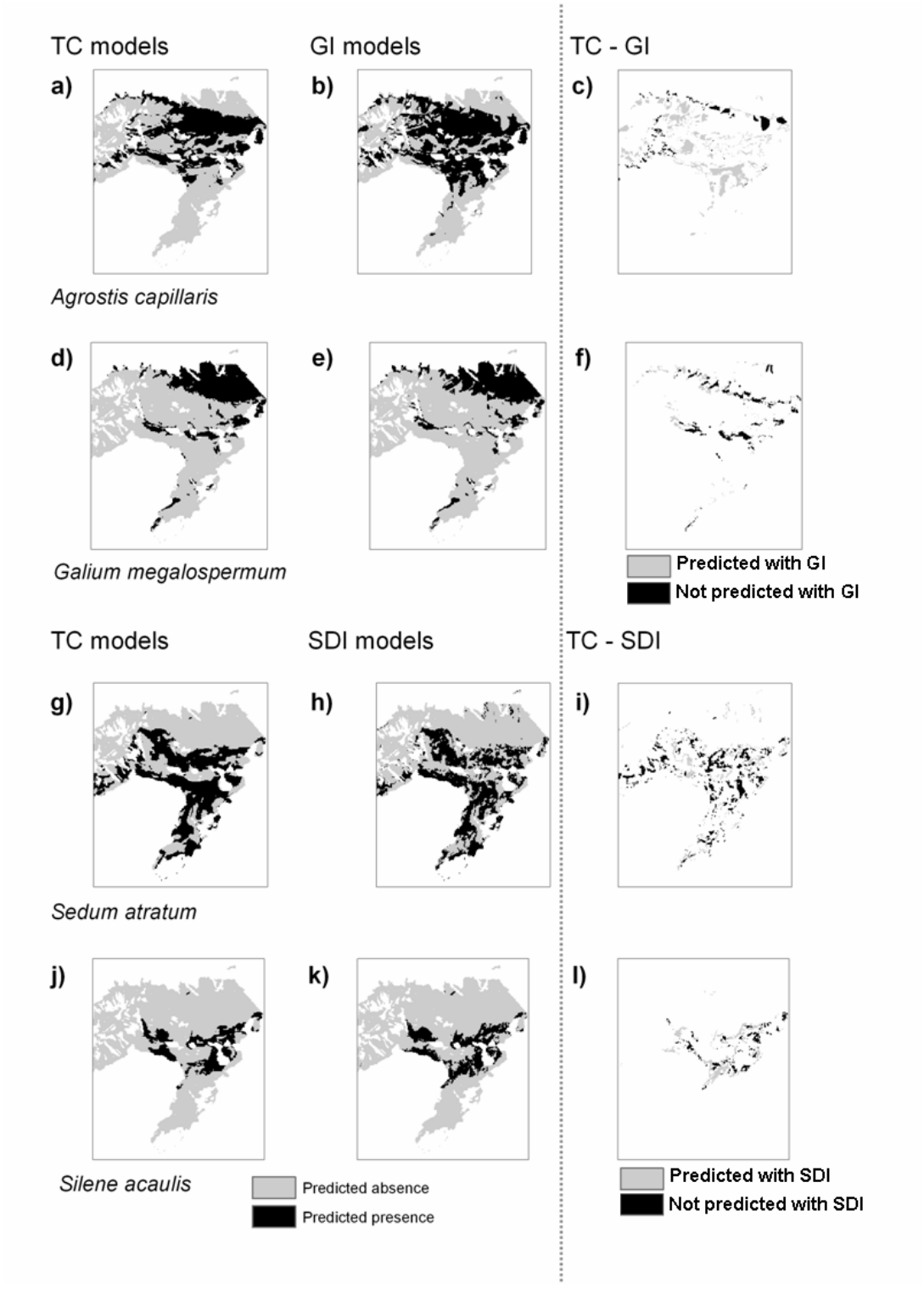


Figure 7 Potential filtered habitat maps for the four exemplar species: (a-c) *Agrostis capillaris*, (d-f) *Galium megalospermum*, (g-i) *Sedum atratum* and (j-l) *Silene acaulis*). The three maps for each species correspond (by column) to: TC models (a, d, g, j), GI or SDI models (b, e, h, k) and the difference between TC and GI or SDI models (c, f, i, l).

Table 6 Performances models (GLM) for the four selected species used for spatial projections.

	GLM models										
	Model fit			Predictive power			ANOVA	Partial regression			
	Dxy cal TC	*Dxy cal* GI	Δ Dxy cal (GI-TC)	*Dxy eval* TC	*Dxy eval* GI	Δ Dxy eval (GI-TC)	*P* value of GI variable	%TC	%GI	%TC+GI	%unexplained
Agrostis capillaris	0.728	0.866	0.138	0.687	0.779	0.092	0.004	0.0	9.5	43.3	47.2
Galium megalospermum	0.876	0.903	0.027	0.689	0.759	0.070	0.047	22.3	2.6	37.8	37.2
	Dxy cal TC	*Dxy cal* SDI	Δ Dxy cal (SDI-TC)	*Dxy eval* TC	*Dxy eval* SDI	Δ Dxy eval (SDI-TC)	*P* value of SDI variable	%TC	%SDI	%TC+SDI	%unexplained
Sedum atratum	0.521	0.683	0.162	0.164	0.520	0.356	0.144	0.0	3.5	19.5	77.0
Silene acaulis	0.947	0.967	0.020	0.704	0.815	0.111	0.065	0.0	1.1	74.8	24.1

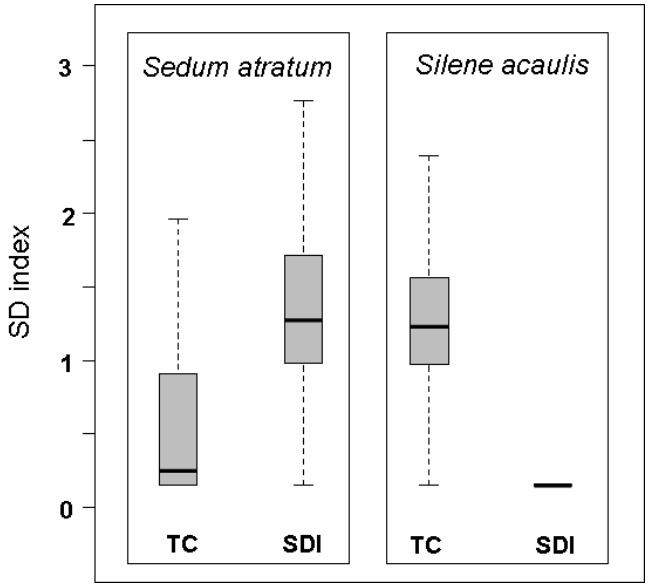


Figure 8 Boxplots of the simulated snow distribution index for the potential occurrences predicted with TC and SDI models for *Sedum atratum* and *Silene acaulis*. Difference between the snow distribution index of occurrences predicted with TC and SDI were significantly different (Wilcoxon test: $P < 0.001$).

Table 7 Difference in spatial projections when adding the GI variable in TC models.

	Agrostis capillaris			*Galium megalospermum*		
	Surface unchanged	% predicted with GI		Surface unchanged	% predicted with GI	
GI categories	km^2	+	-	km^2	+	-
No perturbations	0.67	123.4	0.0	1.29	0.0	17.2
Low perturbations	4.50	11.2	0.0	4.77	1.2	3.7
Medium perturbatinons	2.02	0.0	18.8	2.15	11.4	0.0
High perturbations	1.73	7.9	0.0	1.75	0.8	6.0

Discussion

In this study we quantified the importance of process-based disturbance variables in models of alpine plant species distribution, compared to more commonly used topo-climatic (TC) variables. Two new variables were tested: a geomorphic index (GI) and a snow distribution index (SDI). A general result was that, on average, GI and SDI improved significantly the fit of models, but their predictive power as evaluated by cross-validation was not better or lower than that of TC models. The reduction of variance by GI and SDI variables was only significant for 13.2% and 9.9% of the species, respectively. In addition, partial regression analyses showed that individual contribution of GI and SDI to the deviance explained by models was very weak on average (2.4% for GI and 1.7% for SDI). In contrast, the maximum contribution of GI and SDI to the predictive power of models was important, reaching 11.9% and 35.6% respectively. The independent evaluation of the predictive power in the remote area did not show consistent results with the other methods to calculate the contribution of GI and SDI. Finally, spatial predictions of models suggest that the inclusion of GI into TC models led to a doubling of predicted surfaces for some categories, whereas significant divergence with TC models were observed with SDI models for the two species considered.

Predictive power of process-based disturbance variables

One reason to explain the weak global evaluation scores of models incorporating the GI variable may reside in the way geomorphic maps were obtained. Geomorphic processes were delimited on the map by large adjacent polygons, without transition zones between them. As these abrupt limits do not exist in nature, the use of a continuous perturbation index should prove more powerful than the semi-quantitative index used here. Another problem with geomorphic maps is that their units were not initially defined by geographers in relation to vegetation. A different classification, intended to predict vegetation patterns, could prove beneficial here. This highlights the problem of using existing and available data for another purpose than they were originally designed (Yoccoz et al., 2001).

The low predictive performance of the snow distribution index illustrates the difficulty to capture the complex processes of snow distribution in mountain systems. Patterns of snow accumulation are not only generated by the dominant wind translocation, but also by avalanches or more complex wind translocation pathways. Due to the steep slopes and the high elevation amplitude (up to 1300 m) found within the study area, avalanches could be an important driver that explains patterns of snow distribution in Anzeindaz, but not in Vallon de Nant. Moreover, snowbeds and exposed ridges are landscape features that are also affected by the interaction between topo-climatic processes (i.e. slope, curvature and solar radiations) and wind. Consequently, topo-climatic variables commonly used in SDMs may be sufficient to capture the main geomorphic influences on the vegetation. In addition, the dominant wind that was blown to distribute the snow in the landscape is not necessarily the wind that really shapes the snow after snowfalls at the micro-

scale. Summarizing, the strong influence of topo-climatic variables could explain the low average contribution of both GI and SDI variables when added to TC models, as highlighted by the results from variance partitioning. Indeed, the intensity of geomorphicprocesses, such as screes or solifluction is often partly described by topo-climatic variables. However, what is supposedly improved in geomorphic predictors is that slope, curvature, radiations or precipitation act in combination with edaphic or geologic characteristics. Our results tend to suggest that more refinement is needed in the preparation of geomorphic predictors in alpine environments.

When evaluated in the remote area of Nant, the average predictive power of the GI and SDI models was low. Several study recommended the use of direct or regulator variables instead of indirect variables in SDM, as a way to ensure a better transferability of models in space or time (e.g. Guisan and Zimmermann 2000; Randin et al. 2006). Our study demonstrated the difficulty to generate a set of spatially explicit direct (i.e. ecologically-meaningful) variables that can be generalized to a complex environment such as mountain systems. Moreover, when using existing maps of geomorphology, their quality as predictor of vegetation will strongly depends on the interpretation of the landscape made by geographers. However, to our knowledge, no study formally tested the predictive power of direct variables in a remote area. In addition, when evaluated in Nant, the predictive power of TC and PBD models were the same in average. This may demonstrate that the evaluation in a remote area is a very severe method and for all set of variables.

Expert approaches in a GIS or advanced remote sensing techniques may be used to determine more precisely the intensity of geomorphic perturbation (Tarboton, 1997; Meissl, 2001; Noetzli et al., 2006). For example, the size and the spatial organization of stones in a rock slide may give information about the intensity and the frequency of the perturbation events occurring there. The snow distribution model used in this study was only based on one single climatic variable, which is the dominant wind. Therefore, the simulation of snow distribution used to derive our index only represented one possible realization, without any variation associated to it (i.e. how far from the average prediction?). More complex physical models of snow distribution, based on several simulation runs, on more thorough observed climatic data recorded over longer periods, and on more complex processes of snow formation and melting, should be tested in the future.

Statistical models of species' distribution were fitted using species presence-absence that were derived from abundance data. In further studies, models should be fitted using abundance directly, rather than presence absence. This could sensibly change the results, because abundance is ecologically more informative than simple presence/absence. In this context, using ordinal SDMs, Dirnböck et al. (2003) showed that landuse (a form of human disturbance) could explain up to 50% of variation in species abundance.

Species projections

Contrary to strict but blind quantitative evaluations, spatial projections of GI and SDI models suggest some improvement of TC models by GI and SDI variables. For instance, *Agrostis capillaris* is a grassland species that has a large ecological tolerance. It is mainly found in exploited meadows and pastures, but only in subalpine and alpine grasslands, wetlands, fallows, heathlands and some forests (Aeschimann et al., 2005). The species can be occasionally and locally abundant but is generally never dominant and avoids perturbations. As expected, it was more often predicted by GI models in areas without perturbation (+123%), On the contrary, *Gallium megalospermum* is a species typically found in the calcareous screes of subalpine and alpine zones (Aeschimann et al., 2005). This species was correctly predicted more often in areas with medium perturbations (+11.4%) and less often in areas without perturbations (-17.2%). *Sedum atratum* is found in calcareous snowbeds, whereas *Silene acaulis* prefers south slopes and windy ridges (Braun-Blanquet, 1975). The divergence of SD projections compared to TC models for these species was consistent with field observations since *Sedum* was significantly more predicted in high snow accumulation areas and *Silene* in low accumulation.

Process-based disturbance for future predictions

As shown with spatial projections, process-based disturbance variables in topo-climatic models might remain important factors to include in SDMs when they are used to derive projections of climate change impact on plant species and diversity. For those 10% of species sensitive to GI and SDI variables, geomorphic perturbations and snow distribution can act as facilitators or inhibitors (Raffl et al., 2006), thus acting as barrier or corridor for plant dispersal. Therefore, migration rate of pioneer species from low altitude, which are adapted to mechanical perturbations, could be accelerated, whereas populations of high-elevation species could remain found at low elevation in some abyssal situations where low competition is regulated by perturbation. Moreover, rockfalls caused by the degradation of permafrost are likely to increase in a warmer climate (Gruber et al., 2004), thus leading to changes in connectivity or barriers for sensitive species. For instance, snowbed species currently located at their lower distributional limit may be threatened by climate change, leading to the local extinctions of very specialized species.

Conclusion

In this study, we showed that:

1. Process-based disturbance variables were only important for 10% of the species. The contributions of these variables were highly dependant on the species considered and their ecological characteristics. These geomorphic disturbance variables may thus be regarded as key variables for some species in mountain systems, especially in a context of climate change.
2. Considering them or not can yield important changes in term of surface area predicted by SDMs. These changes in projected surfaces were interpreted as ecologically relevant when considering the two process-based disturbance variables. This is an important issue when preparing climate change projections, because the species' range size maintained in the future will determine the species sensitivity to changing conditions.
3. Models fitted with geomorphic disturbance variables failed to predict in a remote area, thus highlighting the difficulty to obtain spatially-explicit and ecologically meaningful direct variables.
4. Further studies should develop a continuous geomorphic index based on more ecologically meaningful units of geomorphology, and consider using more complex models of snow distribution, even at the cost of longer computer time.

Acknowledgments

This study was supported by the Centre de Conservation de la Nature (SFFN) of the Canton de Vaud, by the Swiss National Science Foundation (SNF Grant No 3100A0-110000) and the European Commission (FP6 projects ECOCHANGE and MACIS). We thank all the people who helped with the field work in the Swiss Western Alps, particularly Stéfanie Maire, Dario Martinoni, Séverine Wydler, Chantal Peverelli, Valentine Hof, Grégoire Vuissoz, Lorenzo De Stefani, Roxane Milleret, Pierre Bermane Favrod-Coune, Daniel Cherix, Anny and Willy Berra. C.R. and A.G. greatly benefited from interactions with colleagues within the Swiss National Centre for Competence in Research (NCCR) "Plant survival in natural and agricultural ecosystems" (http:// www.unine.ch/nccr).

References

Aeschimann, D. and Burdet, H. M., 1994: *Flore de la Suisse et des territoires limitrophes. Le nouveau Binz*. 2 ed. Neuchâtel: Le Griffon, 597 pp.

Aeschimann, D. and Heitz, C., 1996: *Index synonymique de la flore de Suisse et territoires limitrophes*. Genève: CRSF, 317 pp.

Aeschimann, D., Lauber, K., Moser, D. M., and Theurillat, J.-P., 2005: *Flora Alpina*. Belin ed. Paris.

Akaike, H., 1973: Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models. *Biometrika,* 60: 255-265.

Austin, M. P., 1980: Searching for a model for use in vegetation analysis. *Vegetatio,* 42: 11-21.

Austin, M. P., Cunningham, R. P., and Fleming, P. M., 1984: New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. *Vegetatio,* 55: 11-27.

Austin, M. P., 1985: Continuum concept, ordination methods and niche theory. *Annu. Rev. Ecol. Syst.,* 16: 39-61.

Austin, M. P. and Heyligers, P. C., 1989: Vegetation survey design for conservation: gradsec sampling of forests in north-east New South Wales. *Biol. Conserv.,* 50: 13-32.

Billings, W. D. and Bliss, L. C., 1959: An Alpine Snowbank Environment and Its Effects on Vegetation, Plant Development, and Productivity. *Ecology,* 40: 388-397.

Borcard, D., Legendre, P., and Drapeau, P., 1992: Partialling out the spatial component of ecological variation. *Ecology,* 73: 1045-1055.

Bouët, M., 1985: *Climat et météorologie de la Suisse romande*. Payot ed. Lausanne.

Bowman, W. D., 1992: Inputs and Storage of Nitrogen in Winter Snowpack in an Alpine Ecosystem. *Arctic and Alpine Research,* 24: 211-215.

Braun-Blanquet, J., 1975: Fragmenta Phytosociologica Raetica I: Die Schneebodengesellschaften (Klasse der Salicetea herbaceae). *Jahresbericht der Naturforschenden Gesesellschaft Graubünden,* 96: 42-71.

Bretz-Guby, N., 1994: Géomorphologie et végétation à l'étage alpin: L'exemple du Mont-Gelé, Institut de Géographie, Université de Lausanne, Lausanne. Pages pp.

Burga, C. A., 1999: Vegetation development on the glacier forefield Morteratsch (Switzerland). *Applied Vegetation Science,* 2: 17-24.

Dirnbock, T., Hobbs, R. J., Lambeck, R. J., and Caccetta, P. A., 2002: Vegetation distribution in relation to topographically driven processes in southwestern Australia. *Applied Vegetation Science,* 5: 147-158.

Dirnböck, T., Dullinger, S., and Grabherr, G., 2003: A regional impact assessment of climate and land-use change on alpine vegetation. *Journal of Biogeography,* 30: 401-417.

Erschbamer, B., 1989: Vegetation on Avalanche Paths in the Alps. *Vegetatio,* 80: 139-146.

ESRI, 2005: ARCInfo Version 9.1. Redlands, CA: Environmental Systems

Research Institute, Inc.

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

Gehrig-Fasel, J., Guisan, A., and Zimmermann, N. E., 2007: Treeline shifts in the Swiss Alps: climate change or land abandonment? *Journal of Vegetation Science,* 18.

Gottfried, M., Pauli, H., and Grabherr, G., 1998: Prediction of vegetation patterns at the limits of plant life: A new view of the alpine-nival ecotone. *Arctic and Alpine Research,* 30: 207-221.

Gottfried, M., Pauli, H., Reitter, K., and Grabherr, G., 1999: A fine-scaled predictive model for changes in species distribution patterns of high mountain plants induced by climate warming. *Diversity & Distributions,* 5: 241-251.

Greene, E. M., Liston, G. E., and Pielke, R. A., 1999: Simulation of above treeline snowdrift formation using a numerical snow-transport model. *Cold Regions Science and Technology,* 30: 135-144.

Gruber, S., Hoelzle, M., and Haeberli, W., 2004: Permafrost thaw and destabilization of Alpine rock walls in the hot summer of 2003. *Geophysical Research Letters,* 31.

Guisan, A., Theurillat, J. P., and Kienast, F., 1998: Predicting the potential distribution of plant species in an Alpine environment. *Journal of Vegetation Science,* 9: 65-74.
Guisan, A. and Theurillat, J.-P., 2000: Assessing alpine plant vulnerability to climate change: A modeling perspective. *Integrated Assessment,* 1: 307-320.
Guisan, A. and Zimmermann, N. E., 2000: Predictive habitat distribution models in ecology. *Ecological Modelling,* 135: 147-186.
Guisan, A. and Hofer, U., 2003: Predicting reptile distributions at the mesoscale: relation to climate and topography. *Journal of Biogeography,* 30: 1233-1243.
Guisan, A. and Thuiller, W., 2005: Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* 8: 993-1009.
Harrell, F. E., 2001: *Regression Modeling Strategies*. New York. USA: Springer-Verlag, 60-61 pp.
Johnson, P. L. and Billings, W. D., 1962: Alpine Vegetation of Beartooth Plateau in Relation to Cryopedogenic Processes and Patterns. *Ecological Monographs,* 32: 105-&.
Jonasson, S. and Callaghan, T. V., 1992: Root Mechanical-Properties Related to Disturbed and Stressed Habitats in the Arctic. *New Phytologist,* 122: 179-186.
Körner, C., 2003: Alpine plant life. 2nd Ed. *Springer*.
Leathwick, J. R., Burns, B. R., and Clarkson, B. D., 1998: Environmental correlates of tree alpha-diversity in New Zealand primary forests. *Ecography,* 21: 235-246.
Legendre, P., 1993: Spatial Autocorrelation: Trouble or New Paradigm? *Ecology,* 74: 1659-1673.
Liston, G. E. and Sturm, M., 1998: A snow-transport model for complex terrain. *Journal of Glaciology,* 44: 498-516.
Liston, G. E., Winther, J. G., Bruland, O., Elvehoy, H., Sand, K., and Karlof, L., 2000: Snow and blue-ice distribution patterns on the coastal Antarctic Ice Sheet. *Antarctic Science,* 12: 69-79.
Liston, G. E. and Elder, K., 2006: A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). *Journal of Hydrometeorology,* 7: 217-234.
Loye, E. and Pahud, P.-F., 1996: Levé géomorphologique de Solalex-Anzeindaz, Institut de Géographie, Facultés des Lettres, University of Lausanne, Lausanne. Pages pp.
McCullagh, P. and Nelder, J. A., 1989: *Generalized Linear Models. 2nd edition*. London: Chapman and Hall.
Meissl, G., 2001: Modelling the runout distances of rockfalls using a geographic information system, Z.
Geomorphol. *Geomorphology,* 125: 129-137.
Moore, D. M., Lees, B. G., and Davey, S. M., 1991: A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environmental Management,* 15: 59-71.
Nichols, W. F., Killingbeck, K. T., and August, P. V., 1998: The influence of geomorphological heterogeneity on biodiversity II. A landscape perspective. *Conservation Biology,* 12: 371-379.
Noetzli, J., Huggel, C., Hoelzle, M., and Haeberli, W., 2006: GIS-based modelling of rock-ice avalanches from Alpine permafrost areas. *Computational Geosciences,* 10: 161-178.
Pearson, R. G., Dawson, T. P., Berry, P. M., and Harrison, P. A., 2002: SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. *Ecological Modelling,* 154: 289-300.
Phillips, M., 1993: Géomorphologie du Vallon de Nant, Institut de Géographie, Facultés des Lettres, University of Lausanne, Lausanne. Pages pp.
R 2.4.1 A language and environment, 2006: R Development Core Team.
Raffl, C., Mallaun, M., Mayer, R., and Erschbamer, B., 2006: Vegetation succession pattern and diversity changes in a glacier valley, Central Alps, Austria. *Arctic Antarctic and Alpine Research,* 38: 421-428.
Randin, C. F., Dirnbock, T., Dullinger, S., Zimmermann, N. E., Zappa, M., and Guisan, A., 2006: Are niche-based species distribution models transferable in space? *Journal of Biogeography,* 33: 1689-1703.
Randin, C. F., Jaccard, H., Vittoz, P., and Guisan, A., in prep: Importance of « Land use » in predictive habitat modeling of plant species distribution in the Swiss Alps.
Rappaz, F. and Hunziker, J.-l., 1996: Levé géomorphologique d'Anzeindaz, Institut de Géographie, Facultés des Lettres, University of Lausanne, Lausanne. Pages pp.

Schaefer, J. A. and Messier, F., 1995: Scale-Dependent Correlations of Arctic Vegetation and Snow Cover. *Arctic and Alpine Research,* 27: 38-43.

Tappeiner, U., Tappeiner, G., Aschenwald, J., Tasser, E., and Ostendorf, B., 2001: GIS-based modelling of spatial pattern of snow cover duration in an alpine area. *Ecological Modelling,* 138: 265-275.

Tarboton, D. G., 1997: A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research,* 33: 309-319.

van Houwelingen, J. C. and Le Cessie, S., 1990: Predictive value of statistical models. *Statistics in Medicine,* 9: 1303-1325.

Villaret, P., 1973: Flore et végétation d'Anzeinde. *Mém. Soc. Vaud. Sc. Nat.,* 15: 139-145.

Walsh, S. J., Butler, D. R., Allen, T. R., and Malanson, G. P., 1994: Influence of snow patterns and snow avalanches on the alpine treeline ecotone. *Journal of Vegetation Science,* 5: 657-672.

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.

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

Appendix A

List of species used in analysis

Adenostyles glabra
Agrostis capillaris
Agrostis rupestris
Alchemilla conjuncta aggr.
Alchemilla glabra aggr.
Alchemilla xanthochlora aggr.
Androsace chamaejasme
Anthoxanthum odoratum aggr.
Anthyllis vulneraria s.l.
Asplenium viride
Aster bellidiastrum
Bartsia alpina
Botrychium lunaria
Campanula barbata
Campanula scheuchzeri
Carduus defloratus s.str.
Carex sempervirens
Carlina acaulis subsp. caulescens
Cerastium latifolium
Cirsium spinosissimum
Coeloglossum viride
Crepis aurea
Deschampsia cespitosa
Dryas octopetala
Euphorbia cyparissias
Euphrasia minima
Festuca quadriflora
Festuca rubra aggr.
Festuca violacea aggr.
Galium anisophyllon
Galium megalospermum
Gentiana campestris s.str.
Gentiana purpurea
Gentiana verna
Geum montanum
Globularia cordifolia
Gypsophila repens
Helianthemum nummularium s.l.
Helictotrichon versicolor
Hieracium bifidum aggr.
Hieracium murorum aggr.
Hieracium villosum aggr.
Hippocrepis comosa
Homogyne alpina
Leontodon hispidus s.l.
Leucanthemum vulgare aggr.
Ligusticum mutellina
Lotus alpinus
Lotus corniculatus
Luzula multiflora
Myosotis alpestris
Nardus stricta
Pedicularis verticillata
Phleum alpinum aggr.
Phyteuma orbiculare
Plantago alpina
Plantago atrata s.str.
Poa alpina
Poa cenisia
Poa minor
Polygala alpestris
Polygonum viviparum
Potentilla aurea
Potentilla crantzii
Pritzelago alpina s.str.
Pulsatilla alpina s.str.
Ranunculus alpestris
Ranunculus montanus aggr.
Salix retusa
Saxifraga aizoides
Saxifraga oppositifolia
Saxifraga paniculata
Scabiosa lucida
Sedum atratum
Selaginella selaginoides
Sesleria caerulea
Silene acaulis
Silene vulgaris s.l.
Soldanella alpina
Taraxacum officinale aggr.
Thesium alpinum
Thlaspi repens
Thymus praecox subsp. polytrichus
Trifolium pratense s.str.
Trifolium thalii
Trollius europaeus
Vaccinium gaultherioides
Vaccinium myrtillus
Vaccinium vitis-idaea
Veronica aphylla
Viola calcarata

Chapter 4

Very high resolution digital elevation models: do they improve models of plant species distribution?

Lassueur Thierry[1], Joost Stéphane[1], and Randin Christophe[2]

This manuscript has been published in Ecological Modelling **198**: 139-153

[1]Ecole Polytechnique Fédérale de Lausanne (EPFL), Laboratoire de Systèmes d'information Géographique (LaSIG), 1015 Lausanne, Switzerland

[2]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

Abstract

Numerous studies have successfully used topographic variables for the prediction of plant species distributions in mountainous areas. A current question is which scale optimizes the predictive power of these variables. Due to the development of airborne radar/laser scanning, Very High Resolution Digital Elevation Models (VHR DEMs) with a typical resolution of 1 m have emerged, allowing an accurate estimation of surface derivatives on a fine scale. Improved understanding of the relationship between scale and predictive power, through a comparative assessment of the significance of topographic variables computed at different scales, is needed. It should help researchers in selecting an adequate scale for the computation of topographic factors and further also increase the precision of models for predicting plant species distributions, and find direct applications, such as a better estimation of potential refuges for plant species in a climate change scenario. In this context, it is worth exploring the use of VHR DEM's techniques for data acquisition.

In this study, we use a 1 m resolution laser DEM and a standard 25 m resolution DEM to compute 4 terrain variables (slope, aspect, plan curvature and profile curvature) at 6 different scales and compare their predictive ability. Univariate logistic regression models are fitted to presence/absence data of 117 alpine plant species collected at 125 sites in the Western Alps of Switzerland. The results show that slope and aspect, through northness = cos(aspect), are strongly correlated to the presence/absence data of many surveyed species. Using a laser VHR DEM significantly improves the predictive ability of aspect. The predictive power of slope and aspect is optimized when these variables are derived using a calculation window of about 100 × 100 m (slope) and 20 × 20 m (aspect). Curvature variables are much less significant. Studies that use them should begin with a careful analysis of their behaviour according to the scale used to compute them.

Keywords: Habitat modelling; Vegetation models; Topographic predictors; Airborne laserscanning; Scale; Logistic regression

Nomenclature: Aeschimann and Heitz (1996)

Introduction

Topographic variables have been used in many studies as relevant variables for predicting plant species distributions in mountainous areas (Table 1). A current question is which scale, defined as the area within which they are estimated (calculation window), optimizes the predictive ability of these variables. Due to the development of airborne radar/laser scanning, Very High Resolution Digital Elevation Models (VHR DEMs) with a typical resolution of 1 m have emerged, allowing an accurate estimation of surface derivatives on a fine scale. Improved understanding of the relationship between scale and predictive power, through a comparative assessment of the significance of topographic variables computed at different scales is needed. It should help researchers in selecting an adequate scale for the computation of topographic factors and further also increase the precision of models for predicting plant species distributions, and find direct applications, such as a better estimation of potential refuges for plant species in a climate change scenario. In this context, it is worth exploring the use of VHR DEM's techniques for data acquisition.

Improving the proximality, the generality and the precision of models are among the main concerns for ecologists. In order to improve their proximality, statistical distribution models should use a high variety of predicting variables, including resource (nutrient, water or light) and other environmental variables such as temperature or pH (see Austin 1980, Austin et al. 1984, Austin 1985, Austin and Heyligers 1989). Models based on such variables supposed to be causal also meet the generality criterion, in that they theoretically allow to correctly transfer predictions in space and time (Randin et al. 2006). However, environmental variables can be difficult to measure adequately. For example, standard climate stations do not necessarily provide biologically relevant microclimates, or their interpolation will generate a lot of uncertainties (Guisan and Zimmermann 2000). In this case, the lack of precision can be compensated by using indirect environmental variables, i.e. by exploiting correlation between easily observed non-causal and difficult-to-observe causal variables (Guisan et al. 1999). We can use elevation and its derivatives such as slope, aspect (exposure) and curvature, which serve as proxies for multiple environment variables (Gottfried et al. 1998, Gottfried et al. 1999, Guisan and Zimmermann 2000, Jelaska et al. 2003, Kubota et al. 2004), including soil depth, nutrient status and water status. In fact, in a mountainous terrain, many patterns of the environmental conditions as experienced by plants are driven by relief. Snow distribution, wind and energy allocation in structured alpine terrain are the most obvious consequences of topography and its interaction with climate (Billings and Bliss 1959, Körner 2003).

Curvature affects snow accumulation during the winter, which will directly influence growing season duration, the water budget in the soil and the amount of energy for plants. The relative position of plants between ridges and valleys exposes them to different wind systems, which can substantially

modify plant temperature, convective heat loss and evaporative cooling. Mechanical processes, such as abrasion or sculpturing from winds on edges, may also influence plant growth (Körner 2003).

Slope can express a part of the gravitation and dynamic processes that influence the vegetation through snow movement, water flow, erosion and organical deposition. This variable can also be used in combination with aspect to express the quantity of energy received by plants or the variability of diurnal energy patterns (Billings and Bliss 1959), which sometimes can not be detected by the classically interpolated climatic variables.

The strong influence of local microclimate and the predictor weakness of available macroclimate variables explain the use and the success of topographic predictors in many studies. Table 1 lists some of these studies. Despite a tendency to use high resolution in real alpine systems, (Gottfried et al. 1998, Guisan et al. 1998, Guisan et al. 1999, Dirnböck et al. 2003b) there is an obvious lack of consensus about the adequate scale to be used. This directly addresses what Openshaw and Taylor (1979, 1981) identified as the modifiable areal unit problem (MAUP). The MAUP originates from the fact that different ways exist by which a geographical study area can be divided into non-overlapping areal units for the purpose of spatial analysis (Marceau 1999, Marceau and Hay 1999; see references therein for an in depth review of MAUP).

For the specific case of topographic variables, an additional issue is that using different scales for explanatory and response variables may make sense. In fact, we can assume that the conditions favouring plant growth within small plots, which can be considered as points, are not those which are measured at these precise points, but the average conditions of the surrounding area. Table 1 shows that there is no necessary connection between the scale of computation of terrain variables and the plot size. While the latter often is dictated by the detectability of the surveyed species - e.g. 1 m for alpine plant species (Gottfried et al. 1998) - the scale of computation of the topographic variables, which is often coarser, can be implied by the limited resolution of the DEM, but also by deliberate choices.

In this work, we use presence/absence data of 117 alpine plant species collected in the Western Alps of Switzerland to assess and compare the predictive ability of four fundamental topographic predictor variables (slope, aspect, plan curvature and profile curvature). Each surface derivative is computed using six different calculation windows, ranging from 9 × 9 m to 175 × 175 m. The windows are centred on the sampling plots, whose size is constant at 8 × 8 m. Calculations are based on a laser VHR DEM and on a standard 25 m resolution DEM. We use univariate logistic regression models to compare the predictive power of the variables and address the following questions:

1. For each topographic variable, is it possible to determine an adequate scale of computation for its use in vegetation predictive models?

2. Do topographic variables better approximate ecological patterns and increase the power of prediction of potential distribution models when computed with the more precise laser VHR DEM?
3. Given a topographic variable, does the scale of the best predictor change between species?

Table 1 Topographic variables used as predictors in models for predicting plant species distributions in mountainous environments.

Authors	Study area	Predicted distribution type	Model type	Predictor	Plot size	R (m)
Guisan et al. (1998)	Alps, Switzerland	*C. curvula* ssp. *Curvula*	GLM	Curvature, slope, aspect	2 m × 2 m	25
Gottfried et al. (1998)	Tyrol, Austria, 650,000 m^2	Alpine plant species	CA and CCA	Curvature, slope, aspect	1 m × 1 m	1
Gottfried and Pauli (1999)	Tyrol, Austria, 650,000 m^2	Alpine plant species	CCA	Curvature, slope, aspect	1 m × 1 m	1
Guisan et al. (1999)	Nevada mountains, USA	Ten tree and 13 shrub species	GLM and CCA	Curvature, slope, aspect	20 m × 20 m	30
Higgins et al. (1999)	Cape Peninsula, South Africa, 471 km^2	Six alien plant species	GLM	Slope		50
Higgins et al. (1999)	Cape Peninsula, South Africa, 471 km^2	Six alien plant species	GLM	Slope	Individual occurrences	~7 × 7
Zimmermann and Kienast (1999)	Three sub-regions in the Alps, Switzerland, 17,500 km^2	Thirty-four graminoid species	Logistic regression	Slope	4 m × 4 m	50
Guisan and Harrell (2000)	Nevada mountains, USA	Plant cover	Ordinal GLM	Curvature	20 m × 20 m	30
Jelaska et al. (2003)	Medvednica Nature Park, Croatia, 220 km^2	Thirty-seven vascular plant species	GLM, DA, CT	Slope, aspect	1/64 of 10° × 6°	20
Randin et al. (in press)	Western Alps, Switzerland (270 km^2) and Hochschwab, Austria (110 km^2)	Fifty-four plant species (grassland, rock and scree vegetation)	GLM and GAM	Slope		25

Topographic variables are usually derived from an altitude matrix, which is defined by its resolution (R). Attributes are computed locally from data within an $n \times n$ (n odd) window. In this paper, scale is defined as the area within which the variables are estimated, i.e. $(n \times R)^2$. While the resolution of the altitude matrix is usually stated, the information on the window size is missing in most studies. However, this table shows that studies in real alpine systems tend to use fine resolution (between 1 and 50 m), i.e. presumably fine scale.
Curvature is used as surrogate for availability of moisture, wetness index, accumulation of snow, acceleration, deceleration, convergence and divergence of flow, erosion, deposition or rockfall. Slope is used as surrogate for gravitation processes, snow movement, water flow, erosion, rockfall, organical deposition, solar radiation or temperature. Aspect is used as surrogate for solar radiation or weather events (wind-, rain-shadow).

Methods

Study area

The study sites are located in an area of 100 km^2, in the Western Alps of the Canton of Vaud, Switzerland (Fig. 1). The Western Alps are those lower elevation mountains bordering each side of the Central European Alps, usually on calcareous bedrock. Elevation ranges from about 400 m in the bottom of the Rhone Valley up to more than 3000 m at the top of the highest peak of the Western Alps (Diablerets, 3210 m), where important glaciers are still found. The climate of this region is essentially wet and cool. Rainfalls are abundant and increase with elevation.

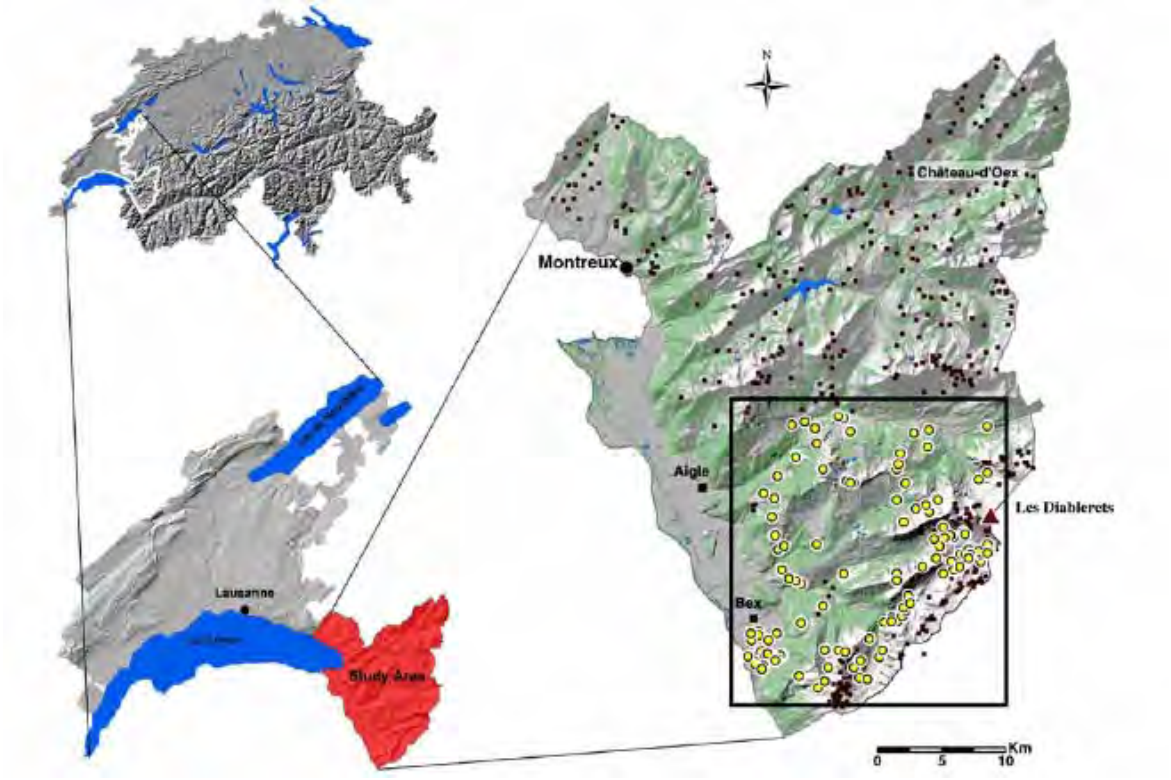


Figure 1 Study area. One hundred twenty five sampling plots in the Western Alps of Switzerland used to assess the predictive power of topographic variables in models of plant species distribution (circular dots in the black frame). The 125 sites are a subset of a total of 550 sites (black squares in the right picture), whose locations were selected by random stratified sampling strategy, as proposed by Hirzel and Guisan (2002).

The study area belongs to a single floristic unit, the medio-european or subatlantic domain (Ozenda 1985), which also corresponds to a single general climate of a suboceanic type. The vegetation is distributed along elevation according to the typical climax vegetation belts of the calcareous Alps: colline with deciduous forests (mainly Beech); montane with mixed forests (Beech-Abies); subalpine with coniferous forest (Picea); alpine with heath, meadow

and grassland vegetation; nival with a sparse vegetation cover of high elevation taxa. Vegetation has been and is still influenced by human activity. Pastures are common in this region, and are observed from the valley bottoms to the subalpine and lower alpine areas. The study boundaries also include two natural preserves.

Vegetation data

Vegetation data consist of presence/absence records of 117 plant species collected in 125 plots within the study area (Fig. 1). The 125 sites are a subset of a total of 550 sites, whose location was selected by random stratified sampling strategy (Hirzel and Guisan 2002). The environmental stratification was based on land-use information and three topographic predictors (elevation, slope and aspect) calculated from a 25 m resolution DEM, with a 3 × 3 focal window. Only 125 sites were used, since laser DEM data were not available for the entire area.

Sampling plots are 8 × 8 m squares. The exact coordinates of their centre were calculated on site using a Trimble GPS system, which allows an accuracy of 1-2 m after a differential post-correction. The frequency of species in the full data set ranges from 14/125 (*Daucus carota*) to 95/125 (*Trifolium pratense s. str*.). Data were collected during the summers of 2002 and 2003. The full list of species is given in Appendix A.

Elevation data

Laser DEM data are generated by airborne laserscanning technology. From the aircraft, a laser beam is deflected at right angle, making it possible to scan a wide stripe of land along the flight line. Distance from the earth's surface is determined by measuring the laser pulse return time (up to 5000 impulses a second). Global Positioning System (GPS), together with Inertial Navigation System (INS), calculates the location of the sensor. In combination with the scan angle, it provides the 3D position of each laser beam's point of reflection on the earth's surface. Typical density of raw data is about 1 point per square meter. Data may be then interpolated to create a very high resolution continuous surface.

The laser data we use in this study were provided in the form of a point cloud, i.e. a file of *xyz* coordinates corresponding to irregularly spaced ground points covering the study area. They derive from an airborne data acquisition campaign launched in 2000 by the Swiss Federal Office of Topography (Swisstopo) within the SAU (Agriculturally productive areas) project (Swisstopo 2004). Areas over an elevation of 2000 m are not flown. At the time when laser data were captured (spring 2000), Swisstopo (2004) reported an average density of laser points of 0.5 pt/m^2 in open areas, and a standard error $\sigma_z = 0.5$

m between the elevation coordinate of the data points and the exact elevation of the ground surface.

For the 25 m DEM, the elevations are interpolated from the contour lines of the 1:25 000 Swiss National Map (Swisstopo 2004).

Computation of the surface derivatives

For ease of surface derivative computation, laser data had first to be interpolated to a regular grid (altitude matrix). The objective was to generate a continuous elevation grid that optimized the informational content of the raw data while at the same time smoothing out the elevation variances inherited from the imprecision of the raw data σ_z. The following had to be considered: the grid resolution, the interpolation method and possible filtering of the altitude matrix.

The methods we used were based on an intensive sensitivity analysis (Laboratoire de Systèmes d'Information Géographique 2004). The elevation matrix was optimized for computations of the surface derivatives within the areas of the study plots (8 × 8 m). The grid resolution was based on the mean density of points as measured within a radius of 5 m from the centre of the plots. We found 1.27 pt/m^2 (median 1.1), and opted thus for a resolution of 1 m. To interpolate data, we assigned to each grid node the elevation of the nearest data point found in its neighbourhood. Extremes were then removed from the data by performing a single-level 2D wavelet decomposition with a wavelet low-pass filter. Wavelet decomposition is a broadly used method in all fields concerned with signal clearance and simplification (for an introduction, see Walnut, 2002). Computations were done in Matlab using the *Daubechies 5* wavelet (MathWorks 2002).

Numerous methods exist to compute slope, aspect and curvature from a regular elevation grid. We used the method of Zevenbergen and Thorne (1987). This broadly used method is considered as one of the best by many authors (for a review, see Burrough and McDonnell 1998), in that it provides values for surface derivatives that are very representative of those on site. Topographic attributes are computed locally for each cell on the altitude matrix from data within an $n \times n$ (n odd) cell kernel or "window". Slope and aspect are the two first order derivatives of the surface. Slope was measured in degrees from the horizon; aspect in degrees clockwise from North. Profile and plan curvature are second differentials of the surface. Curvatures were measured in [1/1000 m].

In order to examine the response of the surveyed species to the scale on which the surface derivatives were evaluated, a large range of window sizes was used. The derivatives were evaluated within windows of size 9 × 9, 15 × 15, 25 × 25 from the laser DEM, and of size 3 × 3, 5 × 5 and 7 × 7 from the 25 m

DEM. Using a 9 × 9 window from the 1 m resolution laser DEM means estimating the surface derivative within an area of 9 × 9 m. Using a 3 × 3 window from the 25 m DEM means estimating the surface derivative within an area of 75 × 75 m. The surface derivatives were computed for the pixel corresponding to the centre of the sites.

Statistical analysis

Univariate logistic regression models were fitted to the data. The likelihood equations were solved in Matlab. For all models, the logit link was of the simple form $g(x) = \alpha + \beta(x)$, with x the topographic variable. Two reasons justified not using more complex forms. First, the model is biologically meaningful: plant species are either responsive to extreme values of the variable in question or indifferent (flat S-spline). Second, even if response curves of certain species to some variables exhibit more complex shapes, we assumed that using the same model (be it S-spline shaped, logistic-gaussian or more complex) for all variables of the same type was a sufficient condition for showing which scale variable, if any, had greater predictive ability than the others.

Because aspect is a circular variable, it was inappropriate to include it in a model without prior transformation. Two main approaches are usually proposed. The first consists in converting aspect into a collection of design variables (Hosmer and Lemeshow 2000), such as N, E, S, W; the second in creating two variables, "northness" = cos(aspect) and "eastness" = sin(aspect) (Gottfried et al. 1999). Though both methods were used, we will only present the results returned from the latter. In fact, it proved by far the best approach. Using a set of design variables not only caused Matlab to reach the maximum number of iterations fixed by the *glmfit* function to solve the likelihood equations for many models involving plant species with a limited number of recorded presences; it also globally tended to underestimate the predictive ability of aspect for the other species models.
With aspect converted into 2 variables and 6 scales used to compute each topographic attribute, this is a total of 30 variables that were tested. In order to simplify the reading, we used the same nomenclature when referring to variables or models (see in Table 2).

Table 2 Topographic variables tested to model plant species occurrences in the Prealps of Switzerland and abbreviation of the models according to the DEM type and scales.

DEM type	Grid resolution (m)	Window size	Equivalent area (m)	Variable type				
				Slope (SLO)	Northness (NS)	Eastness (EW)	Plan curvature (PLC)	Profile curvature (PRC)
Laser	1	9 × 9	9 × 9	SLOlas9	NSlas9	EWlas9	PLClas9	PRClas9
Laser	1	15 × 15	15 × 15	SLOlas15	NSlas15	EWlas15	PLClas15	PRClas15
Laser	1	25 × 25	25 × 25	SLOlas25	NSlas25	EWlas25	PLClas25	PRClas25
Standard	25	3 × 3	75 × 75	SLO25m3	NS25m3	EW25m3	PLC25m3	PRC25m3
Standard	25	5 × 5	125 × 125	SLO25m5	NS25m5	EW25m5	PLC25m5	PRC25m5
Standard	25	7 × 7	175 × 175	SLO25m7	NS25m7	EW25m7	PLC25m7	PRC25m7

For the distribution and correlation analysis, ASP (aspect) is used instead of NS (=cos(aspect)) and EW (=sin(aspect)). For instance, ASPlas15 refers to aspect computed from the laser DEM within a 15 × 15 window.

The significance of the coefficients in the models was assessed using the Wald test. Under the hypothesis that the coefficient in question is zero, the Wald statistics (W) will follow a standard normal distribution (Hosmer and Lemeshow 2000). The mean of the absolute values of the Wald statistics for β returned by the 117 species models associated to each variable (mean $|W|$) was used to assess their significance. The overall mean $|W|$ (i.e. the mean $|W|$ returned by all species models associated with a given type of variable) was also calculated to estimate the overall significance of each type of variable.

In order to assess the difference in terms of predictive power between variables of a same type, we used different methods. For an overall comparison, we used Wilcoxon rank sum tests (see Wilcox 1997). They tested the null hypothesis that the distributions of the populations of $|W|$ returned by two scale models were identical with respect to their median. A small p-value was interpreted as evidence that one of the two scale variables had a greater mean predictive ability among the surveyed plant species. To deepen the analysis of the results, we also plotted the values of W returned by two scale models (see examples in Fig. 2) and used "rejection tables". Rejection tables were designed to detect some possible structures in the results, in particular to estimate the possibility of differences between species in their response to the scale on which each type of variable was computed. A rejection table is presented in Table 5. Each column corresponds to one of the variables tested, each line to one of the surveyed species. The Wald statistics testing the significance of the variables are given at each intersection. To simplify the reading, the results are grouped in separate blocks by type of variable, and the cells corresponding to a rejection of the null hypothesis at the level of significance of = 0.05 grey-shaded. The number of null hypotheses rejected per variable is given at the top of each column, along with the percentage of H_0 rejected. Rejection tables were also used to provide a synthesized view of the results.

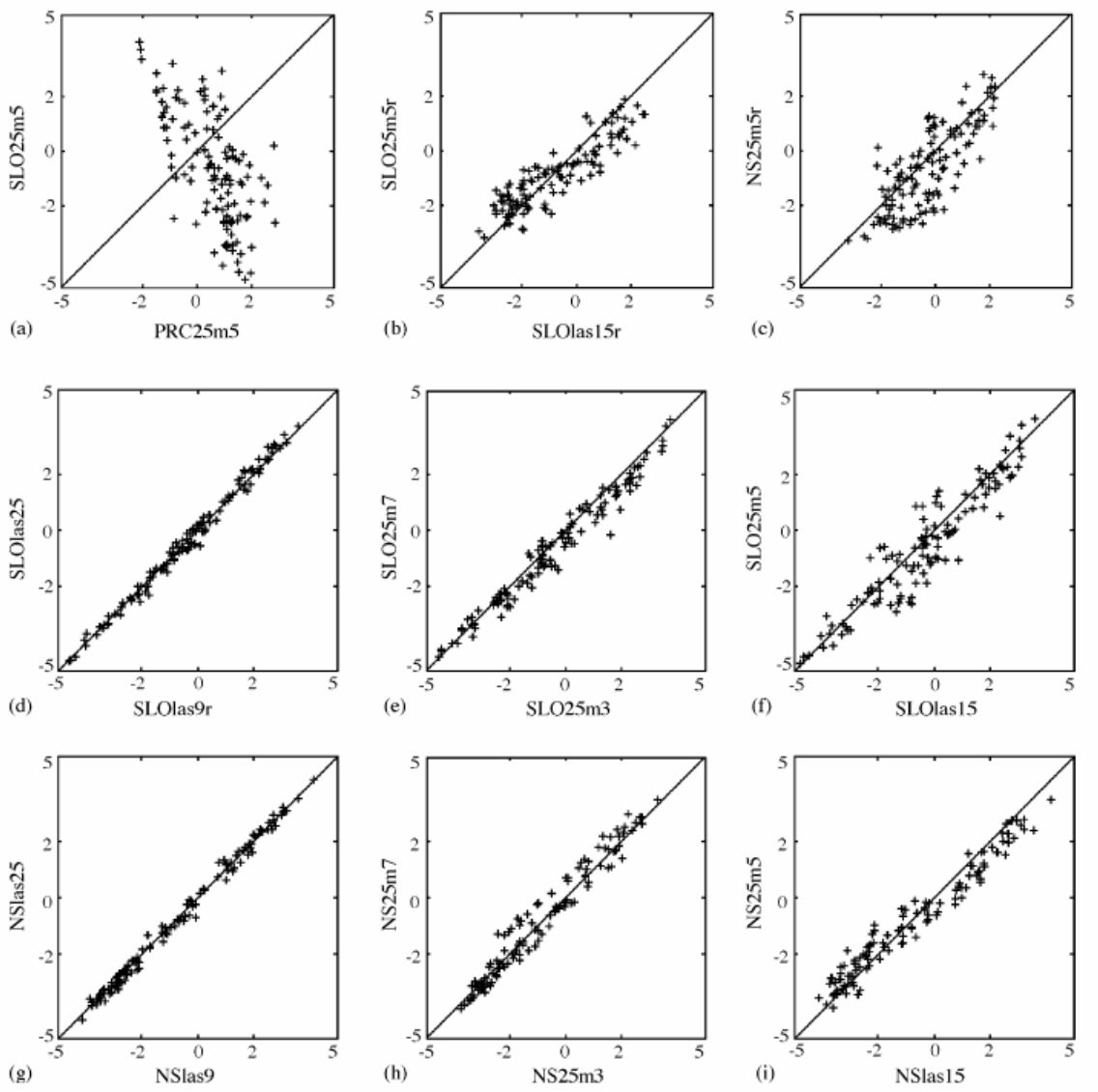


Figure 2 Scatterplots of Wald statistics testing the significance of topographic variables in the Western Alps of Switzerland. The meaning of the variables is given in Table 2. SLOlas15r, SLO25m5r, NSlas15r and NS25m5r correspond to reduced datasets (see explanation in the text).

Results

Distribution and correlation

On all scales, the distribution of SLO (slope) and ASP (aspect) values were rather uniform, even if slopes between 0° and 5° were slightly under-represented, and if sites with north-western and northern aspect (280° to 360°) tended to prevail over others. The distribution of PLC (plan curvature) and PRC (profile curvature) values resembled that of a normal distribution with mean 0. The fact that all key topographic configurations were well represented among the 125 study sites (slope and aspect regularly distributed between [0-40°] and [0-360°], presence of convex, ~null and concave curvatures) ensured high discriminatory power of the logistic regression equations.

Examining the correlation between variables of the same type showed that the different SLO variables were strongly correlated (Table 3). The same for ASP variables. It was still 0.67 between SLOlas9 and SLO25m7, and even came to 0.94 between ASPlas9 and ASP25m7, although these variables were at the extremes of the range of the investigated scales. PLC and PRC variables were much less correlated. The correlation was not negligible between variables computed within windows of adjacent size; but then, as the difference between the scales on which the variables were calculated increased, the correlation became less important.

Table 3 Correlation matrix of topographic variables calculated from laser and standard DEMs at various scales at 125 sites in the Prealps of Switzerland.

		SLO						ASP						
		SLOlas9	SLOlas15	SLOlas25	SLO25m3	SLO25m5	SLO25m7	SLO25m5r	ASPlas9	ASPlas15	ASPlas25	ASP25m3	ASP25m5	ASP25m7	ASP25m5r
SLO															
	SLOlas9	1													
	SLOlas15	0.97	1												
	SLOlas25	0.96	0.97	1											
	SLO25m3	0.77	0.81	0.85	1										
	SLO25m5	0.73	0.76	0.80	0.96	1									
	SLO25m7	0.67	0.70	0.74	0.91	0.98	1								
	SLO25m5r							0.62							
ASP															
	ASPlas9	−0.01	0.03	0.06	0.14	0.15	0.16		1						
	ASPlas15	0.00	0.03	0.07	0.19	0.19	0.20		0.99	1					
	ASPlas25	0.03	0.06	0.07	0.16	0.18	0.19		0.99	0.99	1				
	ASP25m3	0.04	0.07	0.09	0.14	0.16	0.19		0.95	0.95	0.96	1			
	ASP25m5	−0.03	0.01	0.03	0.08	0.12	0.16		0.95	0.95	0.96	0.98	1		
	ASP25m7	−0.07	−0.03	−0.01	0.06	0.08	0.11		0.94	0.95	0.95	0.97	0.99	1	
	ASPlas15r														0.92
FSL															
	FLClas9	−0.05	−0.04	−0.05	−0.10	−0.07	−0.08		−0.09	−0.11	−0.03	−0.01	−0.05	−0.04	
	FLClas15	0.12	0.12	0.12	−0.01	−0.04	−0.08		−0.13	−0.14	−0.10	−0.02	−0.03	−0.02	
	FLClas25	0.14	0.17	0.14	0.02	−0.04	−0.08		−0.10	−0.09	−0.08	−0.05	−0.05	−0.03	
	FLC25m3	−0.02	−0.01	0.03	0.11	0.03	0.02		0.16	0.14	0.14	0.10	0.08	0.11	
	FLC25m5	−0.18	−0.16	−0.13	−0.18	−0.30	−0.26		0.15	0.13	0.15	0.10	0.07	0.07	
	FLC25m7	−0.23	−0.22	−0.20	−0.30	−0.43	−0.40		0.13	0.11	0.11	0.09	0.06	0.05	
FRC															
	FRClas9	−0.05	−0.04	−0.08	−0.14	−0.06	−0.05		−0.13	−0.11	−0.05	0.04	0.05	0.05	
	FRClas15	−0.02	−0.04	−0.06	−0.20	−0.18	−0.18		−0.14	−0.18	−0.12	−0.01	0.02	0.02	
	FRClas25	00.04	0.02	−0.04	−0.19	−0.21	−0.21		−0.15	−0.12	−0.07	0.01	0.04	0.06	
	FRC25m3	−0.08	−0.10	−0.11	−0.32	−0.39	−0.38		−0.10	−0.05	−0.04	−0.01	−0.06	−0.06	
	FRC25m5	−0.16	−0.18	−0.19	−0.39	−0.50	−0.51		−0.09	−0.07	−0.08	−0.05	−0.10	−0.12	
	FRC25m7	−0.21	−0.24	−0.25	−0.44	−0.56	−0.59		−0.08	−0.09	−0.10	−0.08	−0.12	−0.13	

Logistic regression models: overview

SLO and NS models had much greater explanatory power with respect to presence/absence data than EW, PLC and PRC models (Table 4). The values of mean $|W|$ for EW, PLC and PRC models were close to the expected absolute value of a variable that follows a standard normal distribution (0.8), indicating that EW, PLC and PRC were weakly correlated to the presence/absence data of the surveyed species. The results of the hypothesis tests associated with the Wald test for β at the level of significance α = 0.05 are given in Table 5 for each model. Table 5 emphasizes the general lack of significance of EW, PLC and PRC variables in comparison with SLO and NS variables. When considering all EW, PLC and PRC species models, the percentage of H0 rejected (105/2106 = 5.0%) was equal to the probability of committing a type I error.

Differences in mean $|W|$ between SLO models as well as between NS models were small (Table 4). The horizontal structure formed by the grey-shaded cells in Table 5 also shows that the results of the Wald hypothesis tests testing the significance of the variables at α = 0.05 for two scale models often matched for a given species, indicating that the results returned by the different scale models were strongly correlated. In the next sections, we deepen the analysis for each type of variable.

Table 4 Mean of the absolute values of the Wald statistics returned by 117 species models.

		Mean |W|
SLO		
SLOlas	SLOlas9	1.62
	SLOlas15	1.66
	SLOlas25	1.69
SLO25m	SLO25m3	1.77
	SLO25m5	1.78
	SLO25m7	1.74
	Overall mean	1.71
Reduced dataset	SLOlas15r	1.08
	SLO25m5r	1.42
NS		
NSlas	NSlas9	2.12
	NSlas15	2.14
	NSlas25	2.16
NS25m	NS25m3	1.80
	NS25m5	1.83
	NS25m7	1.88
	Overall mean	1.99
Reduced dataset	NSlas15r	1.57
	NS25m5r	1.30
EW		
EWlas	EWlas9	0.71
	EWlas15	0.65
	EWlas25	0.65
EW25m	EW25m3	0.75
	EW25m5	0.75
	EW25m7	0.81
	Overall mean	0.72
PLC		
PLClas	PLClas9	0.71
	PLClas15	0.80
	PLClas25	1.09
PLC25m	PLC25m3	0.79
	PLC25m5	0.83
	PLC25m7	0.91
	Overall mean	0.86
PRC		
PRClas	PRClas9	0.73
	PRClas15	0.69
	PRClas25	0.72
PRC25m	PRC25m3	0.93
	PRC25m5	1.07
	PRC25m7	1.07
	Overall mean	0.87

The meaning of the variables is given in Table 2. SLOlas15r, SLO25m5r, NSlas15r and NS25m5r correspond to reduced data sets (see explanation in the text).

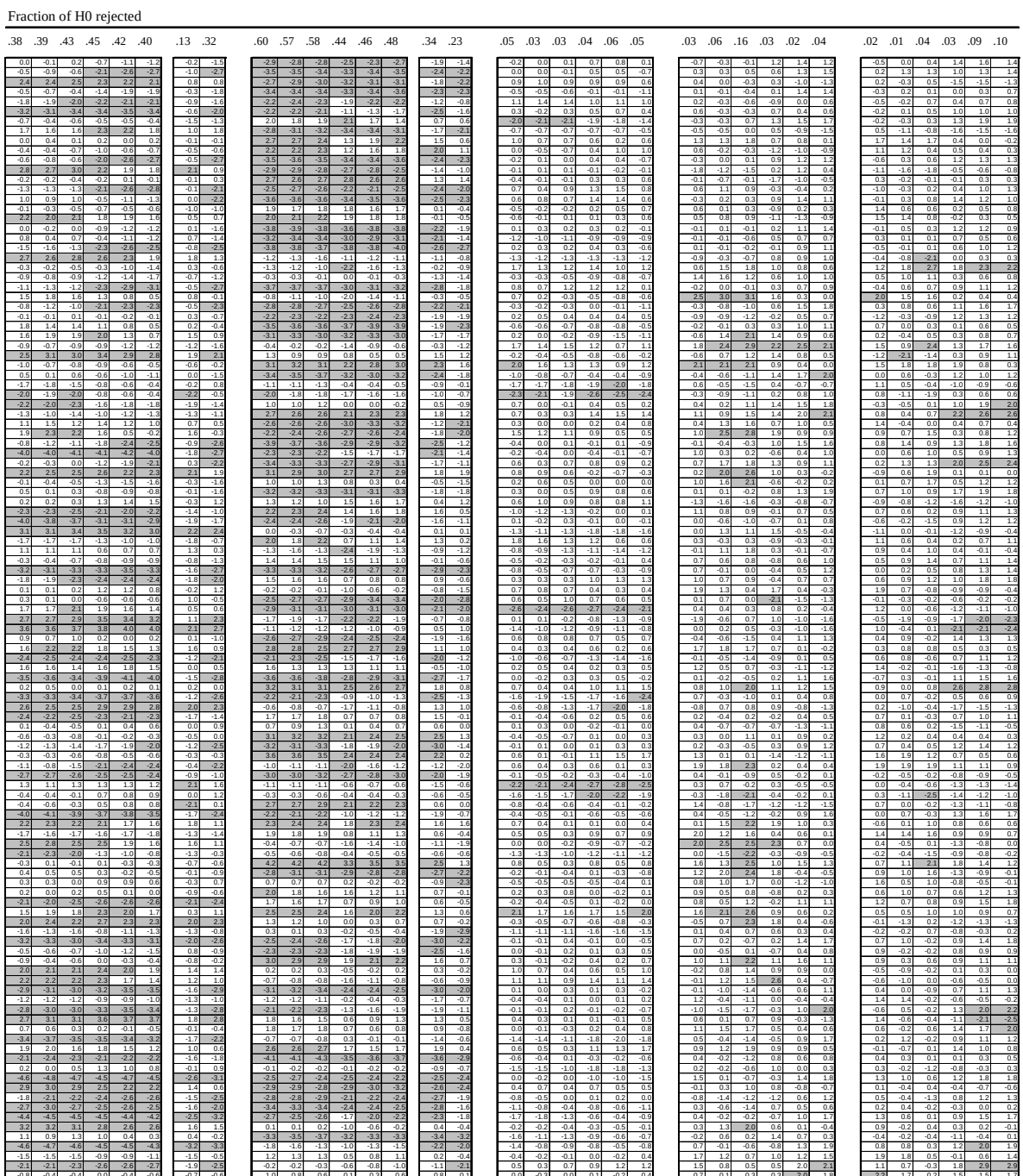
Plan curvature (PLC models)

The results returned by the different PLC scale variables were very close to those that would have been returned by random variables following a standard normal distribution (Tables 4 and 5). PLClas25, with a mean $|W|$ of 1.09, was assessed as a significant variable for 19/117 (= 16%) of the surveyed species at the = 0.05 significance level (Table 5), suggesting that PLClas25, in contrast to the other PLC variables, had some predictive power. The p-values returned by the Mann-Whitney tests comparing the distributions of $|W|$ for PLClas25 models with those returned by the other PLC models were small (Table 6), giving some weight to this assumption.

Profile curvature (PRC models)

For PRC models based on the laser DEM (PRClas models), the results for mean $|W|$ (Table 4) and for the percentage of H0 rejected (Table 5) strongly suggested that these variables had no predictive ability. Comparatively, PRC models based on the 25 m DEM (PRC25m models) returned higher values for mean $|W|$ (Table 4). However, we could not conclude that these variables had an intrinsic predictive power. As shown in Table 3, a non-negligible correlation was found between the raw data of PRC based on the 25 m DEM and SLO variables (esp. those based on the 25 m DEM). This correlation could still be observed in the results for W, as illustrated in Fig. 2a (plot of W for PRC25m5 vs. W for SLO25m5). The results suggested that the small amount of predictive power of PRC25m variables was in part due to their correlation with SLO variables, and thus that none of the PRC scale variables had predictive ability on the surveyed species.

Table 5 Wald statistics testing the significance of topographic variables for 117 alpine plant species.



Mean | W |

1.62	1.66	1.69	1.77	1.78	1.74	1.08	1.42	2.12	2.14	2.16	1.80	1.83	1.88	1.57	1.30	.71	.65	.65	.75	.75	.81	.71	.80	1.09	.79	.83	.91	.73	.69	.72	.93	1.07	1.07
Number of H0 rejected																																	
44	46	50	53	49	47	15	38	70	67	68	52	54	56	40	27	6	4	3	5	7	6	3	7	19	4	2	5	2	1	5	4	10	12
Fraction of H0 rejected																																	
.38	.39	.43	.45	.42	.40	.13	.32	.60	.57	.58	.44	.46	.48	.34	.23	.05	.03	.03	.04	.06	.05	.03	.06	.16	.03	.02	.04	.02	.01	.04	.03	.09	.10

Correspondence between ID and species name is given in Appendix A. The meaning of the variables is given in Table 2. SLOlas15r, SLO25m5r, NSlas15r and NS25m5r correspond to reduced data sets (see explanation in the text). Grey-shaded cells correspond to a rejection of the H0 at a = 0.05 that the variable is not correlated to the vegetation data. This table is referred to as a "rejection table".

Slope (SLO models)

On average, SLO variables based on the laser DEM (SLOlas variables) were less correlated to the presence/absence data of the surveyed species than SLO25m variables (Table 4). However, the results of the Mann-Whitney tests suggested that none of the scales used to compute SLO optimized its predictive ability (P_{min} = 0.25, Table 6). When restricting to SLOlas variables (P_{min} = 0.64, Table 6), the differences observed for a given species between the values of W returned by two scale models were never significant, as illustrated in Fig. 2d (plot of W for SLOlas9 vs. W for SLOlas25). We assumed thus that, even when restricting to a given species, all three SLOlas variables had an equivalent predictive ability, and, for further comparisons, we retained the intermediate variable SLOlas15 as representative of SLO on a fine scale. When restricting to SLO25m variables, we reached the same conclusions (P_{min} = 0.83, Table 6; illustration Fig. 2e), and subsequently retained SLO25m5 as representative of SLO on a local scale.

In Fig. 2f (plot of W for SLOlas15 (fine scale) vs. W for SLO25m5 (local scale)), points are still linearly distributed along the $y = x$ line, illustrating that none of the variables could be assessed as having a greater predictive ability (SLO25m5 better explained the data than SLOlas15 for only 53.8% of the species, P = 0.34 for the Mann-Whitney test comparing the two distributions of $|W|$, Table 6). In order to better identify the effect of scale on the predictive ability of SLO, we computed $\Delta = |\text{SLOlas15} - \text{SLO25m5}|$ at the 125 study sites, retained the 60 sites that returned the highest values of Δ, and built logistic regression models on the basis of the reduced data set (SLOlas15r and SLO25m5r models). Reason for that is: if any scale used to compute SLO optimizes its predictive ability, then retaining the study sites where the greatest differences between topography on fine and local scales are observed should highlight it. We verified that the values of slope were still acceptably distributed between [0-40°] within the reduced sample of study sites. The correlation between SLOlas15r and SLO25m5r was 0.62, instead of 0.76 between SLOlas15 and SLO25m5 (Table 3).

Using only 60 observations clearly weakened the ability of the logistic regression models to detect the significance of the variables (Tables 4 and 5). However, the decrease in mean $|W|$ was much more marked for SLO on a fine scale than for SLO on a local scale (Table 4). For 67.5% of the species, SLO25m5r better explained the data than SLOlas15r. The Mann-Whitney test strongly confirmed that SLO25m5r had a greater predictive ability than SLOlas15r (P = 0.0037, Table 6). This can be seen empirically in Fig. 2b (plot of W for SLOlas15r vs. W for SLO25m5r), by observing that points are more sparsely distributed along the y-axis than along the x-axis. The null hypothesis that the variable was not correlated to the presence/absence data at $\alpha = 0.05$ was also rejected in only 15 cases for SLOlas15r as against 38 cases for SLO25m5r (Table 5). Using a reduced set of data stressed thus the significance of slope on a local scale.

Table 6 Results of Mann-Whitney tests testing the null hypothesis that the distributions of the populations of ½W½ returned by two scale models are identical with respect to their median

SLO	SLOlas15	SLOlas25	SLO25m3	SLO25m5	SLO25m7	SLO25m5r
SLOlas9	0.84	0.64	0.27	0.25	0.34	
SLOlas15		0.79	0.34	0.34	0.42	
SLOlas25			0.50	0.49	0.62	
SLO25m3				0.95	0.83	
SLO25m5					0.83	
SLOlas15r						0.00
NS	**NSlas15**	**NSlas25**	**NS25m3**	**NS25m5**	**NS25m7**	**NS25m5r**
NSlas9	0.88	0.72	0.02	0.04	0.10	
NSlas15		0.88	0.02	0.03	0.07	
NSlas25			0.01	0.02	0.04	
NS25m3				0.86	0.63	
NS25m5					0.78	
NSlas15r						0.01
EW	**EWlas15**	**EWlas25**	**EW25m3**	**EW25m5**	**EW25m7**	
EWlas9	0.42	0.52	0.56	0.62	0.08	
EWlas15		0.96	0.21	0.22	0.01	
EWlas25			0.27	0.23	0.02	
EW25m3				0.99	0.28	
EW25m5					0.20	
PLC	**PLClas15**	**PLClas25**	**PLC25m3**	**PLC25m5**	**PLC25m7**	
PLClas9	0.30	0.00	0.13	0.01	0.00	
PLClas15		0.01	0.70	0.18	0.06	
PLClas25			0.01	0.05	0.20	
PLC25m3				0.20	0.07	
PLC25m5					0.32	
PRC	**PRClas15**	**PRClas25**	**PRC25m3**	**PRC25m5**	**PRC25m7**	
PRClas9	0.48	0.49	0.00	0.00	0.00	
PRClas15		0.86	0.00	0.00	0.00	
PRClas25			0.00	0.00	0.00	
PRC25m3				0.11	0.19	
PRC25m5					0.86	

W is the Wald statistics. The meaning of the variables is given in Table 2. SLOlas15r, SLO25m5r, NSlas15r and NS25m5r correspond to reduced data sets (see explanation in the text).

Aspect (NS and EW models)

Mean |W| for NS models based on the laser DEM (NSlas models) was greater than mean |W| for NS models based on the 25m DEM (NS25m models), suggesting that NSlas variables had a greater predictive ability than NS25m variables (Table 4). This was corroborated by the results of the Mann-Whitney tests: when comparing NSlas models with NS25m models, *P* ranged in fact from 0.01 to 0.10 (Table 6).

As for SLO variables, it was not possible to determine any variable within those based on the laser DEM, and within those based on the 25 m DEM on the other hand, that explained the data consistently better than the other (Fig. 2g and 2h). For further comparisons, we thus retained NSlas15 and NS25m5 as representative of NS on a fine and a local scale respectively.

For the majority of the surveyed plant species (76.9%) NSlas15 was more correlated to the presence/absence data than NS25m5 (Mann-Whitney test P = 0.03, Table 6; illustration Fig. 2i).
Following the same procedure as for SLO, we built logistic regression models using data from the 60 study sites where the highest values of Δ = |ASPlas15 – ASPlas25m5| were observed (NSlas15r and NS25m5r models). We verified that the values of aspect were still acceptably distributed between [0-360°]. The correlation was 0.92 between ASPlas15r and ASP25m5r (Table 3). As for the reduced SLO models, we observed a general decrease in the ability of the logistic regression models to detect the significance. The results of fitting the reduced models corroborated the results obtained from the full data (Table 4). For 70.1% of the species, |W| for the NSlas15r model was greater than |W| for the NS25m5r model (Mann-Whitney test P = 0.014, Table 6). The inverse S-shaped pattern already seen in Fig. 2i was even more marked when plotting W for NSlas15r vs. W for NS25m5r (Fig. 2c). All these results confirmed that NS had a greater overall predictive ability when it was computed on a fine rather than a local scale.

Regarding EW variables, the results strongly suggested that they were not correlated to the vegetation data (Tables 4 and 5).

Discussion

The results showed no significance for plan curvature (PLC) and profile curvature (PRC) in the models, excepting for PLClas25. These results are at first sight surprising because these variables are employed in many studies (Table 1). The altitudinal distribution of the sampling plots may explain the lack of pertinence of these two variables. The upper limit of the subalpine stage (Ozenda 1985) corresponds to the maximum elevation of the laser DEM used in this study (2000 m). Up to this elevation limit the sampling plots are essentially pastures and meadows, whose soils are often deep, fertilized and well differentiated (Delarze and Galland 1998, Dirnböck and Grabherr 2000, Dirnböck et al. 2003b). In such surfaces, moisture, nutrient and competition play a stronger role on the distribution of the species than wind, snow and geomorphologic processes, which interact with the topography (Tilman 1994, Bridge and Johnson 2000). Moreover, because of the altitudinal limitation of the laser DEM, the main summits are not included in the elevation range of the study area. The summit effect increases the role of temperature on plant growth (Körner 2003, Stanisci et al. 2005). At the highest altitudes of the study area, the influence of curvature, which generates microclimatic contrasted situations, is more pronounced because slopes tend to be steeper, climatic vectors stronger (Ozenda 2002, Ozenda and Borel 2003), and consequently plant life becomes more dependant on decoupling from a "hostile" atmosphere (Erschbamer 1989, Körner 2003). Such situations were hardly represented in our dataset.

PLClas25 is the only curvature scale that had some predictive power. Though we cannot exclude that these results were due to chance alone, they suggest that the predictive ability of plan curvature is optimized when this variable is computed on a relatively fine scale (within an area of ~25 × 25 m). At such a scale, plan curvature may reflect intensive hydro-mechanical processes or snow avalanches and debris flow paths in steep slopes (Gottfried et al. 1998, Dirnböck et al. 2003a).

All slope (SLO) variables were significantly correlated to the vegetation data. Using a reduced dataset stressed the significance of slope on a local scale. On such a scale, gravitational processes such as hydromechanical perturbation or snow movements start to have a significant impact on diversity and distribution patterns of the vegetation. Some authors (Gottfried et al. 1998, Dirnböck et al. 2003a) suggest to go further, by using the upslope area, positively correlated to the distances to the upper ridges.

Northness (NS) was a more significant explanatory variable on a fine scale than on a local scale. On a fine scale, northness roughly expresses the contrast of solar radiation between north and south faces. It is also a surrogate for snow cover duration, which significantly limits the growth period along north faces. Northness is a kind of synthetical expression of irradiance (Dirnböck et al.,

2003). Eastness, taken at a larger scale, could have reflected the effects of the west to east directed weather events (wind and rainfall), which are common in the Western Alps (Gottfried et al. 1998, Dirnböck et al. 2003a).

The correlation analyses showed that in the study area, which is representative of a mountainous environment, the values of plan and profile curvatures vary a lot according to the scale used to compute them. This implies that the possibility of detecting a predictive power for these variables strongly depends on the scale. Studies that use curvature as an explanatory variable should begin with a careful analysis of its behaviour according to scale. Until now, only a few studies used curvature at different scales in order to address different terrain processes (Gottfried et al. 1998, Dirnböck et al. 2003a). For slope and aspect, the difference between fine and local scales is much less important. However, the results show that the predictive power of slope and aspect are optimized for computations on a local and a fine scale respectively. They suggest that they are the effective scales that allow slope and aspect to be correlated to the species occurrence. Slope on a fine scale and aspect on a local scale, which were assessed as less significant, may acquire some predictive ability just because they are correlated to the variables on other more relevant scales.

The possibility of differences between species in their response to the scale at which topographic variables are calculated is a delicate issue. There is no scenario in the results, where, for a given type of variable, two scales are assessed as significant while returning much differentiated results. However, in our point of view, assuming that each variable has its most appropriate scale of computation is justifiable. Topographic variables are surrogates for different direct variables that have an impact on different scales on plant growth and species distribution. In that sense, the results of the Mann-Whitney tests used in this study can be interpreted as an evidence for showing what this optimal scale is.

With respect to the MAUP issue, the results discussed above must be restricted to the scale at which the behaviour of the response variables was assessed (McCarty et al. 1982), i.e. plots of 8×8 m.

We deliberately used univariate models, considering that it was a suitable way to discuss the scale issue. However, we do not rule out that any variable, individually identified as not significant, may turn out to be significant in a multivariate context (Franklin 1998, Guisan and Zimmermann 2000). Likewise, slope, aspect and curvature are not the only DEM-derived variables which can be used in habitat distribution models. Upslope area (Dirnböck et al. 2003a), or other variables based on flow accumulation algorithms such as the erosion index or the topographic wetness index (Wilson and Gallant 2000, Dirnböck et al. 2002), may also be used.

To answer our initial questions about the contribution of VHR DEMs to plant species modelling in a mountainous environment, we can say that in general, over the whole surface of the study area, the 25 m DEM has comparatively the same predictive efficiency as the laser DEM, and enables correct prediction over most of the territory: the empirical response curves differ little in comparison with those calculated with a laser model. What is, perhaps, most convincing in our results with the laser data, is that aspect variables consistently improve ecological patterns approximation and increase the models power of prediction on fine scales. However, even if northness is the best predictor of plant presence in our study, this topographic predictor remains a rough alternative to energetic processes direct information (the same comment can be made for gravitational processes approximated by slope variables). Northness will not constitute a totally perfect substitute for radiation processes in a mountainous landscape - because of the strong influence of the surrounding relief - as long as radiation has not been processed on the basis of a laser model including the entire altitudinal range.

Acknowledgments

We would like to thank Christopher Williams, Abram Pointet, Gilles Gachet, Gwenaelle Lelay, Régis Caloz, Antoine Guisan and two anonymous reviewers for their valuable comments.

References

Aeschimann, D., and C. Heitz. 1996. Index synonymique de la flore de Suisse et territoires limitrophes. CRSF, Genève.

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. Vegetatio **42**:11-21.

Austin, M. P. 1985. Continuum concept, ordination methods and niche theory. Annu. Rev. Ecol. Syst. **16**:39-61.

Austin, M. P., R. P. Cunningham, and P. M. Fleming. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio **55**:11-27.

Austin, M. P., and P. C. Heyligers. 1989. Vegetation survey design for conservation: gradsec sampling of forests in north-east New South Wales. Biol. Conserv. **50**:13-32.

Billings, W. D., and L. C. Bliss. 1959. An Alpine Snowbank Environment and Its Effects on Vegetation, Plant Development, and Productivity. Ecology **40**:388-397.

Bridge, S. R. J., and E. A. Johnson. 2000. Geomorphic principles of terrain organization and vegetation gradients. Journal of Vegetation Science **11**:57-70.

Burrough, P. A., and R. A. McDonnell. 1998. Principles of geographical information systems. Oxford University Press, Oxford.

Delarze, R., and P. Galland. 1998. La "Typologie des milieux de suisse". Expérience acquise lors de son utilisation et développement futur. Colloques phytosociologiques **22**.

Dirnböck, T., S. Dullinger, M. Gottfried, C. Ginzler, and G. Grabherr. 2003a. Mapping alpine vegetation based on image analysis, topographic variables and Canonical Correspondence Analysis. Applied Vegetation Science **6**:85-96.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003b. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Dirnböck, T., and G. Grabherr. 2000. GIS assessment of vegetation and hydrological change in a high mountain catchment of the Northern Limestone Alps. Mountain Research and Development **20**:172-179.

Dirnböck, T., R. J. Hobbs, R. J. Lambeck, and P. A. Caccetta. 2002. Vegetation distribution in relation to topographically driven processes in southwestern Australia. Applied Vegetation Science **5**:147-158.

Erschbamer, B. 1989. Vegetation on Avalanche Paths in the Alps. Vegetatio **80**:139-146.

Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. Journal of Vegetation Science **9**:733-748.

Gottfried, M., H. Pauli, and G. Grabherr. 1998. Prediction of vegetation patterns at the limits of plant life: A new view of the alpine-nival ecotone. Arctic and Alpine Research **30**:207-221.

Gottfried, M., H. Pauli, K. Reitter, and G. Grabherr. 1999. A fine-scaled predictive model for changes in species distribution patterns of high mountain plants induced by climate warming. Diversity & Distributions **5**:241-251.

Guisan, A., J. P. Theurillat, and F. Kienast. 1998. Predicting the potential distribution of plant species in an Alpine environment. Journal of Vegetation Science **9**:65-74.

Guisan, A., S. B. Weiss, and A. D. Weiss. 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecology **143**:107-122.

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

Hirzel, A., and A. Guisan. 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling **157**:331-341.

Hosmer, D. W., and S. Lemeshow. 2000. Applied logistic regression, second edition, New York.

Jelaska, S. D., O. Antonic, T. Nikolic, V. Hrsak, M. Plazibat, and J. Krizan. 2003. Estimating plant species occurrence in MTB/64 quadrants as a function of DEM-based variables - a case study for Medvednica Nature Park, Croatia. Ecological Modelling **170**:333-343.

Körner, C. 2003. Alpine plant life. 2nd Ed. Springer.

Kubota, Y., H. Murata, and K. Kikuzawa. 2004. Effects of topographic heterogeneity on tree species richness and stand dynamics in a subtropical forest in Okinawa Island, southern Japan. Journal of Ecology **92**:230-240.
LaSIG. 2004. Laboratoire de Systèmes d'Information Géographique. Pages Travaux de diplôme effectués au LaSIG. http://lasig.epfl.ch/recherche/diplomes.shtml#2004 *in*.
Marceau, D. J. 1999. The scale issue in the social and natural sciences. Canadian Journal of Remote Sensing **25**:357-366.
Marceau, D. J., and G. J. Hay. 1999. Contributions of remote sensing to the scale issues. Canadian Journal of Remote Sensing **25**.
MathWorks. 2002. Wavelet toolbox for use with Matlab: user's guide version 2. *in*. The MathWorks Inc, Natick.
McCarty, H. H., J. C. Hook, and D. S. Knos. 1982. The Measurement of Association in Industrial Geography. Greenwood Press, Westport.
Openshaw, S., and P. J. Taylor. 1979. A million of so correlation coefficients: Three experiments on the modifiable areal unit problem. Pages 127-144 *in* N. Wrigley, editor. Statistical Applications in the Spatial Sciences. Pion, London.
Openshaw, S., and P. J. Taylor. 1981. The modifiable areal unit problem. Pages 60-70 *in* N. W. a. R. J. Bennett, editor. Quantitative Geography: A British View. Routledge and Kegan Paul, London.
Ozenda, P. 1985. La végétation de la chaîne alpine dans l'espace montagnard européen. Masson, Paris.
Ozenda, P. 2002. Perspectives pour une géobiologie des montagnes. Presses polytechniques et universitaires romandes, Lausanne.
Ozenda, P., and J.-L. Borel. 2003. The Alpine Vegetation of the Alps. Ecological Studies **167**:53-64.
Randin, C. F., T. Dirnbock, S. Dullinger, N. E. Zimmermann, M. Zappa, and A. Guisan. 2006. Are niche-based species distribution models transferable in space? Journal of Biogeography **33**:1689-1703.
Stanisci, A., G. Pelino, and C. Blasi. 2005. Vascular plant diversity and climate change in the alpine belt of the central Apennines (Italy). Biodiversity and Conservation **14**:1301-1318.
Swisstopo. 2004. *in*.
Tilman, D. 1994. Competition and Biodiversity in Spatially Structured Habitats. Ecology **75**:2-16.
Wilcox, R. R. 1997. Introduction to robust estimation and hypotheses testing. Academic Press, San Diego.
Wilson, J. P., and J. C. Gallant. 2000. Secondary topographic attributes. Pages 87-131 *in* J. P. W. a. J. C. Gallant, editor. Terrain analysis, principles and applications. John Wiley and Sons, New York.
Zevenbergen, L. W., and C. R. Thorne. 1987. Quantitative-Analysis of Land Surface-Topography. Earth Surface Processes and Landforms **12**:47-56.

Appendix A List of plant species used in this study (nomenclature from Aeschimann and Heitz (1996).

ID	Species name
1	*Daucus carota*
2	*Salvia pratensis*
3	*Brachypodium pinnatum*
4	*Centaurea scabiosa s.l.*
5	*Colchicum autumnale*
6	*Glechoma hederacea s.str.*
7	*Alchemilla coriacea aggr.*
8	*Carex flacca*
9	*Dryas octopetala*
10	*Euphrasia minima*
11	*Knautia arvensis*
12	*Linum catharticum*
13	*Veratrum album subsp. lobelianum*
14	*Vicia sativa s.l.*
15	*Bromus erectus s.str.*
16	*Carex pallescens*
17	*Chaerophyllum hirsutum aggr.*
18	*Medicago lupulina*
19	*Plantago media*
20	*Centaurea jacea s.str.*
21	*Crepis pyrenaica*
22	*Gentiana campestris s.str.*
23	*Hieracium lactucella*
24	*Ranunculus bulbosus*
25	*Thesium alpinum*
26	*Tragopogon pratensis subsp. orientalis*
27	*Vicia cracca s.str.*
28	*Thymus pulegioides s.str.*
29	*Hippocrepis comosa*
30	*Potentilla crantzii*
31	*Pulsatilla alpina s.str.*
32	*Salix retusa*
33	*Sanguisorba minor s.str.*
34	*Vicia sepium*
35	*Carum carvi*
36	*Geum montanum*
37	*Vaccinium gaultherioides*
38	*Carlina acaulis subsp. caulescens*
39	*Helianthemum nummularium s.l.*
40	*Poa pratensis*
41	*Veronica arvensis*
42	*Helictotrichon pubescens*
43	*Hieracium bifidum aggr.*
44	*Luzula multiflora*
45	*Primula veris s.l.*
46	*Campanula rhomboidalis*
47	*Crepis aurea*
48	*Holcus lanatus*
49	*Knautia dipsacifolia s.str.*
50	*Phleum alpinum aggr.*
51	*Thymus praecox subsp. polytrichus*
52	*Crocus albiflorus*
53	*Lolium perenne*
54	*Nardus stricta*
55	*Alchemilla glabra aggr.*
56	*Rhinanthus alectorolophus*
57	*Briza media*
58	*Carduus defloratus s.str.*
59	*Euphorbia cyparissias*
60	*Lathyrus pratensis*
61	*Bartsia alpina*
62	*Cynosurus cristatus*
63	*Potentilla erecta*
64	*Arrhenatherum elatius*
65	*Vaccinium myrtillus*
66	*Bellis perennis*
67	*Phyteuma orbiculare*
68	*Plantago alpina*
69	*Ranunculus nemorosus aggr.*
70	*Selaginella selaginoides*
71	*Achillea millefolium*
72	*Ligusticum mutellina*
73	*Plantago atrata s.str.*
74	*Galium album*
75	*Pimpinella major*
76	*Astrantia major*
77	*Deschampsia cespitosa*
78	*Poa trivialis s.str.*
79	*Aster bellidiastrum*
80	*Poa alpina*
81	*Scabiosa lucida*
82	*Heracleum sphondylium s.l.*
83	*Homogyne alpina*
84	*Silene vulgaris s.l.*
85	*Trollius europaeus*
86	*Hypericum maculatum s.str.*
87	*Potentilla aurea*
88	*Polygonum viviparum*
89	*Sesleria caerulea*
90	*Agrostis capillaris*
91	*Trisetum flavescens*
92	*Ajuga reptans*
93	*Soldanella alpina*
94	*Carex sempervirens*
95	*Anthyllis vulneraria s.l.*
96	*Festuca pratensis s.l.*
97	*Prunella vulgaris*
98	*Rumex acetosa*
99	*Alchemilla conjuncta aggr.*
100	*Campanula scheuchzeri*
101	*Ranunculus acris s.l.*
102	*Ranunculus montanus aggr.*
103	*Plantago lanceolata*
104	*Geranium sylvaticum*
105	*Taraxacum officinale aggr.*
106	*Leucanthemum vulgare aggr.*
107	*Veronica chamaedrys*
108	*Dactylis glomerata*
109	*Cerastium fontanum subsp. vulgare*
110	*Galium anisophyllon*
111	*Lotus corniculatus*
112	*Trifolium repens s.str.*
113	*Alchemilla xanthochlora aggr.*
114	*Anthoxanthum odoratum aggr.*
115	*Festuca rubra aggr.*
116	*Leontodon hispidus s.l.*
117	*Trifolium pratense s.str.*

Chapter 5

Are niche-based species distribution models transferable in space?

Randin Christophe[1], Dirnböck Thomas[2], Dullinger Stefan[3], Zimmermann Niklaus[4], Zappa Massimiliano and Guisan Antoine[1]

This manuscript has been published in Journal of Biogeography **33**: 1689-1703

This paper is part of the Special Issue, Species distribution modelling: methods, challenges and applications, which owes its origins to the workshop on generalized Regression Analyses and Spatial Predictions: Grasping Ecological Patterns from Species to Landscape held at the Centre Pro Natura Aletsch in Riederalp, Switzerland, in August 2004.

[1]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]Federal Environment Agency – Austria, Spittelauer Lände 5, A-1090 Vienna

[3]VINCA - Vienna Institute for Nature Conservation and Analyses, Giessergasse 6/7, A-1090 Vienna

[4]Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf

Abstract

Aim: To assess the geographic transferability of niche-based species distribution models fitted with two modelling techniques.

Location: Two distinct geographic study areas in Switzerland and Austria, in the subalpine and alpine belts.

Methods: Generalized linear and generalized additive models (GLM and GAM) with a binomial probability distribution and a logit link were fitted for 54 plant species, based on topo-climatic predictor variables. These models were then evaluated quantitatively and used for spatially explicit predictions within (internal evaluation and prediction) and between (external evaluation and prediction) the two regions. Comparisons of evaluations and spatial predictions between regions and models were conducted in order to test if species and methods meet the criteria of full transferability. By full transferability, we mean that: (1) the internal evaluation of models fitted in region A and B must be similar; (2) a model fitted in region A must at least retain a comparable external evaluation when projected into region B, and vice-versa; and (3) internal and external spatial predictions have to match within both regions.

Results: The measures of model fit are, on average, 24% higher for GAMs than for GLMs in both regions. However, the differences between internal and external evaluations (AUC coefficient) are also higher for GAMs than for GLMs (a difference of 30% for models fitted in Switzerland and 54% for models fitted in Austria). Transferability, as measured with the AUC evaluation, fails for 68% of the species in Switzerland and 55% in Austria for GLMs (respectively for 67% and 53% of the species for GAMs). For both GAMs and GLMs, the agreement between internal and external predictions is rather weak on average (Kulczynski's coefficient in the range 0.3-0.4), but varies widely among individual species. The dominant pattern is an asymmetrical transferability between the two study regions (a mean decrease of 20% for the AUC coefficient when the models are transferred from Switzerland and 13% when they are transferred from Austria).

Conclusions: The large inter-specific variability observed among the 54 study species underlines the need to consider more than a few species to test properly the transferability of species distribution models. The pronounced asymmetry in transferability between the two study regions may be due to peculiarities of these regions, such as differences in the ranges of environmental predictors or the varied impact of land-use history, or to species-specific reasons like differential phenotypic plasticity, existence of ecotypes or varied dependence on biotic interactions that are not properly incorporated into niche-based models. The lower variation between internal and external evaluation of GLMs compared to GAMs further suggests that overfitting may reduce transferability. Overall, a limited geographical transferability calls for caution when projecting niche-based models for assessing the fate of species in future environments.

Keywords: Austria, generality, generalised additive models (GAM), generalised linear models (GLM), geographic transferability, habitat distribution, model evaluation, predictions, spatial modelling, Switzerland,

Abbreviations: SDM, species distribution model; GLM, generalized linear model; GAM, general additive model; CH, Switzerland; P/A, presence and absence; AT, Austria; GIS, geographic information system, DEM, digital elevation model; IDW, inverse distance weighting; IE, internal evaluation; EE, external evaluation; IP, internal prediction; EP, external prediction; AUC, metric based on the area under the curve of a receiver operating characteristic plot (ROC plot); TI, transferability index

INTRODUCTION

Niche-based species distribution models (Guisan and Theurillat 2000, Guisan and Zimmermann 2000, Guisan and Thuiller 2005) are models that relate observations of species, gathered over a certain period of time, to various attributes of the environment, such as topography, climate or geology. These environmental attributes, or predictors, arranged along a gradient from proximal to distal predictors (Austin 2002), can have direct or indirect effects on species establishment and survival. As species distribution models (SDM) do not consider the time dimension, they are said to be static. Conceptually, they assume the fitted relationship to be an adequate representation of the realized niche of the species under a stable equilibrium constraint (Franklin 1995, Guisan and Theurillat 2000).

In recent years, a large number of studies have used SDMs, developed for individual species at various spatial scales (see e.g. Scott et al. 2002 for many examples). These models are not only of theoretical interest in biogeography, but are also valuable tools in conservation biology and species management (Miller 1986, Carey and Brown 1994, Godown and Peterson 2000, Engler et al. 2004). However, most published SDMs were developed for sections of a species' range and few models were fitted at the global scale (e.g. Jeffree and Jeffree 1994) or used the full extent of a species' geographical range. Examples of the latter include models where regionally endemic species (Peterson et al. 2000) or the whole native range for invasive alien species (Peterson et al. 2003) were considered.

Furthermore, SDMs are usually evaluated and applied within the region in which they were fitted. As a result, their applicability to other parts of a species' geographic range (Thomas and Bovee 1993, Fielding and Haworth 1995, Glozier et al. 1997, Ozesmi and Mitsch 1997, Schröder and Richter 1999, Kleyer 2002) or to other time periods (Schröder and Richter 1999, Araujo et al. 2005b) was rarely assessed. This concept of geographical or temporal cross-applicability of models was defined as transferability (Thomas and Bovee 1993, Glozier et al. 1997, Schröder and Richter 1999, Kleyer 2002) or generality (Fielding and Haworth 1995). Such transferability can be an important feature of SDMs, for instance if they are used for projections into new areas (transferability in space) or for predictions of climate change responses (transferability in time) (Fielding and Haworth 1995).

There could be potential obstacles to model transferability. According to Walter's law of relative habitat constancy (Walter and Breckle 1985), species can shift their apparent habitat – often defined with regard to indirect predictors, such as topography or vegetation structure – to fit to their basic ecological requirements determined by direct, physiologically meaningful predictors (such as sum of temperature, water availability, etc.). As a result, transferability should be limited when indirect predictors are used, as these may fail to express the true habitat requirements of the species in distinct geographic areas. Using direct or resource predictors should allow a more universal and thus more transferable definition of a species' realized niche, but requires these predictors to be available in a spatially-

explicit form (Austin 1980, Austin et al. 1984, Austin 1985, Austin and Smith 1989, Guisan and Zimmermann 2000). Models incorporating spatial or temporal autocorrelation may also be difficult to transfer into distinct geographic areas (Hampe 2004, Araujo et al. 2005b, Guisan and Thuiller 2005).However, even when the abiotic conditions remain constant, changes in the regional species' pools usually occur in distinct parts of a species' range, for instance as a result of different historical influences. Such local changes in biotic pressure are likely to generate local modifications of a species' realized niche (Pulliam 2000), and thus to affect the geographic transferability of locally-fitted models. The existence of ecotypes (climatic, edaphic, geographic, etc.) may also cause problems when transferring a model from one region to another (Walter and Breckle 1985, Joshi et al. 2001).

Moreover, it is well documented that species tend to thrive in a more varied array of habitats at the centre of their distribution and to become rarer and more restricted to specialized habitats towards the margins (e.g. Brown et al. 1995). In addition, landscape-scale population dynamic processes, such as mass effects and source-sink dynamics (Dias 1996, Pulliam 2000, Mouquet and Loreau 2002, 2003), can inflate realized niches where species are abundant, whereas other processes, such as Allee effects, may shrink them where species are rare (Groom 1998, Keitt et al. 2001). Hence, the position of the different study areas within the whole species' range is also likely to affect the way niche models are fitted (Peterson and al. 2000). Furthermore, it has been shown for SDMs of plant species that the spatial predictability depends on particular traits of individual species like those that determine their colonization ability (Dirnböck and Dullinger 2004).

Despite these potential obstacles, several studies provided evidence supporting the idea that niche positions are more than merely regional phenomena (Thompson et al. 1993, Hill et al. 2000, Prinzing et al. 2002). They showed that most species can occupy similar niche positions in distinct regions and that geographic variation of the niche does not usually increase for species that are more susceptible to competitive displacement or ecophysiological stress. This is further supported by several previous studies that successfully tested spatial and temporal transferability of habitat models. However, these studies were restricted to a limited number of habitats and species within a restricted number of taxonomic groups, e.g. arthropods in fens (Schröder and Richter 1999), fishes (Freeman et al. 1997, Glozier et al. 1997) or birds (Fielding and Haworth 1995, Ozesmi and Mitsch 1997). Comparisons across many species (Araujo et al. 2005b) and separate regions should allow more general conclusions.

Testing of model transferability could prove particularly powerful for complementing standard procedures of model evaluation (for best practice techniques (see Maggini et al. 2006; this issue). The use of observations independent from the training dataset has been recommended for a proper evaluation of models (Fielding and Bell 1997, Guisan and Zimmermann 2000). If the training and test datasets are restricted to the same spatial and temporal domains, internal evaluation techniques such as data partitioning or split sample approaches

are sufficient. If the predictions are to be tested for their generality and robustness (accuracy and stability of predictions in a new situation), Fielding and Bell (1997) recommend using a geographically (e.g. Fielding and Haworth 1995) or temporally (e.g. Boyce et al. 2002, Araújo et al. 2005) independent dataset for their external evaluation.

So far, a large number of studies, involving many different modelling techniques, have been conducted to predict the potential distribution of species under changing environmental conditions (Huntley 1995, Sykes and Prentice 1995, Guisan and Theurillat 2001, Bakkenes et al. 2002, Peterson et al. 2002, Dirnböck et al. 2003, Thuiller et al. 2005, Araújo et al. 2006) as well as in regions where such species do not yet occur (Weber 2001). However, testing for proper geographic or temporal transferability of models and the related uncertainty has usually been neglected.

For full model transferability between two regions, we suggest that the three following conditions have to be fulfilled: (1) the internal evaluation of models fitted in region A and B must be similar; (2) a model fitted in region A must at least retain a comparable external evaluation when projected into region B; (3) Internal and external spatial predictions have to match within both regions; here, internal prediction means that a model fitted in region A is used for predictions in region A, whereas external prediction means that a model fitted in region B is used for predictions in region A.

We hypothesize that some modelling techniques are likely to be more robust than others when transferred from one geographic region to another. For instance, the shape of response curves in generalized additive models (GAM; Hastie and Tibshirani 1986), being based on smoothing techniques, is not predefined and thus allows modelling closer to the data (Guisan et al. 2002, Lehmann et al. 2002) than do generalized linear models (GLM; McCullagh and Nelder 1989.), which are based on parametric (often polynomial) response curves. For this reason, and for an identical number of degrees of freedom allowed for each predictor in each technique, we expect GAMs to provide higher fits overall as shown by Moisen and Frescino (2002) but potentially at the cost of being less generalizable than GLMs to other situations, i.e. showing weaker transferability in space and time.

In this study, we propose a quantitative measure (i.e. a new index) and test of transferability. Our data set covers 54 alpine plant species in two distinct geographic regions of Switzerland and Austria. We conducted a cross-region assessment using two modelling techniques - GLMs and GAMs - to answer three main questions:

1. How can transferability best be quantified?
2. Are niche-based models fitted with GLMs and GAMs transferable in space?
3. Does model transferability depend on the modelling technique used (GLM versus GAM)?

METHODS

Study areas

The two study areas cover parts of the alpine and forest-free subalpine zones of the Western Swiss Alps (6°60′ – 7°10′ E and 46°10′ – 46°30′ N) and the North-Eastern calcareous Alps in Austria (14°60′ – 15°50′ E and 47°30′ – 47°50′ N). A comparison of the environmental gradients realised within the two study areas is given in Table 1. Hereafter, the Swiss and Austrian study areas will be abbreviated CH and AT, respectively.

Table 1 Environmental context of the two study areas

	Switzerland (CH)	Austria (AT)
Surface (km^2)	270	110
Number of vegetation plots	402	603
Altitudinal range (m)	1300–3210	1300–2273
Temperature range (°C)	−3.5–6	1–4.5
Temperature degree days (°C day $year^{-1}$)	313–2346	788–1830
Precipitation range (mm)	1400–2400	1500–2500
Moisture index (average of monthly values June–August) (mm day^{-1})	−13–174	−7–160
Global solar radiation (average of monthly values June–August) (kJ m^{-2} day^{-1})	3264–30659	3267–30616
Number of days with snow cover per year (period 1980–2000)	75–347	130–237

CH, Swiss study area; AT, Austrian study area.

In the CH region, elevations reach 3210 m a.s.l. at the top of the Diablerets chain where important glaciers are still found. Along the altitudinal gradient the sequence of vegetation belts is that typical for the calcareous Alps: colline belt with deciduous forests (mainly *Fagus sylvatica* L., European beech); montane belt with mixed forests (*F. sylvatica and Abies alba* Mill., silver fir); subalpine belt with coniferous forest (*Picea abies* (L.) Karsten, Norway spruce); alpine belt with heath, meadow and grassland vegetation; and nival belt with sparse vegetation cover of characteristic high elevation taxa. Vegetation has long been and is still under the influence of human land use: pasturing is common in this region from the valley bottoms up to the subalpine and lower alpine areas.

The AT study area comprises four distinct mountains (Mt. Hochschwab, Mt. Schneealpe, Mt. Rax and Mt. Schneeberg). Summits vary between 1900 and 2300 m a.s.l. Vegetation belts are generally similar to the CH region although the nival zone is lacking. As for CH, summer pastures are abundant (Dirnböck et al. 2003).

Within this study, we focus on the subalpine to nival belts only. Accordingly, we set the lower limit of the study areas at 1300 m a.s.l. in both CH and AT. This altitude corresponds to the average lower elevation limit of the subalpine belt in the Alps (Ozenda 1985).

Species' data

Data on the presence-absence of species originate from two separate data sets. The CH data comprise 402 vegetation plots sampled during the years 2002 and 2003. The data for AT were collected between the years 1994 and 2001 on a total of 603 plots. A random stratified sampling strategy restricted to open non-woody vegetation (grassland, rock and scree vegetation) was applied in both study areas. The plot size was constant in CH (16 m^2), whereas it varied from 5 to 30 m^2 in AT (Dirnböck et al. 2003). From the overall species' pool of these samples, a total of 54 species with more than 30 occurrences in both data sets were selected for modelling.

Environmental predictors

We generated a comparable set of environmental predictors for both study areas, using identical GIS algorithms and types of input data (Table 2). All GIS predictors were calculated with a 25m spatial resolution, as derived from the digital elevation models (DEM) available in each study area.

First, we calculated the slope from the DEMs to account for gravitational processes such as snow avalanches and rockfalls (Guisan et al. 1998, Dirnböck et al. 2003) and their impact upon vegetation. Second, we calculated linear lapse rates for long-term (1961-1990) monthly mean temperature and monthly rainfall taken from the national meteorological networks of Switzerland and Austria (MeteoSuisse and ZAMG). Next, we normalized the monthly values to 0m above sea level, using the regression lapse rates, and interpolated the 0m data to the whole surface of both study areas, using inverse distance weighted interpolations (IDW). Finally, the spatially interpolated values (representing locally adjusted regression intercepts) were re-projected to actual elevations using the 25m DEM of each study area and the regression lapse rates. This method differs from the approach of (Zimmermann and Kienast (1999) in using IDW rather than thin-plate splines for the interpolation. These basic climatic variables were then transformed into three physiologically more meaningful bioclimatic predictors: the growing degree-days (with 0°C threshold), moisture index over the growing season (difference between precipitation and potential evapotranspiration from June to August) and potential global solar radiation during the growing season (see Table 2 and references therein). Additionally, the spatially-distributed hydrological model PREVAH (Gurtz et al. 2003) was used to obtain a physically-based predictor for snow cover duration in both the Swiss and the Austrian study sites.

Table 2 Physiologically meaningful environmental predictors used to model the distribution of species.

Variables	Units	Details	Method	References
Temperature degree days	°C day year^{-1}	Sum of days multiplied by temperature > 0 °C	ARCInfo AML	Zimmermann & Kienast (1999)
Moisture index (average of monthly values June–August)	mm day^{-1}	Monthly average of daily atmospheric H_2O balance	ARCInfo AML	Zimmermann & Kienast (1999)
Global solar radiation (average of monthly values June–August)	kJ m^{-2} day^{-1}	Monthly average of daily global solar radiation	ARCInfo AML	Kumar *et al.* (1997)
Snow cover	days	Number of days with snow cover for the period 1980–2000	PREVAH HRU model	Gurtz *et al.* (1999)
Slope	degrees	Slope inclination	DEM, ARCInfo GRID routine	Anon (2004)

As the latter model has not yet been used in any similar biogeographical study, we provide more details here. PREVAH has previously been used for the simulation of the hydrological behaviour of catchments at different spatial scales (Zappa 2002) and, among other model applications, checked against point observations of snow water equivalent and remotely sensed maps of snow cover distribution (M. Zappa, in prep.). In the present application, the hydrological model is forced by interpolated daily values of observed climatic variables collected by the Swiss and Austrian meteorological agencies (MeteoSuisse and ZAMG). Data for five meteorological variables (precipitation, air temperature, relative sunshine duration, wind speed and water vapour pressure) have been used for the period 1979-2000. Global radiation was estimated as described in (Schulla 1997). Both temperature and radiation are locally adjusted to take into account slope, aspect and shade. 1979 was adopted as the initialization period. The parameters controlling snow accumulation, snowmelt and runoff generation were then calibrated for the 6-year period 1980-1985 and spatially distributed maps of cumulative snow cover duration were finally summarized for the 20-year period 1980-1999.

All these environmental variables are expected to have a major direct ecophysiological impact on plant species (Pearson et al. 2002, Dirnböck et al. 2003, Körner 2003), as required for successful transferability.

All steps of the analyses are summarized in Figure 1. We divided the analytical design into three main parts: (1) model fitting, (2) model evaluation and (3) model prediction.

Model fitting

Two separate models were fitted for each species in the S-Plus 2000 software (MathSoft 1999), with presence/absence values in each regional dataset, using generalized linear models (McCullagh and Nelder 1989.) and generalized additive models (Hastie and Tibshirani 1986) with a binomial variance and a logistic link

function. In both GLMs and GAMs, an AIC-based stepwise procedure in both directions was used to select the most significant predictors (Akaike 1973). Up to second order polynomials (linear and quadratic terms) were allowed for each predictor in GLMs, with the linear term being forced in the model each time the quadratic term was retained. Up to four degrees of freedom were allowed for the smooth functions in GAMs. The fit of GLMs and GAMs was measured with the Nagelkerke R^2 (Guisan and Zimmermann 2000). Model fits within and between models (GLMs and GAMs) and regions (CH and AT) were compared for the 54 species with Wilcoxon signed-rank tests (treating the samples as grouped by species).

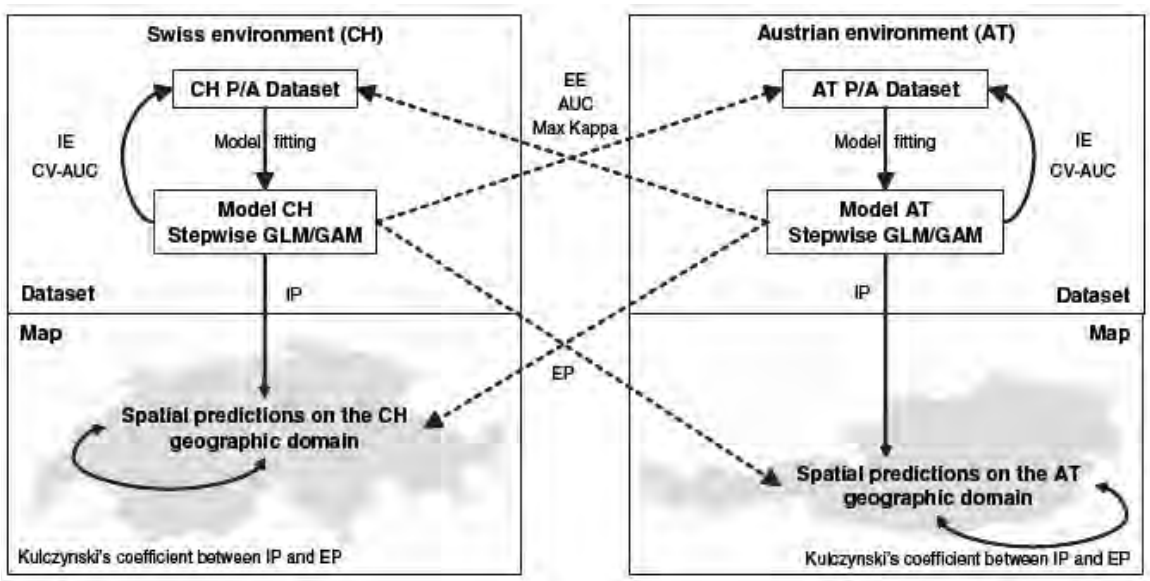


Figure 1 Experimental design. GLM and GAM models of each species were evaluated with the CH and AT presence/absence datasets and then applied to the geographical domains of both regions. (— IE = internal evaluation; — IP = internal prediction; --- EE = external evaluation; --- EP = external prediction)

Spatial predictions

For both GLMs and GAMs, spatial predictions were made over the full geographical domains in the S-Plus software using custom codes, and then mapped using in the ArcGIS 9.0 software (ESRI 2004). Predictions in CH were made both from the model fitted in CH (internal predictions; IP) and from the model fitted in AT (external predictions; EP), and reciprocally for predictions in AT.

Model evaluation

Comparisons of predicted (probability scale) and observed (presence-absence) values were based on the area under the curve (AUC) of a Receiver-operating characteristic plot (ROC; Fielding and Bell 1997) and the Kappa coefficient maximized over the full range of possible probability thresholds (hereafter max Kappa; (Huntley 1995, Guisan et al. 1998). AUC takes values between 0 and 1, with 0.5 meaning no agreement, 0 an inverse relationship (errors better predicted), and 1 a perfect agreement. Max Kappa takes values between -1 and +1, with 0 meaning no correlation, -1 an inverse relationship, and +1 a perfect agreement.

The AUC values were interpreted with the classification of Araujo et al. (2005a) adapted from Swets (1988): $AUC>0.90$; good $0.80>AUC<0.90$; fair $0.70>AUC<0.80$; poor $0.60>AUC<0.70$; fail $0.50>AUC<0.60$ while the Kappa values were interpreted with the ranges of agreement of Araujo et al. (2005a)) adapted from Landis and Koch (1977) excellent $K>0.75$; good $0.40>K<0.75$; and poor $K<0.40$. We assumed that the geographical transferability fail for models which have an internal evaluation (=IE) > 0.7 for the AUC coefficient (>0.4 for the KAPPA) and which have an external evaluation (=EE) < 0.7 for the AUC (< 0.4 for the KAPPA).

The internal evaluation of the models was made by running a 10-fold cross-validation (van Houwelingen and Le Cessie 1990) on the training dataset. During the cross-validation procedure, the original prevalence of the species' presences and absences in the dataset was maintained in each fold.

The external evaluation was made by projecting each model in the other area and comparing predictions with geographically independent observations using AUC and max Kappa. Hence, this represents a fully independent evaluation, as recommended by Fielding and Bell (1997).

Internal and external evaluations were compared within and between models and regions for the 54 species with paired t-tests.

Measuring and testing transferability

According to our previous definition of full model transferability (see Introduction), three conditions have to be fulfilled, which are respectively based on comparisons of model fit, model evaluation and spatial predictions. Whereas comparison of model fit is straightforward (one measure for each model in each region), comparison of evaluation measures is more complex. To achieve this, we developed an index (Equation 1) that numerically assesses the transferability of a species distribution model across two regions:

$$TI = \frac{\frac{1}{2}\left(\left(1-\frac{\left|AUC_{regA \rightarrow regA} - AUC_{regA \rightarrow regB}\right|}{0.5}\right)+\left(1-\frac{\left|AUC_{regB \rightarrow regB} - AUC_{regB \rightarrow regA}\right|}{0.5}\right)\right)}{1+\left|\frac{AUC_{regA \rightarrow regA} - AUC_{regA \rightarrow regB}}{0.5}\right|-\left|\frac{AUC_{regB \rightarrow regB} - AUC_{regB \rightarrow regaA}}{0.5}\right|} \quad \textbf{Equation1}$$

where $AUC_{regA \rightarrow regA}$ means that the model is fitted in region A and evaluated in the same region. The transferability index (*TI*) is based on the decrease of the AUC coefficient when switching from the internal ($AUC_{regA \rightarrow regA}$ and $AUC_{regB \rightarrow regB}$) to the external ($AUC_{regA \rightarrow regB}$ and $AUC_{regB \rightarrow regA}$) evaluation for both regions. The *TI* varies from 0 to 1 and is at its maximum when the difference between IE and EE is null. Note that this index is only based on the AUC evaluation measure, and does not include assessment of reciprocal spatial predictions. Thus, it only provides information on criteria 1 and 2 when assessing the full transferability.

Four potential relationships implying this transferability index and various species or model properties were tested with linear models across the whole set of species, for both GLMs and GAMs: (1) *TI*s as a function of the differences between prevalences in Switzerland and Austria; (2) *TI*s as a function of the differences in degrees of freedom between the models fitted in CH and in AT; (3) *TI*s as a function of the differences between the adjusted deviance of models fitted in CH and in AT; (4) *TI*s as a function of the similarities between models compositions in CH and in AT.

In addition, we tested separately the difference in degrees of freedom between GLMs and GAMs as a function of ([IE – EE] of GLMs) – ([IE – EE] of GAMs) in each region with a linear regression. The prevalence of a species in one region was calculated using the ratio of the occurrences of the species to the total number of observations. The similarity between the model compositions in CH and AT was calculated with the Simple Matching Coefficient (Legendre and Legendre 1983).

To assess the third criterion of full model transferability, the agreement between internal and external prediction maps was calculated with an asymmetric distance version of Kulczynski's coefficient (Legendre and Legendre 1983). The Kulczynski's coefficient (KC) for comparison of map *j* versus *k*, (here IP vs EP for a given region) D_{jk}, is:

$$KC_{jk} = \frac{\sum_{i=1}^{n} X_{ij} - \sum_{i=1}^{n} \min\left(X_{ij}, X_{ik}\right)}{\sum_{i=1}^{n} X_{ij}}$$ **Equation 2**

where X_{ij} and X_{ik} are the habitat suitability for cell i for maps j and k respectively, and n is the number of cells. Similarly, KC_{kj} can be calculated for k vs j, and KC_{kj} is not, in general, equal to KC_{jk}. The Kulczynski's coefficient varies from 0 to 1 and is at its maximum when the difference between IP and EP is null. The Kulczynski's coefficients were then compared within and between models and regions with Wilcoxon signed-rank tests. Overall trends between models and regions across species were assessed using boxplots and standard pairwise tests.

RESULTS

The measures of model fit are, on average, 24% higher for GAMs than for GLMs in both CH and AT regions. On average, GLMs and GAMs have a higher fit in CH than in AT (83% higher for GLMs and 78% for GAMs; Table 3, Fig. 2.). In addition, the range of values for the explained deviance (adj-D^2) is narrower in CH than in AT for both GLMs and GAMs (Fig. 2.). Results of comparisons across regions and across methods are given in Table 4. Hereafter, reference numbers given in the text for each type of comparison (C1 to C6) are those explained in Table 4.

Table 3 Global comparisons, using Wilcoxon signed-rank tests, of the adjusted deviance of the 54 species models between GLMs and GAMs (= IDs 1 and 2) and between CH and AT (= IDs 3 and 4). The IDs correspond to the numbers in Fig. 2.

	ID	*x*	Test	*y*	*P*
Comparisons between GLMs and GAMs	1	GLM.CH	<	GAM.CH	< 0.001
	2	GLM.AT	<	GAM.AT	< 0.001
Comparisons between CH and AT	3	GLM.CH	>	GLM.AT	0.007
	4	GAM.CH	>	GAM.AT	0.001

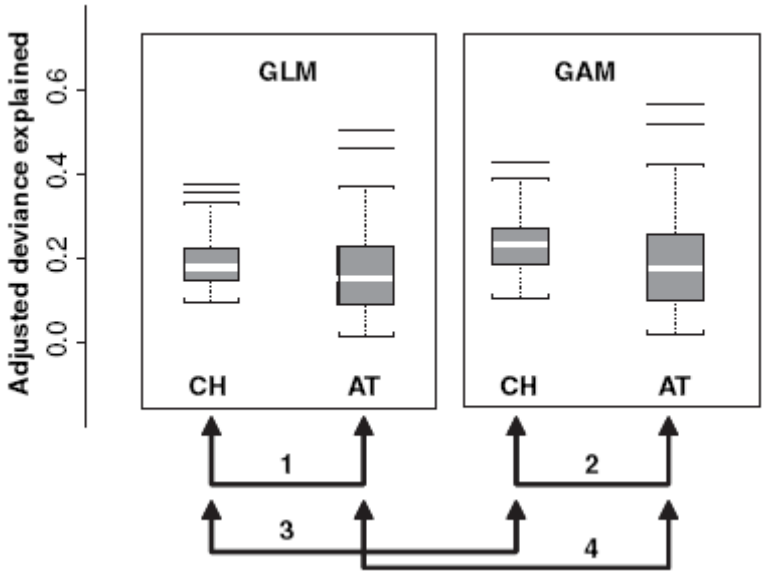


Figure 2 Boxplots of the adjusted D2 explained by GLMs and GAMs in CH and AT for the 54 species. None of the differences in the adjusted D^2 within models between the two regions (arrows 1 and 2) and between GLMs and GAMs within regions is significant (arrows 3 and 4). Wilcoxon signed-rank tests with p-values of 0.898/0.951 and 0.770/0.051 respectively.

Internal evaluations (IE) are systematically higher than external evaluations (EE) for both GLMs and GAMs, between and within regions (C1 and C2). When comparing the two methods (C3) using the Kappa coefficient, IE for GLMs are higher than those for GAMs in AT, whereas no significant difference is observed with the AUC

coefficient. In CH, comparing IE of GLMs and GAMs using either the AUC or the Kappa coefficients yields no significant difference.

Table 4 Comparison by paired t-tests of internal evaluation (IE) and external evaluation (EE) for GLMs and GAMs in CH and AT (e.g. GLMCHAT means a GLM fitted in CH and evaluated with the AT dataset).

	AUC				maxKappa			
	x	Test	*y*	*P*	*x*	Test	*y*	*P*
Transposability								
C1: Comparisons of IE with EE	GLM.CH.ON.CH	>	GLM.CH.ON.AT	< 0.001	GLM.CH.ON.CH	>	GLM.CH.ON.AT	< 0.001
across regions within models	GAM.CH.ON.CH	>	GAM.CH.ON.AT	< 0.001	GAM.CH.ON.CH	>	GAM.CH.ON.AT	< 0.001
	GLM.AT.ON.AT	>	GLM.AT.ON.CH	< 0.001	GLM.AT.ON.AT	>	GLM.AT.ON.CH	< 0.001
	GAM.AT.ON.AT	>	GAM.AT.ON.CH	< 0.001	GAM.AT.ON.AT	>	GAM.AT.ON.CH	0.023
C2: Comparisons of IE with EE	GLM.CH.ON.CH	>	GLM.AT.ON.CH	< 0.001	GLM.CH.ON.CH	>	GLM.AT.ON.CH	< 0.001
within regions and models	GAM.CH.ON.CH	>	GAM.AT.ON.CH	< 0.001	GAM.CH.ON.CH	>	GAM.AT.ON.CH	< 0.001
	GLM.AT.ON.AT	>	GLM.CH.ON.AT	< 0.001	GLM.AT.ON.AT	>	GLM.CH.ON.AT	< 0.001
	GAM.AT.ON.AT	>	GAM.CH.ON.AT	< 0.001	GAM.AT.ON.AT	>	GAM.CH.ON.AT	< 0.001
Methods								
C3: Comparisons of IE	GLM.CH.ON.CH	=	GAM.CH.ON.CH	0.27	GLM.CH.ON.CH	=	GAM.CH.ON.CH	0.48
within regions between models	GLM.AT.ON.AT	=	GAM.AT.ON.AT	0.14	GLM.AT.ON.AT	>	GAM.AT.ON.AT	0.004
C4: Comparisons of IE	GLM.CH.ON.CH	>	GLM.AT.ON.AT	< 0.001	GLM.CH.ON.CH	>	GLM.AT.ON.AT	0.019
across regions within models	GAM.CH.ON.CH	>	GAM.AT.ON.AT	< 0.001	GAM.CH.ON.CH	>	GAM.AT.ON.AT	0.001
Transposability by methods								
C5: Comparisons of EE	GLM.CH.ON.AT	=	GAM.CH.ON.AT	0.82	GLM.CH.ON.AT	>	GAM.CH.ON.AT	0.010
within regions between models	GLM.AT.ON.CH	=	GAM.AT.ON.CH	0.65	GLM.AT.ON.CH	=	GAM.AT.ON.CH	0.99
C6: Comparisons of EE	GLM.CH.ON.AT	=	GLM.AT.ON.CH	0.61	GLM.CH.ON.AT	=	GLM.AT.ON.CH	0.12
across regions within models	GAM.CH.ON.AT	<	GAM.AT.ON.CH	0.015	GAM.CH.ON.AT	<	GAM.AT.ON.CH	0.002

The comparisons of regions (C4) show that IE of GLMs and GAMs are higher in CH than in AT. Furthermore, GLMs fitted in CH have higher EE with the Kappa coefficient than GAMs fitted in CH when both are transferred to AT (C5), whereas no significant differences is observed with the AUC coefficient. However, no such significant differences exist between EE of GLMs and GAMs when they are transferred from AT to CH. The EE of GAMs transferred from AT to CH are higher on average, than those of GAMs transferred from CH to AT (Table 5).

Table 5 Results of the linear regressions correlating the transferability index to the difference of the prevalence of species between the two study regions, the difference of the degrees of freedom between GLMs and GAMs, the difference of the adjusted D2 of the models, and the similarity of model composition between the two regions.

	CH-AT							
	Prevalence		Adjusted D^2		d.f.		Similarity	
Y/X	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*
Transferability index (TI)								
GLM	1.66E-03	7.69E-01	4.19E-04	8.83E-01	1.30E-02	4.10E-01	6.89E-04	8.50E-01
GAM	2.71E-03	7.081E-01	4.23E-02	1.35E-01	8.79E-03	5.00E-01	1.23E-05	9.79E-01

The histograms (Fig. 3) show an important difference between IE and EE of GLMs and GAMs in both regions. When testing the transferability with the AUC metric, the transferability failed for 68% of the species' models fitted in Switzerland and for 55% of the models fitted in Austria for GLMs, and respectively for 67% (CH) and 53% (AT) of the models for GAMs. These values are even higher when evaluated with the Kappa coefficient: 82% of the GLMs fitted in CH failed and 100% of the GAMs, whereas all the models fitted in AT and 91% of the GAMs failed. The distribution of the AUC coefficients on the EE histograms (Fig. 3) shows that, globally, GLMs are only slightly – yet significantly - more robust to transferability than GAM models.

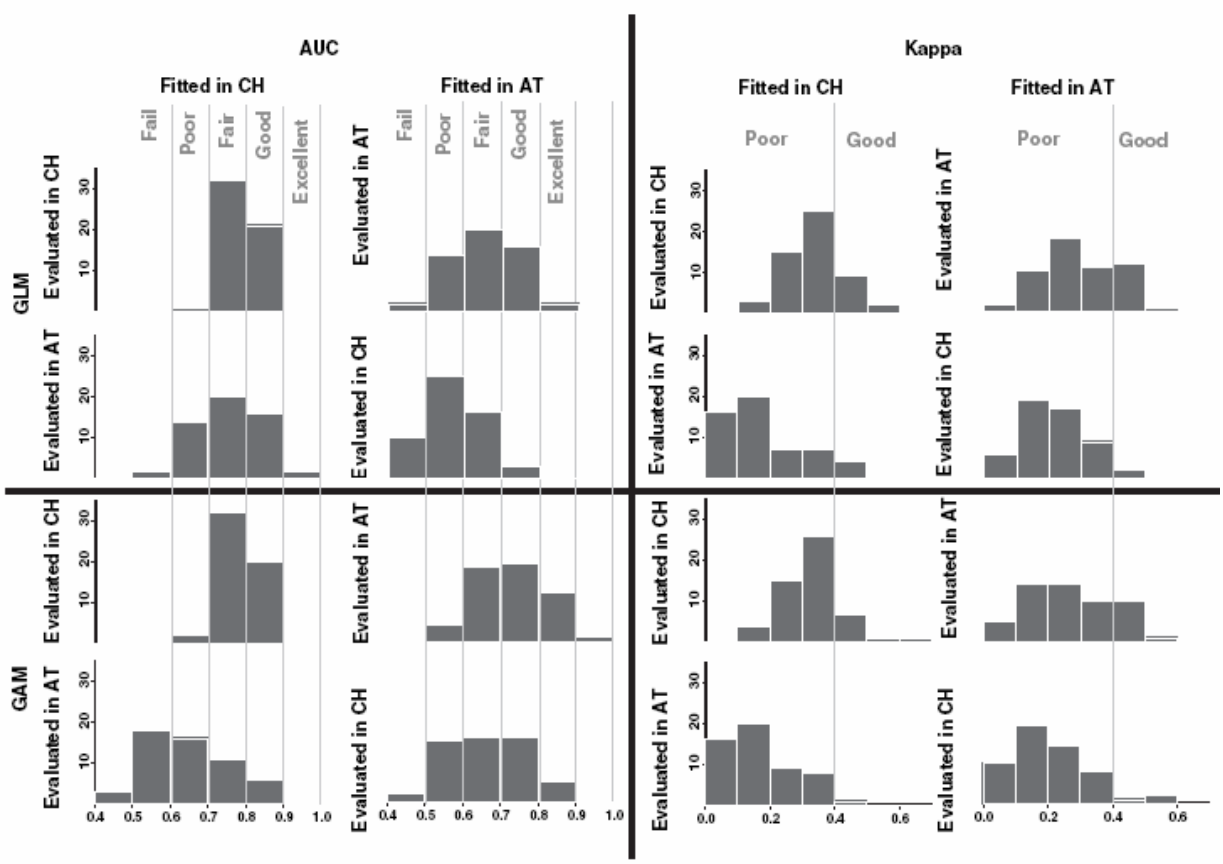


Figure 3 Histograms of the distribution of AUC and Kappa coefficients for the internal and external evaluation in Switzerland and Austria.

The results of the linear regressions in Table 5 show no relationships between the transferability index of GLMs and GAMs and the difference of prevalence, degrees of freedom and adjusted deviance between the two regions. There is also no relation between the TI and the similarity of models, in terms of predictor composition, between CH and AT.

The number of degrees of freedom used in GAMs are higher on average than those in GLMs in CH and AT (Wilcoxon signed-rank tests: p-values<0.001), but the linear regressions show no relationship between the difference in IE and EE between GLM and GAM (i.e. $[IE\text{-}EE]_{GLM} - [IE\text{-}EE]_{GAM}$) and the difference in degrees of freedom between GLMs and GAMs in both regions ($R^2 < 0.01$).

Regarding spatial predictions, the Wilcoxon signed-rank tests on the agreement between prediction maps (Kulczynski's coefficient) show no significant differences between GLMs and GAMs within CH (Fig. 4, arrows 1 below the graph) and within models across regions (Fig. 4, arrows 3 and 4). On the other hand, the agreements between IP and EP of GAMs are higher on average than those of GLMs in AT (Fig. 4, arrows 2, p-value < 0.01). In general, the agreement between internal and

external predictions is rather weak (average value of the coefficient between 0.3 and 0.39) and the variation across species is considerable (from 0 to 0.97).

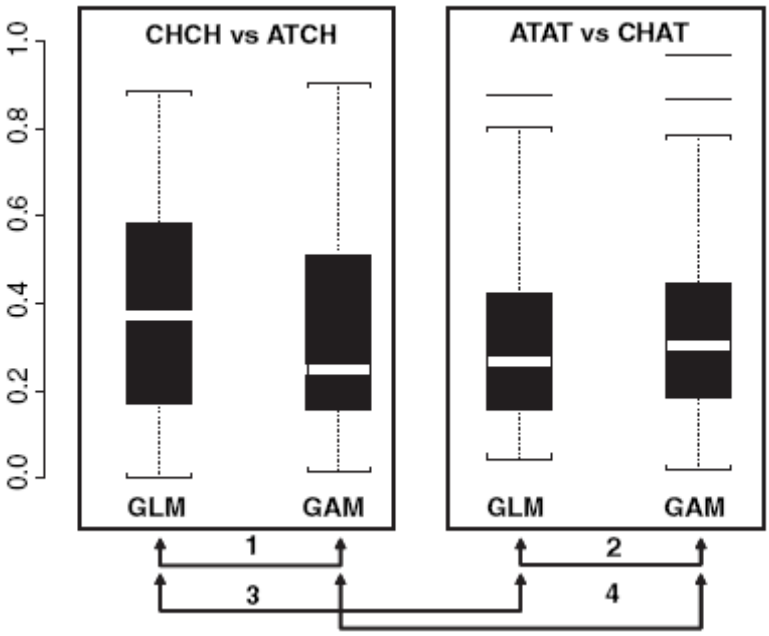


Figure 4 Boxplots of the agreement (Kulczynski's coefficient) between internal and external prediction maps (see figure 1) within the same geographic domain. CHCH versus ATCH means agreement between the model fitted in CH and predicted on the CH geographic domain and the model fitted in AT and predicted on the CH geographic domain. No significant difference was observed in average between the agreements of GLMs and GAMs within CH (arrow 1, Wilcoxon signed-rank test, p-value: 0.179), whereas the Kulczynski's coefficients of GAMs are on average greater than those of the GLMs within AT (arrow 2, Wilcoxon signed-rank test, p-value: 0.002). No significant difference was either observed in average for the agreements of either GLMs or GAMs between the two ranges (arrows 3 and 4, Wilcoxon signed-rank test, p-value: 0.123 and 0.770).

The four examples in Figure 5 illustrate the different types of patterns of GLM transferability across regions. In order to be compared across regions, the four selected species have a similar adjusted deviance in CH and AT. The results of GAM models are not presented because they show exactly the same patterns. Three scenarios are possible. A few species, like *Thesium alpinum* L., meet all the criteria for full transferability (Fig. 5a). For other species, such as *Luzula multiflora* (Retz.) Lej. and *Hypericum maculatum* Crantz s.str., models transfer well from one region to another but not vice versa (asymmetrical transferability; Fig. 5b and 5c). This asymmetric transferability represents the predominant pattern across our 54 species.

Finally, the models of some species are not adequately transferable in either direction (Fig. 5d). However, even when the first two criteria for full transferability are met (see Introduction), the agreement between internal and external predictions remains very low in a significant number of cases (*c.* 10% of the species).

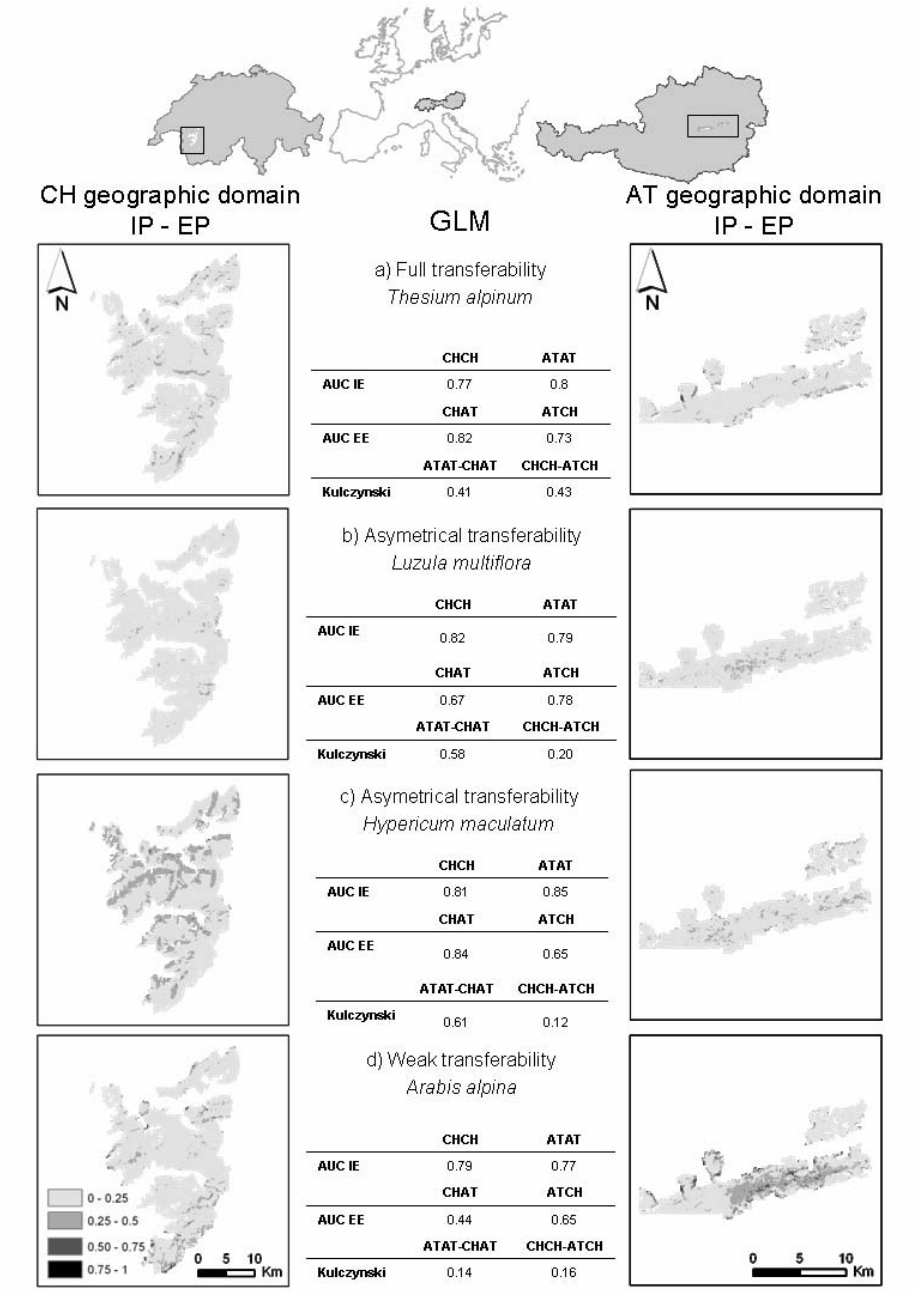


Figure 5 Four examples of GLM transferability across regions. A full transferability gives good external evaluations (AUC) in both regions. *Thesium alpinum* meets all the transferability criteria. Models of other species (*Luzula multifloraa* and *Hypericum maculatum*) are only transferable from one region to another but not vice versa (asymmetrical transferability) whereas the last group of species does not meet transferability criteria in any direction (*Arabis alpina*). The differences between internal and external predictions of GLMs (difference between continuous predictions) are represented on a map for each region. These patterns are the same for the GAMs. The variation of predictions between IP and EP can be very high, even if the conditions for transferability are fully met. For the four examples, the Kappa values and the maps show an important variation between IP and EP.

DISCUSSION

General patterns in transferability

The average transferability of our models across all species is weak compared to other studies (Thomas and Bovee 1993, Schröder and Richter 1999, Mäki-Petäys et al. 2002). The external evaluation failed for more than fifty percent of the species. Considerable discrepancies among species are observed for trans-geographic predictions (i.e. between internal and external prediction maps). Overall, prediction agreement (measured by the Kulczynski's coefficient) across all species and between methods is, on average, weak. Fielding and Haworth (1995) reached the same conclusion, namely that there is a lack of generality in their models. Taken together, these results suggests that care should be taken when these models are to be projected to future climates (see also Araujo et al. 2005b). However, the capacity for transferability is highly species specific. Models for some species meet the criteria for full transferability (e.g. *Thesium alpinum* L.; Fig. 5a), whereas many others do not, showing either asymmetrical transferability or none at all (e.g. *Arabis alpina* L.; Fig. 5d).

Our study thus highlights the importance of testing transferability on a large number of species. Moreover, it reveals a considerable variation in the transferability of SDMs between regions and between modelling techniques. The causes for this limited transferability and high variability cannot be unveiled directly from our results. However, there are aspects of our results that can be used to help identify potential reasons and propose hypotheses for guiding further investigation.

From an inspection of our models we could not find evidence in support of the idea that differences in the degrees of freedoms and in the selected predictors have an effect on the transferability index. Thus, we focus our discussion on the predominant pattern of asymmetrical transferability. Possible causes may be separated in two categories: (i) *environmental* causes, which are specific to differences between geographic regions and (ii) *biotic* causes, which are intrinsic to each individual species being modelled and to the regional species' pool with which it is interacting.

Environmental and biotic limitations to full transferability

With respect to geographic differences, the asymmetrical transferability might depend on the predictor variables in the model and their range in the training region compared to their range in the test region. From a transferability perspective, a model with predictors that cover the same or a wider range in the training region is more likely to give accurate predictions in the test region than the reverse. Hence, differences in the size and in the upper limit of the altitudinal range between the two regions (McPherson et al. 2004) may have truncated the response curves of alpine and subalpine species to temperature-related predictors in AT but not in CH, e.g. true alpine species could have a linear response to degree-days in AT and a

unimodal response in CH. This is equivalent to the problem of extrapolating beyond the realized range of one or more gradients within the training region (Van Horn 2002), which is an important issue in general for the geographical transferability of models.

Transferability is obviously sensitive to where, within the distributional and environmental range of a species, a model was developed and parameterized: the variability of transferability is more pronounced when models are transferred from CH to AT.

Larger scale effects, such as the geographic situation of the study areas (e.g. north-south versus east-west oriented valleys), and differences in land use practices could also affect transferability (Fielding and Haworth 1995). Indeed, Dirnböck *et al.* (2003) showed that for many alpine species land use (history) is a significant predictor of regional distribution patterns. This kind of qualitative variable is, however, extremely difficult to standardize across regions that have different human traditions and agricultural practices.

Shifts in the microclimatic niches of species from one region to another can be understood as a response to climatic differences between regions, i.e. the niche requirements of the species are met by occupying different types of site (Walter and Breckle 1985). However, some differences in local microclimatic conditions may not have been captured by the interpolated environmental variables used in this study. For example, a displacement in the geologic or edaphic (Coudun et al. 2006; this issue) gradient that may be required to recover the water conditions could not have been taken into account with our set of predictors. In our case, the 25m resolution of the predictors may not be appropriate to capture processes like snow accumulations, rock falls or micro-topographic refugees(Gottfried et al. 1998, Lassueur et al. 2006).

Another reason for asymmetrical transferability may be that species whose abundance is highly unequal in the two regions may have their niches deformed in divergent directions. For instance, whereas high regional abundances may drive a niche inflation due to mass effects or source-sink systems (Dias 1996, Pulliam 2000), regional rarity may cause a niche restriction due to Allee effects (Groom 1998) or dispersal limitations (Pulliam 2000; Dirnböck and Dullinger 2004). However, such discussion remains speculative as long as these effects cannot be revealed by specifically designed observational studies of field experiments.

Differences in phenotypic plasticity and the presence of distinct ecotypes in the two regions may also have an influence on transferability ((Joshi et al. 2001)). Solving this would have required genetic analyses to be performed, which was beyond the scope of our study, but it constitutes an interesting direction for future research.

Asymmetrical transferability may also be caused by external biotic factors. Although some authors argue that competitive displacement only rarely affects a species'

geographic range (Thompson et al. 1993, Hill et al. 2000, Prinzing et al. 2002), others suggest that biotic interactions play quite an important role in limiting species' ranges (Zobel 1997, Odland and Birks 1999). More generally, Ozesmi and Mitsch (1997) suppose that transferability is difficult to assess without taking the main interspecific interactions into account. These different points of view may be the result of studies having been taken at different resolutions, and thus may be a question of scale. If the resolution used does not correspond to the one at which competition potentially takes place (i.e. if the competing species can co-occur in the same cell without actually competing for resources), such competitive effects on species' range may not be detectable (Guisan and Thuiller 2005).

As the capacity for transferability seems to depend largely on the species being modelled, further investigations should focus on individual species' properties, such as the relationship between species' traits or full functional types and their transferability ability (Kleyer 2002). In this respect, preliminary results on the relationship between plant traits and the robustness/accuracy of SDMs are promising (Dirnböck and Dullinger 2004, Thuiller 2004a, Thuiller et al. 2004). For instance, by controlling the proportion of the realized niche that a species was able to colonize, the dispersal capacity can influence the distribution of alpine plant species and their niche breadth along environmental gradients, and thus also affect the transferability of models.

Modelling techniques and choice of predictors

The scores of the external evaluation are globally higher for GLMs than for GAMs when models are transferred from CH to AT and the scores of the external predictions are higher for GLMs in both directions of transferability. Moreover, the variability of the external prediction is on average slightly less pronounced with GLMs, which suggests that models fitted with this technique are more robust when transferred than are GAMs. This result is in line with our hypothesis that overfitting can reduce transferability, with GLMs being less prone to overfitting and thus more generalizable. It is, however, in contradiction with results obtained by Araújo et al. (2005b) who found GAM to show superior transferability, in time, to GLMs. Of course, such results are highly dependent upon the way GLMs and GAMs are fitted. If polynomial orders higher than quadratic are allowed, GLMs could well show more pronounced overfitting than GAMs. At least, for the same number of degrees of freedom allowed for each predictor (e.g. third order in GLM and smoothing with three degrees of freedom in GAM), GAMs, as non-parametric models, will always tend to be closer to the data, i.e. more sensitive to peculiarities of particular samples and thus more prone to overfitting.

Another methodological issue with transferability, although very difficult to assess, may be the quality of the predictor variables, and the way each method deals with these. For instance, as GAMs tend to be closer to the data, one could expect them to be more sensitive to measurement or modelling errors in the predictors. However, although these errors may well weaken the models, and thus limit their

transferability, they are unlikely to cause the asymmetrical transferability observed in our study. An additional problem here is that the exact level of error is usually unknown for individual predictors. As a result, a proper error assessment can be difficult to conduct in this type of study. Nevertheless, in future studies, attempts should be made to pay more attention to the quality and the type of predictor variables. According to other studies using similar predictors, those used here are supposedly of a rather proximal nature. However, they still might not be the most proximal and physiologically meaningful ones, and thus may have potentially contributed to cause limited transferability. Further efforts should be made to improve the accuracy of the environmental predictors used to fit this type of model, their spatial resolution and their proximality (in the sense of closeness to causality; (Austin 2002)), especially in complex landscapes such as mountain ecosystems.

Conclusions

Overall, we observed a weak geographic transferability for the 54 SDMs, with considerable variation among species. In this regard, our transferability index proved useful in providing a quantitative estimate of the geographic transferability across regions. Furthermore, none of the differences in the degrees of freedom, in the composition of models, or in species prevalence across regions, seemed to have an effect on the transferability values. The proposed index thus seems useful but requires more thorough testing and evaluation than possible in this single study.

Only a minority of species met the criteria of full transferability while asymmetrical transferability was the predominant pattern. The pronounced variability across species calls for additional multi-species assessments in order to test properly the transferability of SDMs. From our set of species, we suggest that several factors – be they region- or species-specific – may explain the partial transferability, or even total lack of transferability for some species. The slightly better transferability of GLMs compared to GAMs further suggests that overfitting may reduce transferability. This also requires further investigations.

Overall, we conclude that transferability is an important component of model evaluation, particularly when models are to be projected in space or time. Other sources of uncertainty have already been shown to weaken the suitability of SDMs for climate change projections (Thuiller 2004b, Thuiller et al. 2004, Araújo et al. 2005, Araújo et al. 2006, Pearson et al. 2006). Here, we demonstrated that failure to achieve full transferability in space can constitute an additional – yet very important - component of the uncertainty associated with these projections. Their robustness to transferability and the related uncertainty should thus first be estimated and then provided alongside the projections themselves, in order to allow nature managers and conservationists to make informed decisions.

Acknowledgements

We thank MeteoSwiss and the Austrian meteorological office (ZAMG) for providing the meteorological data, Thomas Czáka for GIS support, Nicolas Ray for providing help with the Kulczynski's coefficient, Sophie Rickebusch for useful comments, and all the people who helped with the field work in Switzerland, and particularly: Pascal Vittoz, Stéfanie Maire, Dario Martinoni, Séverine Wydler, Grégoire Vuissoz and Lorenzo De Stefani.

References

Akaike, H. 1973. Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models. Biometrika **60**:255-265.

Araujo, M. B., Pearson, R. G., W. Thuiller, and M. Erhard. 2005a. Validation of species-climate impact models under climate change. Global Change Biology **11**:1504-1513.

Araujo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005b. Validation of species-climate impact models under climate change. Global Change Biology **11**:1504-1513.

Araújo, M. B., W. Thuiller, and R. G. Pearson. 2006. Climate warming and the decline of amphibians and reptiles in Europe. Journal of Biogeography **In press**.

Araújo, M. B., R. J. Whittaker, R. Ladle, and M. Erhard. 2005. Reducing uncertainty in projections of extinction risk from climate change. Global Ecology & Biogeography **In press**.

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. Vegetatio **42**:11-21.

Austin, M. P. 1985. Continuum Concept, Ordination Methods, and Niche Theory. Annual Review of Ecology and Systematics **16**:39-61.

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

Austin, M. P., R. P. Cunningham, and P. M. Fleming. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio **55**:11-27.

Austin, M. P., and T. M. Smith. 1989. A new model for the continuum concept. Vegetatio **83**:35-47.

Bakkenes, M., J. R. M. Alkemade, F. Ihle, R. Leemans, and J. B. Latour. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. Global Change Biology **8**:390-407.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling **157**:281-300.

Brown, J. H., D. W. Mehlman, and G. C. Stevens. 1995. Spatial variation in abundance. Ecology **76**:2028-2043.

Carey, P. D., and N. J. Brown. 1994. The use of GIS to identify sites that will become suitable for a rare orchid, Himantoglossum hircinum L., in future changed climate. Biodiversity Lett. **2**:117-123.

Coudun, C., J. C. Gegout, and C. Piedallu. 2006. Nutritional factors are needed in complement to climate in predicting plant
species distribution: an illustration with Acer campestre L. in France. Journal of Biogeography **In press**.

Dias, P. C. 1996. Sources and sinks in population biology. Trends in Ecology & Evolution **11**:326-330.

Dirnböck, T., and S. Dullinger. 2004. Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. Journal of Vegetation Science **15**:77-84.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Engler, R., A. Guisan, and L. Rechsteiner. 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.

ESRI. 2004. ArcGIS 9.0. *in*, Redlands.

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

Fielding, A. H., and P. F. Haworth. 1995. Testing the generality of bird-habitat models. Conservation Biology **9**:1466-1481.

Franklin, J. 1995. Predictive vegetation mapping : geographic modelling of biospatial patterns in relation to environmental gradients. Progress in Physical Geography **19**:474-499.

Freeman, M. C., Z. H. Bowen, and J. H. Crance. 1997. Transferability of habitat suitability criteria for fishes in warmwater streams. North American Journal of Fisheries Management **17**:20-31.

Glozier, N. E., J. M. Culp, and G. J. Scrimgeour. 1997. Transferability of habitat suitability curves for a benthic minnow, Rhinichthys cataractae. Journal of Freshwater Ecology **12**:379-393.

Godown, M. E., and A. T. Peterson. 2000. Preliminary distributional analysis of US endangered bird species. Biodiversity and Conservation **9**:1313-1322.
Gottfried, M., H. Pauli, and G. Grabherr. 1998. Prediction of vegetation patterns at the limits of plant life: A new view of the alpine-nival ecotone. Arctic and Alpine Research **30**:207-221.
Groom, M. J. 1998. Allee effects limit population viability of an annual plant. American Naturalist **151**:487-496.
Guisan, A., T. C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling **157**:89-100.
Guisan, A., and J.-P. Theurillat. 2001. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**.
Guisan, A., and J. P. Theurillat. 2000. Equilibrium modeling of alpine plant distribution: how far can we go? Phytocoenologia **30**:353-384.
Guisan, A., J. P. Theurillat, and F. Kienast. 1998. Predicting the potential distribution of plant species in an Alpine environment. Journal of Vegetation Science **9**:65-74.
Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models? Ecology Letters **8**.
Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.
Gurtz, J., M. Zappa, K. Jasper, H. Lang, M. Verbunt, A. Badoux, and T. Vitvar. 2003. A comparative study in modelling runoff and its components in two mountainous catchments. Hydrological processes **17**:297-311.
Hampe, A. 2004. Bioclimate envelope models: what they detect and what they hide. Global Ecology and Biogeography **13**:469-471.
Hastie, T., and R. Tibshirani. 1986. Generalized Additive Models. Statistical Sciences **1**:297-318.
Hill, M. O., D. B. Roy, J. O. Mountford, and R. G. H. Bunce. 2000. Extending Ellenberg's indicator values to a new area: an algorithmic approach. Journal of Applied Ecology **37**:3-15.
Huntley, B. 1995. Plant-Species Response to Climate-Change - Implications for the Conservation of European Birds. Ibis **137**:S127-S138.
Jeffree, E. P., and C. E. Jeffree. 1994. Temperature and the biogeographical distributions of species. Funct. Ecol. **8**:640-650.
Joshi, J., B. Schmid, M. C. Caldeira, P. G. Dimitrakopoulos, J. Good, R. Harris, A. Hector, K. Huss-Danell, A. Jumpponen, A. Minns, C. P. H. Mulder, J. S. Pereira, A. Prinz, M. Scherer-Lorenzen, A. S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis, and J. H. Lawton. 2001. Local adaptation enhances performance of common plant species. Ecology Letters **4**:536-544.
Keitt, T. H., M. A. Lewis, and R. D. Holt. 2001. Allee effects, invasion pinning, and species' borders. American Naturalist **157**:203-216.
Kleyer, M. 2002. Validation of plant functional types across two contrasting landscapes. Journal of Vegetation Science **13**:167-178.
Körner, C. 2003. Alpine plant life. Springer, Berlin.
Landis, J. R., and G. G. Koch. 1977. Measurement of Observer Agreement for Categorical Data. Biometrics **33**:159-174.
Lassueur, T., S. Joost, and C. F. Randin. 2006. Very high resolution digital elevation models: Do they improve models of plant species distribution? Ecological Modelling **198**:139-153.
Legendre, L., and P. Legendre. 1983. Numerical Ecology. Elsevier, New York.
Lehmann, A., J. M. Overton, and M. P. Austin. 2002. Regression models for spatial prediction: their role for biodiversity and conservation. Biodiversity and Conservation **11**:2085-2092.
Maggini, R., A. Lehmann, N. E. Zimmermann, and A. Guisan. 2006. Improving generalized regression analysis for the spatial prediction of forest communities. Journal of Biogeography **In press**.
Mäki-Petäys, A., A. Huusko, J. Erkinaro, and T. Muotka. 2002. Transferability of habitat suitability criteria of juvenile Atlantic salmon (Salmo salar). Canadian Journal of Fisheries and Aquatic Sciences **59**:218-228.
MathSoft. 1999. S-PLUS 2000 Guide to Statistics. *in*. Data Analysis Products
Division, MathSoft, Seattle, WA.

McCullagh, P., and J. A. Nelder. 1989. Generalized Linear Models. 2nd edition. Chapman and Hall, London.
McPherson, J. M., W. Jetz, and D. J. Rogers. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? Journal of Applied Ecology **41**:811-823.
Miller, R. I. 1986. Predicting Rare Plant-Distribution Patterns in the Southern Appalachians of the Southeastern USA. Journal of Biogeography **13**:293-311.
Moisen, G. G., and T. S. Frescino. 2002. Comparing five modelling techniques for predicting forest characteristics. Ecological Modelling **157**:209-225.
Mouquet, N., and M. Loreau. 2002. Coexistence in metacommunities: The regional similarity hypothesis. American Naturalist **159**:420-426.
Mouquet, N., and M. Loreau. 2003. Community patterns in source-sink metacommunities. American Naturalist **162**:544-557.
Odland, A., and H. J. B. Birks. 1999. The altitudinal gradient of vascular plant richness in Aurland, western Norway. Ecography **22**:548-566.
Ozenda, P. 1985. La végétation de la chaîne alpine dans l'espace montagnard européen. Masson, Paris.
Ozesmi, U., and W. J. Mitsch. 1997. A spatial habitat model for the marsh-breeding red-winged blackbird (Agelaius phoeniceus L) in coastal Lake Erie wetlands. Ecological Modelling **101**:139-152.
Pearson, R. G., T. P. Dawson, P. M. Berry, and P. A. Harrison. 2002. SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. Ecological Modelling **154**:289-300.
Pearson, R. G., W. Thuiller, M. B. Araujo, E. Martinez-Meyer, G. Midgley, L. Brotons, C. McClean, L. J. Miles, P. Segurado, T. E. Dawson, and D. C. Lees. 2006. Model-based uncertainty in species' range prediction. Journal of Biogeography **In press**.
Peterson, A. T., and e. al. 2000. Geographic analysis of conservation priorities using distributional modelling and complementarity: Endemic birds and mammals in Veracruz, Mexico. Biol. Conserv. **93**:85-94.
Peterson, A. T., S. L. Egbert, V. Sanchez-Cordero, and K. P. Price. 2000. Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. Biological Conservation **93**:85-94.
Peterson, A. T., M. A. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R. H. Buddemeier, and D. R. B. Stockwell. 2002. Future projections for Mexican faunas under global climate change scenarios. Nature **416**:626-629.
Peterson, A. T., M. Papes, and D. A. Kluza. 2003. Predicting the potential invasive distributions of four alien plant species in North America. Weed Science **51**:863-868.
Prinzing, A., W. Durka, S. Klotz, and R. Brandl. 2002. Geographic variability of ecological niches of plant species: are competition and stress relevant? Ecography **25**:721-729.
Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters **3**:349-361.
Schröder, B., and O. Richter. 1999. Are habitat models transferable in space and time? Zeitschrift für Ökologie und Naturschutz **8**:195-205.
Schulla, J. 1997. Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen. Institute of Geography ETH, Zürcher Geographische Schriften, Heft 69, Zürich.
Scott, J. M., P. J. Heglund, J. B. Haufler, M. Morrison, M. G. Raphael, W. B. Wall, and F. Samson, editors. 2002. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.
Swets, J. A. 1988. Measuring the Accuracy of Diagnostic Systems. Science **240**:1285-1293.
Sykes, M. T., and I. C. Prentice. 1995. Boreal Forest Futures - Modeling the Controls on Tree Species Range Limits and Transient Responses to Climate-Change. Water Air and Soil Pollution **82**:415-428.
Thomas, J. A., and K. D. Bovee. 1993. Application and Testing of a Procedure to Evaluate Transferability of Habitat Suitability Criteria. Regulated Rivers-Research & Management **8**:285-294.

Thompson, K., J. G. Hodgson, J. P. Grime, I. H. Rorison, S. R. Band, and R. E. Spencer. 1993. Ellenberg numbers revisited. Phytocoenologia **23**:277–289.
Thuiller, W. 2004a. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology **10**:2020-2027.
Thuiller, W. 2004b. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology **in press**.
Thuiller, W., M. B. Araujo, R. G. Pearson, R. J. Whittaker, L. Brotons, and S. Lavorel. 2004. Biodiversity conservation - Uncertainty in predictions of extinction risk. Nature **430**:U1-U1.
Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the United States of America **102**:8245-8250.
Van Horn, B. 2002. Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. Pages 63-72 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.
van Houwelingen, J. C., and S. Le Cessie. 1990. Predictive value of statistical models. Statistics in Medicine **9**:1303-1325.
Walter, H., and S. W. Breckle. 1985. Ecological systems of the biosphere. I. Ecological principles in global perspective, Springer-Verlag edition, Berlin, Germany.
Weber, E. 2001. Current and potential ranges of three exotic goldenrods (Solidago) in Europe. Conservation Biology **15**:122-128.
Zappa, M. 2002. Multiple-response verification of a distributed hydrological model at different spatial scales. Dissertation No. 14895, ETH Zurich.
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.
Zobel, M. 1997. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence. TREE **12**:266-269.

Chapter 6

Extinction risk of plant species in mountain regions; contrasted views from continental and local scales?

Randin Christophe[1], Engler Robin[1], Zappa Massimiliano[2], Normand Signe[3], Zimmermann Niklaus[2], Pearman Peter[1], Vittoz Pascal[4], Thuiller Wilfried[5] and Guisan Antoine[1]

This manuscript will be submitted to Global Change Biology

[1]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf

[3]Department of Biological Sciences, University of Aarhus, Ny Munkegade, Bygn. 1540, DK-8000 Aarhus C, Denmark

[4]Faculty of Geosciences and Environment (FGSE), University of Lausanne, CH-1015 Lausanne, Switzerland

[5]Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France

Abstract

Due to their conic shape and the reduction of area with increasing elevation, mountain ecosystems were early identified as potentially very sensitive to global warming. Mountain systems may experience unprecedented rates of warming during the next century, two or three times higher than that records of the 20th century. In this context, species distribution models (SDM) have become important tools for rapid assessment of the impact of accelerated land use and climate change on the distribution plant species.

Depending on the geographic extent, elevation range and spatial resolution, different plant species extinction rates are predicted by SDMs. However, no comparison of predictions across scale has been attempted so far, even if SDMs are commonly used as raw material for preparing IPCC reports and eventually guiding nature managers in actions aiming to anticipate biodiversity losses in ecosystems.

Here we assess whether predictions made at two spatial scales – local (LS; 25m pixels) and European (ES; 10′ (16km) pixels) – result in similar species extinction patterns under four scenarios of climate change by the end of the 21st century for 78 subalpine and alpine species in two distinct regions of the Swiss Alps.

We demonstrated that models at ES predicted higher extinction rates than did those at LS under any of the four climate change scenarios. Linear regressions revealed that there is a good agreement between forecasted extinction mismatch between ES and LS and observed elevation range calculated from fine grained data within big pixels. The percentage of species extinct at ES but not at LS reached 86% depending on the mountain system considered. In addition, comparisons on the species level also revealed that alpine and nival species were subject to higher mismatch in regions with an extended nival zone. Our results clearly suggest that elevation range as measured from the local scale within a big pixel was the main driver for the mismatching between predictions of extinctions at the two scales.

If projections at local scale may better reflect the possibility for species to track their environmental optimum to higher elevation, it may also lead to unrealistic projections for low-elevation species. Thus, new approaches that combine models fitted at different scales should be tested. Further studies are also required to better explore which extent and resolution are best for deriving future spatial projections in mountain areas, balancing local peculiarities (site conditions, variability) and larger scale generalities (niche requirements). Decision makers (e.g. IPCC) could then make choices on the basis of projections conducted both at continental, regional and local scales.

Keywords: species distribution model, BIOMOD, mountain, Swiss Alps, scale, Zermatt, Diablerets, Europe, climate change.

Nomenclature: Aeschimann and Heitz (1996)

Introduction

Due to their conic shape and the reduction of area with increasing elevation, mountain ecosystems were early identified as potentially very sensitive to global warming (Guisan et al. 1995, Theurillat et al. 1998, Diaz et al. 2003, Beniston 2006). Furthermore, a recent global assessment of mountain regions sensitivity to climate change indicated that these areas will experience unprecedented rates of warming, two or three times higher than that records of the 20th century (Nogués-Bravo et al. 2006). High-elevation vegetation is expected to be most affected by long-term climate change (Guisan et al. 1995, Beniston et al. 1996, Guisan and Theurillat 2000, Walther 2003), as climate and other abiotic factors usually are more important for plant species distributions than biotic interactions at high altitude(Grabherr et al. 1995, Körner 2003). The first biological impacts of global warming are already apparent, as evidenced by the rapid upward expansion of alpine and nival species' ranges,on many summits in the Alps (Braun-Blanquet 1957, Hofer 1992, Grabherr et al. 1994, Pauli et al. 1996, Walther et al. 2005, Vittoz et al. 2006, Pauli et al. 2007).

In the last decade, species distribution models (SDM; Guisan and Zimmermann 2000, SDM; Guisan and Thuiller 2005) have become important tools for rapid assessment of the impact of accelerated landuse and climate change on the distribution of organisms (Bakkenes et al. 2002, Thomas et al. 2004, Thuiller et al. 2005). In this context, modelling plant species' distributions is particularly appropriate because valid absences are easier to obtain for sessile species than for mobile species, e.g. from plot surveys where exhaustive species lists are drawn (Guisan and Thuiller 2005). Furthermore, as sessile organisms, plants are likely to have more difficulties tracking shifts in suitable habitats under transient climate change than more mobile species such as mammals or birds. So far, only few studies have used SDMs to assess the possible impact of climate change on plant species in mountain environments.

In a recent large-scale SDM study, Thuiller et al. (2005) forecasted that certain mountains regions of Europe (e.g. mid-elevation Alps) could be disproportionably sensitive to climate change with up to 60% species loss per unit area (10' x 10'). This study modeled the distribution of 1350 species from Atlas Florae Europaeae (AFE; Lahti and Lampinen 1999) over whole Europe at a coarse spatial resolution (model fitting at c. 50 km x 50 km; spatial projection at c. 10' x 10'). The climate change scenarios used in this study were those developed by the Intergovernmental Panel on Climate Change (IPCC) mapped for Europe by the Climate Research Unit (CRU; New et al. 2002; http://www.cru.uea.ac.uk/, CRU; Mitchell et al. 2003, Mitchell and Jones 2005) with the A1 "business-as-usual" scenario yielding the most extreme predictions of extinction by year 2080. In a study of 85 subalpine and alpine non-woody open-habitat plants conducted at a much higher resolution (20 m x 20 m) in the Austrian Alps, Dirnböck et al. (2003) predicted that up to 40-50% species could potentially go extinct due to climate and land-use changes. The climate change scenarios used here were regionalized climatic simulations of the

ECHAM4 GCM for the year 2050 and land-use changes. In a sensitivity study using 62 alpine or nival species, Guisan and Theurillat (2000) predicted smaller rates of strict extinction (100% habitat loss), between 2 to 5%, but nearly 40% of the species were nevertheless predicted to loose over 90% of their suitable habitats.

Thus, depending on the geographic extent, elevation range and spatial resolution, different extinction rates are predicted. However, no comparison of predictions across scale has been attempted so far. As predictions based on SDMs are commonly used as raw material for preparing IPCC reports and eventually guiding nature managers in actions aiming to anticipate biodiversity losses in ecosystems, there is an urgent need to assess whether predictions at different scales are comparable.

In particular, what needs to be assessed is whether the discrepancy between predictions results from different species being considered in separate assessments, or if the discrepancy persists for the same set of species when modeled at different scales. A discrepancy between predictions may for instance result from the fact that different scales may reflect topography in different ways, thus translating into divergent predictions. More generally, the same environmental variables may not bear the same meaning for a species when calculated at local or global scales (Guisan and Thuiller 2005). Taking the studies that were previously discussed, the mean temperature interpolated from a few local stations at a 20 m (Dirnböck et al. 2003) resolution does certainly not carry the same information than the mean temperature within a 10' x 10' cell (Thuiller et al. 2005) where variation in elevation is not taken into account. Such differences may explain differences between predictions, as the position of a species' optimum along a gradient of mean temperature may change between the two scales. Besides spatial resolution, geographic extent is also likely to affect predictions. In this regard, a clear advantage when calibrating models over larger extents is the ability to capture a larger part – if not all – of the species' realized niche. At local scale, the extent may not be sufficient to sample all possible conditions a species can tolerate, resulting in a partial view of its niche and truncated response curves (Van Horn 2002). In this case, future predictions of species' distributions may be entailed with errors (Thuiller et al. 2004). Finally, the technique used to fit the models may also play a role, by being variably efficient in capturing the niche accurately.

Here we assess whether predictions made at two spatial scales – local (LS; 25m pixels) and European (ES; 10' (16km) pixels) – result in similar species extinction patterns under scenarios of climate change by the end of the 21st century for a common set of 78 subalpine and alpine species. If not, we ask whether the mismatch between LS and ES predictions may result from: (1) The particular topographic configuration in the study areas at LS, (2) The relative distance between the climatic niche measured at LS and ES, (3) Truncated response curves along climatic gradients at either of the two scales but most likely at LS, (4) The choice of modelling techniques and (5) Difference in model performance between the two scales of models fitted under current conditions.

Material & Method

Study areas

Species distribution models were parameterised for Europe (34 °N – 72 °N, 11 °W – 32 °E; 5.7×10^6 km^2) and for two local study areas in the Western Alps: Diablerets and Zermatt (Fig.1). The Diablerets study area (Fig.1) covers nearly all mountain massifs of the Western Alps of the Canton de Vaud (Swiss state; 6°60′ to 7°10' E; 46°10′ to 46°30′ N) and has a surface of 700 km^2. Elevation ranges from 375 m in Montreux to 3210 m on the top of the Diablerets massif. Temperature and precipitation vary respectively from 8°C and 1200 mm at 600 m to -5°C and 2600 mm at 3000 m (Bouët 1985). The soil parent material is mainly calcareous. Vegetation has been and is still heavily influenced by human activity. Pastures are common in this region and are found from the bottom in the Rhone valley to the subalpine and alpine zones. At lower altitude, grasslands exist under diverse management. The Zermatt study area is located in the Central Alps of the Canton of Valais (Swizerland; 7°58′ to 7°91' E; 45°92′ to 46°06′ N) and covers a surface of 243 km^2. Its elevation ranges from 1480 at the bottom of the valley near the village of Zermatt to 4634 m on the top of the Mt Rose massif. The climate is dry, warm and highly insolated (low cloud cover) in summer and comparably cold and humid during the winter. This area also encompasses the second, the fifth and the eighth biggest glaciers of Switzerland (Gornerletscher, 68.9 km^2; Findelgletscher, 19.1 km^2, Zmuttgletscher, 16.7 km^2; World Glacier Monitoring Service WGMS, http://www.geo.unizh.ch/wgms), representing at this time 43% of the study site.

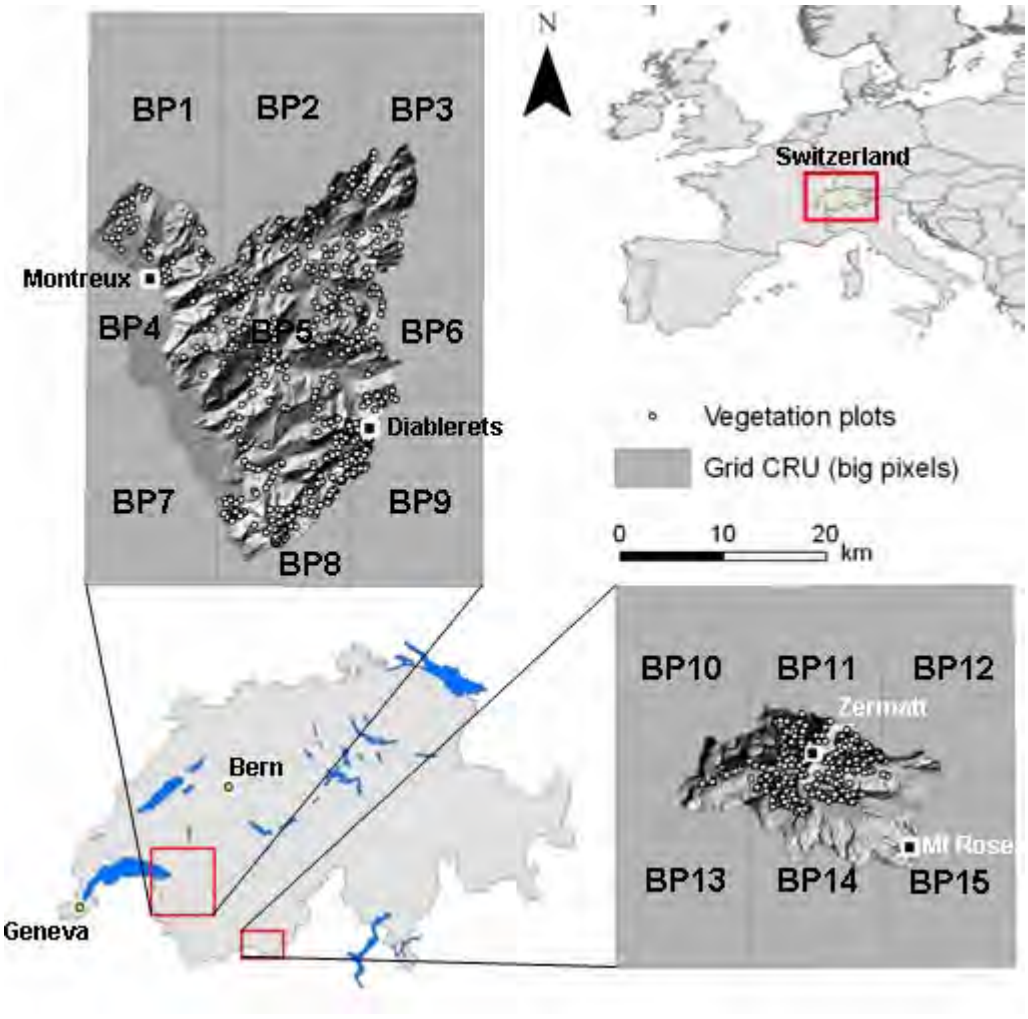


Figure 1 The study area of Diablerets and Zermatt and their location in Switzerland and Europe. Big pixels used for projections at the European scale are indicated for each study area.

Species data

Species distribution data for Europe were extracted from the Atlas Florae Europaeae (AFE; Lahti and Lampinen 1999), which uses a mapping unit of c. 50 x 50 km (AFE cell) based on the Universal Transverse Mercator (UTM) projection and the Military Grid Reference. Sample size was 2089 AFE cells. Species distributions at the local scale were based on species occurrences in 550 and 1511 vegetation plots in respectively the Diablerets and Zermatt study areas. Plots were restricted to open non-woody vegetation (grassland, rock and scree vegetation).

The European species dataset consisted in 78 species that occurs in >19 cells of AFE and of the Diablerets or the Zermatt study area. The Diablerets and Zermatt species

datasets were composed of 42 and 51 species respectively, out of the 78 AFE species in common between the two scales (Appendix).

Climatic predictors

We used seven climatically-derived variables expected to have a major direct ecophysiological impact on plant species (see Prentice et al. 1992 for examples of variables, see Körner 2003 for autoecology of alpine plants): growing degree days (GDD>5°C), mean annual temperature, minimum temperature of the coldest month, mean annual, winter and summer precipitation, and potential evapotranspiration. For the European scale these climatic variables were obtained from the Climatic Research Unit (http://www.cru.uea.ac.uk) at a 10′ resolution. Mean values were averaged for the standard period 1961-1990. These data were then aggregated by averaging all 10′ pixels within each 50 x 50 km AFE cell to match the resolution of the European species dataset. The Diablerets and the Zermatt study areas were respectively captured by nine and six pixels of this coarse grid (hereafter called "big pixels", BP), respectively.

We generated a comparable set of environmental predictors for the Diablerets and Zermatt study areas at a local scale resolution (i.e. 25 m). We first calculated linear lapse rates for long-term (1961–90) monthly mean temperature and monthly rainfall taken from the national meteorological networks of Switzerland (Meteoswiss). Next, we normalized the monthly values to sea level (0 m a.s.l.), using the regression lapse rates, and interpolated the 0-m data to the whole surface of both study areas using inverse distance weighted interpolations (IDW). Finally, the spatially interpolated values (representing locally adjusted regression intercepts) were reprojected to actual elevations using a 25-m DEM and the regression lapse rates (for details, see Zimmermann and Kienast 1999). Additionally, the spatially distributed hydrological model PREVAH (Gurtz et al. 1999, Zappa and Gurtz 2003) was used to obtain a physically based predictor for potential evapotranspiration in both study sites.

Climate change scenarios

Future climate projections for the 2080 time period (averages for 2070-2099) were derived using a general circulation model (GCM) provided by the UK Hadley Center for Climate Prediction and Research (HadCM3; Carson 1999). We used the four different scenarios (A1, A2, B1, B2) provided by the IPCC (Intergovernmental Panel on Climate Change, Nakicenovic and Swart 2000) included in the GCM.

We used the 10′ climatic grids provided by the GCM for spatial projections on the European scale. For projections on the local scale of the four IPCC scenarios, we calculated monthly mean anomalies between the standard period 1961-1990 and the future time period 2070-2099 for temperature and precipitation based on the 10′ grids of the CRU. These anomalies were then downscaled to 25 m-resolution using bilinear interpolation and added to the environmental predictors on the local scale at present climatic conditions.

Niche based-modelling and estimates of extinction risk

Species distribution models were calibrated and projected on both scales using the BIOMOD package (Thuiller 2003). For each species, Generalised Linear Models (GLM), Generalised Additive Models (GAM) and Generalised Boosted Model (GBM) were calibrated on a random sample of the initial data (70%) and evaluated on the remaining dataset with the area under the curve (AUC) of a Receiver Operating Characteristic (ROC) plot (Fielding and Bell 1997). Parametric Generalised Linear Models (GLM) is so far the most common modeling technique used in ecology (e.g. Hill et al. 1999, e.g. Bakkenes et al. 2002). Semi-parametric Generalised Additive Models are now increasingly used (Yee and Mitchell 1991, Frescino et al. 2001, Guisan et al. 2002, Thuiller et al. 2006a). Generalised Boosted Models (Friedman et al. 2000) are somehow newer in ecology (Leathwick et al. 2006, Thuiller et al. 2006b) but appeared to be the best performing across 17 techniques tested in Elith *et al.* (2006).

At the European scale, niche-based models were calibrated on a 50x50 km UTM grid. The modeled species distributions were then projected back onto the fifteen 10' pixels encompassing the two study areas for current and future climate according to Araujo *et al.* (2005). At the local scale, models were calibrated on the vegetation plots for the 42 species of Diablerets and the 51 species of Zermatt. Spatial projections were made over the full geographic extent of each study area using climatic maps at 25-m resolution. A mask based on forests, roads, urbanized areas and rivers was first applied to avoid nonsense predictions at the local scale. Second, species were projected into three different land use categories (grassland, rock and scree) based on the observations in the sampling plots. The distribution of a species was not projected in categories where it species was never observed.

Projected distributions of presence-absence were derived from the probability values predicted by niche-based models, by maximizing the percentage of presence and absence from the calibration dataset correctly predicted by models under current conditions (Pearce and Ferrier 2000, Thuiller 2003).

A species was considered extinct by 2100 when it was predicted to loose 100% of its big or 25-m pixels at the European and local scale, respectively. We compared the percentage of species predicted extinct at ES and LS within each BP under the four climate change scenarios. The percentage of species extinct within a BP was given by the ratio between the number of species predicted to be extinct and the total number of species in the pixel (Eq.1).

(# species extinct / (# species extinct + # species not extinct)) * 100 **(Eq. 1)**

To examine if predictions of extinctions was similar at ES and LS, we calculated the number of species extinct at ES and the number of species extinct at ES but not at LS under the four scenarios. This information was summarized for the two study areas separately. The percentage of "extinction mismatch" between the numbers of

species predicted to go extinct at each of the two scales was then calculated as the number of species predicted extinct at ES but not at LS divided by the number of species extinct at ES (Eq.2).

(# species extinct at ES but not at LS / # species extinct at ES) * 100. **(Eq. 2)**

Assessing the possible causes of mismatch between predicted extinction rates at the European and local scales

To asses the reasons for a possible mismatch between the predicted extinction at European and local scale, we calculated a series of measurements, which, from theory, could explain the observed discrepancy. Disagreement between the predictions at the two scales could result from differences in the range of altitudes represented in each 10' pixel or from divergences between the estimate of climatic conditions at European and local scales. We therefore examined, in a first step, the altitudinal distribution of the 25m pixels and their distribution across vegetation zones within each big pixel. The limits of the vegetation zones were defined differently between Diablerets and Zermatt, taking the effect of continentally into account (Ozenda 1985, Aeschimann and Burdet 1989). We then controlled if there was a relationship between the "extinction mismatch" within a BP and the percent overlay between the study area and the BP, minimum, mean, maximum elevation and range of elevation of 25m pixels. Secondly, within each big pixel we tested the agreement between the estimates of mean annual temperature and precipitation based on the low and high resolution climatic data available for the European and local scale, respectively.

However, the "extinction mismatch" may also depend on the altitudinal optimum at which a species grows, as species at higher elevations are expected to be most negatively affected by climate change. We derived an index reflecting the elevation optimum for the 78 species. This index was a weighted average of the altitudinal distribution of species given in the Flora Alpina (Aeschimann et al. 2005), split into five categories, varying from 1 for strictly collinean species to 5 for strictly nival species. We then examined the "extinction mismatch" within each of these categories. Likewise, species extinction risk could depend on the position of their optimum along climatic gradients. We therefore calculated the position of the environmental optimum of each species at the LS and ES. For each study area, the species calibration datasets at ES and LS were concatenated, centered and scaled for the seven climatic variables so that the multidimensional space defined by the environmental variables were the same for the two scales. Finally, a Principal Component Analysis (PCA) was applied to the difference between the centroids of the species' position at LS and ES. The relationship between the difference in species niche position on the LS and ES and their "extinction mismatch" was tested under the four climate change scenarios with linear regressions.

As proposed by Thuiller *et al.* (2004), truncated response curves can result in spurious future predictions of species' distributions. Therefore, we estimated the

number of species with truncated response curves at LS and ES with Huisman-Olff-Fresco models (HOF-models; Huisman et al. 1993, Oksanen and Minchin 2002). HOF models include a hierarchical set of five models of increasing complexity: (I) flat curve with no response, (II) monotone increasing curve, (III) monotone increasing curve reaching a 'plateau', (IV) symmetric unimodal curve and (V) skewed unimodal response curve. Models I-III represents truncated responses, while models VI-V represent symmetric or skewed unimodal responses. We assessed the effect of a truncated response curve along the local and European scale temperature gradient on the "extinction mismatch", by testing (Wilcoxons signed-rank tests) if species with truncated response curves along the LS and/or ES gradients had a higher degree of "extinction mismatch" than species without truncated response curves at both scales.

Finally, we assessed the difference in AUC for each species between the LS and ES on the evaluation dataset. The relationship between the species mismatching and the difference in AUC was estimated with a linear regression.

Results

Raw predictions of extinctions and mismatching

Our results showed a clear mismatch between the number of extinctions of mountainous plant species in year 2100 when species distribution models were calibrated at either European or local scale. The mismatch between the amount of species predicted to go extinct at the ES and LS was higher than 60% for both study areas (Fig. 2). Mismatching was more important in Zermatt (Fig.6b) for all scenarios. The percentage of mismatching was highest for B2 but the number of species extinct on ES but not on LS was the highest for A1 in both study areas.

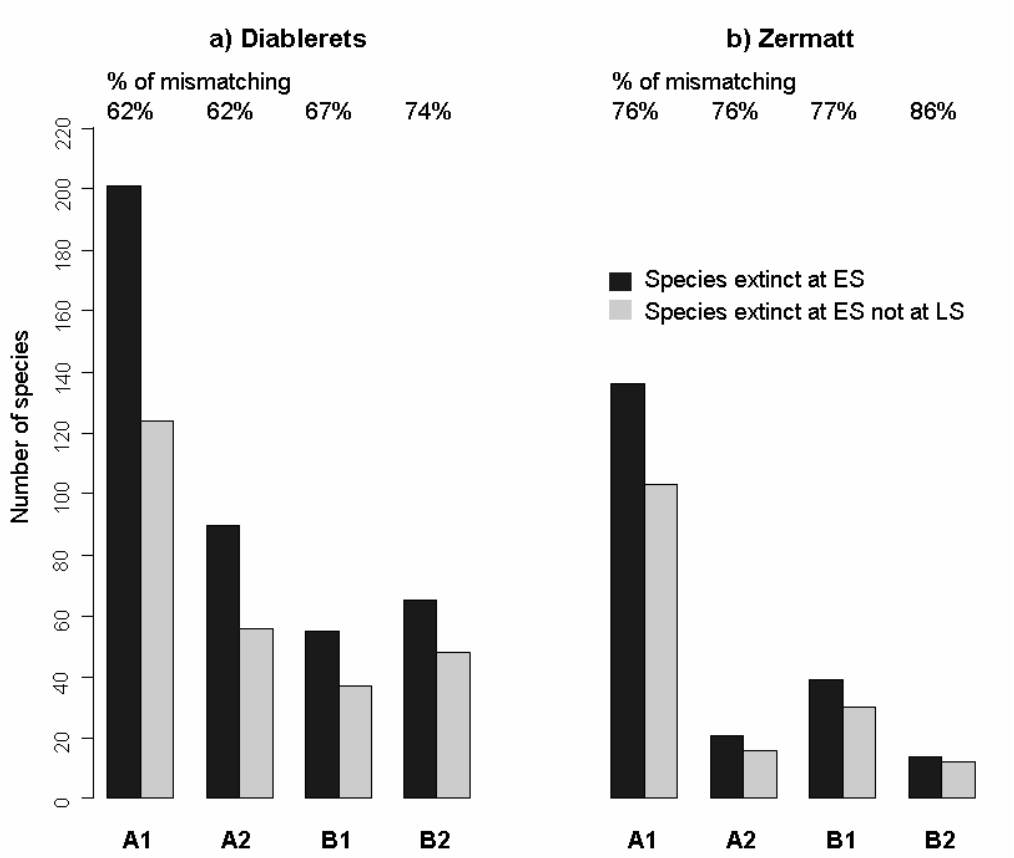


Figure 2 Number of species extinct at ES (columns in black) and number of species extinct at ES but not at LS (column in grey). The percentage of mismatching is the ratio of species extinct at ES but not at LS over species extinct at ES. Big pixels are pooled by study area.

The percentage of species extinctions within big pixels under A1 reached a maximum of 83.3% at ES and 28.6% at LS in Diablerets and 54.2% and 7.8% at ES and LS respectively in Zermatt (Table 1). Under B2, species extinctions were low at both scales in Zermatt, with a maximum of 8.3% at ES and 5.9% at LS. On the

contrary, the extinctions decreased at ES (max.28.6%) but remained high at LS (23.8%) under B2.

Table 1 Percentage of species extinction within big pixels at the European (ES) and local scales (LS).

BP	A1 ES	A1 LS	A2 ES	A2 LS	B1 ES	B1 LS	B2 ES	B2 LS
BP1	76.2	14.3	38.1	11.9	19.0	14.3	26.2	11.9
BP2	61.9	28.6	26.2	31.0	9.5	14.3	23.8	23.8
BP3	42.9	26.2	16.7	28.6	11.9	23.8	14.3	23.8
BP4	83.3	23.8	47.6	28.6	31.0	19.0	28.6	23.8
BP5	38.1	26.2	16.7	31.0	11.9	14.3	11.9	19.0
BP6	38.1	21.4	16.7	19.0	11.9	11.9	14.3	11.9
BP7	64.3	7.1	28.6	7.1	16.7	9.5	11.9	7.1
BP8	38.1	21.4	11.9	21.4	11.9	9.5	14.3	14.3
BP9	35.7	16.7	11.9	19.0	7.1	11.9	9.5	14.3
BP10	52.3	3.9	6.8	3.9	13.6	3.9	4.5	3.9
BP11	54.2	7.8	10.4	3.9	12.5	3.9	6.3	3.9
BP12	40.5	3.9	2.4	3.9	11.9	3.9	2.4	3.9
BP13	52.2	3.9	10.9	3.9	15.2	3.9	6.5	3.9
BP14	45.5	3.9	2.3	5.9	13.6	5.9	2.3	5.9
BP15	54.2	3.9	12.5	3.9	18.8	3.9	8.3	3.9

Linear regressions (Fig.3a) showed no significant positive relationships between species extinctions on the ES and the LS within a BP under the A1 scenario. No significant relationships neither were observed for the A2 and B1 scenarios (pval >0.05). On the other hand, there was a strong and significant relationship between extinctions at ES and LS under the B2 scenario (Fig.3b). Under the A1 scenario, the percentage of extinction at ES were higher than at LS for all BP in both study areas (Table 1). In addition, all the BP in Zermatt and BP1, BP2, BP4 and BP7 in Diablerets showed the highest disagreement.

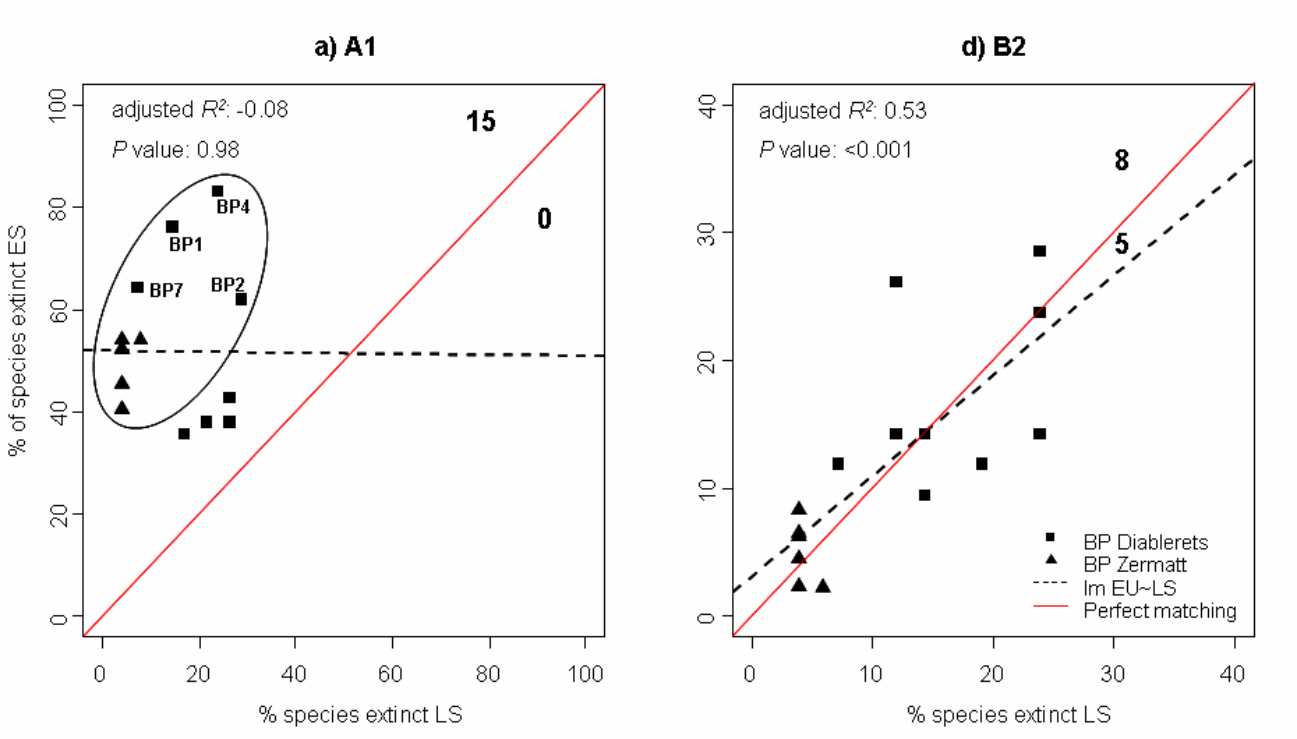


Figure 3 Relationships between extinctions at ES and the LS within each BP for the two extremes scenarios A1 and B2. There is no significant relationship (linear models; P<0.05) between the percentage of species extinct at LS and ES but the explained variance (adjusted R^2) is higher for the low emission scenario B2.

Topographic and climatic differences between European and local scale

Elevation was restricted to colinean and montane zones for BP1 and BP7 in Diablerets (Fig.4). For BP2 and BP3, elevation range was narrow and limited to the subalpine-alpine zones and subalpine zone, respectively.

All the big pixels of Zermatt (BP10 to BP15) had pixels on the 25-m resolution in the nival zone and higher than 3500 m.

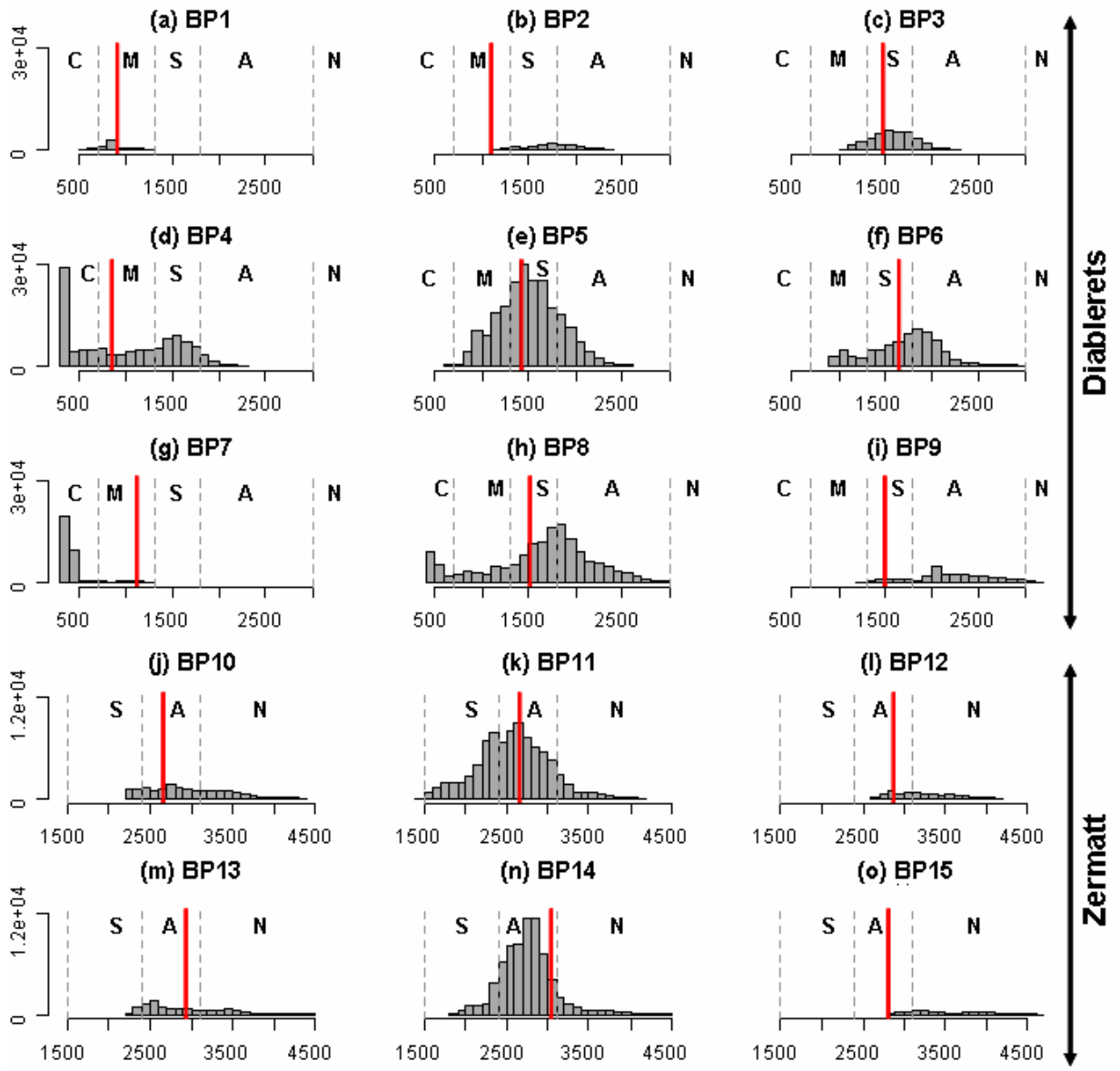


Figure 4 Frequency distribution of elevation (with a 100 m interval) for the 25-m pixels within each BP (big pixel) of Diablerets and Zermatt, and within each altitudinal zones (C=collinean, M=mountainous, S=subalpine, A=alpine, N=nival). The distribution of altitudinal zones reflects the local climatic variation within each study area. The vertical red line indicates the mean elevation of all 25-m pixels contained within a big pixel.

A good relationship was observed within each big pixel (BP) between values of annual mean temperature measured under present climate at European (ES) and local scale (LS) averaged over the BP (adjusted R^2=0.79 and P<0.001; Fig. 5a). On the contrary, no significant relationship could be obtained for annual precipitation measured within BP at the two scales (adjusted R^2=0.00 and P=0.986; Fig. 5b). At ES, precipitation was systematically overestimated and all very similar for BP in Zermatt. In contrast, precipitation was underestimated but again all similar for BP in

Diablerets compared to LS projections (Fig. 5b). BP7 in Diablerets and BP15 in Zermatt both had marginal positions with regards to other BPs.

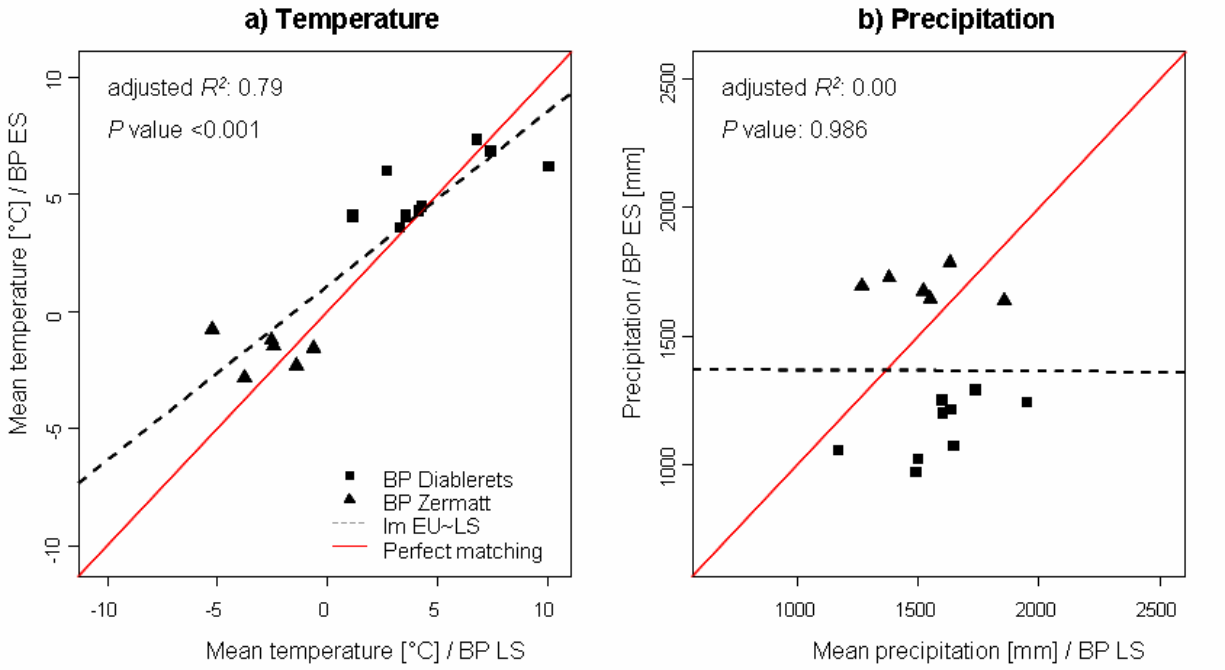


Figure 5 Relationships between (a) projections of the mean annual temperature within the big pixels (BP) at the European scale (ES) and at the local scale (LS) and (b) projections of the mean annual precipitation within big pixels at ES and LS. The mean temperature and mean precipitation at LS are calculated by averaging the values of all 25-m pixels. The diagonal represents a perfect agreement between scales and the dashed lines are regressions lines.

Hypotheses of mismatching

Linear regressions showed no significant relationships between mismatching within big pixels and surface, min, mean and max elevation of big pixels (P values >0.05). However there was a strong and significant relationship between mismatching and elevation range within big pixels under the four climate change scenarios (Fig.6). In addition, the stronger the climate change scenario, the higher the level of significance.

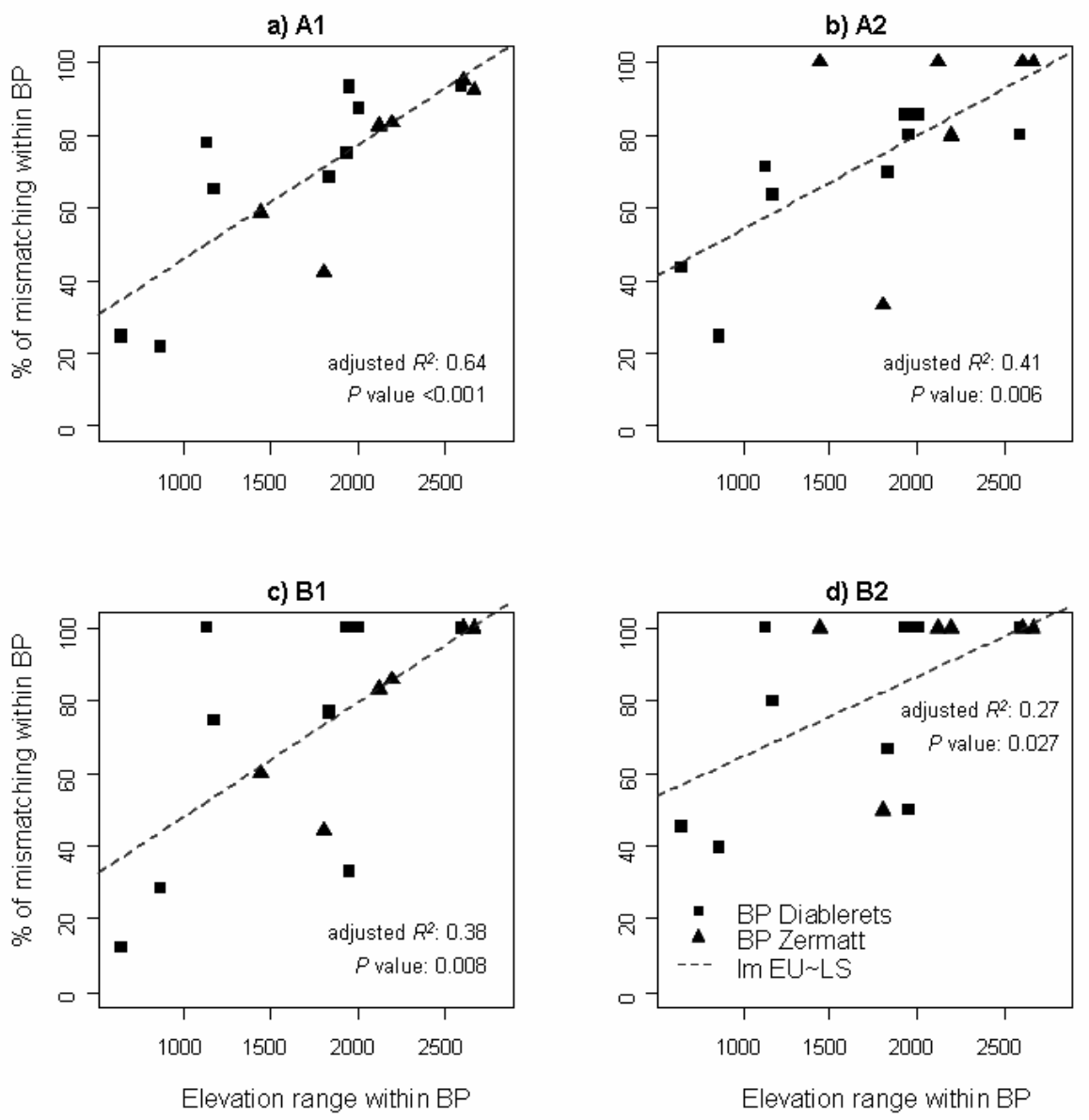


Figure 6 Relationship between the percentage of mismatching and the elevation range within a BP. Dashed lines represent regression trends.

The percentage of extinction mismatch per elevation type was high in general for species with optimum in the subalpine and alpine zones in both study areas (Fig.7). There was no mismatching for species of the nival zone in Diablerets. On the contrary, the mismatching was important for species of the nival zone in Zermatt.

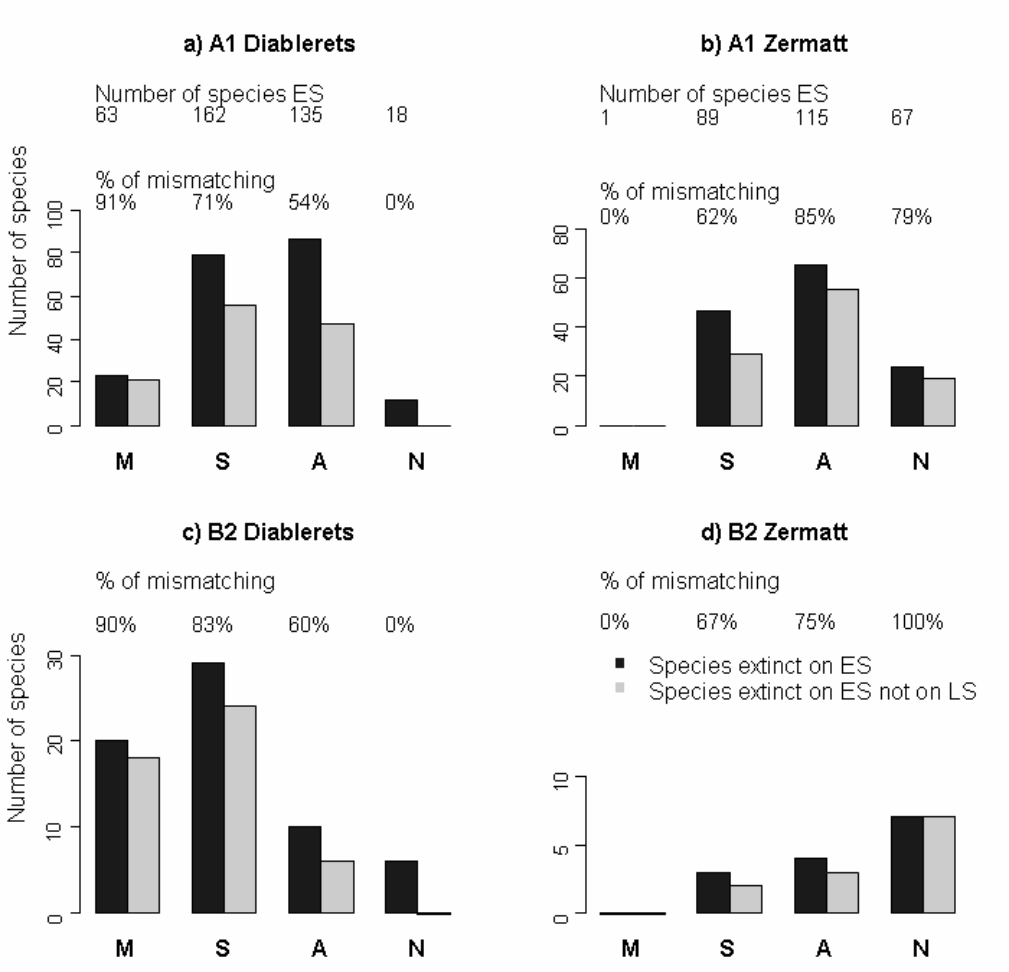


Figure 7 Number of species extinct at ES (columns in black) and number of species extinct at ES but not at LS (columns in grey) per optimum of elevation (M=Mountainous, S=Subalpine, A=Alpine, N=Nival). The two extreme scenarios A1 and B2 are represented for each study area. Big pixels in each study area are pooled. The total number of species present under current climate conditions at the ES is represented on the first top line of each graph per elevation type. The percentage of mismatching (percentage of species extinct at ES but not at LS) is indicated on the second top line.

Linear regressions showed no significant relationship (*P* values >0.05 for A1, A2, B1 and B2) between mismatching and difference in niche position or model quality between LS and ES. Additionally, Wilcoxon tests showed no significant difference (*P* values >0.05 for A1, A2, B1 and B2) of mismatching between species with a truncated response curve of temperature on one or both scales and species without truncated response curve. No difference (Wilcoxon test; *P* values>0.05 for A1, A2, B1 and B2) either was found between species with different modelling techniques at ES and LS and species with the same modelling technique at both scales. Finally, there was no relationship (linear regression; P values > 0.05 for A1, A2, B1 and B2) between mismatching and difference in model performance between LS and ES.

Discussion

In this study we assessed how projections of potential extinction rates by the end of the 21st century at local (LS) and European scale (ES) compare for a common set of species. We demonstrated that models at the European scale (ES) predicted higher extinction rates than did those at the local scale (LS) under any of the four climate change scenarios (A1, A2, B1, B2). The mismatching between scales was higher in Zermatt than in Diablerets. Linear regressions revealed that there is a good agreement between forecasted extinction mismatch and observed elevation range calculated from fine grained data within big pixels. In addition, comparisons on the species level also revealed that alpine and nival species were subject to higher mismatch in Zermatt than in Diablerets. These results clearly suggest that elevation range as measured from the local scale within a big pixel was the main driver for the mismatching between predictions of extinctions at the two scales. Alternative hypotheses failed to explain such mismatch.

Our results highlight the importance of conducting studies at local scale. Indeed, many species were predicted to go extinct at the European scale where no extinction was forecasted at local scale. Since local scale predictions were calibrated from comparably accurate models, they are more likely to represent the expected future extinctions than do ES models. A reason for this is that the number of possible refuges is much better captured at a 25m resolution than at a 10' resolution (corresponding to ~16km in the Swiss Alps). By this – and according to Walter's law of the relative constancy of habitat (Walter and Breckle 1985) – species may shift rapidly to sites with new combinations of the modified climatic gradients at the local scale to meet their niche requirement under climate change. This is obviously not easy to forecast at the scale of a European big pixel. In addition, ES predictions do not allow identifying local extinctions (e.g. extinctions in a particular mountain massif) and thus, potential loss of connectivity between local populations. Yet, we also advise caution regarding local predictions. In our study, the set of predictor variables at the local scale may not fully cover the species niche requirements since it does not capture processes such as snow accumulations, rockfalls, avalanche paths (Dirnbock and Dullinger 2004, Lassueur et al. 2006) or human-induced perturbations such as land use practices (Dirnböck et al. 2003) and intensity, which potentially limit species distribution. Furthermore, the full realized niche of the species may not be captured at the local scale. This may lead to truncated response curves for low-elevation species and spurious predictions of species' future distributions after a major climate change. Indeed, species could be predicted to either migrate infinitely to higher elevations or stay predicted at the lowest elevation under climate change if no limit is set in one of the directions of the gradient (Van Horn 2002, Thuiller et al. 2004).

Overall, the main reason for divergence in model predictions is the importance of the elevation gradient and the topographic diversity expressed at the local scale within each single big pixel. Climatic differences along elevation gradients, as apparent at the 25m resolution, allow plant species to follow their climatic

requirement and to track their climatic optimum under climate change. In contrast, models at 10' resolution only reflect mean climatic conditions within a big cell and thus only capture a fraction of the climatic gradient along altitude. Moreover, our results suggested that GCMs may not express hygric continentality from the precipitation regime, since precipitation is over- or under-estimated in both study area, thus increasing the divergence between predictions. Continentality is one of the main climatic drivers in mountain systems (Beniston 2006), and without accurate representation of it by GCMs/RCMs, model forecasts of species extinction risks will be hampered. As highlighted by Nogués-Bravo *et al.* (2006), the coarse spatial resolution of GCMs used at the European scale does not enable to capture the complex and topographically driven patterns of temperature and other regional climate features. As further proposed by the later authors, regional and local projections based on RCM (Regional Climate Model) or statistical downscaling techniques should be considered in future research in order to better reflect local patterns of climate change within individual mountain ranges.

The divergence between predictions at the two scales was higher in Zermatt, which is not surprising when considering the high elevation ranges available in Zermatt. The shift of plant species to higher elevation was often limited by landcover constraints, which allow only specialized species to colonize rock and scree surfaces. However, our future projections considered glaciers as unsuitable for plant establishment and stable surfaces across time. Ice-covered surfaces represent 43% of the study area of Zermatt and are likely to decrease seriously during the 21st century. Yet, the newly available gravel surfaces will offer no new environmental conditions compared to the currently available surfaces to upward shifting plants.

Subalpine and alpine species were predicted to be less threatened by climate change in Zermatt than in Diablerets, and comparison of extinctions of subalpine and alpine species between scales showed a higher divergence in Zermatt than in Diablerets. Our results highlight the importance of assessing the impact of climate change independently for distinct regions in mountain systems. However, even though divergence across study areas and scales was high, our predictions were made assuming unlimited dispersal from present to future conditions. Considering dispersal under climate change scenario (Carey 1996, Dullinger et al. 2004) based on transient dynamic models may result in less diverging predictions between regions and scales by limiting unrealistically long dispersal or fast migration (Randin et al. in prep).

At both the European and the local scale, the set of purely climatic predictors used in our study may only partly express the true habitat requirements of species. Using more proximal predictors would allow a finer and ecologically more meaningful characterization of the species' realized niche (Austin 1980, Austin et al. 1984, Austin 1985, Austin and Heyligers 1989, Guisan and Zimmermann 2000), thus leading to more accurate predictions with less over-predicted surfaces. However, even with a better description of the niche, genotypic and phenotypic adaptations occur differently in distinct parts of a species' range (Randin et al. 2006). Such local

changes in biotic pressure are likely to generate local modification of the species' realized niches (Pulliam 2000), and will thus affect comparisons between models fitted in the entire niche of the species at the European scale and model fitted locally (Randin et al. 2006).

Conclusions

Despite the coarse resolution of studies at continental scale (e.g. Thuiller et al. 2005), coarse-resolution data are likely to remain attractive, due to the increasing number of distribution atlases becoming available (Araujo et al. 2005). Local field studies are expensive and time consuming, especially in mountainous terrain. The numbers of species collected at the LS is often limited with regards to the thousands of species available in broader scale atlases. If projections at local scale may better reflect the possibility for species to track their environmental optimum to higher elevation, it may also lead to unrealistic projections for low-elevation species. Thus, new approaches that combine models fitted at different scales should be tested. Further studies are also required to better explore which extent and resolution are best for deriving future spatial projections in mountain areas, balancing local peculiarities (site conditions, variability) and larger scale generalities (niche requirements). Decision makers (e.g. IPCC) could then make choices on the basis of projections conducted both at continental, regional and local scales.

Acknowledgments

This study was supported by the Centre de Conservation de la Nature (SFFN) of the Canton de Vaud, by the Swiss National Science Foundation (SNF Grant No 3100A0-110000) and the European Commission (FP6 projects ECOCHANGE and MACIS). We thank all the people who helped with the field work in the Swiss Western Alps, particularly Stéfanie Maire, Dario Martinoni, Séverine Wydler, Chantal Peverelli, Valentine Hof, Grégoire Vuissoz, Lorenzo De Stefani, Roxane Milleret, Pierre Bermane Favrod-Coune, Daniel Cherix, Anny and Willy Berra. C.R. and A.G. greatly benefited from interactions with colleagues within the Swiss National Centre for Competence in Research (NCCR) "Plant survival in natural and agricultural ecosystems" (http:// www.unine.ch/nccr).

References

Aeschimann, D., and H. M. Burdet. 1989. Flore de la Suisse et des territoires limitrophes. Le nouveau Binz, 1 edition. Le Griffon, Neuchâtel.

Aeschimann, D., and C. Heitz. 1996. Index synonymique de la flore de Suisse et territoires limitrophes. CRSF, Genève.

Aeschimann, D., K. Lauber, D. M. Moser, and J.-P. Theurillat. 2005. Flora Alpina, Belin edition, Paris.

Araujo, M. B., W. Thuiller, P. H. Williams, and I. Reginster. 2005. Downscaling European species atlas distribution to a finer resolution: implications for conservation planning. Global Ecology & Biogeography **14**:17-30.

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. Vegetatio **42**:11-21.

Austin, M. P. 1985. Continuum concept, ordination methods and niche theory. Annu. Rev. Ecol. Syst. **16**:39-61.

Austin, M. P., R. P. Cunningham, and P. M. Fleming. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio **55**:11-27.

Austin, M. P., and P. C. Heyligers. 1989. Vegetation survey design for conservation: gradsec sampling of forests in north-east New South Wales. Biol. Conserv. **50**:13-32.

Bakkenes, M., J. R. M. Alkemade, F. Ihle, R. Leemans, and J. B. Latour. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. Global Change Biology **8**:390-407.

Beniston, M. 2006. Mountain weather and climate: A general overview and a focus on climatic change in the Alps. Hydrobiologia **562**:3-16.

Beniston, M., D. G. Fox, S. Adhikary, R. Andressen, A. Guisan , J. I. Holten, J. Innes, J. Maitima, M. F. Price, and L. Tessier. 1996. Impacts of climate change on mountain regions. Chapter 5. *in* IPCC, editor. Intergovernmental panel on climate change. Second assessment report. Cambridge University Press, Cambridge.

Bouët, M. 1985. Climat et météorologie de la Suisse romande, Payot edition, Lausanne.

Braun-Blanquet, J. 1957. Ein Jahrhunder Florenwandel am Piz Linard (3414 m). Bull. Jard. Bot. Etat Brux. **volume jubilaire Walter Robyns**:221-232.

Carey, P. D. 1996. DISPERSE: A cellular automaton for predicting the distribution of species in a changed climate. Global Ecology and Biogeography Letters **5**:217-226.

Carson, D. J. 1999. Climate Modelling: achievements and prospects. Quarterly Journal of the Royal Meteorological Society **125**:1:27.

Diaz, H. F., M. Grosjean, and L. Graumlich. 2003. Climate variability and change in high elevation regions: Past, present and future. Climatic Change **59**:1-4.

Dirnbock, T., and S. Dullinger. 2004. Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. Journal of Vegetation Science **15**:77-84.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Dullinger, S., T. Dirnbock, and G. Grabherr. 2004. Modelling climate change-driven treeline shifts: relative effects of temperature increase, dispersal and invasibility. Journal of Ecology **92**:241-252.

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

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

Frescino, T. S., T. C. Edwards, and G. G. Moisen. 2001. Modeling spatially explicit forest structural attributes using Generalized Additive Models. Journal of Vegetation Science **12**:15-26.

Friedman, J. H., T. Hastie, and R. Tibshirani. 2000. Additive Logistic Regression: a Statistical View of Boosting. Annals of Statistics **28**:337-374.
Grabherr, G., M. Gottfried, A. Gruber, and H. Pauli. 1995. Patterns and current changes in alpine plant diversity. Pages 167-181 *in* A. Chapin and C. Körner, editors. Arctic and alpine biodiversity: patterns, causes and ecosystem consequences. Springer, Heidelberg.
Grabherr, G., M. Gottfried, and H. Pauli. 1994. Climate effects on mountain plants. Nature **369**:448.
Guisan, A., T. C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling **157**:89-100.
Guisan, A., J. I. Holten, R. Spichiger, and L. Tessier, editors. 1995. Potential Ecological Impacts of Climate Change in the Alps and Fennoscandian Mountains. Ed. Conserv. Jard. Bot. Geneva, Geneva, Switzerland.
Guisan, A., and J.-P. Theurillat. 2000. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**:307-320.
Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters **8**:993-1009.
Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.
Gurtz, J., A. Baltensweiler, and H. Lang. 1999. Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountain basins. Hydrol. Process. **13**:2751-2768.
Hill, J. K., C. D. Thomas, and B. Huntley. 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. Proc. R. Soc. Lon. B **266**:1197-1206.
Hofer, H. R. 1992. Veränderungen in der Vegetation von 14 Gipfeln der Berninagebietes zwischen 1905 und 1985. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich **58**:39-54.
Huisman, J., H. Olff, and L. F. M. Fresco. 1993. A Hierarchical Set of Models for Species Response Analysis. Journal of Vegetation Science **4**:37-46.
Körner, C. 2003. Alpine plant life, 2 edition. Springer, Berlin.
Lahti, T., and R. Lampinen. 1999. From dot maps to bitmaps - Atlas Flora Europaeae goes digital. Acta Botanica Fennica **162**.
Lassueur, T., S. Joost, and C. F. Randin. 2006. Very high resolution digital elevation models: Do they improve models of plant species distribution? Ecological Modelling **198**:139-153.
Leathwick, J. R., J. Elith, M. P. Francis, T. Hastie, and P. Taylor. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology-Progress Series **321**:267-281.
Mitchell, T. D., T. R. Carter, P. D. Jones, M. Hulme, and M. New. 2003. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1990-2000) and 16 scenarios (2001-2100). 55, Tyndall Centre Working Paper.
Mitchell, T. D., and P. D. Jones. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology **25**:693-712.
Nakicenovic, N., and R. Swart, editors. 2000. Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
New, M., D. Lister, M. Hulme, and I. Makin. 2002. A high-resolution data set of surface climate over global land areas. Climate Research **21**:1-25.
Nogués-Bravo, D., M. Araújo, M. P. Errea, and J. P. Martinez-Rica. 2006. Exposure of global mountain systems to climate warming during the 21st century. GLOBAL ENVIRONMENTAL CHANGE.
Oksanen, J., and P. R. Minchin. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling **157**:119-129.
Ozenda, P. 1985. La végétation de la chaîne alpine dans l'espace montagnard européen. Masson, Paris.
Pauli, H., M. Gottfried, and G. Grabbherr. 1996. Effects of climate change on mountain ecosystems. Upward shifting of alpine plants. World Resource Review **8**:382-390.

Pauli, H., M. Gottfried, K. Reiter, C. Klettner, and G. Grabherr. 2007. Signals of range expansions and contractions of vascular plants in the high Alps: observations (1994-2004) at the GLROIA master site Schrankogel, Tyrol, Austria. Global Change Biology **13**:147-156.

Pearce, J., and S. Ferrier. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. Ecological Modelling **128**:127-147.

Prentice, C., W. Cramer, S. P. Harrison, R. Leemans, R. A. Monserud, and A. M. Solomon. 1992. A global biome model based on plant physiology and dominance, soil properties and climate. Journal of Biogeography **19**:117-134.

Pulliam, H. R. 2000. On the relationship between niche and distribution. Ecology Letters **3**:349-361.

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

Randin, C. F., R. Engler, P. Vittoz, T. Czáka, M. Beniston, N. E. Zimmermann, and A. Guisan. in prep. Simulating the dispersal of plant species under transient climate change in a mountain environment: towards more realistic predictions? .

Theurillat, J.-P., F. Felber, P. Geissler, J.-M. Gobat, M. Fierz, A. Fischlin, P. Küpfer, A. Schlüssel, C. Velluti, and G.-F. Zhao. 1998. Sensitivity of plant and soil ecosystems of the Alps to climate change. Pages 225-308 *in* P. Cebon, U. Dahinden, H. C. Davies, D. Imboden, and C. C. Jäger, editors. Views from the Alps. Regional perspectives on climate change. MIT Press, London.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. Van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from climate change. Nature **427**:145-147.

Thuiller, W. 2003. BIOMOD - Optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biology **9**:1353-1362.

Thuiller, W., O. Broennimann, G. Hughes, J. R. M. Alkemade, G. F. Midgley, and F. Corsi. 2006a. Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. Global Change Biology **12**:424-440.

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

Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the United States of America **102**:8245-8250.

Thuiller, W., G. F. Midgley, M. Rouget, and R. M. Cowling. 2006b. Predicting patterns of plant species richness in megadiverse South Africa. Ecography **29**:733-744.

Van Horn, B. 2002. Approaches to habitat modelling: the tensions between pattern and process and between specificity and generality. Pages 63-72 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.

Vittoz, P., S. Jutzeler, and A. Guisan. 2006. Flore alpine et réchauffement climatique: observation de trois sommets valaisans à travers le 20ème siècle. Bulletin de la Murithienne **123/2005**:49-59.

Walter, H., and S. W. Breckle. 1985. Ecological systems of the biosphere. I. Ecological principles in global perspective, Springer-Verlag edition, Berlin, Germany.

Walther, G.-R., S. Beißner, and C. A. Burga. 2005. Trends in the upward shift of alpine plants. Journal of Vegetation Science **16**:541-548.

Walther, G. R. 2003. Plants in a warmer world. Perspectives in Plant Ecology, Evolution and Systematics **6**:169-185.

Yee, T. W., and N. D. Mitchell. 1991. Generalized additive models in plant ecology. J. Veg Sci. **2**:587-602.

Zappa, M., and J. Gurtz. 2003. Simulation of soil moisture and evapotranspiration in a soil profile during the 1999 MAP-Riviera Campaign. Hydrology and Earth System Sciences **7**:903-919.

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.

Appendix

Species found in AFE and in the Diablerets or Zermatt dataset.

Species AFE-Diablerets

Aconitum napellus aggr.
Arabis alpina s.str.
Arabis caerulea
Arabis hirsuta
Asplenium viride
Biscutella laevigata
Botrychium lunaria
Cardamine pratensis
Cerastium fontanum subsp. vulgare
Cerastium latifolium
Cerastium arvense s.l.
Cystopteris fragilis
Draba aizoides
Gypsophila repens
Pritzelago alpina s.str.
Juniperus communis subsp. nana
Picea abies
Polygonum bistorta
Polystichum lonchitis
Polygonum viviparum
Pulsatilla alpina s.str.
Ranunculus aconitifolius
Ranunculus alpestris
Ranunculus bulbosus
Ranunculus acris s.l.
Ranunculus montanus aggr.
Ranunculus repens
Rumex acetosa
Rumex alpinus
Rumex alpestris
Rumex crispus
Sagina saginoides
Salix retusa
Salix reticulata
Selaginella selaginoides
Silene acaulis
Silene dioica
Silene vulgaris s.l.
Stellaria graminea
Thesium alpinum
Trollius europaeus
Urtica dioica

Species AFE-Zermatt

Arabis alpina s.str.
Arenaria ciliata
Arabis caerulea
Arabis ciliata
Biscutella laevigata
Botrychium lunaria
Cardamine resedifolia
Cerastium latifolium
Cerastium pedunculatum
Cerastium cerastoides
Cerastium arvense subsp. strictum
Cerastium uniflorum
Dianthus carthusianorum subsp. vaginatus
Dianthus sylvestris
Draba aizoides
Draba siliquosa
Draba dubia
Draba fladnizensis
Draba hoppeana
Equisetum variegatum
Erysimum rhaeticum
Gypsophila repens
Herniaria alpina
Pritzelago alpina subsp. brevicaulis
Juniperus communis subsp. nana
Juniperus sabina
Minuartia recurva
Minuartia sedoides
Minuartia verna
Oxyria digyna
Pinus cembra
Polygonum viviparum
Pulsatilla halleri
Pulsatilla alpina subsp. apiifolia
Pulsatilla vernalis
Ranunculus glacialis
Ranunculus montanus aggr.
Rumex acetosella s.str.
Salix breviserrata
Salix foetida
Salix herbacea
Sagina saginoides
Salix retusa
Salix reticulata
Salix serpillifolia
Selaginella selaginoides
Silene nutans s.str.
Silene rupestris
Silene vulgaris s.str.
Thesium alpinum
Thalictrum foetidum

Chapter 7

Simulating dispersal under transient climate change in a mountain environment: towards more realistic projections?

Randin Christophe[1], Engler Robin[1], Vittoz Pascal[4], Czáka Thomas[1,2], Beniston Martin[3], Zimmermann Niklaus[4] and Guisan Antoine[1]

This manuscript will be submitted to Global Change Biology

[1]University of Lausanne, Department of Ecology and Evolution, Faculty of Biology and Medicine, Bâtiment Biophore, 1015 Lausanne, Switzerland

[2]Faculty of Geosciences and Environment (FGSE), University of Lausanne, CH-1015 Lausanne, Switzerland

[3]Climate Change and Climate Impacts group, (C3i), University of Geneva, Switzerland

[4]Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf

Abstract

While many modeling studies have forecasted the possible impact of climate change on plant distribution, only few of them have attempted to take dispersal mechanisms into account. Indeed, most studies published so far are based on static species distribution models, using the assumption of either universal (unlimited) or no dispersal. However, depending on the rate of climatic change, the fragmentation of the landscape and the individual dispersal capabilities of the species, these assumptions are likely to reveal inaccurate, possibly leading to overestimation of future species' distributions. As a result, the concepts of "potentially suitable" and "potentially colonizable" habitats are expected to differ significantly.

To test this hypothesis, we used MIGCLIM, a dispersal model simulating plant dispersal and migration under transcient climate change, which implements various dispersal parameters (e.g. dispersal distance, barriers to dispersal, long distance dispersal). Four simulations (unlimited and no dispersal, short and long distance dispersal) were carried out for 284 species in the Western Swiss Alps, under four climate change scenarios. Our results indicate that (i) the future potential distributions generated using MIGCLIM can significantly differ from those predictions ignoring dispersal, but also that (ii) the uncertainty related to the temperature warming scenario can be greater than that related to the dispersal parameters.

We showed with realistic dispersal simulations that, under the most optimistic climate change scenario (B1), 53.9% of the species may loose more than 80% of their initial habitat by the end of the 21st century. As a major result, simulations based on realistic dispersal distances for >300 species yielded results on average closer to unlimited dispersal simulations than to the no-dispersal scenario. Secondarily, long distance dispersal simulations yielded similar results to short distance simulations, but this may depend on the way both estimates were obtained from literature. The simulations over the entire 21st century also showed that, due to the possibility of a large number of species shifting their distribution to higher elevation, the bulk of extinctions would mainly occur between 2040 and 2100. However, landuse changes during the first part of the century may contribute to decrease or increase this estimated rates of extinction.

Keywords: generalized linear models (GLM), habitat distribution, cellular automaton, seed dispersal, vegetation, Western Swiss Alps, climate change.

Nomenclature: Aeschimann and Heitz (1996)

Introduction

During the last three decades, strong evidences have accumulated to suggest that increased atmospheric concentration of greenhouse gases – and particularly carbon dioxide and methane, likely caused by human activities – have already begun to modify the climate (Houghton et al. 1983, Houghton and Woodwell 1989, Woodwell et al. 1998). On average for the Earth, temperature already increased by 0.61 (+/-0.18K) between 1861 and 2000 (Folland et al. 2001, Root et al. 2003). This warming likely result from the parallel CO_2 increase from 280 to 370 ppmv observed since the "Industrial Revolution", at a rate 100 times faster than over the last 20'000 years (Berger 2002).

In this context, plants can react in three different ways (Theurillat and Guisan 2001, Körner 2003): (i) stay (persist) in the modified climate, (ii) migrate to more suitable climates, and (iii) extinct. If plants cannot stand or adapt to these rapid climate changes, there remain only two possibilities for them, either migrate or disappear. Many studies show that species are already moving both polewards in latitude and upwards in elevation (Grabherr et al. 1994, Parmesan 1996, Shugart et al. 2001, Klanderud and Birks 2003, Parmesan and Yohe 2003). Migration can only occur if the dispersal strategies of the species allow it, and if no migration barrier is present. Successful migration could be defined as situations where species can migrate and colonise a substantial amount of new suitable locations, whereas unsuccessful migration would represent situations where the species fail to migrate to newly suitable habitats, for instance due to the presence of physical barriers (valley edge, large areas of unsuitable substrate, or forests). If at the same time, the species cannot adapt to the modified climatic conditions at its current locations, then drastic range reduction or even local extinction – or quasi extinction – is to be expected. Quasi extinction means that some species could still persist in a few little scattered locations under the form of relict populations, but these may be very vulnerable to any disturbance or further climate change. Extinction is particularly expected for species having a long maturation time (i.e. before it can produce new propagules), slow growth rate, or a weak dispersal capacity (Bradshaw and McNeilly 1991, Klanderud and Birks 2003). Indeed, a few extinctions have already been observed during the last 30 years (Parmesan and Yohe 2003, Root et al. 2003).

Due to their conic shape and the reduction of area with increasing elevation, mountains ecosystems were thus early identified as potentially very sensitive to global warming (Beniston 1994, Guisan et al. 1995, Beniston et al. 1996, Theurillat et al. 1998, Diaz et al. 2003). Rebetez (2002) already reported for the Swiss Alps a temperature increase twice the global average in the Northern hemisphere (i.e. +1.35K during the twentieth century both at low and high altitudes). More recently, a global assessment of mountain sensitivity to climate change further indicated that these areas will experience a rate of warming two or three times higher than that recorded during the 20th century (Nogués-Bravo et al. 2006). High-elevation vegetation in particular is expected to be most affected by a long-term climate change (Grabherr et al. 1995, Beniston et al. 1996, Guisan and Theurillat 2000,

Theurillat and Guisan 2001, Walther 2003). The first biological impacts of global warming (the so-called biological "fingerprints of climate change") are already apparent, as shown by the rapid expansion of alpine and nival species' range upward in elevation, recently evidenced for many summits in the Alps (Braun-Blanquet 1957, Hofer 1992, Grabherr et al. 1994, Pauli et al. 1996, Walther et al. 2005, Vittoz et al. 2006).

In the last decade, static species distribution models (SDM; Guisan and Zimmermann 2000, Guisan 2005) have become important tools to provide a rapid assessment of the impact of accelerated land use and climate change on the distribution of organisms. A broad range of projections have already been carried-out to forecast future plant species distributions, at large scale (e.g. Huntley et al. 1995, Bakkenes et al. 2002, Thuiller et al. 2005). However, so far, only few of these studies have included dispersal to draw model projections (e.g. Carey 1996, Iverson and Prasad 1998, Iverson et al. 2004, Broennimann et al. 2006). Instead, most of them have used the assumption of universal dispersal (i.e. a species has unlimited dispersal, its future distribution becoming the entire area projected by the habitat distribution model; Guisan and Theurillat 2000, Bakkenes et al. 2002, Dirnböck et al. 2003). While this hypothesis might provide good approximations for plants that are man-dispersed or have high dispersal ability, it is likely to overestimate the future distribution of many other species, because: (i) human-driven habitat fragmentation has rendered the landscape increasingly impassable (Pitelka et al. 1997) and (ii) the speed of the expected global warming is predicted to be rapid – one or more orders of magnitude faster than past climate changes (Etterson and Shaw 2001) – and thus may, at least for certain species, require migration rates much faster than those observed during post-glacial time (Malcolm et al. 2002). This means that not all species will be able to keep pace with climate change. Furthermore, such rapid change may not allow sufficient time for species to adapt (Etterson and Shaw 2001). Because the universal dispersal assumption represents a too optimistic "best case" scenario, some authors (e.g. Thomas et al. 2004, Thuiller et al. 2004, Thuiller et al. 2005) also provide a "worst case" without dispersal for each species, as a way to provide a lower bound to projections, or set the same maximum dispersal rate for all the species (Broennimann et al. 2006). However, the difference between these extreme projections yields large uncertainties (e.g. Thuiller 2004). Reducing such uncertainty calls for the inclusion of dispersal processes when addressing the issue of climate change impact on plant distribution. So far, only one study included limitations by dispersal in SDM projections at a fine scale in a mountain environment (Dullinger et al. 2004) and only for one species, but no study assessed the effect of dispersal limitation on a large number of species in mountain systems when projecting SDM under climate change.

One generally distinguishes between the normal dispersal, or short-distance dispersal, with most of seeds falling in a restricted distance, predictable with distance kernel, and long-distance dispersal relying on more stochastic events. As reported by Pearson (2006), stochastic long distance dispersal was necessary to explain the rapid migration of trees during the last postglacial warming. Moreover,

Clark (1998) and Iverson *et al.* (2004) both stated that long distance may have more impact on the overall rate of migration under a future climate change than movement of the concentrated front. Stochastic long distance dispersal events should then be taken into account and estimated, as well as realistic short and mean distance of dispersal.

In this study we assessed the extinctions of mountain plants under four climate change scenarios for the 21st century based on realistic dispersal distances of seeds. We tested this by simulating both short and long distance dispersal events into a cellular automaton in the Western Swiss Alps. More specifically, it was to answer the following questions:

1) Are the realistic dispersal simulations closer to distribution models assuming no dispersal of unlimited dispersal?
2) What is the temporal evolution of extinctions?
3) Does extinction risk of mountain plants estimated with different dispersal simulations relate to species traits like their initial distribution extent, their dispersal capacity or their elevation optimum?

Material & Method

Study areas

The Diablerets study area (Fig. 1) covers nearly all mountains of the Western Alps of the Canton de Vaud (Switzerland; 6°60' to 7°10' E; 46°10' to 46°30' N) and has a surface of 700 km^2. Elevation ranges from 375 m in Montreux to 3210 m on the top of the Diablerets massif. Annual temperature and precipitation vary respectively from 8°C and 1200 mm at 600 m to -5°C and 2600 mm at 3000 m (Bouët 1985). The soil parent material is mainly calcareous. Vegetation has been and is still very influenced by human activity: the pasture is common in this region, and is observed from the bottom in the Rhone valley to the subalpine and lowers alpine zones at higher elevations. Meadows of diverse intensity of management are mainly found at lower elevations.

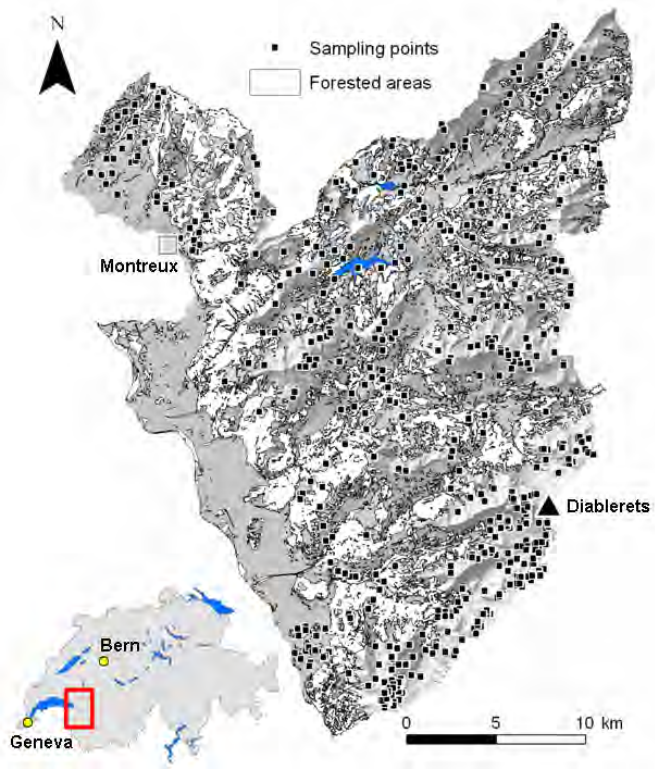


Figure 1 Study area of Diablerets and location in Switzerland.

Species dataset

During summers 2002-2004, 550 vegetation plots were sampled following a random-stratified sampling strategy restricted to non-woody vegetation (grassland, rock and scree). Stratification was done by elevation, slope, aspect and simplified classes of geology. Presence of all species was recorded on 64m^2 plots. We registered the exact coordinates of the centre of each plot using a Trimble GPS

system, with a mean accuracy of less than 1 m after a differential post-correction. We conserved only the 284 species with 20 or more presence points (Appendix A) to avoid loss of robustness (Lehmann et al. 2002).

Climatic predictors

Climate data were derived from the network of national meteorological stations at different altitudes. Long-term monthly means for average temperature (°C) and sum of precipitation (mm) for the period 1961-1990 were used. A digital elevation model (DEM) was used to spatially interpolate the climatic data. Method of computation and description of the variable are developed in Zimmermann *et al.* (1999). In this study, we used three climatic and two topographic variables at a resolution of 25 m (Table 1). These variables are expected to be of greatest ecophysiological significance (Pearson et al. 2002, Körner 2003). Degree-days of the growing season were derived from interpolated daily temperatures using a threshold value of zero degree. Moisture index is calculated as the difference between precipitation and evapotranspiration, and thus expresses the amount of soil water potentially available at a site. The sum of mean daily values for the months of June, July and August were used and represents the warmest months of the year.. The potential global solar radiations were calculated over the year. Topographic position and slope were derived from the 25-m resolution DEM. Positive values of topographic position express relative ridges, tops and exposed sites, whereas negative values indicate sinks, valleys or toe slopes.

Table 1 Table of the topographic and climatic variables used in this study to model the distribution of species

Variables	Units	Details	Method	References
Temperature degree days	°C * day * year^{-1}	Sum of days multiplied by temperature > 0 C	ARCInfo AML	Zimmermann & Kienast (1999)
Moisture index (average of monthly values June–August)	mm * day^{-1}	Monthly average of daily atmospheric H_2O balance	ARCInfo AML	Zimmermann & Kienast (1999)
Global solar radiation (sum of monthly values)	kJ * m^{-2} * day^{-1}	Monthly average of daily global solar radiation	ARCInfo AML	Zimmermann & Kienast (1999)
Slope	degrees	Slope inclination	DEM, ARCInfo, GRID routine	ESRI (2005)
Topographic position	uniteless	Integration of topographic features (ridge, slope, toe slope) at various spatial scales	ARCInfo AML	Zimmermann (1996)

Climate change scenarios

We used the four different scenarios (A1, A2, B1, B2) provided by the IPCC (Intergovernmental Panel on Climate Change, Nakicenovic and Swart 2000) included in the GCM .Future climate projections for the 2001-2100 time period were derived using the general circulation model (GCM) provided by the UK Hadley Center for Climate Prediction and Research (HadCM3; Carson 1999).

The 10' climatic grids provided by the GCM were used to calculate monthly mean anomalies between the standard period 1961-1990 and 20 future time periods of five years interval from 2001 to 2100 for temperature and precipitation. These anomalies were then downscaled to 25 m-resolution grids using a bilinear interpolation algorithm before being added to the environmental predictors of the study area at present climatic conditions.

Niche based-modelling

For each species, Generalised Linear Models (GLM), Generalised Additive Models (GAM), Generalised Boosted Model (GBM) and Random Forests (RF) were calibrated on the presence / absence dataset. The predictive ability of each modeling technique was evaluated with a pseudo-independant internal evaluation by running a 10-fold cross-validation evaluation (van Houwelingen and Le Cessie 1990) on the training data set. During the cross-validation procedure, the original prevalence of the species presences and absences in the data set was maintained in each fold. Comparisons of predicted (probability scale) and observed (presence–absence) values were based on the area under the curve (AUC) of a receiver-operating characteristic plot (ROC; Fielding and Bell 1997)

Parametric Generalised Linear Models (GLM) is so far the most common modeling technique used in ecology (e.g. Hill et al. 1999, Bakkenes et al. 2002). Semi-parametric Generalised Additive Models are now increasingly used (Yee and Mitchell 1991, Frescino et al. 2001). Random forests (Breiman 2001) have been rarely used for predicting plant species distribution in ecology (Prasad et al. 2006). Generalised Boosted Model (Friedman et al. 2000) appeared to be the most performing technique when tested by Elith *et al.* (2006).

Spatial projections

The modeling technique that produced on average the highest AUC values over the entire species dataset was selected for spatial projections. AUC values between modeling techniques were compared with Wilcoxon matched pair signed rank test. Projected distributions of presence-absence were derived from probability values of niche-based models by maximizing the percentage of presence and absence correctly predicted under current condition (Thuiller 2003).

Projected distributions of presence-absence were derived from the probability values of the niche-based models by maximizing the percentage of presence and absence correctly predicted under current conditions (Pearce and Ferrier 2000, Thuiller 2003).

A mask based on forests, roads, urbanized areas and rivers was then applied to filter the model predictions, as a way to avoid nonsense projections at the local scale. Each species was further associated with one or several of three categories of open vegetation: grassland, rock and scree. The latter step was based on the empirical

observations from the sampling plots. The rule used was that a species was never predicted in a category where it was never observed in the calibration dataset.

Dynamic simulation of seed dispersal using the MigClim software

MIGCLIM (Engler and Guisan in revision) is a cellular automaton that allows simulating the transient dispersal of plant species in the landscape, while implementing a climate change scenario at the same time. Different parameters such as seed dispersal distance and kernel shape, generation time, long distance dispersal events or barriers to dispersal can be specified and customized to best fit each species' dispersal behaviour.

Because of the high number of species modelled in this study (i.e. 284), very accurate data on each species' dispersal behaviour could not be obtained. Indeed, such data does not exist for the very large majority, if not all, of our species. To overcome this lack of accurate data, we classified our species into seven categories and assigned them dispersal parameters based on the "dispersal types" defined by Vittoz *et al.* (in revision). This procedure allowed us to derive estimated dispersal distances for each species with a sufficient precision to achieve significant improvement over dispersal-ignoring models.

Dispersal parameters assigned to each category are given in table 2. The decrease in colonization probability with distance from source cell was chosen as to reflect a negative exponential seed dispersal shadow. This is a common seed dispersal kernel shape (Willson 1993), although many other exist (e.g. Clark et al. 1999, Nathan and Muller-Landau 2000, Greene et al. 2004). Long distance dispersal (LDD) events that allow the plant to disperse randomly at distances larger than its dispersal kernel (SDD; short distance dispersal, 99% of the seeds) were also assigned to each category. The frequency with which an occupied 25 m cell could generate LDD events was set to 0.01. Further parameters added to increase realness in species dispersal behaviour were generation time (i.e. time required for a pixel from when it gets colonized until it can act itself as source pixel) and landscape fragmentation. Landscape fragmentation parameters are composed of "barrier" cells that impede dispersal through them (but not LDD) and "filter" cells that represent permanent unsuitable surfaces but that are not barriers to dispersal.

Table 2 Parameters used in the MIGCLIM model for each dispersal category (1 to 7).

	Dispersal categories						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Pixel resolution	5 m			12.5 m	25 m		
Dispersal event frequency *	5 years		1 year				
Dispersal distance	1 m	4 m	20 m	40 m	100 m	200 m	1 km
LDD dispersal distance	5 m	20 m	100 m	200 m	1 km	2 km	10 km
LDD frequency **	0.002	0.002	0.0004	0.0025	0.01	0.01	0.01
Generation Time	Parameter set individually for each species.						
Barriers	Parameter set individually for each species (e.g. forest)						
Filter	Urban areas, glaciers, grassland, rock or scree						
Number of species	52	10	55	19	24	83	41
Number of pixels	30'000'000			4'800'000	1'200'000		

For each species, two dispersal scenarios, without (SDD) and with LDD were run for a 100 year period (2001-2100) under each of the four IPCC climate change scenarios (i.e. A1, A2, B1, B2). Simulations without dispersal (ND) and with an unlimited dispersal (UD) were additionally run to compare the two realistic simulations (SDD and LDD) with the two extremes cases. The initial distribution of a species was defined as its potential habitat under current climatic conditions (i.e. 1961-1990 average) and change in habitat suitability reflecting climate change was implemented every fifth year. Dispersal was simulated each year, except for categories 1 and 2, because they have too short dispersal distance to be modeled even on a 5 m resolution grid. Consequently, the dispersal distances of these categories were multiplied by five and dispersal was simulated every fifth year. Finally, because they involve a part of randomness, each simulation was repeated five times and averaged.

Altitudinal structure of the study area

To help interpreting our simulations, we represented the altitudinal distribution of 25m-resolution pixels used for the spatial projection and their distribution across the vegetation belts. These limits were defined as proposed by Aeschiman & Burdet (1994) for the Western Swiss Alps. We also calculated the surfaces of three land cover classes used for spatial projections (grasslands, rock and scree) within each vegetation zone.

Species extinctions by 2100 and evolution of extinctions across the 21st century

We fist calculated the percentage of species extinct in 2100 at the end of the four simulations (ND, SDD, LDD, UD) under the four climate change scenarios (A1, A2, B1 and B2). The percentage of species extinct at the end of simulations was calculated with three different thresholds. The 100%, 90% and 80% thresholds assumed that a given species should loose respectively 100%, 90% or 80% of its initial surface to be considered as extinct or become committed to extinction (i.e. due to insufficient remaining population size). The percentage of initial surface lost at the end of simulations was calculated for each species.

We then calculated the percentage of species extinct for each five years time period of the simulation from 2005 to 2100, assuming that a given species should loose 100% of its initial surface to be extinct.

Relationship between species extinction risk and species traits

We assessed whether species extinction risk estimated with the four dispersal simulations could be related to species traits. First, we tested the relationship between the percentage of initial surface of habitat lost by species in 2100 and the initial surface of habitat to estimate the vulnerability of species by their extent of initial distribution.

Second, we developed an index of migration capacity (IMC; Eq. 1) to numerically assess the relationship between the dispersal capacity of species and the extinction risk. This index is based on the ratio of the log of maximum dispersal distance for 99% of the seeds (see Table 2) of a given species divided by the time for this species before producing seeds after establishment.

$$IMC = \frac{\text{Log}(\text{Max Dispersal}_{99\% \text{ of seeds}})}{\text{Time to produce seeds}} \qquad \textit{(Eq.1)}$$

The relationship between this index of migration capacity and the percentage of initial surface remaining for each species at the end of SDD and LDD simulations was assessed with linear regression models under the four climate change scenarios.

Third, we derived an index reflecting the elevation optimum of the 284 species. This elevation index was calculated by a weighted average of the altitudinal distribution of each species, as obtained from the Flora Alpina atlas (Aeschimann et al. 2004). This index was then split into five categories, from 1 for strictly collinean (lowest elevation) species to 5 for strictly nival (highest elevation) species. We then calculated the percentage of species extinction observed in each of these categories under the four climate change scenarios and the four dispersal simulations. We also

tested the relationship between the percentage of initial habitat lost by species and the elevation optimum with linear regression models.

Results

Testing models performance

Boxplots of evaluation with AUC for the four modelling techniques showed that GLM performed slightly better than GAM, GBM and RF (Fig. 2). In addition, Wilcoxon matched pair signed rank test showed that AUC values are significantly higher for GLM (*P* values < 0.001) when compared to GAM, GBM and RF. GLMs were then used for the following analysis.

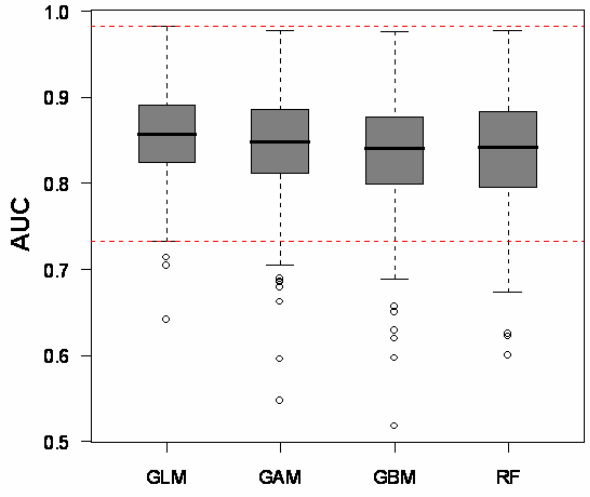


Figure 2 Boxplots of the metric based on the area under the curve of a receiver-operating characteristic plot (AUC) on the evaluation dataset (10-folds cross-validation) for GLM, GAM, GBM and RF. Wilcoxon matched pair signed rank test showed that AUC values are significantly higher for GLM (*P* values < 0.001) when compared to GAM, GBM and RF. Relative position of GLM model whiskers are the two vertical lines.

Altitudinal structure of the study area

The distribution of surface used for spatial prediction and simulation along elevation showed those available surfaces were mostly distributed at the lowest elevation of the collinean zone and in the subalpine zone (Fig. 3a). Grassland are more important in the subalpine zone and scree and rock surfaces are mainly distributed in the subalpine and alpine zone (Fig. 3b). For all land cover categories, only few surfaces are available in the nival zone.

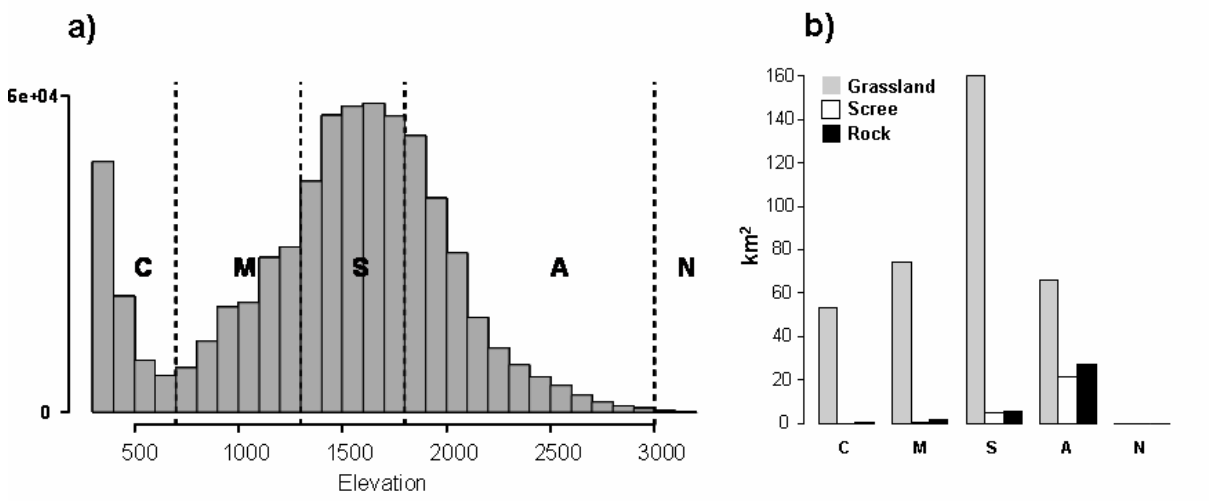


Figure 3 Frequency distribution of elevation (with a 100 m interval) of 25m-pixels used for the spatial prediction (a) and surface occupied by land cover classes within the study area (b). The distribution of the vegetation zones (C=colinean; M=moutainous; S=subalpine; A=alpine; N=nival) are indicated on (a) and (b).

Species extinctions by 2100 and evolution of extinctions across the 21st century

The percentage of species extinction with the two realistic dispersal simulations under the most pessimistic scenario A1 reached 28% (80 species for small distance dispersal (SDD) and 81 for long distance dispersal (LDD; Fig. 4a and Table 3). These values were closer to species extinctions with universal dispersal simulations (UD: 15.8%) than simulations without dispersal (ND: 53.2%). The lowest values of species extinction for realistic dispersal were observed for the B1 scenario (4.6% for SDD and LDDM; Fig. 4c), again closer to UD (14.8%) than ND (0.8%).

Table 3 Number and percentage of species extinct in 2100 under the four climate change scenarios and the four different dispersal strategies (ND = No dispersal; UD = unlimited dispersal; LDD = Long distance dispersal; SDD = Short distance dispersal) and for the three thresholds of extinction.

		100% treshold		90% treshold		80% treshold	
		# species	% species extinct	# species	% species extinct	# species	% species extinct
A1	ND	151	53.2	253	89.1	258	90.8
	SDD	81	28.5	214	75.4	220	77.5
	LDD	80	28.2	214	75.4	220	77.5
	UD	45	15.8	193	68.0	213	75.0
A2	ND	108	38.0	243	85.6	249	87.7
	SDD	51	18.0	187	65.8	202	71.1
	LDD	49	17.3	186	65.5	202	71.1
	UD	20	7.0	158	55.6	182	64.1
B1	ND	42	14.8	195	68.7	209	73.6
	SDD	13	4.6	131	46.1	153	53.9
	LDD	13	4.6	130	45.8	153	53.9
	UD	1	0.4	71	25.0	106	37.3
B2	ND	56	19.7	207	72.9	221	77.8
	SDD	23	8.1	143	50.4	166	58.5
	LDD	23	8.1	143	50.4	165	58.1
	UD	6	2.1	89	31.3	123	43.3

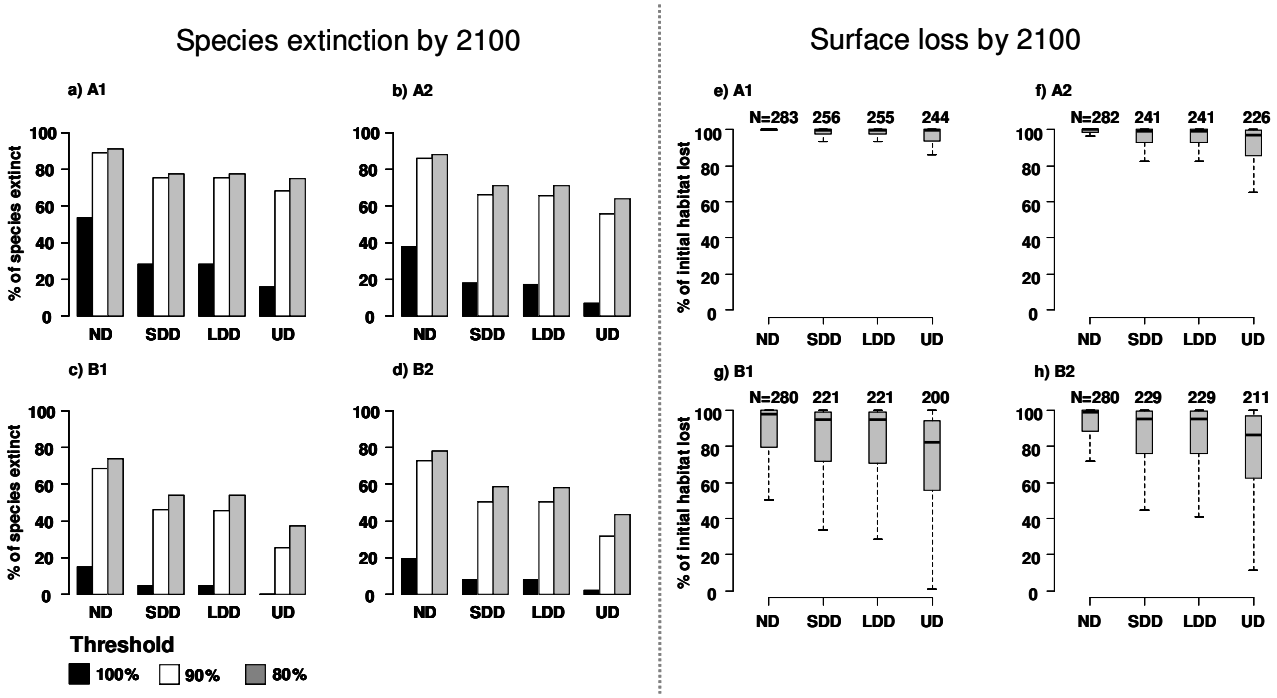


Figure 4 Percentage of species extinct in 2100 under the four climate change scenarios and the four different dispersal strategies (a-d; ND = No dispersal; UD = unlimited dispersal; LDD = Long distance dispersal; SDD = Short distance dispersal). Percentage of species extinct at the end of simulations was calculated with three different thresholds. The 100%, 90% and 80% thresholds assumed that a given species should loose 100%, 90% or 80% of its initial surface to be considered as extinct. Boxplots of percentage of initial habitat lost in 2100 (e-h) under the four climate change scenarios and the four simulations. The number of species losing surface in 2100 is indicated on the top of each boxplot.

When considering the two most realistic dispersal simulations (SDD and LDD), at least 53.9% of the species lost 80% of their initial surface (B1 scenario; Table 3; Fig. 4e-h). This percentage remained important for the most optimistic dispersal simulation (B1 with UD; 37.3%).

Evolution of species extinction across time showed that extinction started to increase only since 2040 for ND, SDD, LDD and UD under A1 scenario (Fig. 5a). Under the three other scenarios (Fig. 5b-d), increases of extinctions were predicted after 2080 for SDD, LDD and UD. The evolution of extinctions of SDD and LDD simulations was closer to UD than ND.

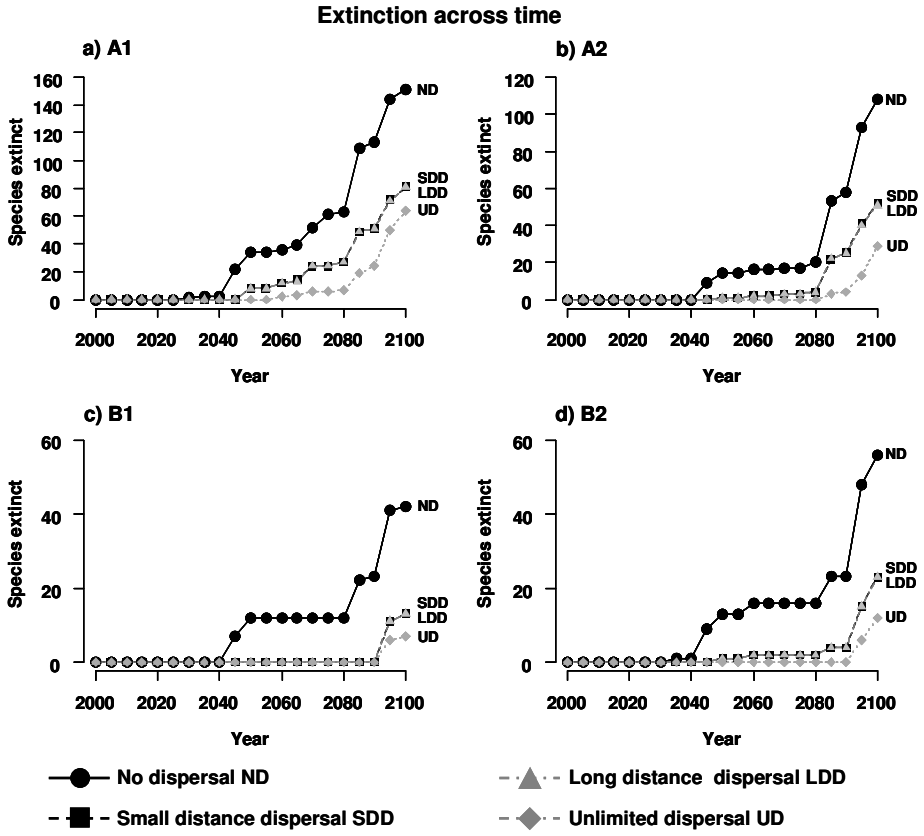


Figure 5 Evolution of the number of species extinct across time under the four climate change scenarios and the four dispersal simulations.

The percentage of species with a decreasing surface of their initial habitat in 2100 was important. At least 70.1% of species may loose a part or the totality of their habitat at the end of the 21st century (B1 with UD simulation) and 77.5% with SDD and LDD simulations (Table 4).

Table 4 Percentage of species with a decrease of their initial habitat in 2100 under the four climate change scenarios (ND = No dispersal; UD = unlimited dispersal; LDD = Long distance dispersal; SDD = Short distance dispersal).

	%species with habitat surface deacreasing in 2100			
	A1	**A2**	**B1**	**B2**
ND	99.6	99.3	98.6	98.6
SDD	89.8	84.5	77.5	80.3
LDD	89.4	84.5	77.5	80.3
UD	85.6	79.2	70.1	73.9

Relationship between habitat lost in 2100 and other factors

We observed a positive relationship between initial habitat lost and initial surface under all scenarios when including realistic distance dispersal in the simulations (Fig. 6 b,c,f,g,j,k,n and o).

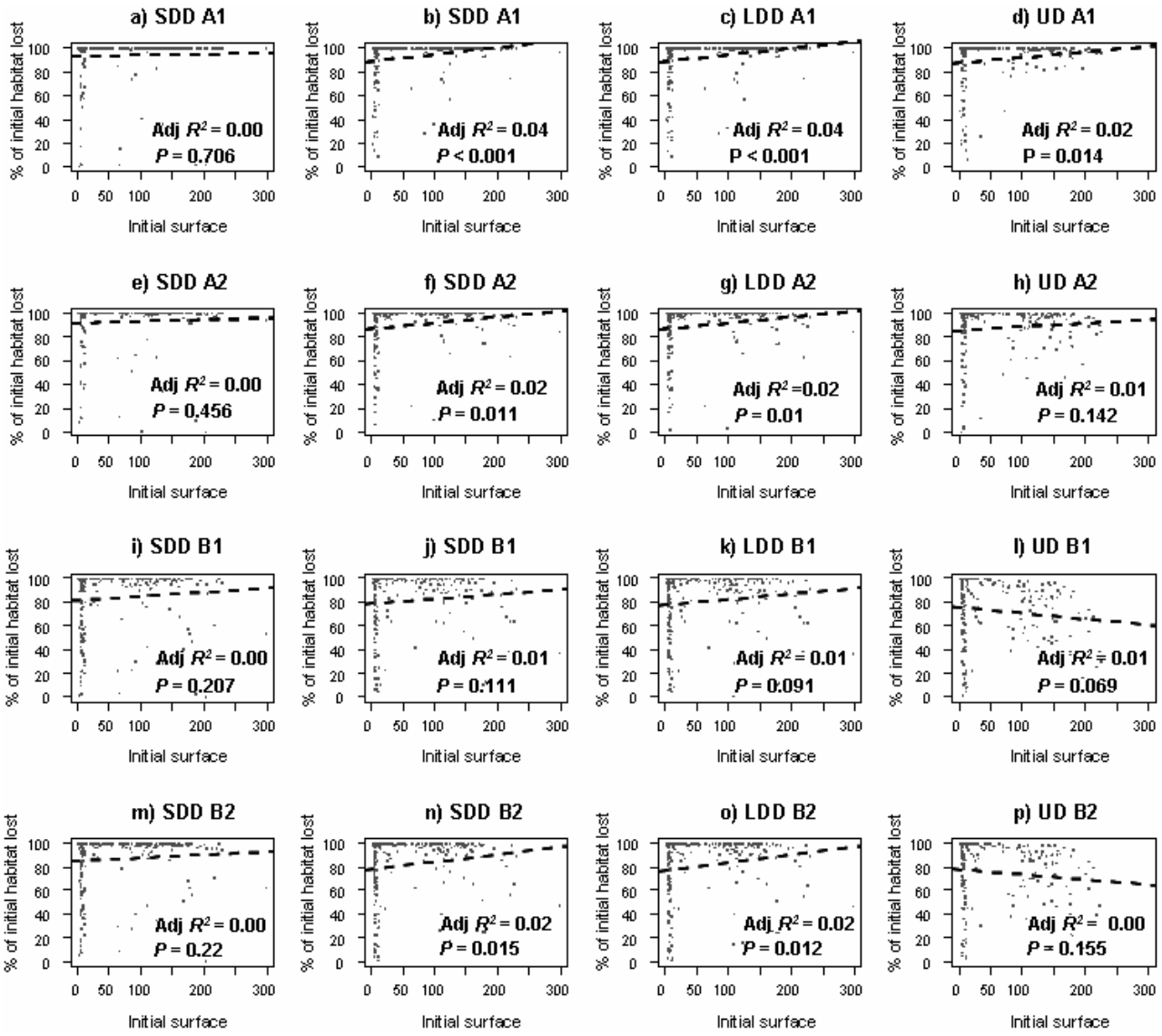


Figure 6 Relationship between the percentage of initial surface of habitat lost by species in 2100 and initial surface of habitat of under the four climate change scenarios (a-d: A1; e-h: A2; i-l: B1; m-p: B2) and the four dispersal simulations.

Overall, there was a significant negative relationship between the migration index and the percentage of initial habitat lost. Species with a high capacity of dispersal were less affected by surface loss (Fig. 7).

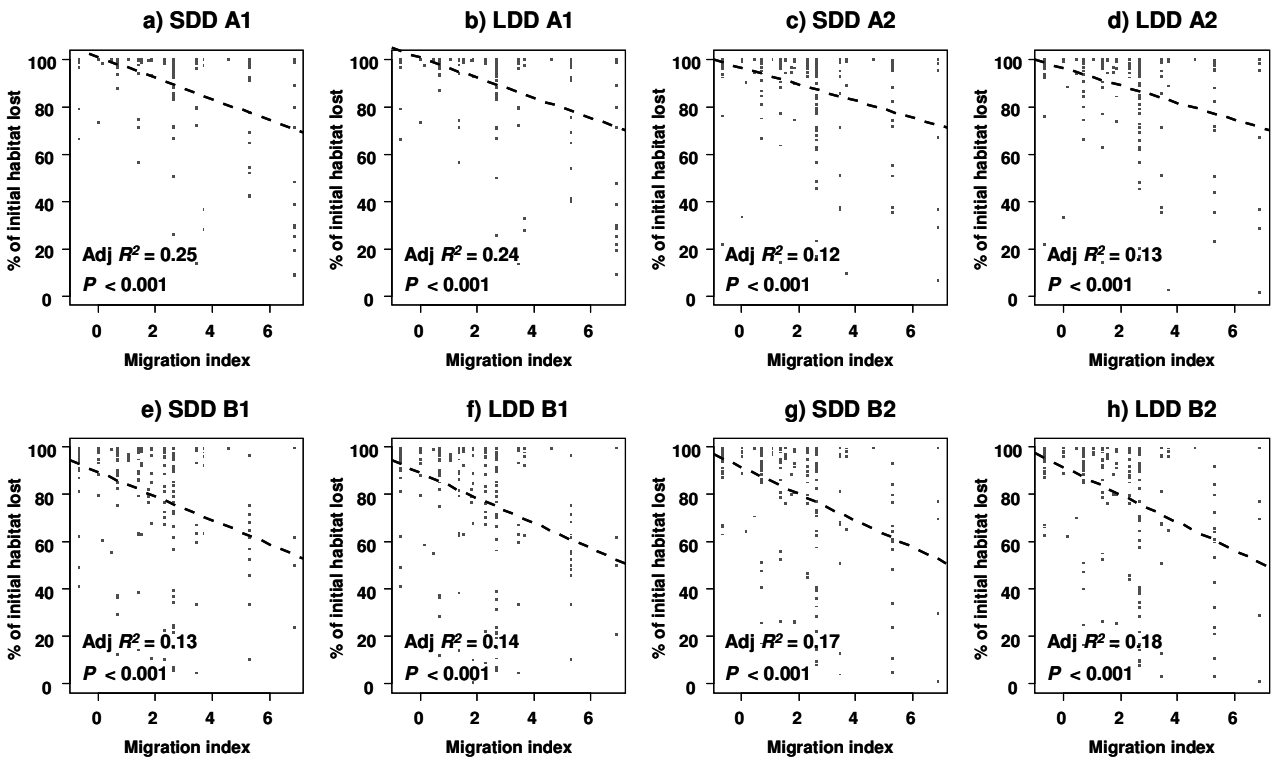


Figure 7 Relationship between the percentage of initial habitat lost in 2100 and the migration index under the four scenarios and with the four dispersal simulations.

The relationships between the migration index and the loss of potential initial habitat were similar between LDD and SDD simulations and were all significant under the four scenarios. The agreement was the highest for A1. However, an important variation along the linear trend was observed for all the scenarios and dispersal simulations.

Extinction per elevation optimum

The distribution of extinction within elevation optimum of species showed that alpine species will be the more affected under the four scenarios and the four dispersal simulations (Fig. 8). In addition, there was a positive relationship between the percentage of initial habitat lost and the elevation optimum of species (Fig. 9).

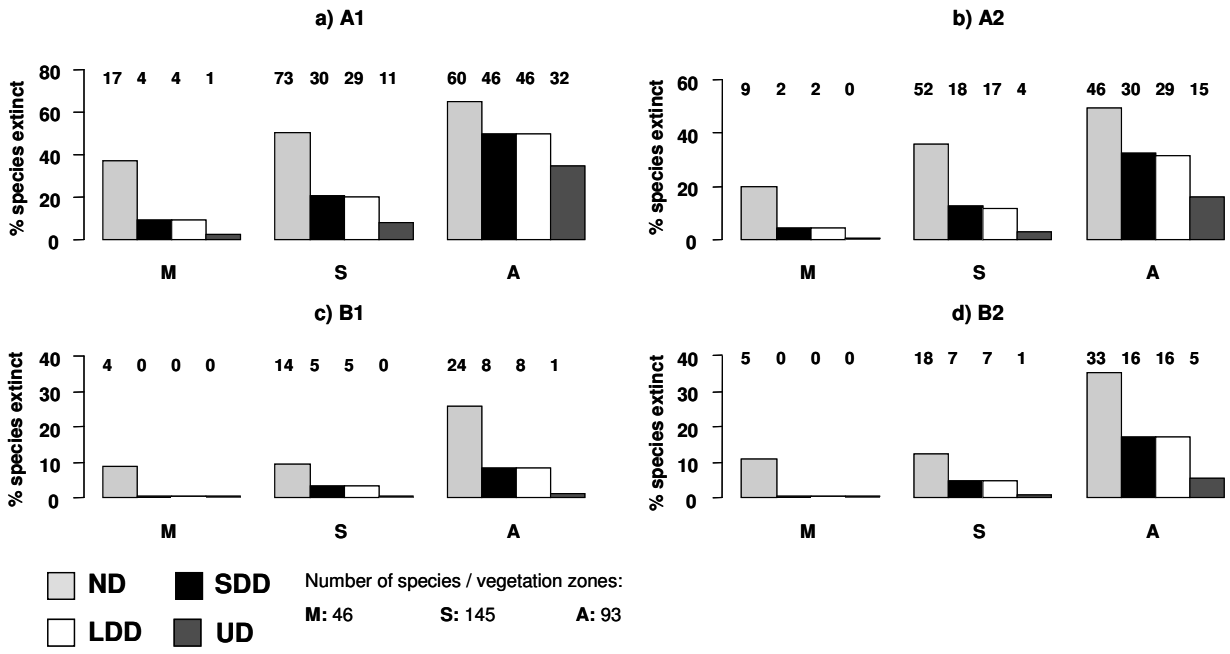


Figure 8 Percentage of species extinction per optimum of elevation classified by the vegetation zones (M=moutainous; S=subalpine; A=alpine).

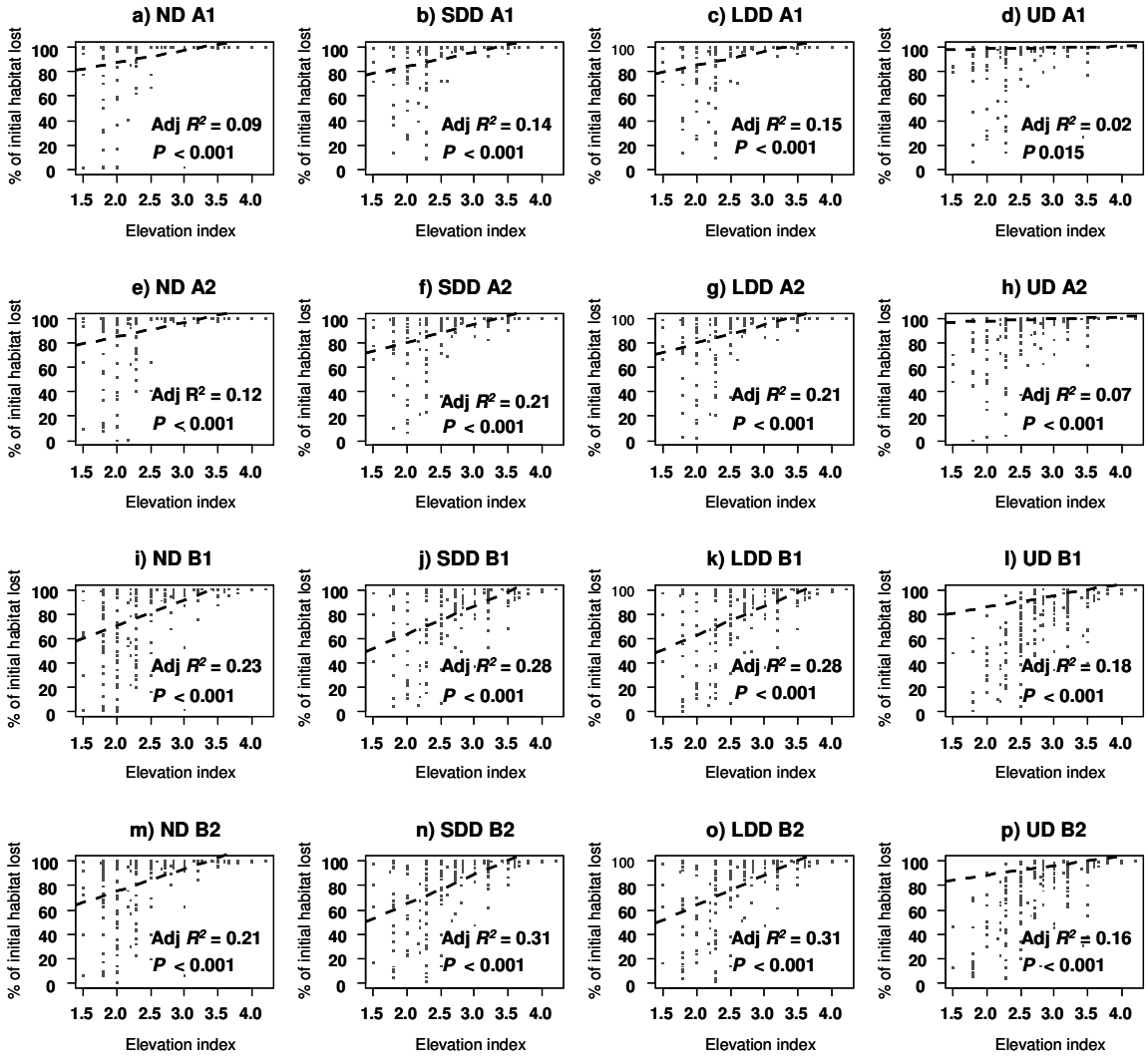


Figure 9 Relationship under the four scenarios and dispersal simulations between the percentage of initial habitat lost in 2100 and the elevation index.

Discussion

In this study we assessed the extinctions of mountain plants under four climate change scenarios for the 21st century, based on realistic seed dispersal distances. We showed that predictions of species extinctions produced with realistic distances of seed dispersal were closer to predictions with unlimited dispersal. The two most realistic dispersal simulations showed that 53.9% of the species considered in our study may loose more than 80% of their initial habitats at the end of the century with the most optimistic climate change scenario (B1; Table 3). When evaluating the evolution of extinctions across the 21st century, we showed that extinction rate would increase considerably after 2040 whatever the dispersal distance chosen under the A1 scenarios. However, the large increase in extinctions would only appear after 2080 with short distance (SDD) and long distance (LDD) dispersal simulations under the three other climate change scenarios (A2, B1, B2; Fig. 5). A significant trend was observed between species' extinction and the size of their initial distribution for SDD and LDD (Fig. 6) and we showed a significant relationship between the distribution of a species at the end of simulation and its index of migration capacity (Fig. 7). Overall, alpine plant species with dispersal distance smaller than 20 meters were shown to be particularly sensitive to climate change, with a likely complete extinction predicted.

Predictions of extinctions from short and long distance dispersal simulations were closer to simulations with unlimited dispersal than simulations with no dispersal (Fig. 4). This demonstrates that predictions of previous studies using static distribution models to assess the impact climate change in mountain regions (Guisan and Theurillat 2000, Dirnböck et al. 2003) give trustful results but very probably underestimate real extinction by using unlimited dispersal.

In a recent large-scale SDM study, Thuiller *et al.* (2005) forecasted that certain mountains regions of Europe (e.g. mid-elevation Alps) could be disproportionably sensitive to climate change, with up to 60% of species loss. This study was conducted using 1350 species from the Atlas Florae Europaea (AFE; Lahti and Lampinen 1999) over the entire extent of Europe at a coarse spatial resolution (model fitting at 50 km x 50 km; spatial projections at 16 km x 16 km), using the A1 scenario derived from the HadCM3 general circulation model (GCM). These predictions of extinction are four times higher than those predicted by unlimited simulations in our study (15.8%), two times higher than simulations with realistic dispersal (28.2%), but close to the simulation without dispersal (53.2%). However, caution should also be taken here when comparing these different results. Our study area did not encompass the entire surface covered by pixels of European projections and the lists of species used at the European scale were not representative of our study area.

In this study, we were able to identify critical periods of the 21st when alpine plants may experience massive extinctions (Fig. 5). To our knowledge, no study assessed so far the evolution of species extinctions across time, with such a large number of

species and at such spatial accuracy. Projections of SDM under current conditions are usually shifted to future conditions that represent averages for several decades (Bakkenes et al. 2002, Dirnböck et al. 2003, Thomas et al. 2004, Thuiller et al. 2005). Our results revealed that important species loss will start to occur since 2040 for the most pessimistic A1 scenario and only since 2080 for the three other scenarios (A2, B1, B2). This temporal estimate is a novel and important information for monitoring and conservation planning. As to conservation issues, the identification of areas that will be most affected by extinction in 2040 and in 2080 could be identified and preserved from direct land use impact, to limit future losses. As to monitoring issues, our findings show that no extinctions could be detected under ongoing climate change during the next 40 to 80 years, due to the inertia of species that will still have the possibility to move upwards to higher elevation.

Within our study area, alpine species will be more affected by climate change, whatever the scenario (Fig. 8). However, the shape of the study area, the elevation range and the land cover type at high elevation considerably influence the prediction of extinction. This is because elevation of mountains in the Western Alps is limited to 3000 meters, just reaching the limit of the nival zone. It has been shown by Randin *et al.* (in prep) (chapter 6) that less extinctions are predicted in mountains with high elevation range and for species with important surfaces at the nival zone. On the other hand, local extinctions should be considered as potentially important if new refuges are far from areas of extinctions. Moreover, in the worst climate change scenario (A1), some montane and subalpine species are committed to extinction as well (Fig. 8). Mountain summits is certainly not the cause of their extinction, but they are unable to migrate quick enough to keep in pace with climate change.

Surprisingly, extinction rate was very similar in our results with or without considering long-distance dispersal (Fig. 4). Available data for LDD are almost inexistent for the considered species and LDD uses generally unexpected dispersal vectors (Higgins et al. 2003) and thus is unpredictable. We retained a 1% frequency for LDD (Higgins et al. 2003, Soons and Ozinga 2005) and a distance 5 times longer than habitual dispersal. However, these values were not sufficient to make a significant difference with short-dispersal distance. Higher migration rates than what we used for LDD were sometimes recorded but they belong to invading species (e.g. Pysek and Hulme 2005) with broad ecological requirements or to post-glacial migration with possible unknown factors (Ronce 2001, Pearson 2006). Takahashi and Kamitani (2004) who observed migration through a forested area recorded values of the same order as ours: quick migration for zoochorous seeds (category 6 in table 2) but very low displacement for winged seeds or seeds without particularities (categories 1-3). If we consider that the retained values in our models for LDD are correct, the absence of difference with SDD can be explained by the fragmented landscape: in montane and subalpine level grasslands are separated by larges forests (Fig. 1) and in alpine level, the high diversity of topography and disturbances situations separates strongly sites with similar ecological conditions. This means that very long dispersal distances or more frequent LDD events will be

necessary to plants to be able to keep in pace with climate change. Future researches will be necessary to better understand what are the obstacles to migration or what is the necessary migration rate to overpass them.

Conclusion

We showed with realistic dispersal simulations that, under the most optimistic climate change scenario (B1), 53.9% of the species may loose 80% or more of their initial habitat by the end of the 21st century. As a major result, simulations based on realistic dispersal distances for >300 species yielded results on average closer to unlimited dispersal simulations than to the no-dispersal scenario. Secondarily, long distance dispersal simulations yielded similar results to short distance simulations, but this may depend on the way both estimates were obtained from literature. The simulations over the entire 21st century also showed that, due to the possibility of a large number of species shifting their distribution to higher elevation, the bulk of extinctions would mainly occur between 2040 and 2100. However, landuse changes during the first part of the century may contribute to decrease or increase this estimated rates of extinction.

Acknowledgements

This study was supported by the Centre de Conservation de la Nature (SFFN) of the Canton de Vaud, by the Swiss National Science Foundation (SNF Grant No 3100A0-110000) and the European Commission (FP6 projects ECOCHANGE and MACIS). We thank all the people who helped with the field work in the Swiss Western Alps, particularly Stéfanie Maire, Dario Martinoni, Séverine Wydler, Chantal Peverelli, Valentine Hof, Grégoire Vuissoz, Lorenzo De Stefani, Roxane Milleret, Pierre Bermane Favrod-Coune, Daniel Cherix, Anny and Willy Berra. C.R. and A.G. greatly benefited from interactions with colleagues within the Swiss National Centre for Competence in Research (NCCR)'Plant survival in natural and agricultural ecosystems' (http:// www.unine.ch/nccr).

References

Aeschimann, D., and H. M. Burdet. 1994. Flore de la Suisse et des territoires limitrophes. Le nouveau Binz, 2 edition. Le Griffon, Neuchâtel.

Aeschimann, D., and C. Heitz. 1996. Index synonymique de la flore de Suisse et territoires limitrophes. CRSF, Genève.

Bakkenes, M., J. R. M. Alkemade, F. Ihle, R. Leemans, and J. B. Latour. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. Global Change Biology **8**:390-407.

Beniston, M. 1994. Mountain Environment in Changing Climates, Routledge edition. Routledge, London.

Beniston, M., D. G. Fox, S. Adhikary, R. Andressen, A. Guisan , J. I. Holten, J. Innes, J. Maitima, M. F. Price, and L. Tessier. 1996. Impacts of climate change on mountain regions. Chapter 5. *in* IPCC, editor. Intergovernmental panel on climate change. Second assessment report. Cambridge University Press, Cambridge.

Berger, A. 2002. Global warming 2001. Journal de Physique IV (Proceedings) **12**:19-26.

Bouët, M. 1985. Climat et météorologie de la Suisse romande, Payot edition, Lausanne.

Bradshaw, A. D., and T. McNeilly. 1991. Evolutionary response to Global Climatic Change. Annals of Botany **68**:5-14.

Braun-Blanquet, J. 1957. Ein Jahrhunder Florenwandel am Piz Linard (3414 m). Bull. Jard. Bot. Etat Brux. **volume jubilaire Walter Robyns**:221-232.

Breiman, L. 2001. Random Forests. Mach Learn **45**:5-32.

Broennimann, O., W. Thuiller, G. Hughes, G. F. Midgley, J. R. M. Alkemade, and A. Guisan. 2006. Do geographic distribution, niche properties and lifeforms explains vulnerability of plants to climate changes? . Global Change Biology **12**:1-15.

Carey, P. D. 1996. DISPERSE: A cellular automaton for predicting the distribution of species in a changed climate. Global Ecology and Biogeography Letters **5**:217-226.

Carson, D. J. 1999. Climate Modelling: achievements and prospects. Quarterly Journal of the Royal Meteorological Society **125**:1:27.

Clark, J. S. 1998. Why trees migrate so fast: confronting theory with dispersal biology and the paleorecords. The American Naturalist **152**:204-224.

Clark, J. S., M. Silman, R. Kem, E. Macklin, and J. H. Islamber. 1999. Seed Dispersal Near and Far: Patterns across temperate and tropical forests. Ecology **80**:1475-1494.

Diaz, H. F., M. Grosjean, and L. Graumlich. 2003. Climate variability and change in high elevation regions: Past, present and future. Climatic Change **59**:1-4.

Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography **30**:401-417.

Dullinger, S., T. Dirnbock, and G. Grabherr. 2004. Modelling climate change-driven treeline shifts: relative effects of temperature increase, dispersal and invasibility. Journal of Ecology **92**:241-252.

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

Engler, R., and A. Guisan. in revision. MigClim: Predicting plant distribution and dispersal in a changing climate. Global Change Biology.

Etterson, J. R., and R. G. Shaw. 2001. Constraint to Adaptive Evolution in Response to Global Warming. Science **294**:151-154.

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

Folland, C. K., N. A. Rayner, S. J. Brown, T. M. Smith, S. S. P. Shen, D. E. Parker, I. Macadam, P. D. Jones, R. N. Jones, N. Nicholls, and D. M. H. Sexton. 2001. Global temperature change and its uncertainties since 1861. Geophysical Research Letters **28**:2621-2624.

Frescino, T. S., T. C. Edwards, and G. G. Moisen. 2001. Modeling spatially explicit forest structural attributes using Generalized Additive Models. Journal of Vegetation Science **12**:15-26.

Friedman, J. H., T. Hastie, and R. Tibshirani. 2000. Additive Logistic Regression: a Statistical View of Boosting. Annals of Statistics **28**:337-374.

Grabherr, G., M. Gottfried, A. Gruber, and H. Pauli. 1995. Patterns and current changes in alpine plant diversity. Pages 167-181 *in* A. Chapin and C. Körner, editors. Arctic and alpine biodiversity: patterns, causes and ecosystem consequences. Springer, Heidelberg.

Grabherr, G., M. Gottfried, and H. Pauli. 1994. Climate Effects on Mountain Plants. Nature **369**:448-448.

Greene, D. F., C. D. Canham, K. D. Coates, and P. T. Lepage. 2004. An evaluation of alternative dispersal functions for trees. Journal of Ecology **92**:758-766.

Guisan, A. 2005. Niche-based models as tools to assess climate change impact on the distribution and diversity of plants in mountain reserves. *in* Proceedings of the 2nd and 3rd GLOCHAMORE workshops in L'Aquila (Italy) and Granada (Spain). UNESCO-MRI, Paris, France.

Guisan, A., J. I. Holten, R. Spichiger, and L. Tessier, editors. 1995. Potential Ecological Impacts of Climate Change in the Alps and Fennoscandian Mountains. Ed. Conserv. Jard. Bot. Geneva, Geneva, Switzerland.

Guisan, A., and J.-P. Theurillat. 2000. Assessing alpine plant vulnerability to climate change: A modeling perspective. Integrated Assessment **1**:307-320.

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

Higgins, S. I., R. Nathan, and M. L. Cain. 2003. Are long-distance dispersal events in plants usually caused by nonstandard means of dispersal ? Ecology **84**:1945-1956.

Hill, J. K., C. D. Thomas, and B. Huntley. 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. Proc. R. Soc. Lon. B **266**:1197-1206.

Hofer, H. R. 1992. Veränderungen in der Vegetation von 14 Gipfeln der Berninagebietes zwischen 1905 und 1985. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich **58**:39-54.

Houghton, R. A., J. E. Hobbie, J. M. Melillo, B. Moore, B. J. Peterson, G. R. Shaver, and G. M. Woodwell. 1983. Changes in the Carbon Content of Terrestrial Biota and Soils between 1860 and 1980 - a Net Release of Co2 to the Atmosphere. Ecological Monographs **53**:235-262.

Houghton, R. A., and G. M. Woodwell. 1989. Global Climatic-Change. Scientific American **260**:36-44.

Huntley, B., P. M. Berry, W. Cramer, and A. P. McDonald. 1995. Modelling present and potential future ranges of some European higher plants using climate response surfaces. Journal of Biogeography **22**:967-1001.

Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecological Monographs **68**:465-485.

Iverson, L. R., M. W. Schwartz, and A. M. Prasad. 2004. How fast and far might tree species migrate in the eastern United States due to climate change? Global Ecology and Biogeography **13**:209-219.

Klanderud, K., and H. J. B. Birks. 2003. Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. The Holocene **13**:1-6.

Körner, C. 2003. Alpine plant life. 2nd Ed. Springer.

Lahti, T., and R. Lampinen. 1999. From dot maps to bitmaps - Atlas Flora Europaeae goes digital. Acta Botanica Fennica **162**.

Lehmann, A., J. R. Leathwick, and J. M. Overton. 2002. Assessing New Zealand fern diversity from spatial predictions of species assemblages. Biodiversity and Conservation **11**:2217-2238.

Malcolm, J. R., A. Markham, R. P. Neilson, and M. Garaci. 2002. Estimated migration rates under scenarios of global climate change. Journal of Biogeography **29**:835-849.

Nakicenovic, N., and R. Swart, editors. 2000. Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Nathan, R., and H. C. Muller-Landau. 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends in Ecology & Evolution **15**:278-285.

Nogués-Bravo, D., M. Araújo, M. P. Errea, and J. P. Martinez-Rica. 2006. Exposure of global mountain systems to climate warming during the 21st century. GLOBAL ENVIRONMENTAL CHANGE.

Parmesan, C. 1996. Climate and species' range. Nature **382**:765-766.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature **421**:37-42.

Pauli, H., M. Gottfried, and G. Grabbherr. 1996. Effects of climate change on mountain ecosystems. Upward shifting of alpine plants. World Resource Review **8**:382-390.

Pearce, J., and S. Ferrier. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. Ecological Modelling **128**:127-147.

Pearson, R. G. 2006. Climate change and the migration capacity of species. Trends in Ecology & Evolution **21**:111-113.

Pearson, R. G., T. P. Dawson, P. M. Berry, and P. A. Harrison. 2002. SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species. Ecological Modelling **154**:289-300.

Pitelka, L. F., R. H. Gardner, J. Ash, S. Berry, H. Gitay, I. R. Noble, A. Saunders, R. H. W. Bradshaw, L. Brubaker, J. S. Clark, M. B. Davis, S. Sugita, J. M. Dyer, R. Hengeveld, G. Hope, B. Huntley, G. A. King, S. Lavorel, R. N. Mack, G. P. Malanson, M. McGlone, I. C. Prentice, and M. Rejmanek. 1997. Plant migration and climate change. American Scientist **85**:464-473.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems **9**:181-199.

Pysek, P., and P. E. Hulme. 2005. Spatio-temporal dynamics of plant invasions: Linking pattern to process. Ecoscience **12**:302-315.

Randin, C. F., W. Thuiller, R. Engler, P. Vittoz, S. Poulsen, M. Zappa, P. B. Pearman, and A. Guisan. in prep. Extinction risk of plant species in mountain regions: contrasted views from continental and local scales.

Rebetez, M. 2002. La Suisse se réchauffe: effet de serre et changement climatique, Presse polytechniques et unversitaires romandes edition, Lausanne.

Ronce, O. 2001. Understanding plant dispersal and migration. Trends in Ecology & Evolution **16**:663-664.

Root, T. L., J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds. 2003. Fingerprints of global warming on wild animals and plants. Nature **421**:57-60.

Shugart, H. H., N. H. F. French, E. S. Kasischke, J. J. Slawski, C. W. Dull, R. A. Shuchman, and J. Mwangi. 2001. Detection of vegetation change using reconnaissance imagery. Global Change Biology **7**:247-252.

Soons, M. B., and W. A. Ozinga. 2005. How important is long-distance seed dispersal for the regional survival of plant species ? Diversity and Distributions **11**:165-172.

Takahashi, K., and T. Kamitani. 2004. Effect of dispersal capacity on forest plant migration at a landscape scale. Journal of Ecology **92**:778-785.

Theurillat, J.-P., F. Felber, P. Geissler, J.-M. Gobat, M. Fierz, A. Fischlin, P. Küpfer, A. Schlüssel, C. Velluti, and G.-F. Zhao. 1998. Sensitivity of plant and soil ecosystems of the Alps to climate change. Pages 225-308 *in* P. Cebon, U. Dahinden, H. C. Davies, D. Imboden, and C. C. Jäger, editors. Views from the Alps. Regional perspectives on climate change. MIT Press, London.

Theurillat, J. P., and A. Guisan. 2001. Potential impact of climate change on vegetation in the European Alps: A review. Climatic Change **50**:77-109.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. Van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from climate change. Nature **427**:145-147.

Thuiller, W. 2003. BIOMOD - Optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biology **9**:1353-1362.
Thuiller, W., M. B. Araujo, R. G. Pearson, R. J. Whittaker, L. Brotons, and S. Lavorel. 2004. Uncertainty in predictions of extinction risk. Nature **430**:34.
Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the United States of America **102**:8245-8250.
van Houwelingen, J. C., and S. Le Cessie. 1990. Predictive value of statistical models. Statistics in Medicine **9**:1303-1325.
Vittoz, P., and R. Engler. in revision. Seed dispersal distances: a typological system for data analyses and models. Botanica Helvetica.
Vittoz, P., S. Jutzeler, and A. Guisan. 2006. Flore alpine et réchauffement climatique: observation de trois sommets valaisans à travers le 20ème siècle. Bulletin de la Murithienne **123/2005**:49-59.
Walther, G.-R., S. Beißner, and C. A. Burga. 2005. Trends in the upward shift of alpine plants. Journal of Vegetation Science **16**:541-548.
Walther, G. R. 2003. Plants in a warmer world. Perspectives in Plant Ecology, Evolution and Systematics **6**:169-185.
Willson, M. F. 1993. Dispersal Mode, Seed Shadows, and Colonization Patterns. Vegetatio **108**:261-280.
Woodwell, G. M., F. T. Mackenzie, R. A. Houghton, M. Apps, E. Gorham, and E. Davidson. 1998. Biotic feedbacks in the warming of the earth. Climatic Change **40**:495-518.
Yee, T. W., and N. D. Mitchell. 1991. Generalized additive models in plant ecology. J. Veg Sci. **2**:587-602.
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.

Appendix A

Species used for simulations

Achillea atrata	*Anthriscus sylvestris*	*Homogyne alpina*
Aconitum napellus aggr.	*Cirsium acaule*	*Holcus lanatus*
Achillea millefolium	*Cirsium eriophorum s.str.*	*Pritzelago alpina s.str.*
Adenostyles glabra	*Cirsium palustre*	*Hypericum maculatum s.str.*
Agrostis alpina	*Cirsium spinosissimum*	*Hypericum perforatum s.str.*
Agrostis rupestris	*Colchicum autumnale*	*Hypochaeris radicata*
Agrostis schraderiana	*Coeloglossum viride*	*Juncus effusus*
Agrostis stolonifera	*Crocus albiflorus*	*Juniperus communis subsp. nana*
Agrostis capillaris	*Crepis aurea*	*Knautia arvensis*
Ajuga reptans	*Crepis biennis*	*Knautia dipsacifolia s.str.*
Alchemilla conjuncta aggr.	*Crepis pyrenaica*	*Laserpitium latifolium*
Alchemilla coriacea aggr.	*Cruciata laevipes*	*Lathyrus pratensis*
Alchemilla xanthochlora aggr.	*Cynosurus cristatus*	*Leontodon autumnalis*
Alchemilla decumbens aggr.	*Cystopteris fragilis*	*Leontodon hispidus s.l.*
Alchemilla glabra aggr.	*Daucus carota*	*Linum alpinum*
Androsace chamaejasme	*Dactylis glomerata*	*Linum catharticum*
Anemone narcissiflora	*Daphne mezereum*	*Ligusticum mutellina*
Anthoxanthum odoratum aggr.	*Deschampsia cespitosa*	*Linaria alpina s.str.*
Anthyllis vulneraria s.l.	*Doronicum grandiflorum*	*Lotus alpinus*
Aposeris foetida	*Draba aizoides*	*Lotus corniculatus*
Arabis alpina s.str.	*Dryas octopetala*	*Lolium perenne*
Arabis caerulea	*Elyna myosuroides*	*Luzula campestris*
Arrhenatherum elatius	*Epilobium anagallidifolium*	*Luzula multiflora*
Arabis hirsuta	*Erigeron uniflorus*	*Luzula alpinopilosa*
Arnica montana	*Euphorbia cyparissias*	*Luzula sylvatica*
Astrantia major	*Euphrasia hirtella*	*Medicago lupulina*
Asplenium viride	*Euphrasia minima*	*Myosotis alpestris*
Athamanta cretensis	*Euphrasia rostkoviana s.str.*	*Myosotis arvensis*
Bartsia alpina	*Euphrasia salisburgensis*	*Myosotis decumbens*
Aster bellidiastrum	*Festuca arundinacea s.l.*	*Nardus stricta*
Stachys officinalis s.str.	*Festuca pratensis s.l.*	*Nigritella rhellicani*
Bellis perennis	*Festuca quadriflora*	*Traunsteinera globosa*
Biscutella laevigata	*Festuca rubra aggr.*	*Dactylorhiza maculata*
Botrychium lunaria	*Festuca ovina aggr.*	*Oxytropis jacquinii*
Bromus erectus s.str.	*Festuca violacea aggr.*	*Parnassia palustris*
Briza media	*Fragaria vesca*	*Pedicularis verticillata*
Bromus hordeaceus	*Galium album*	*Pedicularis foliosa*
Brachypodium pinnatum	*Galium anisophyllon*	*Petasites paradoxus*
Campanula barbata	*Galium megalospermum*	*Phleum rhaeticum*
Carum carvi	*Galium pumilum*	*Phleum alpinum aggr.*
Campanula cochleariifolia	*Galeopsis tetrahit*	*Phleum hirsutum*
Carduus defloratus s.str.	*Gentiana bavarica*	*Phyteuma orbiculare*
Carex ferruginea	*Gentiana clusii*	*Phleum pratense*
Carex flacca	*Gentiana acaulis*	*Phyteuma spicatum*
Carex nigra	*Gentiana lutea*	*Picea abies*
Carex montana	*Geum montanum*	*Pimpinella major*
Carex ornithopoda	*Gentiana campestris s.str.*	*Pimpinella saxifraga aggr.*
Carex panicea	*Gentiana purpurea*	*Plantago alpina*
Cardamine pratensis	*Geum rivale*	*Plantago atrata s.str.*
Campanula rhomboidalis	*Geranium sylvaticum*	*Plantago lanceolata*
Campanula rotundifolia	*Geum urbanum*	*Plantago major s.str.*
Carex pallescens	*Gentiana verna*	*Plantago media*
Campanula scheuchzeri	*Globularia cordifolia*	*Poa alpina*
Carex sempervirens	*Glechoma hederacea s.str.*	*Poa minor*
Carex sylvatica	*Globularia nudicaulis*	*Poa supina*
Carlina acaulis subsp. caulescens	*Gymnadenia conopsea*	*Potentilla aurea*
Calamagrostis varia	*Gypsophila repens*	*Polygonum bistorta*
Carex caryophyllea	*Helianthemum nummularium s.l.*	*Poa cenisia*
Cerastium fontanum subsp. vulgare	*Hedysarum hedysaroides*	*Polygala chamaebuxus*
Centaurea jacea s.str.	*Helictotrichon pubescens*	*Potentilla crantzii*
Cerastium latifolium	*Heracleum sphondylium s.l.*	*Potentilla erecta*
Centaurea montana	*Helictotrichon versicolor*	*Polygala alpestris*
Centaurea scabiosa s.l.	*Hieracium lactucella*	*Polystichum lonchitis*
Cerastium arvense s.l.	*Hieracium bifidum aggr.*	*Poa pratensis*
Chaerophyllum hirsutum aggr.	*Hippocrepis comosa*	*Potentilla sterilis*
Chaerophyllum aureum	*Hieracium villosum aggr.*	*Poa trivialis s.str.*
Leucanthemum vulgare aggr.	*Hieracium murorum aggr.*	*Polygonum viviparum*

Species used for simulations

Polygala vulgaris s.str.
Primula auricula
Primula elatior s.str.
Prunella grandiflora
Primula veris s.l.
Prunella vulgaris
Pulsatilla alpina s.str.
Ranunculus aconitifolius
Ranunculus alpestris
Ranunculus bulbosus
Ranunculus acris s.l.
Ranunculus nemorosus aggr.
Ranunculus montanus aggr.
Ranunculus repens
Rhinanthus alectorolophus
Rhododendron ferrugineum
Rhinanthus minor
Rosa pendulina
Rumex acetosa
Rumex alpinus
Rumex alpestris
Rumex crispus
Rubus idaeus
Saxifraga paniculata
Acinos alpinus
Saxifraga androsacea
Sagina saginoides
Salix retusa
Sanguisorba minor s.str.
Saxifraga moschata s.l.
Saxifraga oppositifolia
Salix reticulata
Clinopodium vulgare
Clinopodium vulgare
Saxifraga aizoides
Scabiosa lucida
Sedum atratum
Sesleria caerulea
Senecio doronicum
Selaginella selaginoides
Silene acaulis
Silene dioica
Silene vulgaris s.l.
Soldanella alpina
Solidago virgaurea s.l.
Stellaria graminea
Taraxacum alpinum aggr.
Taraxacum officinale aggr.
Thymus praecox subsp. polytrichus
Thesium alpinum
Thymus pulegioides s.str.
Tofieldia calyculata
Trifolium badium
Trollius europaeus
Trisetum flavescens
Trifolium repens s.str.
Trifolium medium
Trifolium montanum
Tragopogon pratensis subsp. orientalis
Trifolium pratense s.str.
Trifolium thalii
Tussilago farfara
Urtica dioica
Vaccinium gaultherioides
Valeriana montana
Vaccinium myrtillus
Valeriana officinalis aggr.
Valeriana tripteris
Vaccinium vitis-idaea
Veronica alpina
Veronica aphylla
Veronica arvensis
Veronica chamaedrys
Veronica officinalis
Veronica persica
Veratrum album subsp. lobelianum
Veronica serpyllifolia s.l.
Viola biflora
Viola calcarata
Vicia cracca s.str.
Vicia sativa s.l.
Viola hirta
Vicia sepium

Synthesis & Perspectives

In my study, I developed and tested new predictor variables for species distribution models (SDM), specific to current and future geographic projections of plant species in a mountain system, using the Western Swiss Alps as model region. Since meso- and micro-topography are relevant to explain geographic patterns of plant species in mountain environments, I assessed the effect of scale on predictor variables and geographic projections of SDM. I also developed a methodological framework of space-for-time evaluation to test the robustness of SDM when projected in a future changing climate. Finally, I used a cellular automaton to run dynamic simulations of plant migration under climate change in a mountain landscape, including realistic distance of seed dispersal. Results of future projections for the 21st century were also discussed in perspective of vegetation changes monitored during the 20th century.

Climate change impacts on mountain vegetation

Overall, I showed in this study that, based on the most severe A1 climate change scenario and realistic simulations of plant dispersal (chapter 7), species extinctions in the Western Swiss Alps could affect nearly one third (28.5%) of the 284 species modeled by 2100. With the less severe B1 scenario, only 4.6% of species are predicted to become extinct. However, even with B1, 54% (153 species) may still loose more than 80% of their initial surface. Thus, caution should be taken when considering only complete extinctions (100% loss of surface) for decision planning (e.g. agriculture or reserve design see Araujo and Williams 2000, Araujo et al. 2004). Indeed, species with small surfaces of remaining habitats predicted are more likely to react very quickly and become extinct if experiencing further extreme climatic events such as long episodes of dryness (Lüdi and Zoller 1949) or heat waves (Jolly et al. 2005). Hence, the proportion of remaining surface should be estimated for dominant or relevant species (i.e. for agriculture) to allow decision makers to better anticipate extinctions at the local to regional scale. Such assessment could be particularly relevant in the Western Alps where some middle elevation subregions (e.g. Montreux, Bex or Pays d'Enhaut) are intensively exploited by mountain agriculture or may become federal or regional natural regional preserves (PNR Muverans and Pays d'Enhaut).

Static and dynamic simulations also showed that plants from different vegetation zones may react differently to climate warming. In the Swiss Western Alps, alpine plants should be more threatened than subalpine plants (chapter 7). Other results (chapter 6) suggested that mountain systems with higher elevation range, such as the Zermatt region, may have less alpine plants threaten by climate change.

My study also highlighted the importance of regional assessments of climate change impact on mountain vegetation. Indeed, predictions at the continental scale may fail to predict local refugees or local extinctions, as well as loss of connectivity between local populations (chapter 6). On the other hand, migrations of low-elevation species to higher altitude may be difficult to predict at the local scale.

SDM in mountain systems under climate change

I also demonstrated that the improvement of SDM by land use and process-based geomorphological disturbance variables was species-specific (chapter 2 and 3). Species-specific response was also observed when testing the effect of scale at which topographic variables optimized the predictive power of SDM (chapter 4). This response will also depend on the position of plant species distribution along the elevation gradient because human land use has a lower impact at high elevation whereas geomorphic processes, snow cover or microtopography are more important at higher altitude Therefore, migration of plant species under climate change may be influenced differentially along the elevation gradient. The potential responses of plant species to the different drivers of mountain system tested in my study are summarized in figure 1.

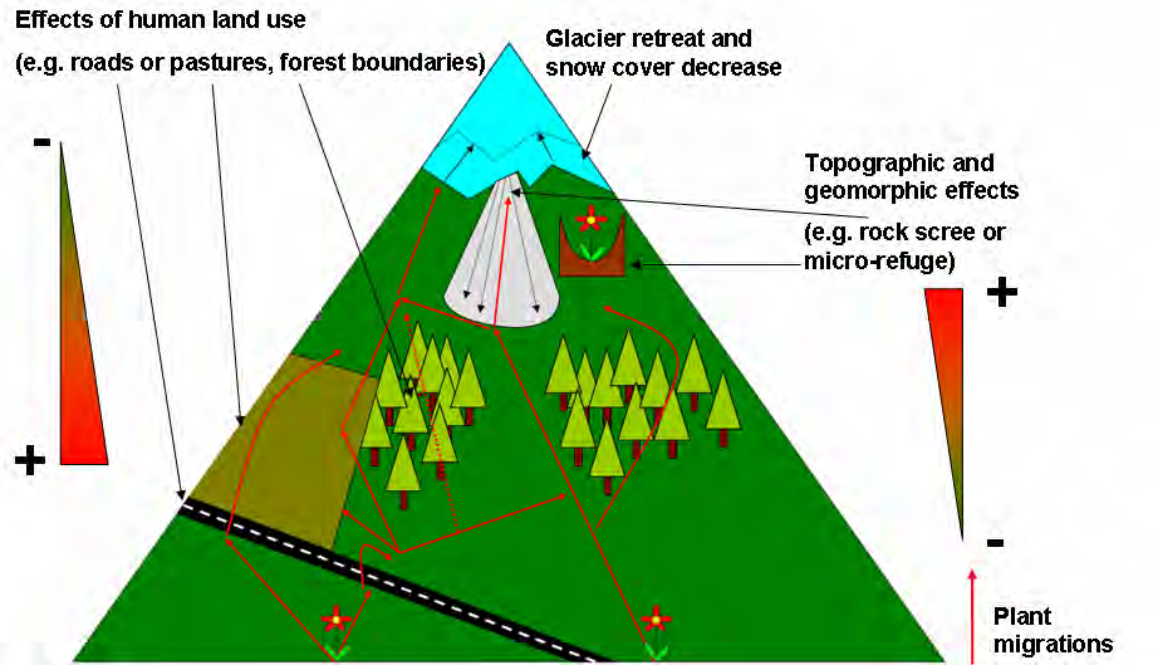


Figure 1 Potential response of plant species to different drivers of mountain system along the elevation gradient and under climate change.

I demonstrated with dynamic simulations that the dispersal capacity of species will significantly influence the ability of species to track suitable habitat or to go extinct under climate change (chapter 7). I also showed that human land use, geomorphic processes and snow cover may act as facilitators or inhibitors of species movements, and consequently act potentially as barriers for seed establishment and/or dispersal (Fig.1). These factors could thus either accelerate or decelerate the rate of extinction of sensitive species, but are themselves also likely to be influenced by climate change. Thus, socio-economic and climate simulations (Rutherford et al. submitted) should be additionally used to determine future rates of landscape change or

changes in landscape connectivity, and increase the quality of future predictions of species distribution. In addition, model of geomorphologic perturbations (Tarboton 1997, Meissl 2001, Noetzli et al. 2006) as well as physical model of snow cover (Gurtz et al. 2003) or glacier retreat (Cook et al. 2003) should help to determine the spatial connectivity (Fig.1) of alpine landscape for geographic projections of plant species distribution in the future. In my study, I incorporated static landuse, geomorphologic and snow cover layers (chapters 2 and 3). However, these variables cannot be easily projected in a future changed climate. Therefore, dynamic models of land cover change should be coupled with SDM to increase the quality of predictions. As shown in figure 2, open vegetation units - and particularly grasslands - currently occupy an important surface of the subalpine zone in the Alps of the Canton de Vaud. Urban density is still low in the collinean and mountain zones. Thus, this current land cover distribution is likely to change with pasture abandonment and urban development, leading to an increasingly fragmented landscape.

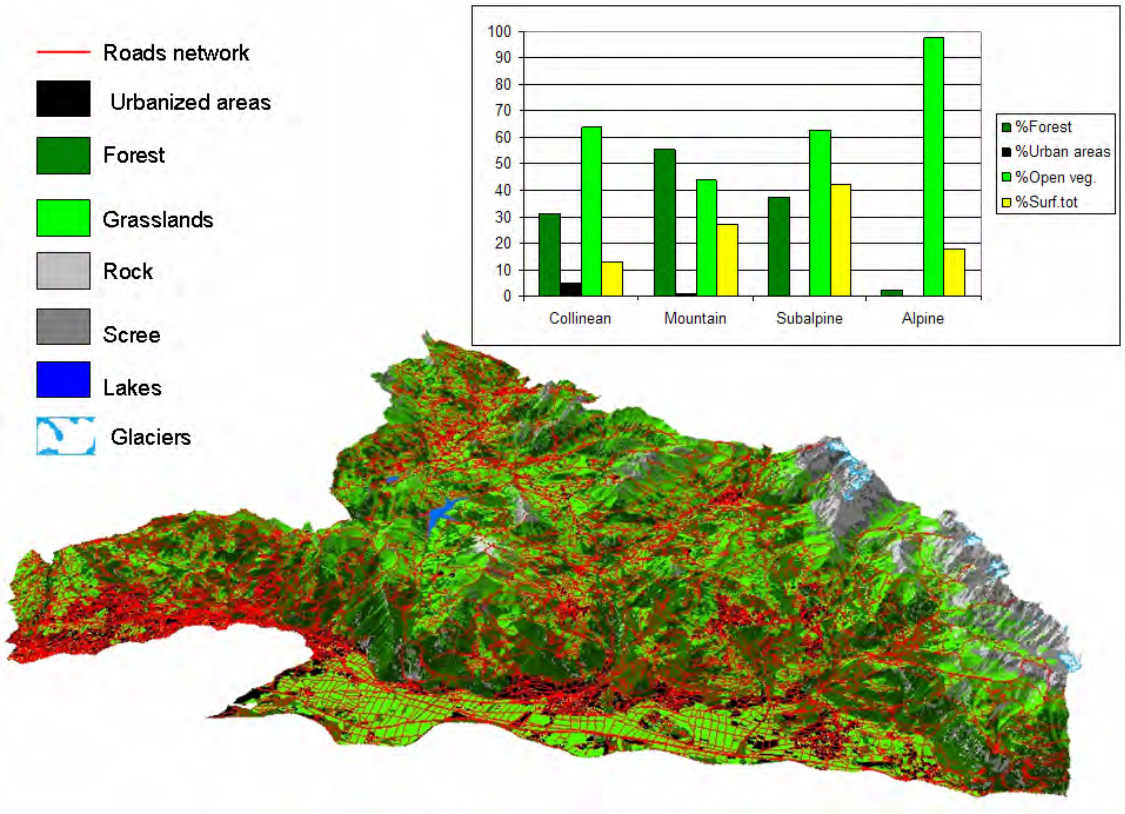


Figure 2 Actual geographic distribution of the land cover in the Western Swiss Alps of the state of Vaud.

When testing the predictive power of topographic variables derived at various scales in a mountainous environment, I showed that each of these variables had a different optimum scale. The predictive power of slope was optimized for scales larger than 100m, likely reflecting hydromechanical perturbations, whereas aspect

was optimized for scale around 20 m, which we interpreted as a proxy measure for the amount of energy received by plants. The validity of these results is unfortunately limited to a maximum neighborhood radius of 2000 m. Nevertheless, several authors suggested that at higher elevation, the influence of topography, and in particular micro-topography, is stronger, because plant life becomes more dependant on decoupling from climate at higher altitude (Fig. 1; Erschbamer 1989, Körner 2003). In an unpublished study (Randin et al. in prep)., we demonstrated that high-elevation life forms (mostly cushions and dwarf shrubs) were more sensitive to variables derived at a 2 m-resolution than at a 25 m-resolutionModel fit and predictive power of models for cushion plants and shrubs at 2 m-resolution were higher than for models at 25m-resolution (Fig. 3), thus showing that high resolution variables better define the niche and distribution of these life forms. As a result, current simulations under future climate change scenarios at coarser resolution may underestimate potential microclimatic refugees available at high elevation for these species. Further studies should then evaluate the effect of high resolution on the predictive power of models for a large number of species and especially evaluate the difference in predicted area between models fitted with predictors derived from fine and coarse resolution DEM.

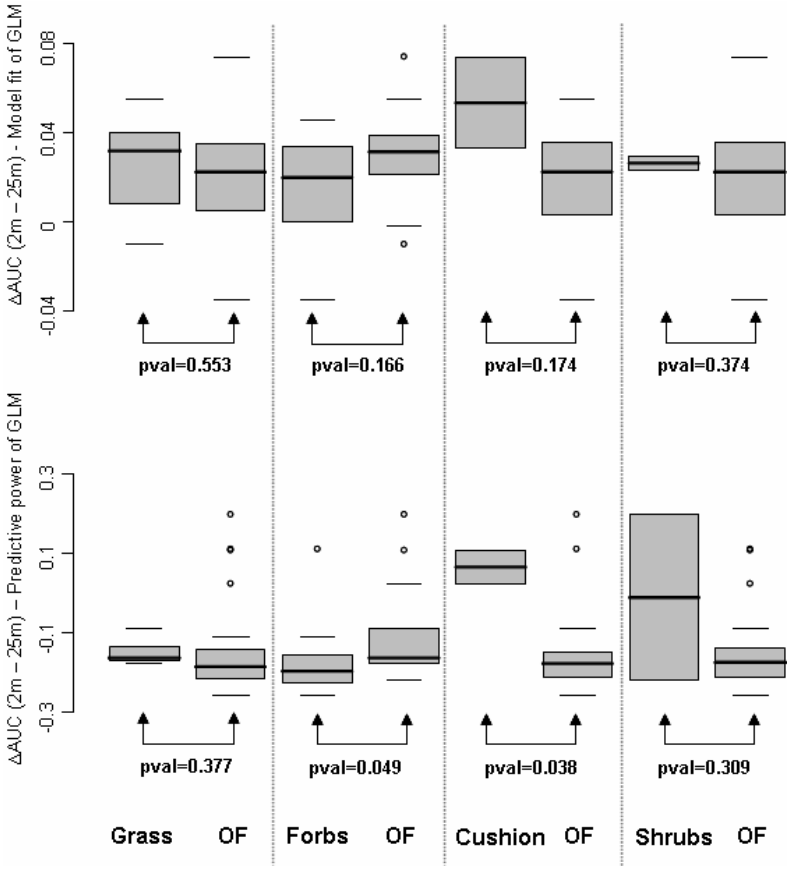


Figure 3 Boxplots of AUC of models at 2m – AUC of models at 25 m for a given life form versus all other forms (OF) for 31 species. P values of comparisons with Wilcoxon tests are indicated below the boxplots.

The monitoring study in the Vallon de Nant (chapter 1) showed that subalpine vegetation may indeed react more slowly than alpine and nival vegetation (Braun-Blanquet 1957, Hofer 1992, Grabherr et al. 1994, Pauli et al. 1996, Walther et al. 2005, Vittoz et al. 2006) under a warming climate. Reasons could be that inter-specific competition is stronger within the subalpine belt than at higher elevations, thus highlighting the diminishing importance of biotic interactions with increasing elevation and showing again the importance of improving SDM by the integration of mechanistic competition or species assembly rules at fine scale, as recommended by Pearson and Dawson (2003), Ferrier and Guisan (2006) or Guisan et al. (2006), and implemented by Silvertown et al. (1992) and Dullinger et al. (2005).

Hence, vegetation changes seems to be already under way in the Alps, as monitoring studies already suggest an increase of species on alpine and nival summits, with species from lower elevation. On the other hand, results from simulations including seed dispersal (chapter 7) forecasted no species extinctions before 2040 under all climate change scenarios. This may suggest that, due to topographic configuration, conic form and important elevation gradients, ongoing warming in mountain systems could first be in time-lagged but then respond perhaps more strongly to climate change than low elevation regions.

In Chapter 6 and 7, I focused on extinctions, but did not take species colonization from low altitude into account in the simulations. Shifts in the distribution of species from low to high elevations are difficult to predict at fine scale. As shown in figure 4, species from low altitude are likely to have truncated response curves (Van Horn 2002) in one of the main climatic gradient (e.g. temperature) if the extent of the study area is not sufficient to sample all possible conditions these species can tolerate. As proposed by Thuiller et al. (2004), truncated response curves can result in spurious predictions of species' future distributions after a major climate change. Indeed, species could be predicted to either migrate infinitely to higher elevations or stay predicted at the lowest elevation under climate change if no limit is set in one of the directions of the gradient (i.e. sigmoid curve "plateauing" after a threshold value; Fig. 4).

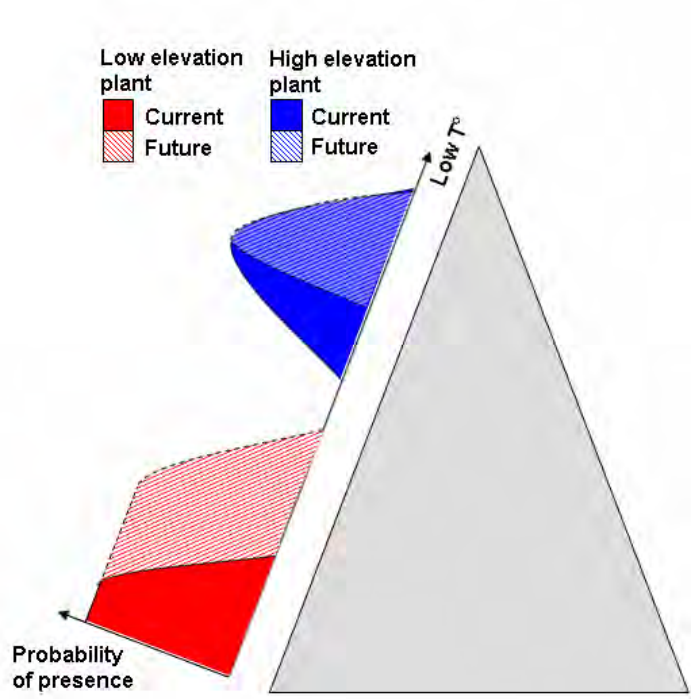


Figure 4 Potential effect of a truncated response curve of the temperature on a low elevation species when the complete distribution of a species was not sampled.

For further studies using SDM at regional scale where a limited environmental extent is suspected to occur, species with truncated response curves should be first identified with methods such as Huisman-Olff-Fresco models (HOF-models; Huisman et al. 1993, Oksanen and Minchin 2002). Second, the quality of the models for these species should be improved with data from larger scale when projecting models in space or time, for example, by combining data from European database with local data (e.g. Pearson and Dawson 2004).

Test of space-for-time transferability between Switzerland and Austria allowed demonstrating that successful transferability is highly variables among species. The next step should be to identify if some species characteristics can favor or limit transferability of models in space or time. One characteristics of particular interest is the genetic intra-specific variability. In an unpublished study (Hautier et al. in prep), we demonstrated that *Poa alpina* invested more resources in sexual reproduction at low altitude and more in clonal propagation at higher elevation, and that these resulted in ecotypic differentiations that seemed genetically fixed (suggesting limited gene flow between natural populations). This could also suggest that we may not be modelling exactly the same species along the elevation gradient.

General conclusion

Overall, my study confirms that mountain vegetation is particularly sensitive to climate change, with 54% of the species considered threatened of 80% of initial habitat loss with the less sever and more realistic simulations. Heavy extinctions of alpine plants may start already in 2040, but the latest in 2080.

This study also highlights the importance of fine scale studies to understand and model the main drivers of alpine systems that act on vegetation distribution. Some of these drivers, such as human land use, geomorphologic disturbances or snow cover may act synergetically with climate change as facilitator or inhibitor of plant movement in the landscape.

In a globalized world with fast environmental changes occurring, future SDM studies should incorporate more dynamic scenarios of land cover change and mechanistic dispersal processes, this way allowing deriving safer projections in a complex and increasingly fragmented landscape. Ongoing European research programmes such as the FP6 ECOCHANGE project will help improving SDM by assessing the effects of long distance dispersal and environmental niche variability or by developing high resolution variables and better land use change scenarios.

References

Araujo, M. B., M. Cabeza, W. Thuiller, L. Hannah, and P. H. Williams. 2004. Would climate change drive species out of reserves? an assessment of existing reserve-selection methods. Global Change Biology **10**.

Araujo, M. B., and P. H. Williams. 2000. Selecting areas for species persistence using occurrence data. Biological Conservation **96**:331-345.

Braun-Blanquet, J. 1957. Ein Jahrhunder Florenwandel am Piz Linard (3414 m). Bull. Jard. Bot. Etat Brux. **volume jubilaire Walter Robyns**:221-232.

Cook, K. H., X. S. Yang, C. M. Carter, and B. N. Belcher. 2003. A modeling system for studying climate controls on mountain glaciers with application to the Patagonian Icefields. Climatic Change **56**:339-367.

Dullinger, S., T. Dirnbock, R. Kock, E. Hochbichler, T. Englisch, N. Sauberer, and G. Grabherr. 2005. Interactions among tree-line conifers: differential effects of pine on spruce and larch. Journal of Ecology **93**:948-957.

Erschbamer, B. 1989. Vegetation on Avalanche Paths in the Alps. Vegetatio **80**:139-146.

Ferrier, S., and A. Guisan. 2006. Spatial modelling of biodiversity at the community level. Journal of Applied Ecology **43**:393-404.

Grabherr, G., M. Gottfried, and H. Pauli. 1994. Climate Effects on Mountain Plants. Nature **369**:448-448.

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

Gurtz, J., M. Zappa, K. Jasper, H. Lang, M. Verbunt, A. Badoux, and T. Vitvar. 2003. A comparative study in modelling runoff and its components in two mountainous catchments. Hydrological processes **17**:297-311.

Hautier, Y., C. F. Randin, J. Stöcklin, and A. Guisan in prep. Trade-off between sexual and clonal reproduction across altitude in Poa alpina L.

Hofer, H. R. 1992. Veränderungen in der Vegetation von 14 Gipfeln der Berninagebietes zwischen 1905 und 1985. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich **58**:39-54.

Huisman, J., H. Olff, and L. F. M. Fresco. 1993. A Hierarchical Set of Models for Species Response Analysis. Journal of Vegetation Science **4**:37-46.

Jolly, W. M., M. Dobbertin, N. E. Zimmermann, and M. Reichstein. 2005. Divergent vegetation growth responses to the 2003 heat wave in the Swiss Alps. Geophysical Research Letters **32**.

Körner, C. 2003. Alpine plant life. 2nd Ed. Springer.

Lüdi, W., and H. Zoller. 1949. Einige Beobachtungen über die Dürrenschäden des Sommers 1947 in der Nordschweiz und am Schweizerischen Jurarand. Ber. Geobot. Inst. ETH Stiftung Rübel Zürich:69-84.

Meissl, G. 2001. Modelling the runout distances of rockfalls using a geographic information system, Z.

Geomorphol. Geomorphology **125**:129-137.

Noetzli, J., C. Huggel, M. Hoelzle, and W. Haeberli. 2006. GIS-based modelling of rock-ice avalanches from Alpine permafrost areas. Computational Geosciences **10**:161-178.

Oksanen, J., and P. R. Minchin. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling **157**:119-129.

Pauli, H., M. Gottfried, and G. Grabbherr. 1996. Effects of climate change on mountain ecosystems. Upward shifting of alpine plants. World Resource Review **8**:382-390.

Pearson, D., and T. P. Dawson. 2004. Modelling species distribution in Britain: a hierarchical integration of climate and land-cover data. Ecography **27**:285-298.

Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? Global Ecology and Biogeography **12**:361-371.

Rutherford, G. N., N. E. Zimmermann, P. Bebi, and P. J. Edwards. submitted. Assessing land-use statistics to model land cover change in a mountainous landscape. Ecological Modelling.

Silvertown, J., S. Holtier, J. Johnson, and P. Dale. 1992. Cellular Automaton Models of Interspecific Competition for Space - the Effect of Pattern on Process. Journal of Ecology **80**:527-534.
Tarboton, D. G. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research **33**:309-319.
Thuiller, W., L. Brotons, M. B. Araujo, and S. Lavorel. 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. Pages 63-72 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.
Vittoz, P., S. Jutzeler, and A. Guisan. 2006. Flore alpine et réchauffement climatique: observation de trois sommets valaisans à travers le 20ème siècle. Bulletin de la Murithienne **123/2005**:49-59.
Walther, G.-R., S. Beißner, and C. A. Burga. 2005. Trends in the upward shift of alpine plants. Journal of Vegetation Science **16**:541-548.

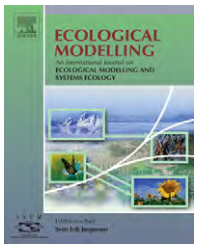
Annexes

Chapter 8

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel



Evaluating the ability of habitat suitability models to predict species presences

Alexandre H. Hirzel[a,b,*], *Gwenaëlle Le Lay*[a], *Véronique Helfer*[a], *Christophe Randin*[a], *Antoine Guisan*[a]

[a] *Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland*
[b] *Division of Conservation Biology, Institute of Zoology, University of Bern, CH-3012 Bern, Switzerland*

ARTICLE INFO

Article history:
Published on line 7 July 2006

Keywords:
Niche-based modelling
Model evaluation
Cross-validation
Generalised linear models (GLM)
Alpine plants
Swiss Alps

ABSTRACT

Models predicting species spatial distribution are increasingly applied to wildlife management issues, emphasising the need for reliable methods to evaluate the accuracy of their predictions. As many available datasets (e.g. museums, herbariums, atlas) do not provide reliable information about species absences, several presence-only based analyses have been developed. However, methods to evaluate the accuracy of their predictions are few and have never been validated. The aim of this paper is to compare existing and new presence-only evaluators to usual presence/absence measures.

We use a reliable, diverse, presence/absence dataset of 114 plant species to test how common presence/absence indices (Kappa, MaxKappa, AUC, adjusted D^2) compare to presence-only measures (AVI, CVI, Boyce index) for evaluating generalised linear models (GLM). Moreover we propose a new, threshold-independent evaluator, which we call "continuous Boyce index". All indices were implemented in the BIOMAPPER software.

We show that the presence-only evaluators are fairly correlated ($\rho > 0.7$) to the presence/absence ones. The Boyce indices are closer to AUC than to MaxKappa and are fairly insensitive to species prevalence. In addition, the Boyce indices provide predicted-to-expected ratio curves that offer further insights into the model quality: robustness, habitat suitability resolution and deviation from randomness. This information helps reclassifying predicted maps into meaningful habitat suitability classes. The continuous Boyce index is thus both a complement to usual evaluation of presence/absence models and a reliable measure of presence-only based predictions.

1. Introduction

Models predicting the spatial distribution of species (Boyce and McDonald, 1999; Guisan and Zimmermann, 2000; Manly et al., 2002; Pearce and Boyce, 2006) – sometimes called resource selection function or habitat suitability models – are currently gaining interest. As they often help both in understanding species niche requirements and predicting species potential distribution, their use has been especially promoted to tackle conservation issues, such as managing species distribution, assessing ecological impacts of various factors (e.g. pollution, climate change), risk of biological invasions or endangered species management (Scott et al., 2002; Guisan and Thuiller, 2005). These models statistically relate field observations to

* *Corresponding author at*: Laboratory of Conservation Biology, Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. Fax: +41 21 692 4105.
E-mail address: Alexandre.Hirzel@unil.ch (A.H. Hirzel).

0304-3800/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2006.05.017

a set of environmental variables, presumably reflecting some key factors of the niche, like climate, topography, geology or land-cover. They produce spatial predictions indicating the suitability of locations for a target species, community or biodiversity. Different types of modelling techniques are used to fit different types of biological information recorded at each sample site: (1) *presence-only*: occurrences of the target species are recorded; (2) *presence/absence*: each sample site is carefully monitored so as to assert with sufficient certainty whether the species is present or absent. With plants, for instance, it is commonly done by listing exhaustively all species present in each sample site. The reliability of absences depends on the species' characteristics (e.g. biology, behaviour, history) (Hirzel et al., 2001), their local abundance and ease of detection (Kéry, 2002), and the survey design (Mackenzie and Royle, 2005). More rarely, data record information about species' abundance or demography (e.g. growth rate, survival).

Although models based on presence-only and presence/absence data provide the same kind of predictions (e.g. habitat suitability scores), they generally cannot use the same technique. This is because presence-only methods cannot contrast their predictions with the characteristics of places where the species is absent. This partly explains why presence/absence methods have known a greater development. These differences, and the lack of absences, make comparison of the two model types difficult (Zaniewski et al., 2002).

Assessing the predictive power of a model is of paramount importance, both for theoretical and applied issues. However, while presence/absence models have received a lot of attention and many evaluators are available for them (Fielding and Bell, 1997), evaluation of presence-only models is lagging behind. There is therefore a crucial need for reliable presence-based evaluation measures, as well as an assessment of how they compare to the presence/absence measures.

The main problem of presence-only evaluation measures is the lack of absences to counterbalance the presences. It is thus difficult to discriminate a model predicting presence everywhere from a more contrasted model. Attempts to solve this problem have followed two main approaches: (1) a first approach is to generate pseudo-absences and then apply the standard presence/absence techniques (e.g. Zaniewski et al., 2002; Anderson et al., 2003). (2) A second approach is to assess how much the model predictions differ from random expectation (e.g. Boyce et al., 2002; Hirzel et al., 2002; Reutter et al., 2003). In this category, the index recently proposed by Boyce et al. (2002) offers new insights. We tested it thoroughly and derived a new evaluator from it, which does not depend on the choice of boundaries between habitat suitability classes. A third original approach, proposed by Ottaviani et al. (2004), is based on compositional analysis. However, it is restricted to cases where evaluation data are in the form of polygons or large mapping units (e.g. large grid cells in an atlas), and thus does not apply here.

In this paper, we present various presence-only evaluation measures. To validate them, we build 114 presence/absence models chosen for the reliability of their absences and evaluate them with presence-only and presence/absence evaluators. We test correspondence between them and discuss how the new "Boyce indices" can improve the interpretation and utilisation of habitat suitability models.

2. Materials and methods

We define a habitat suitability (HS) map as composed of cells (or pixels) whose quantitative values range from 0 to 1. These values indicate how close the local environment is to the species' optimal conditions, higher values standing for the most suitable areas. This map may result from any statistical analysis (Guisan and Zimmermann, 2000; Pearce and Boyce, 2006). The models' evaluation consists in quantifying how accurately the map is predicting the presence and absence of the species (Buckland and Elston, 1993; Manel et al., 2001), as given by a set of *evaluation points*. This set may consist either of verified presences and absences, or of verified presences only. Optimally, this data set should be completely independent from the data used to calibrate the model, e.g. collected on other areas (Randin et al., in press). However, due to time and money constraints, most studies have only one dataset and have to split it between a calibration and an evaluation sets. This is the method we use in this paper.

2.1. Evaluation indices

Most measures currently used in the literature are based on presence/absence information. Their first step generally consists of choosing a habitat suitability threshold (often 0.5) supposed to separate unsuitable areas (HS below threshold) where the species should be absent, from suitable areas (HS above threshold) where it should be present. From this Boolean map, one builds the *confusion matrix*, which counts how many presence and absence evaluation points occur in the suitable and unsuitable areas (Fig. 1). Many evaluators are based on this matrix (see Fielding and Bell, 1997), the commonest being the Kappa index K (Cohen, 1960):

$$K = \frac{N\sum x_{ii} - \sum x_{i\cdot}x_{\cdot i}}{N^2 - \sum x_{i\cdot}x_{\cdot i}} \quad (1)$$

where x and N are counts of evaluation points as defined in Fig. 1. K varies from −1 to 1, high values indicating a good agreement between prediction and data, and 0 corresponds to random agreement.

These methods depend strongly on the suitability cut-off threshold, which is often chosen arbitrarily. Alternatively, one may use threshold-independent methods, like K_{max} (or MaxKappa, Guisan et al., 1998) and area under the curve (AUC, Zweig and Campbell, 1993; Fielding and Bell, 1997). The K_{max} index is the highest Kappa obtained when varying the thresh-

		Observed		Margin sums
		Presence	Absence	
Predicted	Presence	x_{11}	x_{12}	$x_{1\bullet}$
	Absence	x_{21}	x_{22}	$x_{2\bullet}$
Margin sums		$x_{\bullet 1}$	$x_{\bullet 2}$	N

Fig. 1 – Contingency table of the model predictions against the actual observations. The x_{ij} represent counts of evaluation points, with $N = \Sigma x_{ij}$.

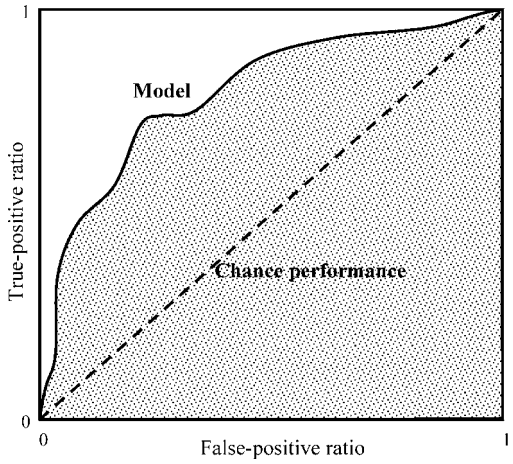


Fig. 2 – Curve of the true presence fraction ($=x_{11}/x_{.1}$) against the false presence fraction ($=x_{12}/x_{.2}$) computed for all possible cut-off points between 0 and 1. The AUC is the "area under the curve" and practically varies between 0.5 (not different from random expectation) and 1 (best model).

old from 0 to 1. The AUC is obtained by plotting, for each threshold in this range, the proportion of true positive $x_{11}/x_{.1}$ against the proportion of false positive $x_{12}/x_{.2}$ and by computing the area under the curve thus defined (Fig. 2). The AUC varies between 0 (worse-than-random model), 0.5 (random model) and 1 (best discriminating model).

When absence data are unreliable or unavailable, the model evaluation should be assessed for presences only. For this purpose, one possibility is to compare the model results to what would be expected from chance alone. Two simple evaluators are the absolute validation index (AVI) and contrast validation index (CVI) (Hirzel and Arlettaz, 2003; Hirzel et al., 2004). The AVI is the proportion of presence evaluation points falling above some fixed HS threshold (e.g. 0.5); it varies from 0 to 1. The CVI is the AVI minus the AVI of a model predicting presence everywhere (chance model), and varies from 0 to 0.5. As for the Kappa index, this approach suffers from having to choose an arbitrary threshold.

Boyce et al. (2002) proposed a way to relieve somewhat the threshold constraint. Their method consists in partitioning the habitat suitability range into b classes (or bins), instead of only two. For each class i, it calculates two frequencies: (1) P_i, the *predicted frequency* of evaluation points:

$$P_i = \frac{p_i}{\sum_{j=1}^{b} p_j} \quad (2)$$

where p_i is the number of evaluation points predicted by the model to fall in the habitat suitability class i and $\sum p_j$ is the total number of evaluation points; (2) E_i, the *expected frequency* of evaluation points, i.e. the frequency expected from a random distribution across the study area. This is given by the relative area covered by each class:

$$E_i = \frac{a_i}{\sum_{j=1}^{b} a_j} \quad (3)$$

where a_i is the number of grid cells belonging to habitat suitability class i, or area covered by the class i, and Σa_j is the overall number of cells in the whole study area.

Finally, for each class i, the *predicted-to-expected* (P/E) ratio F_i is given by

$$F_i = \frac{P_i}{E_i} \quad (4)$$

If the habitat model properly delineates the species suitable areas, a low suitability class should contain fewer evaluation presences than expected by chance, resulting in $F_i < 1$. Conversely, high suitability classes should have F_i increasingly higher than 1. The plot of P/E against the mean habitat suitability of each class thus provides a handy interpretation tool. In such a context, a good model is expected to show a monotonically increasing curve, i.e. F_i increase as suitability increases. Boyce et al. (2002) measure this monotonic increase by the Spearman rank correlation coefficient between F_i and i. This "Boyce Index" B_b varies from −1 to 1. Positive values indicate a model whose predictions are consistent with the presences distribution in the evaluation dataset, values close to zero mean that the model is not different from a chance model, negative values indicate an incorrect model, which predicts poor quality areas where presences are more frequent.

The main shortcoming of the Boyce index is its sensitivity to the number of suitability classes b and to their boundaries (Boyce et al., 2002; personal observations). To fix this problem, we derived a new evaluator based on a "moving window" of width W (say $W = 0.1$) instead of fixed classes. Computation starts with a first class covering the suitability range [0,W] whose P/E ratio is plotted against the average suitability value of the class, $W/2$. Then, the moving window is shifted from a small amount upwards and the P/E is plotted again. This operation is repeated until the moving window reaches the last possible range [1 − W, 1]. This provides a smooth P/E curve, on which a "continuous Boyce index" $B_{cont(W)}$ is computed.

One of the main differences between the Boyce indices and the classical evaluators is that they require the HS prediction to be computed on the whole study area. For maps with low number of pixels, the whole information can be imported into a statistical application; however, in most cases their computation must be done within a GIS, or a program having direct access to the GIS data files. We have thus implemented all these evaluators into the free software BIOMAPPER (Hirzel et al., 2006), which works directly on GIS files and controls the free statistical application R (R Development Core Team, 2005) to compute the GLMs.

2.2. *Cross-validation*

All the evaluation methods presented above provide a single measure of the model predictive power. k-fold cross-validation is a resampling approach that allows assessment of the robustness of this measure (Van Houwelingen and Le Cessie, 1990; Fielding and Bell, 1997; Hastie et al., 2001). Moreover, it also enables one to evaluate a model even when the species dataset is small, as it ensures an optimal use of the data to calibrate and evaluate the model. Cross-validation consists of randomly dividing the dataset into k independent

partitions, using $k-1$ of them to calibrate the model, and computing the evaluator on the left-out partition. This procedure is repeated k times, each time leaving out another partition. This produces k estimations of the evaluator, allowing assessment of its central tendency and variance (in this study, we used median and 90%-confidence interval). The number of partitions typically varies between 3 and 10, depending on the number of species points. This method assumes that the k partitions are independent. See Hastie et al. (2001) for more details on cross-validation processes.

2.3. Test dataset

In order to test and compare the above evaluators, reliable data were required. We used data from 539 non-forest mountain vegetation plots located in a ~700 km^2 area in the external calcareous Alps of Canton de Vaud (6°60′–7°10′E and 46°10′–46°30′N, altitude ranging from 375 to 3210 m) in Switzerland. The sampling points were randomly stratified by classes of elevation, slope and aspect. The plants of each sampling point were exhaustively inventoried (Randin et al., in press). Among these plants, we selected those species that fulfilled three criteria: (i) species are easily detectable in the field during the sampling period, so that absences cannot be due to the species being undetected; (ii) species cannot be confounded with another sister species; (iii) species are at least present in 10 vegetation plots. The points (i) and (ii) guarantied reliable presences and absences. We ended up with 114 species that were present in at least 10 cells and at most 249. These data were collected during three field-sampling periods, in the summers 2002–2004.

2.4. Environmental variables

The environmental variables we used to fit the models are known to have a major direct ecophysiological impact on plant species (Pearson et al., 2002; Dirnböck et al., 2003; Körner, 2003). They were all calculated with a 25 m × 25 m spatial resolution, as derived from the digital elevation models (DEM) available in the study area (MNT25, Swisstopo). We calculated slope from the DEM to account for gravitational processes acting upon vegetation. The TOPO index indicates local convexity of the topography, which is partly correlated with water accumulation, snow persistence, nitrogen enrichment or wind protection. Basic climatic variables were obtained by spatial interpolation of climatic weather stations (monthly data for the period 1961–1990), and then transformed into three physiologically meaningful bioclimatic variables: degree-days (with 0 °C as threshold of plant growth), moisture index over the growing season (June–August) and potential global solar radiation of the growing season (see Table 1 and references therein). Details on these variables can be obtained in Randin et al. (in press).

2.5. Habitat suitability modelling

We used generalised linear models (GLM, McCullagh and Nelder, 1989) with a binomial probability distribution and a logit link to compute the habitat suitability maps based on presence/absence data. In a first step, we computed generalised linear models (GLMs) (as implemented in R, R Development Core Team, 2005) for the whole presence/absence dataset. To prevent the models' sensitivity to species prevalence, we weighted the absence points so as to have a presence/absence ratio of 1:1. The independent variables were those listed in Table 1. For each species, the relevant variables of the model were selected by a stepwise procedure based on the AIC criterion (Akaike, 1973; S-Plus, 1999); whenever a variable was retained in its squared form, we forced its linear form to be also included (T. Hastie, personal communication). The retained model was then fixed and, in a second step, we applied a k-fold cross-validation process (with $k=5$) using only the retained variables. To ensure a similar presence/absence ratio between all partitions, the random selection of the k partitions was done independently for the presence and absences data. Presences and absences were weighted as in the full model. This produced five GLM models and five HS maps for each species.

2.6. Evaluation index comparisons

Each HS map resulting from the cross-validation was evaluated by D^2_{adj}, K, K_{max}, AUC, AVI and CVI. In comparison, to investigate the sensitivity and the significance of the Boyce index, we tested it with several number of classes b: B_2, B_4, B_5 and B_{10}, as well as, in its continuous form, with the corresponding class sizes: $B_{cont(0.5)}$, $B_{cont(0.25)}$, $B_{cont(0.2)}$ and $B_{cont(0.1)}$. B_{10} corresponds to the number of classes used by Boyce et al. (2002). We tested B_2 as it was expected to be more comparable to two classes evaluators like K, AVI and CVI. Moreover, the GLM fit to the calibration data was evaluated by the adjusted explained deviance D^2_{adj}, which corresponds to the amount of deviance explained by the model corrected by the effective number of degrees of freedom used to build the model (Guisan

Table 1 – Environmental variables used to model habitat suitability of the 114 plants

Variables	Details	References
Moisture (mm day^{-1})	Monthly average of daily atmospheric water balance from July to September	Zimmermann and Kienast (1999)
Degree-days (°C day)	Number of days with mean temperature above 0 °C time their mean temperature	Prentice et al. (1992)
Global solar radiation (kJ m^{-2} day^{-1})	Monthly average of daily global solar radiation from July to September	Kumar et al. (1997)
Slope (degrees)	Slope inclination	ArcInfo (2004)
Topo	Topographic convexity	Zimmermann (unpublished)

and Zimmermann, 2000). We computed the median and 90%-confidence interval of each evaluator for the five HS maps. We finally plotted the medians of each evaluator against each other for all species and computed their Pearson's correlation coefficient ρ. We validated the Boyce indices by comparison to AUC and K_{max}, as they are common threshold-independent presence/absence evaluators.

3. Results

The chosen species cover a wide spectrum of ecological niche types and sample size. The quality of their habitat suitability models range from very bad to excellent. All the investigated evaluation measures convey similar information, with Pearson correlation coefficients greater than 0.5 in most cases (Table 2a). In particular, for the models where more than 50 presence points were available, most evaluators show more than 70% of correlation (Table 2b).

Except for those based on very wide classes (B_2 and $B_{cont(0.5)}$), Boyce indices are highly correlated together, and are moreover highly consistent with presence/absence evaluators. They tend to be more correlated to AUC than to K_{max}; this tendency is stronger for the whole dataset (Table 2a) than for the 48 most prevalent species (Table 2b). In spite of having the same class pattern, B_2 and $B_{cont(0.5)}$ are but weakly correlated to K, AVI and CVI; they both give the poorest correlations and will not be considered further. The Boyce indices most consistent with the AUC and K_{max} were those with the largest number of classes (or smallest window sizes), B_{10}, $B_{cont(0.1)}$, B_5 and $B_{cont(0.2)}$. We tested more than 10 classes but the variance of the P/E curves becomes too high (results not shown). To simplify the discussion, we will from now on only consider one evaluator of each type: D^2_{adj}, AUC, K_{max}, B_{10}, $B_{cont(0.1)}$ and CVI (correlation graphs shown in Fig. 3). Although all Boyce indices are highly correlated with K, we chose not to consider this evaluator as it is not threshold-independent.

Fig. 4 shows the sensitivity of these evaluators to species prevalence. K_{max} is sensitive to species prevalence, tending to give higher values to common-species models, while D^2_{adj} and CVI tend to give higher scores to low prevalence models. The other evaluators are insensitive to prevalence (Fig. 4).

Table 2a – Pearson's correlation coefficients between pairs of evaluation indices (median of the k-fold values) on all species (n = 114)

	Presence/absence evaluation indices				Presence-only evaluation indices									
	N_1 [a]	D^2_{adj}	AUC	K	K_{max}	B_2	B_4	B_5	B_{10}	$B_{cont(0.1)}$	$B_{cont(0.2)}$	$B_{cont(0.25)}$	$B_{cont(0.5)}$	AVI	CVI
N_1 [a]	1.00														
D^2_{adj}	−0.40	1.00													
AUC	−0.24	**0.86**	1.00												
K	0.19	0.54	0.63	1.00											
K_{max}	0.47	0.38	0.55	0.79	1.00										
B_2	0.17	0.34	0.47	0.65	0.56	1.00									
B_4	0.11	0.63	0.63	0.74	0.64	0.53	1.00								
B_5	0.02	0.73	0.73	0.74	0.62	0.49	**0.81**	1.00							
B_{10}	−0.19	**0.84**	0.79	0.76	0.56	0.42	0.80	**0.89**	1.00						
$B_{cont(0.1)}$	−0.16	**0.83**	**0.80**	0.80	0.60	0.46	**0.83**	**0.90**	**0.96**	1.00					
$B_{cont(0.2)}$	0.02	0.68	0.69	**0.83**	0.67	0.60	**0.87**	**0.86**	**0.85**	**0.91**	1.00				
$B_{cont(0.25)}$	0.05	0.62	0.64	**0.82**	0.66	0.63	**0.82**	0.79	0.79	**0.85**	**0.98**	1.00			
$B_{cont(0.5)}$	−0.10	0.40	0.50	0.62	0.46	0.65	0.54	0.47	0.52	0.57	0.67	0.71	1.00		
AVI	−0.17	0.80	0.77	0.77	0.55	0.57	0.73	0.76	**0.83**	**0.86**	**0.80**	0.77	0.62	1.00	
CVI	−0.32	0.80	0.74	0.67	0.43	0.52	0.66	0.69	0.79	**0.80**	0.74	0.71	0.64	**0.96**	1.00

Correlations ≥ 0.8 are emphasised.

[a] Number of presence points.

4. Discussion

On the range covered by the 114 studied plant species, and according to the environmental characteristics of our study area, all evaluators convey correlated information. This is an important result meaning that the presence-only evaluators can be trusted.

4.1. *Evaluator comparisons*

As expected, the quality of the 114 GLM models fitted varies greatly. Overall, the tested evaluators agree about their ranking, in particular among prevalence-insensitive indices (AUC, B_{10}, $B_{cont(0.1)}$). The results show that Boyce indices tend to give poor results when computed on a small number of classes. In particular, B_2 and $B_{cont(0.5)}$ have a low correlation with almost all evaluators, including those based on a fixed threshold as

Table 2b – As in Table 2a but only on species with more than 50 presence points (n = 48)

	Presence/absence evaluation indices					Presence-only evaluation indices							
	N_1[a]	D^2_{adj}	AUC	K	K_{max}	B_2	B_4	B_5	B_{10}	$B_{cont(0.1)}$	$B_{cont(0.2)}$	$B_{cont(0.25)}$	$B_{cont(0.5)}$	AVI	CVI
N_1[a]	1.00														
D^2_{adj}	−0.10	1.00													
AUC	−0.07	**0.93**	1.00												
K	0.00	**0.90**	**0.81**	1.00											
K_{max}	0.32	**0.85**	**0.87**	**0.82**	1.00										
B_2	−0.06	0.45	0.57	0.45	0.46	1.00									
B_4	0.19	0.62	0.56	0.71	0.70	0.35	1.00								
B_5	0.07	0.74	0.70	**0.80**	0.76	0.43	0.77	1.00							
B_{10}	−0.05	**0.89**	**0.81**	**0.91**	0.79	0.42	0.78	**0.88**	1.00						
$B_{cont(0.1)}$	−0.07	**0.90**	**0.84**	**0.93**	**0.82**	0.47	0.79	**0.88**	**0.98**	1.00					
$B_{cont(0.2)}$	−0.04	0.78	0.74	**0.86**	0.76	0.49	**0.82**	**0.82**	**0.89**	**0.93**	1.00				
$B_{cont(0.25)}$	−0.06	0.75	0.72	**0.85**	0.73	0.50	0.75	0.74	**0.85**	**0.89**	**0.98**	1.00			
$B_{cont(0.5)}$	−0.38	0.52	0.52	0.52	0.41	0.32	0.41	0.39	0.53	0.58	0.63	0.69	1.00		
AVI	−0.02	**0.91**	**0.83**	**0.97**	**0.82**	0.41	0.64	0.75	**0.89**	**0.91**	**0.82**	**0.82**	0.55	1.00	
CVI	−0.29	**0.85**	0.76	**0.86**	0.65	0.37	0.54	0.65	**0.82**	**0.84**	0.78	0.80	0.69	**0.91**	1.00

Correlations ≥ 0.8 are emphasised.

[a] Number of presence points.

Table 3 – Median value taken by the evaluators (minimum and maximum in brackets)

Evaluator	Value
D^2_{adj}	0.34 (0.07, 0.75)
AUC	0.88 (0.63, 0.99)
K	0.27 (−0.04, 0.70)
K_{max}	0.49 (0.10, 0.82)
B_{10}	−0.07 (−0.67, 0.65)
$B_{cont(0.1)}$	−0.07 (−0.66, 0.61)
CVI	0.24 (−0.12, 0.98)

AVI, CVI and *K*. As the Boyce indices are based on the Spearman rank correlation, they are more sensitive to larger number of classes. However, classes cannot be added indefinitely as the variance among cross-validation partitions increases as their width decreases. Ten classes (class width = 0.1) seems to be the optimum, advocating for B_{10} and $B_{cont(0.1)}$. We prefer the later as it does not depend on any particular class cutting thresholds.

The agreement between presence/absence and presence-only measures tends to be lower when the species prevalence is below 10% (<50 presences) (Tables 2). This is because a low number of presences prevent presence-only evaluators from assessing the overall quality of the model, whilst presence/absence evaluators can still rely on the fit between predicted and observed absences. Therefore, when presences are scarce, presence/absence evaluators often give an intermediate score to the model on the base of absence predictions, whilst presence-only evaluators assess the model as poor (cf. evaluator ranges in Table 3 and Fig. 5). This redemption of the model by the absences may be acceptable if the cost of overlooking suitable habitat is not too high. It is important to take such evaluation scale shifts into account for wildlife management applications.

Why some models better predict absences than presences may come from various causes. First, unreliable species data may bring too much noise for a proper niche modelling. In our case, as the species were carefully selected for the reliability of their presence/absence dataset, we can mostly rule out this effect. Second, the model accuracy depends on the environmental variables relevance for the species. Although we chose good general predictors for plants (Dirnböck et al., 2003), some more specialised variables are obviously missing for the badly modelled species. This was expected, and actually sought, as we wanted the model quality to cover as wide a palette as possible. Thus, when the environmental variables are irrelevant to the species niche, the model cannot efficiently predict presences.

Most measures evaluate how well a model can predict absence and presence. By contrast, the Boyce indices assess the model ability to consistently predict several levels of suitability. A complementary evaluation would be to apply the same Boyce approach to the absences. In that case, one would expect negative *P/E* curves, i.e. negative Spearman ranks. The combination of these two facets of the Boyce indices seems promising, as it would bring the power of continuous evaluation to presence/absence models.

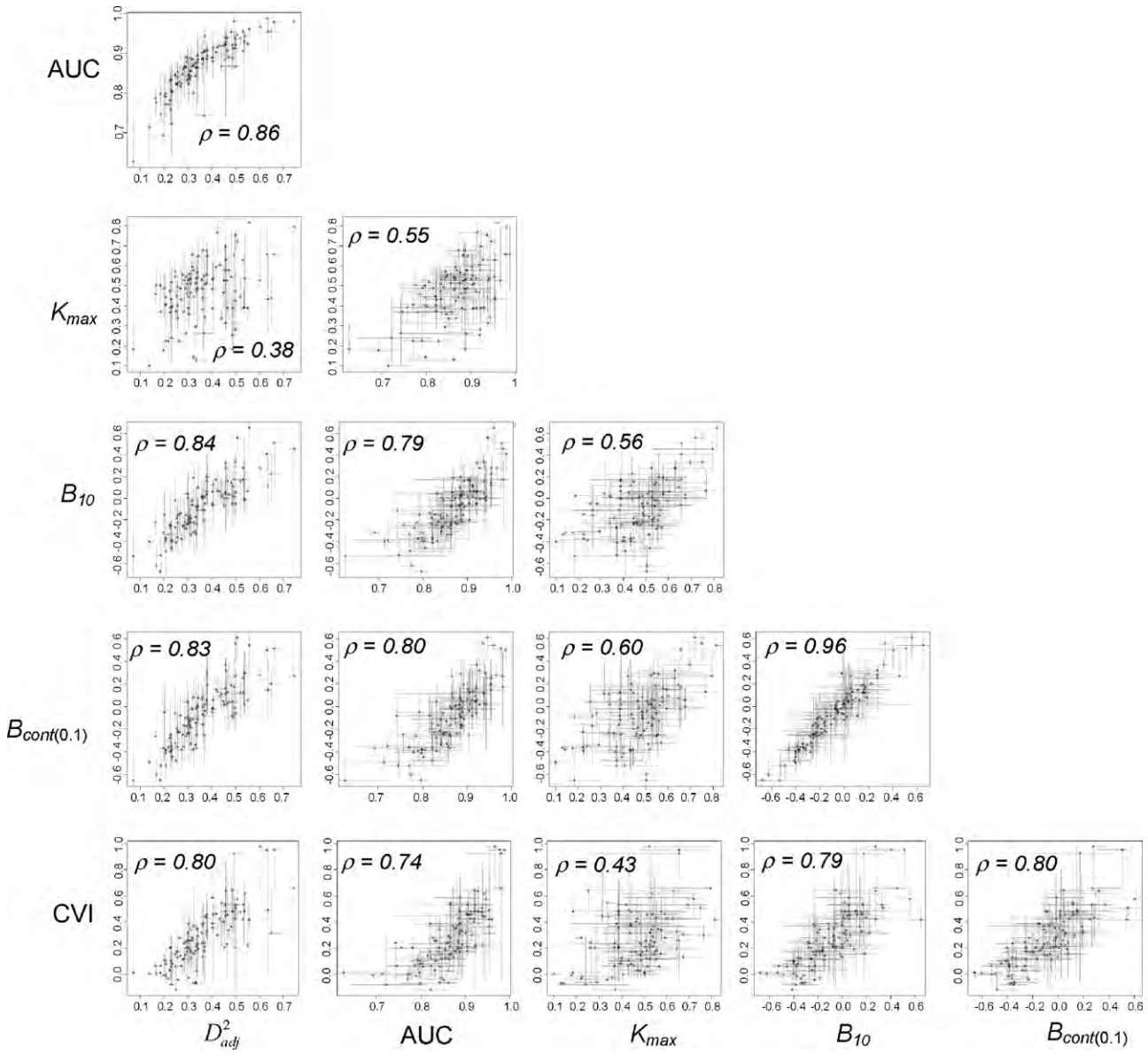


Fig. 3 – Relationship between the main evaluators. Each point represents the median value of the evaluators computed on the cross-validation partitions. The grey lines represent the interquartile range. The Pearson's correlation coefficient is indicated for each pair of evaluators.

4.2. *Interpreting the P/E curves*

While an evaluation index gives a summary of the model prediction ability, the continuous P/E curves provide a wealth of valuable insights into the model accuracy, offering three levels of information.

First, the variance among the cross-validation curves gives information about model *robustness* all along the HS range. The narrowness of the confidence interval reflects the model sensitivity to particular calibration points. As the variance often fluctuate along the curve, one can thus determine which parts of the model are the most accurate. For instance, a model could provide trustable prediction for low suitability regions (good absence prediction) but be more variable about high suitability (bad presence prediction). Such information provides a finer understanding of the weaknesses of the model. The manager can then take these weaknesses into account when applying the predictions to management decisions. Alternatively, these weaknesses may give clues about what parts of the predictions (e.g. absences) must be improved. In this study, we indicate robustness by using cross-validation and providing a confidence interval around the evaluators.

The second information level is related to the actual shape of the P/E curve. An ideal model would have a linear P/E curve and could thus predict habitat suitability with an infinitely fine resolution. It means that the suitability index is really proportional to the probability of use, as defined by Manly et al. (2002). Real curves however may exhibit non-linear (e.g. exponential) or staircase shapes. Wherever the local slope is flat or negative, the corresponding range of HS may be pooled into one

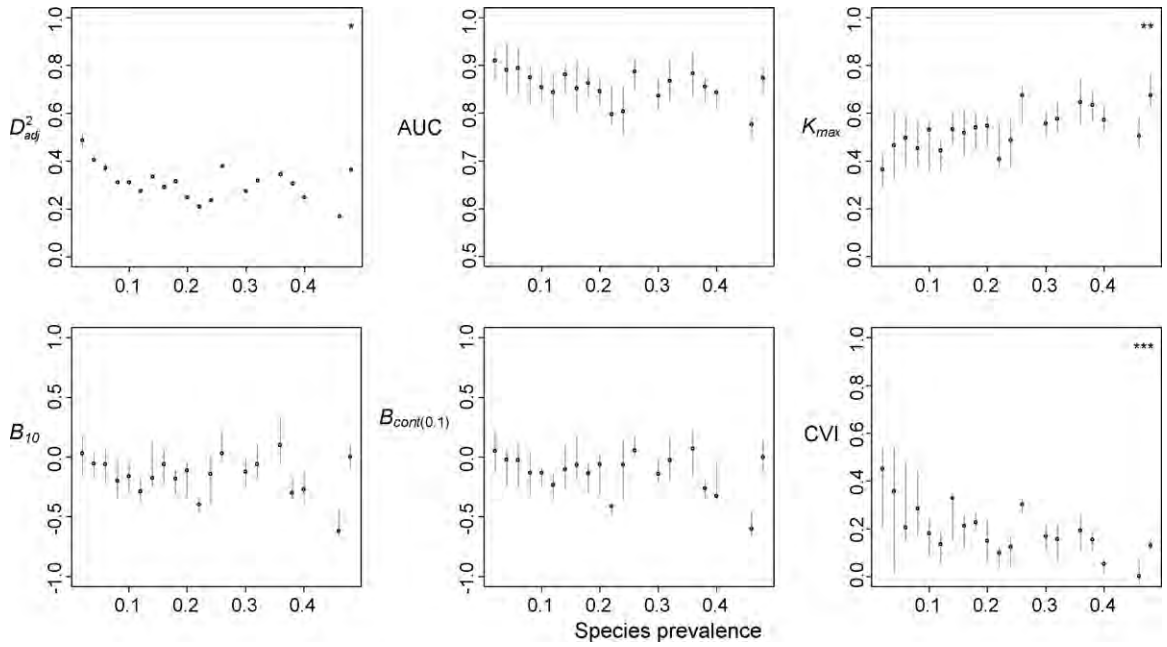


Fig. 4 – Evaluator sensitivity to species prevalence (proportion of presence points in the dataset). The points represent the index median value and the grey lines the interquartile range. t-test significance of the Pearson's correlation coefficient indicated by stars: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

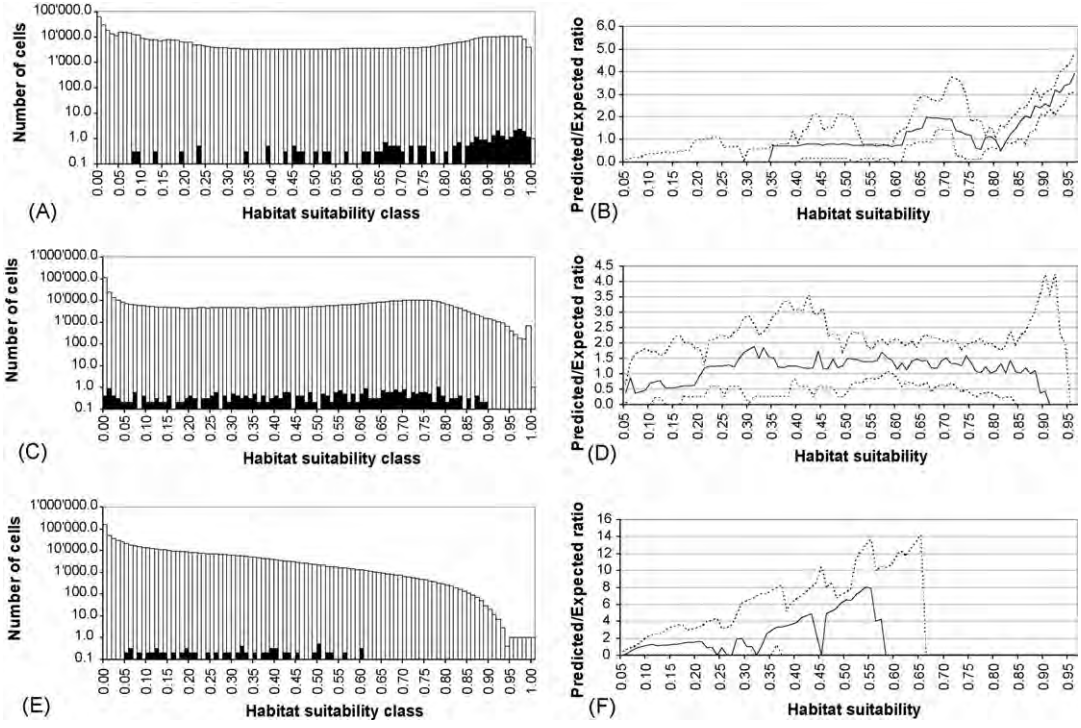


Fig. 5 – Typical examples of the three models: (A and B) good model; (C and D) random model (confidence interval brackets the 1-line all along); (E and F) bad model (presences fall in low suitability areas). In the left-hand column, the white bars show the number of cells belonging to each habitat suitability class, while black bars are the number of cells with asserted presence in these classes. These histograms average the results of the five cross-validation HS maps, where vertical scales are logarithmic. In the right-hand column, the predicted/expected curves computed by a moving window of width 0.1 (plain line = median, dashed lines = 90%-confidence interval). More details on the models include: (A and B) *Lathyrus pratensis*, $K_{max} = 0.68$, AUC = 0.92, $B_{10} = 0.82$, $B_{cont(0.1)} = 0.76$; (C and D) *Alchemilla xanthochlora*, $K_{max} = 0.68$, AUC = 0.87, $B_{10} = 0.00$, $B_{cont(0.1)} = 0.00$; (E and F) *Gypsophila repens*, $K_{max} = 0.37$, AUC = 0.78, $B_{10} = -0.26$, $B_{cont(0.1)} = -0.40$.

class without loss of information (Fig. 5). This means that any departure from the straight line actually decreases the *resolution* of the model predictions, i.e. its ability to distinguish many different classes of suitability. Model resolution is partly contained in the Boyce indices as they get penalised whenever the *P/E* curve goes down. However, being based on the Spearman correlation coefficient, they cannot discriminate between monotonic curves (e.g. linear, exponential and sigmoid curves would all get the same score).

The third information level refers to the maximum value reached by the *P/E* curve. This value reflects how much the model differs from chance expectation, or *deviation from randomness*. This score reflects the model ability to differentiate the species niche characteristics from those of the studied area. This measure must be taken with care as our experience has shown that it highly depends on the species niche breadth, the extent of the study area, the scale of the study (i.e. the environmental variables resolution), and the relevance of the chosen environmental variables. For instance, with a mountain species, a model built at the whole country level with climatic variables is bound to get a higher information index than one focusing on the mountain ranges; however, the country-wide model is obviously not better than the mountain-wide one. Practically, one must use this deviation from randomness only to compare models applied to the same species and the same study area.

4.3. *Reclassifying HS maps*

Most habitat suitability models (including GLMs) generate maps showing continuous gradients of suitability. This kind of output obviously conveys more information than a sheer presence/absence map and is more convenient for wildlife management support. However, the present study has shown that a continuous scale is often misleading. Even good predictive models suffer from uncertainty, making the use of a full continuous HS scale spurious. A reclassified map showing only a few classes may be more honest about its actual informative content. The problem of choosing objectively the HS class boundaries immediately arises. For binary reclassifications – presence/absence – methods already exists: the K_{max} approach readily provides the HS threshold that maximise the *K* index. The AUC approach allows weighting the risks of over- and under-prediction and computing the optimal threshold accordingly (Fielding and Bell, 1997). For more than two classes, the *P/E* curves provide a handy support for choosing (1) the number of classes and (2) their boundaries. The optimal number of classes may be defined by looking at the confidence interval around the continuous *P/E* curve (e.g. Fig. 5B), the goal being of finding how many HS classes (on the horizontal axis) may be defined while minimizing their overlap in *P/E* ratio (vertical axis). The continuous *P/E* curves also allow choosing the class boundaries objectively (Fig. 6). A first, natural boundary is defined by the $P/E = 1$ line: where *P/E* confidence interval is lower than 1, the model is predicting less presences than expected by chance, and the opposite when it is greater than 1; this may be used to distinguish unsuitable, marginal (random and uncertain) and suitable habitat (Fig. 6). Another natural threshold is the HS below which no presence ever occurs, suggesting uninhabitable conditions, which is equivalent to the threshold defining the minimal predicted area (MPA_{100}) of Engler et al. (2004). Additional thresholds may be placed at the steps of the curve. The *P/E* ratios also provide a reproducible HS partitioning scheme: one could define HS classes predicting twice more presences than expected, thrice more, and so on. Fig. 6 illustrates this process of HS partitioning. Note that with categorical variables, it would make little sense to have more HS classes than categories.

In conclusion, our work has shown that evaluating a habitat suitability model based only on presences is possible and a valuable exercise. Among the presence-only evaluators, the continuous Boyce index $B_{cont(0.1)}$ was most accurate for characterizing predictive capability among our sample of 114 plant distributions. Accordingly, we suggest that it is both a complement to usual evaluation of presence/absence models (e.g. GLMs and GAMs, Guisan et al., 2002) and a reliable measure of presence-only based predictions (e.g. ecological niche factor analysis: Hirzel et al., 2002; or resource selection functions: Manly et al., 2002). Such measures could also prove useful to evaluate presence/absence-based models when an accurate

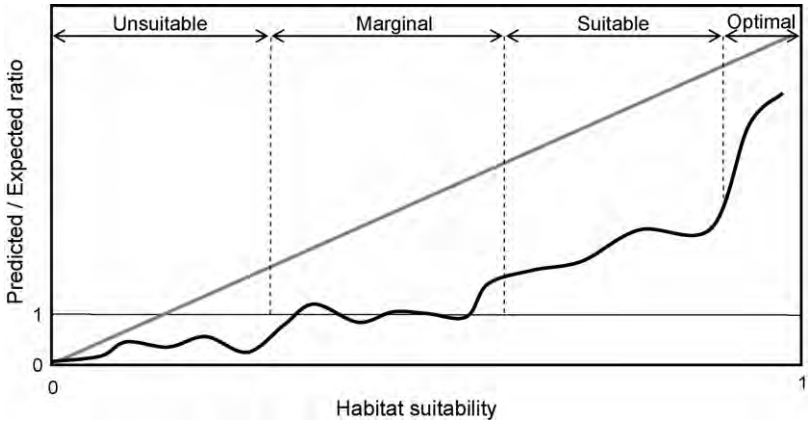


Fig. 6 – Predicted/expected curve shapes. An ideal model would give a straight *P/E* curve (plain grey line). Actual models often exhibit an irregular increase (black plain curve). The curve shape and its confidence interval (dashed curves) may be used to define the boundaries of habitat suitability classes (as suggested by the vertical dashed lines). The horizontal thin line at $P/E = 1$ would be the curve of a completely random model.

prediction of presences is crucial, as in the case of detecting new populations of threatened species (e.g. Engler et al., 2004). Moreover, we stress the need to reclassify habitat suitability maps so as to provide more honest and relevant predictions. The P/E curves described offer a handy support for this reclassification.

Acknowledgements

We wish to thank Mark S. Boyce, Gretchen G. Moisen and all the participants of the Riederalp Workshop, Switzerland, 2004, for stimulating discussions about the model evaluation, Patrick Patthey, Julie Jacquiéry and Pietro Persico for the first explorations of the Boyce index. We also wish to thank Jane Elith and two anonymous reviewers who helped improve this article. We are grateful to Fabien Fivaz for the help with R scripts as well as to all those who contributed to the field work: Pascal Vittoz, Stéfanie Maire, Dario Martinioni, Simone Peverelli, Chantal Peverelli, Séverine Wydler, Lorenzo De Stefani and Roxane Milleret.

REFERENCES

Akaike, H., 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika 60, 255–265.

Anderson, R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol. Model. 162, 211–232.

ArcInfo, 2004. ArcInfo Version 9.0. Environmental Systems Research Institute Inc., Redlands, CA.

Boyce, M.S., McDonald, L.L., 1999. Relating populations to habitats using resource selection functions. Trends Ecol. Evol. 14, 268–272.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. Ecol. Model. 157, 281–300.

Buckland, S.T., Elston, D.A., 1993. Empirical models for the spatial distribution of wildlife. J. Appl. Ecol. 30, 478–495.

Cohen, J., 1960. A coefficient of agreement of nominal scales. Educ. Psychol. Measure. 20, 37–46.

Dirnböck, T., Dullinger, S., Grabherr, G., 2003. A regional impact assessment of climate and land-use change on alpine vegetation. J. Biogeogr. 30, 401–417.

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. J. Appl. Ecol. 41, 263–274.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38–49.

Guisan, A., Edwards Jr., T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol. Model. 157, 89–100.

Guisan, A., Theurillat, J.P., Kienast, F., 1998. Predicting the potential distribution of plant species in an Alpine environment. J. Veg. Sci. 9, 65–74.

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.

Hastie, T., Tibshirani, R., Friedman, J.H., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.

Hirzel, A.H., Arlettaz, R., 2003. Modelling habitat suitability for complex species distributions by the environmental-distance geometric mean. Environ. Manage. 32, 614–623.

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.

Hirzel, A.H., Hausser, J., Perrin, N., 2006. Biomapper 3.2. Lab. for Conservation Biology, University of Lausanne, Lausanne, http://www.unil.ch/biomapper.

Hirzel, A.H., Helfer, V., Métral, F., 2001. Assessing habitat-suitability models with a virtual species. Ecol. Model. 145, 111–121.

Hirzel, A.H., Posse, B., Oggier, P.-A., Glenz, Y.C., Arlettaz, R., 2004. Ecological requirements of a reintroduced species, with implications for release policy: the bearded vulture recolonizing the Alps. J. Appl. Ecol. 41, 1103–1116.

Kéry, M., 2002. Inferring the absence of a species: a case study of snakes. J. Wildlife Manage. 66, 330–338.

Körner, C., 2003. Alpine Plant Life. Springer, Berlin.

Kumar, L., Skidmore, A.K., Knowles, E., 1997. Modelling topographic variation in solar radiation in a GIS environment. Int. J. Geogr. Inf. Sci. 11, 475–497.

Mackenzie, D.I., Royle, J.A., 2005. Designing occupancy studies: general advice and allocating survey effort. J. Appl. Ecol. 42, 1105–1114.

Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. J. Appl. Ecol. 38, 921–931.

Manly, B.F., McDonald, L.L., Thomas, D.L., McDonald, T.L., Erickson, W.P., 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies, 2nd ed. Kluwer Academic Publishers, Dordrecht.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, 2nd ed. Chapman & Hall, London.

Ottaviani, D., Lasinio, G.J., Boitani, L., 2004. Two statistical methods to validate habitat suitability models using presence-only data. Ecol. Model. 179, 417–443.

Pearce, J.L., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. J. Appl. Ecol. 43, 405–412.

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. Ecol. Model. 154, 289–300.

Prentice, C.I., Cramer, W., Harrison, S.P., Leemans, R., Monerud, R.A., Solomon, A.M., 1992. A global biome model based on plant physiology and dominance, soil properties and climate. J. Biogeogr. 19, 117–134.

R Development Core Team, 2005. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Randin, C., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., in press. Are species distribution models transferable in space? J. Biogeogr.

Reutter, B.A., Helfer, V., Hirzel, A.H., Vogel, P., 2003. Modelling habitat-suitability on the base of museum collections: an example with three sympatric *Apodemus* species from the Alps. J. Biogeogr. 30, 581–590.

Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B., 2002. Predicting Species Occurences: Issues of Scale and Accuracy. Island Press, Washington.

S-Plus, 1999. S-PLUS 2000 Guide to Statistics. Data Analysis Products Division, MathSoft, Seattle, WA.

Van Houwelingen, J.C., Le Cessie, S., 1990. Predictive value of statistical models. Stat. Med. 9, 1303–1325.
Zaniewski, A.E., Lehmann, A., Overton, J.M., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecol. Model. 157, 261–280.
Zimmermann, N.E., Kienast, F., 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. J. Veg. Sci. 10, 469–482.
Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (Roc) plots—a fundamental evaluation tool in clinical medicine. Clin. Chem. 39, 561–577.

Chapter 9

Journal of Applied Ecology 2006

Modelling ecological niches with support vector machines

JOHN M. DRAKE,* CHRISTOPHE RANDIN† and ANTOINE GUISAN†

**National Center for Ecological Analysis and Synthesis, 735 State St, Suite 300, Santa Barbara, CA 93101, USA; and*
†Department of Ecology and Evolution, University of Lausanne, Biology Building, CH-1015 Lausanne, Switzerland

Summary

1. The ecological niche is a fundamental biological concept. Modelling species' niches is central to numerous ecological applications, including predicting species invasions, identifying reservoirs for disease, nature reserve design and forecasting the effects of anthropogenic and natural climate change on species' ranges.
2. A computational analogue of Hutchinson's ecological niche concept (the multi-dimensional hyperspace of species' environmental requirements) is the support of the distribution of environments in which the species persist. Recently developed machine-learning algorithms can estimate the support of such high-dimensional distributions. We show how support vector machines can be used to map ecological niches using only observations of species presence to train distribution models for 106 species of woody plants and trees in a montane environment using up to nine environmental covariates.
3. We compared the accuracy of three methods that differ in their approaches to reducing model complexity. We tested models with independent observations of both species presence and species absence. We found that the simplest procedure, which uses all available variables and no pre-processing to reduce correlation, was best overall. Ecological niche models based on support vector machines are theoretically superior to models that rely on simulating pseudo-absence data and are comparable in empirical tests.
4. *Synthesis and applications*. Accurate species distribution models are crucial for effective environmental planning, management and conservation, and for unravelling the role of the environment in human health and welfare. Models based on distribution estimation rather than classification overcome theoretical and practical obstacles that pervade species distribution modelling. In particular, ecological niche models based on machine-learning algorithms for estimating the support of a statistical distribution provide a promising new approach to identifying species' potential distributions and to project changes in these distributions as a result of climate change, land use and landscape alteration.

Key-words: ecological modelling, habitat, presence-only data, species distribution

Journal of Applied Ecology (2006)
doi: 10.1111/j.1365-2664.2006.01141.x

Introduction

Methodological advances in species distribution modelling have been rapid (Guisan & Zimmermann 2000; Scott *et al*. 2002). While the practical and intellectual benefits of obtaining well-tested models for species' distributions are numerous, including forecasting species' range shifts from climate change (Thomas *et al*. 2004) and invasion by introduced species (Peterson 2003; Drake & Bossenbroek 2004), testing evolutionary hypotheses (Graham *et al*. 2004), identifying reservoirs for disease (Peterson *et al*. 2002), and planning for conservation in a dynamic landscape (Ferrier 2002), modelling species' niches is complicated by conceptual and technical difficulties and by data limitations (Guisan & Thuiller 2005). Recent advances in machine-learning techniques for statistical pattern recognition might be used to overcome many of these obstacles, which generally result from assumptions about the statistical distribution of data or restrictive parametric modelling paradigms. We studied the accuracy and reliability of ecological niche models built with support vector machines (SVM) for estimating the support of a statistical distribution (Schölkopf *et al*. 2001; Tax 2001; Tax

Correspondence: John M. Drake, National Center for Ecological Analysis and Synthesis, 735 State St, Suite 300, Santa Barbara, CA 93101, USA (e-mail drake@nceas.ucsb.edu).

& Duin 2004). We show that the SVM framework performs comparably or is superior to other methods with only moderate amounts of data while avoiding common problems and limitations.

The most common obstacles to conventional parametric and non-parametric statistical methods for modelling species' distributions are: (i) autocorrelated observations resulting from the inherent spatial distribution of ecological systems, spatial autocorrelation in species' actual distributions, and haphazard rather than designed sampling; and (ii) observations only of species' occurrences without complementary observations of species' absences. Autocorrelated observations result in inflated *P*-values for hypothesis testing when modelling techniques are based on parametric statistics, and have the potential to introduce bias in estimated models. One approach to this problem in a parametric setting is to add to a generalized linear model (GLM; e.g. logistic model) terms to model the spatial correlation (Augustin, Mugglestone & Buckland 1996; He, Zhou & Zhu 2003). Other studies have taken a similar approach with semi-parametric regression techniques, such as generalized additive models (GAM; Leathwick & Austin 2001). However, these methods place further demands on already sparse data and extrapolate poorly.

Strictly speaking, the second obstacle, lack of data confirming species' absences, renders modelling approaches based on classification/discrimination impossible (Robertson, Caithness & Villet 2001; Hirzel *et al*. 2002). Previous studies have sought to overcome this problem by simulating observations of species' absences (sometimes called pseudo-absences) from data domains in which there are no observations of species' occurrences (Engler, Guisan & Rechsteiner 2004). While remarkably robust models have been developed using this approach (Anderson, Lew & Peterson 2003), a method that does not rely on such heuristics would be useful. Further, it is not clear that these procedures can be used in a setting that is not already information rich, where background knowledge of species' ecologies can guide modelling heuristics (Anderson, Lew & Peterson 2003), although these are precisely the cases where species distribution models are most useful, for instance for forecasting species invasions or range shifts from climate change. Finally, classification models fitted to simulated data are generally ecologically uninformative or cumbersome to interpret (Keating & Cherry 2004). The aim of this study was to introduce a technique that overcomes these obstacles.

A promising alternative to conventional classification-based species distribution models is to use methods designed for modelling one type of data only (Robertson, Caithness & Villet 2001; Hirzel *et al*. 2002; Brotons *et al*. 2004; Phillips, Dudík & Schapire 2004). Many such techniques may be found in the literature on statistical pattern recognition, where a frequent goal is to separate statistical outliers from observations drawn from a high-dimensional distribution (Schölkopf *et al*. 2001; Tax 2001; Tax & Duin 2004). Indeed, rather than estimating the full probability distribution, in such situations it may be simpler (and more robust) to model just the support of the distribution, the set of points where the (unknown) probability density is greater than zero (Schölkopf *et al*. 2001). Sometimes support estimation is called one-class classification (Tax 2001). While many different methods for estimating statistical distributions might be optimized for one-class classification (Tax 2001; Tax & Duin 2004), methods based on SVM have been particularly successful in applications where data represent a large set of variables (Tax 2001, table 4·2; Tax & Duin 2004). SVM use a functional relationship known as a kernel to map data onto a new hyperspace in which complicated patterns can be more simply represented (Müller *et al*. 2001). The choice of kernel is typically based on theoretical properties, while any kernel parameters are optimized using computational techniques such as cross-validation. Because SVM are not based on characteristics of statistical distributions there is no theoretical requirement for observed data to be independent, thereby overcoming the problem of autocorrelated observations, although model performance will be affected by how well the observed data represent the range of environmental variables. Further, SVM are more stable, require less model tuning, and have fewer parameters than other computational optimization methods such as neural networks (Lusk, Guthery & DeMaso 2002). Finally, computational complexity is minimal and standard algorithms can be used for optimization. Thus, implementation is straightforward in familiar scientific computing environments such as R (http://www.r-project.org/, accessed 16 February 2006) and MATLAB (Mathworks Inc., Natick, MA). In contrast to genetic algorithms (Stockwell & Peters 1999; Drake & Bossenbroek 2004), the solution is deterministic, resulting in both faster computation and repeatable results. Thus, the potential gains from using support vector machines for ecological niche modelling are great, including reliable and accurate forecasting, feasible computation and a high level of ecological interpretability (Guo, Kelly & Graham 2005).

Methods

STUDY AREA

The study area and data in our analysis were derived from a project to generate forecasts of the effects of climate change on the distribution and diversity of plant species in alpine areas (MODIPLANT project; http://ecospat.unil.ch, accessed 16 February 2006). The study area included all mountains of the pre-Alps of the Canton de Vaud, a Swiss state (6°60′–7°10′E and 46°10′–46°30′N), with a total area of 564 km^2 (Fig. 1). Altitude ranged from 400 m to 3200 m a.s.l. The bedrock in the area is mainly calcareous. The climate is temperate, and generally wet with abundant rain and snowfalls. The sequence of vegetation along the altitudinal gradient is typical of the calcareous Alps, with deciduous forests at the lowest altitudes, mixed forests at middle altitudes, and coniferous

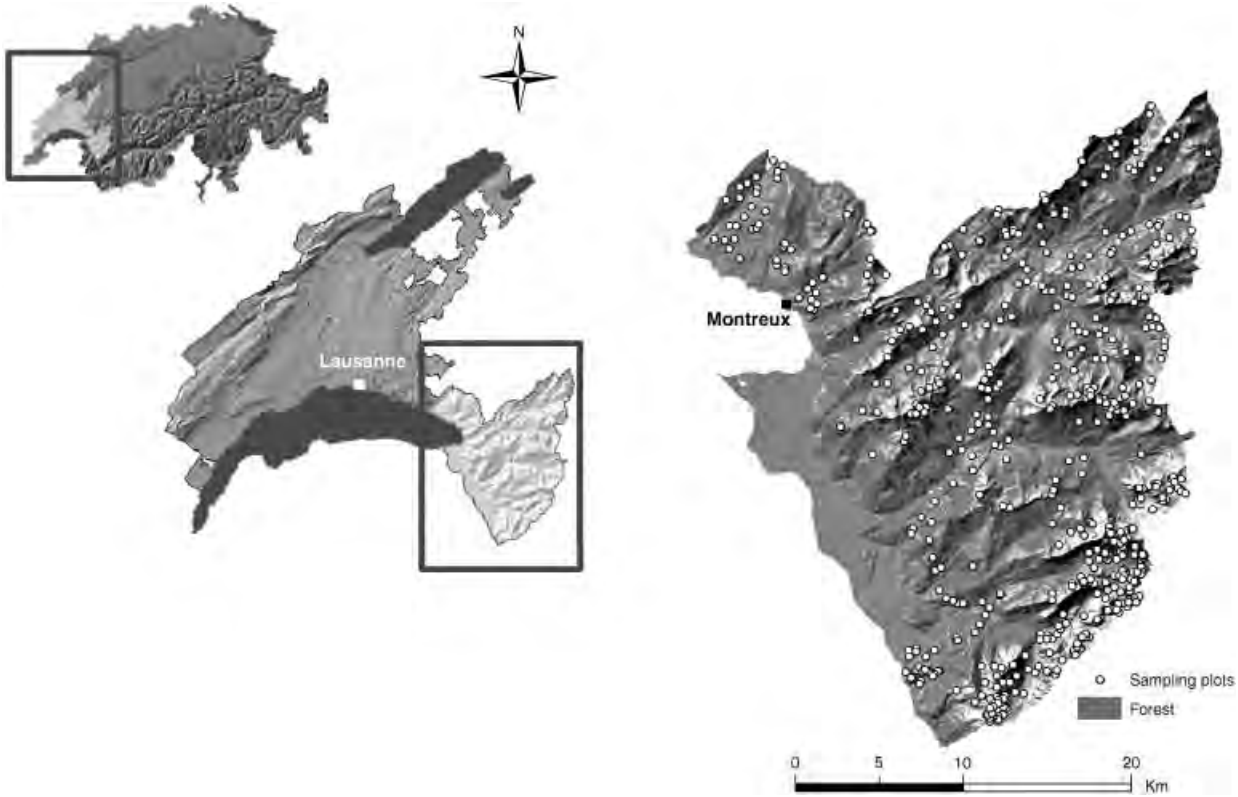


Fig. 1. Location of sampling plots in the pre-Alps of the Canton de Vaud, Switzerland, for 106 species used to compare accuracy and reliability of ecological niche models.

forest at the highest altitudes below the tree line (subalpine belt). Past and current human practices (such as pasturing) have created large gaps in the previously continuous forest cover at all altitudes. Above the tree line, heaths, meadows and grasslands alternate to define the mosaic of alpine vegetation (Randin *et al*. in press).

SPECIES DATA

The data set of species observations comprised 550 vegetation plots of 64 m^2 (8×8 m) that were sampled in summer (May through to mid-September; vegetation season) during the period 2002–04. A stratified random design, restricted to open non-woody vegetation (grasslands, rock and scree vegetation) and considering an equal number of replicates per stratum (Hirzel & Guisan 2002), was applied to the entire study area. We selected 106 species for modelling (see Table S1 in the Supplementary material), according to the following criteria: species must (i) have been represented by 10 or more observations, (ii) be taxonomically stable to avoid errors of identification and (iii) be easily identifiable in the field (i.e. have a high detectability) to ensure a low level of omission errors, which would bias estimates of model accuracy.

ENVIRONMENTAL DATA

All environmental variables were measured or computed at the 25-m spatial resolution of the digital elevation model (DEM; MNT25 Level 2, © 2001 Swisstopo (DV83.4), Federal Office of Topography, Wabern, Switzerland) from which they were derived. These were mainly topographic and climate variables. Slope was calculated from the DEM to account for gravitational and disturbance processes acting upon vegetation. A topographic index was calculated using an ArcInfo AML (Arc Macro Language) code (N. E. Zimmermann, http://www.wsl.ch/staff/niklaus.zimmermann, accessed 16 February 2006). This variable reflected the relative position of plants between ridges and valleys, which exposes them to different microclimates (wind, temperature and radiation). Maps of cumulative monthly precipitation and average monthly temperature were calculated by interpolating measurements from the Swiss meteorological network (1961–90) on the DEM (Federal Office of Meteorology and Climatology, Zurich). Lapse rates along the altitudinal gradient were used to normalize the monthly average to 0 m a.s.l., where the interpolations were performed, and then projected back to their actual altitude using the same lapse rates. This method differs slightly from the approach by Zimmermann & Kienast (1999) in using inverse-weighted difference rather than thin-plate splines for interpolation. The resulting climate maps were then used to derive predictors with putative causal relations to individual fitness (see Table S2 in the Supplementary material). Additionally, the spatial hydrological model PREVAH (Gurtz *et al*. 2003) was used to obtain a predictor for snow cover duration at all study sites, based on interpolated

daily values of five meteorological variables, precipitation, air temperature, relative sunshine duration, wind speed and water vapour pressure, measured during the period 1979–2000 (Randin *et al*. in press). All environmental variables were expected to have direct physiological effects on the plant species (Pearson *et al*. 2002; Dirnböck *et al*. 2003; Körner 2003).

MODEL TRAINING

Models were fit using only observations of species' presences. Prior to analysis, data were linearly rescaled to the interval between 0 and 1 to avoid potential problems with numerical stability. Most variables were symmetrically distributed and centred within the sample space, providing good coverage of the data domain (see Figure S1 in the Supplementary material). Therefore, no further transformations were performed. Because our goal was to estimate accuracy and reliability of ecological niche models fitted to observations of presence only, records of species' absences were excluded from data used for model training but were retained for model testing. Observations of species' presences were then randomly assigned to one of two subsets for model training (80%) and testing (20%). The final model testing data set was obtained by combining these observations (i.e. 20% of the total) with all records of species absence.

Several environmental variables were highly correlated (see Table S3 in the Supplementary material). Although our procedure did not assume uncorrelated predictors, we considered three different approaches to see how dimensionality and bias might be reduced by eliminating correlated variables. Our baseline, method 1, was to use all nine environmental predictors with no additional pre-processing.

Statistical learning exploits the variability in a set of observations and is often compromised by poorly scaled data. The initial rescaling we performed was expected to resolve problems concerning the stability of the numerical algorithms, but did not address the possibility that the data were clumped, reducing variation for the algorithm to exploit. Motivated by this problem, Tax & Juszczak (2003) suggest reducing the dimensionality of a data set through '*k*-whitening', by mapping the observed data onto the principal components of the training data. Accordingly, a computational analogue of principal components analysis (Kernel-PCA; Schölkopf & Smola 2001) was estimated and the principal components accounting for a specified fraction of the total variance were retained, rescaling the data to a feature space centred on zero and unit covariance. *k*-whitening may be thought of as a mapping that relates a full data set to a new, lower-dimensional space, which is retained and used in subsequent analyses. Thus, method 2 comprised constructing a new data set by *k*-whitening the full data set used in method 1, and then training the support vector machine on the new data set. Accounting for 99% of the total variance, the *k*-whitened data set resulted in reducing the number of dimensions used in model training from nine to five.

Finally, we considered whether or not we could reduce dimensionality by simply eliminating variables from the original data set altogether. As all of the variables related to solar radiation were highly correlated, we chose to eliminate all except SRADYY, which was relatively symmetrically distributed over the sample space. We further dropped MIND07 because it was highly correlated with PRECYY. Thus, method 3 consisted of using only the original (non-*k*-whitened) variables SLOPE, TOPO, SRADYY and PRECYY.

All analyses were conducted in MATLAB 7·0·1 (MathWorks Inc., Natick, MA). For model kernel we used the Gaussian radial basis function (RBF), which is a standard kernel for classification tasks and relies on tuning only one parameter. We optimized the models subject to the target false negative rate of 0·1 using the consistency criterion of Tax & Müller (2004), which minimizes overfitting by increasing the complexity of the model until the cross-validation error on the training data is greater than expected based on random sampling. Specifically, a model was determined to be inconsistent when the estimated false negative rate exceeded the target false negative rate by two standard deviations from the binomial distribution. Analyses were performed using a MATLAB toolbox for statistical pattern recognition (Duin *et al*. 2004) and the MATLAB Data Description Toolbox (Tax 2005).

MODEL TESTING

Models were tested using only data that were not used for model training. We were not able to obtain a consistent model for every species for each protocol. For species for which a consistent model could not be found we used as the best model the support vector machine defined by the Gaussian RBF with a tuning parameter equal to the maximum Euclidean distance between observations in the (rescaled) data set. Otherwise the model optimized by the consistency criterion of Tax & Müller (2004) was retained as the best model. This convention for 'best' model where consistency could not be obtained did not greatly affect our analysis as consistent models were almost always obtained for species represented by greater than 40 observations.

Having trained a best model for each species according to each method, we applied it to the data in the model-testing data set to estimate the following performance criteria: the false-positive rate, the false-negative rate, precision (the ratio of the correct positive predictions to the total number of positive predictions in the testing data set, also called positive predictive power), and recall (the ratio of the number of correct predictions to the total number of observations in the testing data set). We also calculated a composite measure of accuracy that accounted for both precision and recall (Tax 2005):

$$f_1 = \frac{2(\text{precision})(\text{recall})}{\text{precision} + \text{recall}}$$

Finally, in analyses such as this where maximizing the rate of true-positive classifications comes at the cost of false-positive classifications, it is customary to compute the receiver-operator curve (ROC), which represents the true-positive rate as a function of the false-positive rate (Fielding & Bell 1997). The ROC summarizes this fundamental trade-off and can be used to compare different modelling methods or data sets. However, as ROC for different models can intersect, the area under the curve (AUC) is commonly computed (Bradley 1997). This quantity permits comparison of different studies using a single number. AUC ranges from 0 (incorrect classification of all examples) to 1 (correct classification of all examples).

Results

Overall, we found that methods 1 and 2 performed similarly over the different measures and were superior to method 3. However, consistent models were more often obtained with methods 1 and 3 than with method 2, so that method 1 most reliably provided the best results. We obtained consistent models with method 2 for 58 out of 106 species (54·7%). In contrast, consistent models were obtained with method 1 for 87 (82·0%) species and with method 3 for 80 (75·5%) species. Logistic regression showed that, for all three methods, the likelihood of obtaining a consistent model for any given species increased significantly with the number of observations in the data set (method 1, $P < 0{\cdot}0001$; method 2, $P < 0{\cdot}0001$; method 3, $P = 0{\cdot}0001$; see Figure S2 in the Supplementary material).

Error rates and summary performance criteria were also affected by the number of observations with which we trained the model (Fig. 2; see Figures S3–S8 in the Supplementary material). We computed Spearman rank-order correlation between each measure of accuracy except f_1 (which, in one case, was undefined) and sample size, by pre-processing method, using only species for which we obtained consistent models. We consistently found highly significant relationships (see Table S4 in the Supplementary material). Surprisingly, AUC was not significantly correlated with sample size.

To see if performance differed significantly among protocols, for each measure of performance we used two-way ANOVA with pre-processing method as a fixed effect and individual species' identities as a random effect, using only species for which we obtained consistent models. Using species identity as a factor accounted both for the effect of sample size (which was shown to significantly affect performance) and for differences among species in their ability to be modelled with the observed environmental variables. Not surprisingly, species identity had a significant effect on each measure of performance ($P < 0{\cdot}0001$). There was no evidence for an effect of modelling method on recall ($P = 0{\cdot}248$) or false-negative rate ($P = 0{\cdot}248$), which was the target of optimization and so was expected to be approximately the same across methods. Modelling method did have a significant effect on false-positive rate ($P < 0{\cdot}0001$), precision ($P < 0{\cdot}0001$), f_1 ($P < 0{\cdot}0001$) and AUC ($P < 0{\cdot}0001$). The group mean performances for each pre-processing method across species showed that, where consistent models could be obtained, methods 1 and 2 typically performed similarly and were superior to method 3 (Fig. 2).

To see if accuracy was driven by the idiosyncratic response of different sampling locations, i.e. if some locations were consistently unpredictable while others were consistently predictable, we compared the predictions (method 1) for each species at each sampling location in the testing data with known occurrence, and

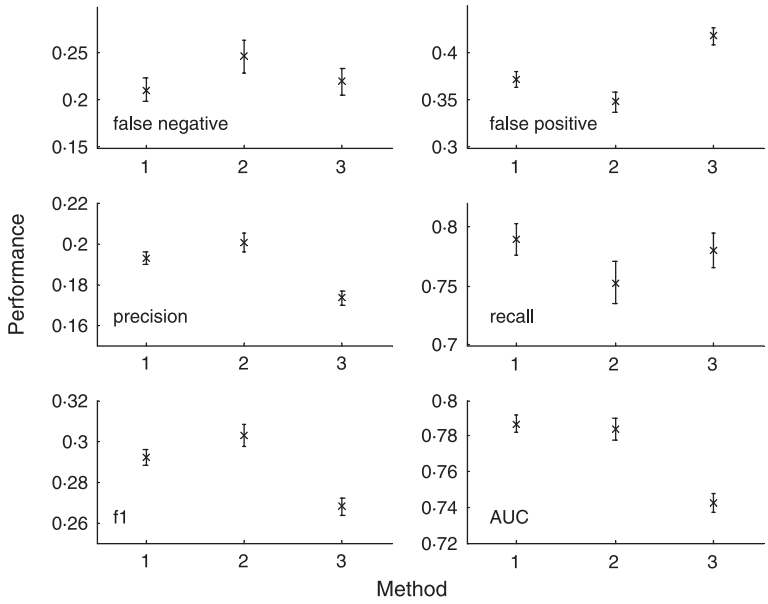


Fig. 2. Means and standard errors of performance criteria by modelling method. Method 1 is for models without pre-processing, using nine environmental variables. Method 2 is for models pre-processed with *k*-whitening, using nine environmental variables. Method 3 is for models using only four environmental variables.

performed two-way ANOVA with species and sampling site as factors. Both effects were highly significant ($P < 0{\cdot}0001$) but explained a small portion of the overall variation ($R^2 = 0{\cdot}203$, partial $R^2_{\text{species}} = 0{\cdot}154$, partial $R^2_{\text{location}} = 0{\cdot}049$).

Discussion

We used a class of recently developed machine-learning algorithms (support vector machines) to model species' distributions using only data concerning species' occurrences. This method assumes only that observations reasonably represent the range of habitable environments; in particular, independence is not assumed. Thus, where data are available only concerning species presence these methods are theoretically superior to classification/discrimination techniques (Hirzel *et al*. 2002). We note that SVM can also be used when there are observations of both habitat and non-habitat (i.e. confirmed absence of species) and indeed can be optimized to make use of as many observations of absence and presence as are available, without requiring balanced observations. We emphasize that the SVM models the support of the statistical distribution of environments from which the species presence observations are drawn, an environmental hyperspace. Thus, the interpretation of the SVM model as an ecological niche is consistent with the classical definition of a niche as a multidimensional environmental space (Hutchinson 1957). Logistic regression (Keating & Cherry 2004), MAXENT (Phillips, Dudík & Schapire 2004), ENFA (Hirzel *et al*. 2002) and other models based on probability densities (Robertson, Caithness & Villet 2001) represent the relative frequency of habitat use and are therefore more closely related to the idea of resource utilization or resource selection (Schoener 1989; Boyce *et al*. 2002; Keating & Cherry 2004).

Of course, theoretical warrant for using support vector machines to model habitat would be unimportant if models performed poorly in independent validation. We used independent observations of species' presences and absences to estimate model accuracy. Summary measures of model performance were generally high. For instance, using our best procedure (method 1) the average AUC obtained was 0·79. For comparison, an analysis of 30 bird species in the Catalan region (Brotons *et al*. 2004) obtained an average AUC of 0·74 on independent data for a model fit using ENFA with only observations of species' presences (Hirzel *et al*. 2002), and 0·82 for logistic regression fit to both species' presences and absences. Zaniewski, Lehmann & Overton (2002) used generalized additive models (GAM) with a logistic link function and binomial distribution fit to both presences and absences of 43 fern species sampled at 19 875 plots in New Zealand to obtain an average AUC of 0·86. Thus, our results are comparable with published results obtained with data for only species presence and data comprising both presence and absence.

We also studied pre-processing approaches that might be taken to increase model performance. Method 1 used no pre-processing or data reduction. Method 2 pre-processed training data using the technique of *k*-whitening. Method 3 relied on a restricted data set in which highly correlated variables were removed from the model training data set. We found that when consistent models could be obtained, method 1 resulted in models with the highest recall and lowest false-negative rate. In contrast, method 2 resulted in models with the highest precision and lowest false-positive rate. Methods 1 and 2 performed similarly as evaluated according to the summary measures f_1 and AUC. In comparison, method 3 performed poorly overall. Consistent models were obtained using method 1 much more frequently than using method 2. Thus, method 1 appears to be the most reliable method in general. Finally, we observed that the relative performance of method 3 compared with methods 1 and 2 indicates that useful information can be obtained by the addition of more environmental variables, even if they are highly correlated.

Finally, we studied how model performance depends on the sample size of the training data set. For 106 species with all three methods we were almost always able to identify a consistent model when the model training data set contained at least 40 observations, which we suggest is the minimum number of observations with which models should be trained in practice. Not unexpectedly, measures of accuracy, such as error rates and precision, were also related to sample size (see Table S4 in the Supplementary material). Minimum sample sizes for modelling and heuristics about how sample size should scale with the number of environmental variables are important topics for research. Curiously, when all species were considered together, AUC was not significantly related to sample size, although the lowest observed AUC scores were always obtained for species represented by fewer than 30 observations (see Figures S6–S8 in the Supplementary material). These results are promising and indicate that often the most accurate models can be obtained with relatively modest data sets. Indeed, models obtained for species with only 40–50 observations routinely performed as well as models for species represented by more than 100 observations. Only precision seemed to increase continuously over the entire range (see Figures S3–S5 in the Supplementary material). These results are about the same as for GARP, which is another machine-learning algorithm and on average obtains near maximal accuracy with 50 observations (Stockwell & Peterson 2002). In contrast, to obtain similar accuracy with logistic regression required 100 observations (Stockwell & Peterson 2002).

An important unanswered question is how many environmental variables are required to predict accurately species' potential distributions, whether with support vector machines or any other technique. In our study, the method with the greatest number of variables (method 1) and no pre-processing provided the best results. An underlying worry is that the higher

dimensionality of this method leads to a model that is overfit and would generalize poorly. We overcame this obstacle by fixing a target error rate and tuning models using cross-validation, which estimates the generalization error directly. Thus, our comparison of the different methods was designed to create the fairest comparison: each method was optimized to achieve (approximately) the same generalization error. Three lines of evidence point to success at achieving this fair comparison. First, we failed to detect an effect of modelling method on the false-negative rate when the model was tested with independent data. Thus, the true generalization error was consistent across methods. Secondly, if correlation among environmental variables had led to overfitting, method 2 would have performed best as the algorithm would only have been trained on the information contained in the first few principal components of the data. Finally, in image-recognition experiments (classifying digital images of hand-written numerals) Tax & Müller (2004) found that the optimized model was sometimes not complex enough when non-target observations (i.e. species' absences) were too close to the training data. Therefore, if anything, there is reason to suspect that our models are underfit rather than overfit. Indeed, simply including more environmental variables, rather than developing more sophisticated ways of reducing dimensionality, might result in the greatest improvements to accuracy. Our analysis was limited by the availability of relevant systematically collected data that had been geo-referenced to the particular sampling sites where our species' distribution records were collected. Future studies could certainly include many more variables as the computational cost that would be imposed is minor. Indeed, we suggest that the computational complexity of the SVM approach is one of its primary features. This aspect could be exploited in many ways that await development. Some obvious possibilities are that different 'submodels' obtained from subsets of the data corresponding to classes of environmental variables (biotic vs. abiotic, chemical vs. physical, etc.) could be compared to explore how these differently affect species' distributions; modelling could be embedded in a non-parametric bootstrap to obtain confidence bounds on the estimated distributions; and resampling schemes could be devised to test hypotheses about niche differentiation, partitioning and competitive exclusion or facilitation. These possibilities, together with the relatively strong performance already shown by this approach, should motivate further research, resulting in both methodological improvements and applications in many areas.

Species distribution modelling is a part of many ecological applications, including forecasting species invasions, devising protocols for biodiversity monitoring, designing nature reserves and planning for habitat conservation, managing vector-borne and environmentally mediated disease, and cultivating renewable resources (e.g. aquaculture and timber). Finally, species distribution modelling is often a fundamental component of projects aimed at understanding the consequences of anthropogenic climate change, such as the MODIPLANT initiative that generated the data used in this study. As SVM are stable algorithms that can deal with large sets of predictors at once, they may prove particularly useful in this arena. In conclusion, these results support the continued use of SVM for ecological niche modelling. Where data are available concerning only species' presences and not species' absences, support vector machines are theoretically superior to classification techniques that rely on simulation of pseudo-absence data. We have shown that support vector machines are also comparable with such models when validated by independent observations of both species presence and absence.

Acknowledgements

We thank Niklaus E. Zimmermann, Felix Kienast and Massimiliano Zappa for making their climate maps available. We additionally thank Stéphanie Maire, Dario Martinoni, Pascal Vittoz, Simone Peverelli and many others for their help with the field sampling of species data. Comments by A. T. Peterson and an anonymous referee resulted in a considerably improved paper. This work was conducted while J. M. Drake was a Postdoctoral Associate at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant No. DEB-94-21535), the University of California at Santa Barbara, and the State of California. The Swiss Office for the Environment (Climate Change Section) and the Conservation Centre of the Canton de Vaud both provided funding to support the field work.

References

Anderson, R.P., Lew, D. & Peterson, A.T. (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, **162**, 211–232.

Anonymous (2004) *ArcInfo*, Version 9·0. Environmental Systems Research Institute Inc., Redlands, CA.

Augustin, N.H., Mugglestone, M.A. & Buckland, S.T. (1996) An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology*, **33**, 339–347.

Boyce, M.S., Vernier, P.R., Nielsen, S.E. & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, **157**, 281–300.

Bradley, A. (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, **30**, 1145–1159.

Brotons, L., Thuiller, W., Araújo, M.B. & Hirzel, A.H. (2004) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.

Dirnböck, T., Dullinger, S. & Grabherr, G. (2003) A regional impact assessment of climate and land-use change on alpine vegetation. *Journal of Biogeography*, **30**, 401–417.

Drake, J.M. & Bossenbroek, J.M. (2004) The potential distribution of zebra mussels in the United States. *Bioscience*, **54**, 931–941.

Duin, R.P.W., Juszczak, P., Paclik, P., Pekalska, E., de Ridder, D. & Tax, D.M.J. (2004) *Prtools4, a Matlab Toolbox for Pattern*

Recognition. Delft University of Technology, Delft, the Netherlands.

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.

Ferrier, S. (2002) Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology*, **51**, 331–363.

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.

Graham, C.H., Ron, S.R., Santos, J.C., Schneider, C.J. & Moritz, C. (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781–1793.

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.

Guo, Q., Kelly, M. & Graham, C.H. (2005) Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling*, **182**, 75–90.

Gurtz, J. *et al.* (2003) A comparative study in modelling runoff and its components in two mountainous catchments. *Hydrological Processes*, **17**, 297–311.

He, F., Zhou, J. & Zhu, H. (2003) Autologistic regression model for the distribution of vegetation. *Journal of Agricultural, Biological and Environmental Statistics*, **8**, 205–222.

Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, **157**, 331–341.

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

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

Keating, K.A. & Cherry, S. (2004) Use and interpretation of logistic regression in habitat-selection studies. *Journal of Wildlife Management*, **68**, 774–789.

Körner, C. (2003) *Alpine Plant Life*. Springer, Berlin, Germany.

Kumar, L., Skidmore, A.K. & Knowles, E. (1997) Modelling topographic variation in solar radiation in a GIS environment. *International Journal for Geographical Information Science*, **11**, 475–497.

Leathwick, J.R. & Austin, M.P. (2001) Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, **82**, 2560–2573.

Lusk, J.J., Guthery, F.S. & DeMaso, S.J. (2002) A neural network model for predicting bobwhite quail abundance in the Rolling Red Plains of Oklahoma. *Predicting Species Occurrences: Issues of Accuracy and Scale* (eds J.M. Scott, P.J. Heglund, M. Morrison, M. Raphael, J. Haufler, B. Wall & F.B. Samson), pp. 345–355. Island Press, Covello, CA.

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K. & Schölkopf, B. (2001) An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, **12**, 181–202.

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.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology*, **78**, 419–433.

Peterson, A.T., Sanchez-Cordero, V., Beard, C.B. & Ramsey, J.M. (2002) Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerging Infectious Diseases*, **8**, 662–667.

Phillips, S.J., Dudík, M. & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the Twenty-First International Conference on Machine Learning*. ACM Press, New York, NY.

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

Robertson, M.P., Caithness, N. & Villet, M.H. (2001) A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. *Diversity and Distributions*, **7**, 15–27.

Schoener, T.W. (1989) The ecological niche. *Ecological Concepts* (ed. J.M. Cherrett), pp. 79–113. Blackwell Scientific Publications, Oxford, UK.

Schölkopf, B. & Smola, A. (2001) *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA.

Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J. & Williamson, R.C. (2001) Estimating the support of a high-dimensional distribution. *Neural Computation*, **13**, 1443–1471.

Scott, J.M., Heglund, P.J., Morrison, M., Raphael, M., Haufler, J., Wall, B. & Samson, F.B. (2002) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Island Press, Covello, CA.

Stockwell, D. & Peters, D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, **13**, 143–158.

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

Tax, D.M.J. (2001) *One-class classification: concept-learning in the absence of counter-examples*. Thesis. Delft University of Technology, Delft, the Netherlands.

Tax, D.M.J. (2005) *Data Description Toolbox (Manual) V·1·1·2*. Delft University of Technology, Delft, the Netherlands.

Tax, D.M.J. & Duin, R.P.W. (2004) Support vector data description. *Machine Learning*, **54**, 45–66.

Tax, D.M.J. & Juszczak, P. (2003) Kernel whitening for one-class classification. *International Journal of Pattern Recognition and Artificial Intelligence*, **17**, 333–347.

Tax, D.M.J. & Müller, K.-R. (2004) A consistency-based model selection for one-class classification. *Proceedings 17th International Conference on Pattern Recognition (22–26 August 2004, Cambridge UK)* (eds J. Kittler, M. Petrou & M. Nixon), pp. 363–366. IEEE Computer Society, Los Alamitos, CA.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., Ferreira de Siqueira, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, 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**, 145–147.

Zaniewski, A.E., Lehmann, A. & Overton, J.M.C. (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, **157**, 261–280.

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

Received 17 June 2005; final copy received 14 November 2005
Editor: Rob Freckleton

Supplementary material

The following supplementary material is available as part of the online article (full text) from http://w.w.w.blackwell-synergy.com.

Table S1. Species for which ecological niche models were trained using support vector machines (n = 106).

Table S2. Predictor variables used to model ecological niches.

Table S3. Pearson correlations for environmental variables used to model ecological niches ranked by the absolute value of the correlation coefficient.

Table S4. Spearman rank-order correlations between performance criteria and sample size.

Figure S1. Rescaled histograms of nine predictor variables used to model ecological niches.

Figure S2. Frequency with which consistent models could be obtained.

Figure S3. Performance of method 1.

Figure S4. Performance of method 2.

Figure S5. Performance of method 3.

Figure S6. Summary measures of performance for method 1.

Figure S7. Summary measures of performance for method 2.

Figure S8. Summary measures of performance for method 3.

Chapter 10

A shift in niche conservatism:

From sacred cow to testable hypothesis

For review in Trends in Ecology & Evolution

P.B. Pearman*, A. Guisan*, O. Broennimann, C.F. Randin

University of Lausanne, Department of Ecology and Evolution, CH-1015 Lausanne, Switzerland

Email: <peter.pearman@unil.ch>, <antoine.guisan@unil.ch>

* Authors contributed equally, others in order of authorship

Abstract

Niche conservatism is often assumed when discussing, explaining or predicting biogeographic patterns in ecology and evolution. This assumption is implicit when species distribution models are used to predict extinction rates and biodiversity loss under climate change, or to predict the risk and spread of biological invasions. Nonetheless, studies in evolutionary ecology demonstrate that niches can diverge among populations and species, and show other patterns as well. These niche changes could impact predictions of biodiversity loss and species invasions that are obtained from species distribution models.

At the species level, niche shift may arise from a process of niche drift, which is random with respect to phylogenetic tree topology and can serve as a null model. In contrast, the relationship of tree nodes to niche characteristics of species may show signatures of selection manifest in community organization. We show that niche conservatism provides an additional null hypothesis, with the alternative being a niche shift.

Confusion concerning niche shifts, lack of their consideration, and failure to detect them may owe to many factors, including insufficient emphasis on appropriate hypothesis testing.

Here, we develop a hypothesis testing framework for diagnosing niche conservatism and shifts. This framework may help to improve foundations for the use of species distribution models in ecology and evolution.

Introduction

The idea that some aspects of species' niches may be immutable over long periods [1-3] influences the current development of ecology and evolution, and has generated considerable debate among ecologists and evolutionary biologists [4]. For example, Peterson's seminal 1999 paper [2] provides evidence for niche conservatism (NC, the absence of niche change or a constraint in rate of niche change) regarding climatic variables during allopatric speciation.

This result, and others on the stability of climatic characteristics of the niche, is widely used to support diverse application of species distribution models (SDM's). In this way, NC plays an influential and high-impact roll in development of biogeographical understanding of species diversification and geographic distribution [5-11], forecasting species extinction and biodiversity loss in response to climate change [12-15] and predicting establishment and spread of invasive species [16-20]. Understanding the occurrence and frequency with which ecological systems fall along the spectrum between niche conservatism and complete shift is crucial; if niches are not static over time and/or space, then species can potentially colonize new niches. These are likely not predictable from contemporary environmental data, contemporary distribution, or native range and may weaken or invalidate predictive niche-based models. Recent findings suggest that, due to ecological and/or evolutionary processes [21-23], niche shifts occur in previously unrecognized situations [6,9,24]. The conditions under which the NC assumption is violated or, perhaps more importantly, holds still lack adequate examination. Consensus among evolutionary biologists, biogeographers, and ecologists regarding the validity of assuming NC for predictive modelling is still needed.

Our goal is to contribute to the discussion of niche conservatism as a basis for predictive modelling. We expand on recent work to help clarify how the idea of niche conservatism influences contemporary research programs in ecology and other integrative sciences. We develop a foundation for consensus by building on theory and tests of the NC null hypothesis, and the niche shift alternative. We (1) review recent work on understanding implications of differing niche definitions for studying NC (different theoretical definitions, potential heuristic definitions), (2) develop the role of hypothesis testing and statistical power in addressing NC and identify factors that may affect detection of niche shifts and their quantification (scale and resolution), and finally (3) we discuss the implications of decisions on existence (or absence) of NC for predicting species invasions and species' response to climate change. We do this by asking: (1) How have NC and shifts been addressed? (2) How can we recognize a niche shift when we see one? (3) What are the implications of decisions and statistical power on the absence, detection, and magnitude of niche shifts? We pay particular attention to the role of niche-based species distribution modeling in studying niche shifts and their potential effects on attempts to predict effects of climate change and risk of biological invasion. We also identify potentially productive research directions.

Niche conservatism in the literature

Defining 'niche'

The lack of consensus on NC is likely related to insufficient or vague operational definition of niche [25,26] and of the theoretical aspects of NC to be tested (e.g. [2]). Several definitions of the niche are in use. Leibold [26] juxtaposes two niche concepts, one driven by the environmental *requirements* of a species (as defined by Grinnell and Hutchinson), the other by the *impact* that the species can have on its environment (as defined by Elton, MacArthur or Levins). These two concepts have distinct theoretical bases. The former, often called the *environmental niche* (see [27]), is embedded within an autecological and ecophysiological approach to the niche. The latter definition, often called the *trophic niche* (see [28,29]), addresses resource consumption, trophic levels and food web theory (i.e. the role and position of a species in a biocenose). Within these two groups, one may distinguish between Grinnell's view of a species occupying all of its suitable habitats the *fundamental niche*, and Hutchinson's view of a species being excluded from a part of its fundamental niche by biotic interactions. The resulting *realised niche* is what is actually observed in nature (Figure 1a) [29,30]. Populations that are outside of the realized niche fail to maintain positive population growth rates [31].

The distinction between fundamental and realized niches also applies to the trophic niche. A change in realized trophic niche occurs when a species experiences a change in trophic position resulting, say, when a competitor in a foodweb is missing in a given community. It remains unclear whether the realized niche should be excised from the fundamental niche using constraining, negative interactions, or whether positive interactions modulating the response to environmental or other biotic factors (e.g. facilitation, mutualism) impact the fundamental niche. Considering such positive interactions in the niche framework is a crucial yet unaddressed question [32]. These distinctions among niche concepts are particularly important in the context of testing for NC. A putative niche shift may result from a change in the realized niche only, for instance due to change in biotic interactions (e.g. release from natural enemies) expanding the realized niche (Fig. 1). This may result, for example, in differences in observed niche of a species in its native and invaded ranges, without change in the species' fundamental niche [33]. Similarly, absence of an observed change in environmental niche does not imply stasis in the trophic niche, because changes along unquantified niche dimensions are always possible.

Niche conservatism: the evidence

A wide range of studies support the existence of niche conservatism. Comparison of the effectiveness of environmental niche models in reciprocally predicting the distribution of sister taxa and non-sister taxa suggests that niches tend to be conserved during speciation [2] (but see [5]). Niche comparisons at the continental or intercontinental scales have supported the existence of niche conservatism in the genus *Fagus*, Pleistocene mammals, Argentine ants and in herbaceous plants [3,14,34-36]. These studies compare the ecological aspects of species distributions at different points in time [14] or across continents [3,34-36], and have concluded that significant similarity in the distributions (geographical or in environmental space) of pairs or larger groups of species indicates niche conservatism. The migration of tree species in response to

Quaternary climate change suggests that it is easier to migrate than to adapt (despite the potential presence of some level of local adaptation [37]). Additionally, the positive correlation between areas of high species richness and high speciation rates in three clades of plankton has suggested niche conservatism [38,39]. One theoretical result is that natural selection should generally favor adaptation to source habitats because higher fitness in source habitats inside the realized niche and low rates of dispersal from sink to source habitats lead to little selection for adaptation to sink habitats [40]. Niche conservatism occurs when the rate of adaptation of sink populations is slower than their rates of extinction, which depend primarily on the magnitude of gene flow [2]. Thus, niche conservatism is expected and may occur in a range of systems (Table.1).

Niche shifts: the evidence

Theoretical evidence. Theoretical support and empirical evidence for niche shifts are found in diverse fields of ecology and evolution. Some models of speciation suggest rapid niche change [41] and newly evolved species may be associated with occupancy of new environments that initially act as demographic or evolutionary 'sinks' [40,42]. While strong differences in fitness between sink and source habitats favor niche conservatism [43,44], weak differences, temporally correlated variation in fitness within sinks, and geneflow from source to sink may, under some restricted conditions, promote niche expansion [43]. The effects of dispersal and selection interact to determine whether the niche (and potentially the range) of a species contracts or expands rapidly [45] (see review [46]). For example, local processes such as predation and competition may in theory influence the evolution of host shifts and species radiations [47]. Species interactions in food webs can promote parallel niche shifts and niche convergence [47] (see empirical examples [48,49]). In summary, theory has described ecological and evolutionary processes that result in niche shifts.

Empirical evidence. Empirical evidence for niche shifts comes from studies of invasive species ecology [33,50], biological response to climate change [51], phylogenetic analysis [6,10,52], and community ecology [29,53]. Host shifts and host race formation (trophic niche) occur in sympatric populations and sister species of insects, and may be reinforced by predation and fitness differences on the different hosts [54-56]. Similarly, the occurrence of host shifts and adaptation of pathogens is well-documented [57,58]. Rapid trophic niche evolution and associated morphological character change has been observed in sympatric finches [59]. Insects may expand the range of habitats they use and evolve increased rates of dispersal in response to climate change [12]. Plant populations in several classic studies expand in edaphic niche in response to gradients of the concentration of toxic substances in the soil, because strong selection for tolerance outweighs the homogenizing effects of gene flow [60-63]. Niche breadth may be maintained through interplay of selection and gene flow in other systems as well [53,63,64]. Niche shifts may also occur during a species' invasion of new habitats and geographical areas by plants [21,31,33] and by animals [50] (Figure 1b). For example, in the invasions of fire ants [50] and spotted Knapweed [33] in North America, niche evolution is likely a factor enabling successful invasion. Niche shifts may be facilitated by evolution of new life history strategies [65]. Thus, the niche is perhaps best viewed as being *potentially* evolutionarily labile, rather a fixed property of a species (Figure 1c).

Experimental evidence. Niche shifts have also been studied experimentally. Invasive exotic species may undergo rapid adaptation and population differentiation, as shown

in both field [66] and laboratory [67] experiments. To some degree, this intraspecific variation in niche characteristics can be captured by niche modelling (but see [68]). For example, Wright et al. [69] conduct a sampling program on a regular grid to examine the presence/absence of a California annual plant, *Collinsia sparsiflora*, in an area of patchy serpentine soils. They use five environmental variables that when incorporated into niche-based SDM's [70] provide predictive information on local species distribution. The authors then report using a randomized block design, F2 seeds from four other populations, and the probabilities of species presence from the SDM's to predict the probability of experimental plants survival from planting to flowering. Wright et al. find that their models (which unavoidably are based only on presence/absence of *C. sparsiflora* on serpentine soils) predict the flowering performance of two nearby populations from serpentine soil, and the distribution of four unrelated syntopic species. The models do not, however, provide predictive information on the performance of two other populations representing the non-serpentine ecomorph of *C. sparsiflora*. Wright et al. [69] also show that the SDM's sufficiently describe the serpentine ecomorph's fundamental niche by including an additional treatment in which potential competitors of *C. sparsiflora* are removed. These results are consistent with those of a large-scale transplant experiment involving *Pinus contorta* that demonstrated populations are locally adapted to climatic conditions, but that individuals within each population are excluded from optimal environments by competitors [71]. Experimental approaches are required for answering questions about the relationship between the realized and fundamental niche.

Phylogenetic evidence. Work in phylogenetics and evolutionary ecology, coupled with niche-based modelling, identifies niche changes at multiple spatial and temporal scales [6,9,24]. Studies analyze the degree of niche change within clades by combining phylogenetic hypotheses with niche information acquired from morphology, behavior, habitat preferences, or climatic variables [5] (Box 1). For example, Knouft et al. [6] examined the degree to which phylogenetic similarity parallels niche similarity in Cuban *Anolis* lizards. These authors. recognize that a negative correlation between evolutionary distance and niche similarity would suggest that niches tend to be conserved within a lizard clade. They report use of a pruned phylogenetic tree of eleven Cuban *Anolis* species to calculate the sum of the branch length distances (patristic distance) between pairs of individuals, each representing a distinct species. Knouft et al. then conduct variable reduction of 19 climatic variables using principal components analysis (PCA) and determine the species' niche similarity in terms of climatic niche dimensions. The authors. find no significant correlation between niche similarity and phylogenetic distance, indicating no detectable trend towards niche conservatism, in the sense of restricted niche divergence, in this group. Thus, niche shifts have been found using multiple approaches and in various systems.

Testing for niche conservatism/shift

Niche-based species distribution models

While the use of field experiments has enabled study of changes in both the fundamental and realized niches of species [66-69,71], niche-based species distribution models [70] are increasingly used to address questions concerning niche conservatism in natural systems and at spatial scales that make experimentation infeasible (e.g., [6,33,50]). One way to model and test for potential niche shifts in response to climate

change or speciation events is to predict past species distributions from models fitted under current climate [5,14,72,73] (hindcasting), or similarly to predict current distributions from models fitted using fossil records (forecasting) [74,75]. Results from these studies are mixed, with some studies claiming support for niche conservatism [14,76] and others able to quantify the relationship between performance of distribution models and observed niche shifts [51]. Similarly, niche-based species distribution models have contributed to understanding changes in the realized niche during invasion by exotic species [33,50] (but see [36]) and niche shifts associated with species diversification [5,6,10] (but see [2]). Niche models and other physiological and ecological characteristics of species, when used in combination with hypotheses on phylogenetic relationships, allow tests of niche conservatism and shift at multiple levels within phylogenies (above). These approaches deserve further development.

Diagnosing niche shift/conservatism

How can we better diagnose a niche shift or confirm niche conservatism? This is a question that needs resolution in such a way that it holds for all definitions of 'niche', fundamental, realized, or another operational conceptualization. No one set of essential criteria exists for confirming the existence of niche shifts, given the variation in taxonomic level, temporal scale and geographic scope in which testing for NC or shift has been explored. While the niche definition and study system may demand a particular approach to, there are likely several requisites when using SDM's and other methods (Box 2). First, testing for either niche shift or conservatism requires the alternative situation be defined a priori. The criteria for rejecting the null condition must exist prior to conducting a statistical test. For example, testing for NC over time requires specification of how much shift is acceptable before NC is rejected. Second, the sources of relevant uncertainty need to be acknowledged and quantified. Uncertainty due to random or systematic errors may come from niche modelling based on incomplete species distribution data [77], variation in the efficiency of modelling algorithms, uncertainty in the evolutionary relationships among taxa, or sampling error (Figs. 2, 3). Understanding of the potential for decision error (Type I and Type II) must be established in order for an outcome to be based on valid statistical inference (Fig. 3). These considerations place testing for niche conservatism/shift using SDM's within a standard framework of statistically based decision-making.

Hypothesis testing framework

Given that there are a number of ways to misdiagnose NC, a valid hypothesis testing framework will help make tests of NC/shift more rigorous. For example, one possibility is to expect that ecological similarity decreases with increasing time since species divergence. Observation of a significantly positive level of ecological or range similarity between two species or the same species at two times or in two place [14,34,76] is likely insufficient in itself to conclude that NC has occurred. This is because random levels of association between niches across time are not the null expectation. In a useful example, Losos et al. test for niche conservatism by determining whether *Anolis* species are less distant ecologically than predicted given their time since divergence [24]. Losos et al. establish a null expectation of a non-zero rate of divergence due to random processes (*niche drift*). Thus, the existence of niche conservatism may be supported by observation of significantly less divergence than would occur randomly over time, the significance of which is established through a randomization test.

Positive evidence for niche conservatism may in some cases require rejecting the alternative of some rate of divergence that is based on a null-model. In contrast, omission of some sources of uncertainty, such as in phylogenetic tree topology, may lead to inaccurate significance levels.

Another issue arises when using niche-based distribution modelling to predict invasion risk and potential range extent of introduced, potentially invasive species. How much difference between actual and predicted range is sufficient to support the existence of a niche shift in an invasive species? For invasive species, change in niche position leads to increasing discrepancy between the predicted and actual ranges [51]. Thus, a potential null hypothesis is that niche conservatism exists (no niche change), with the alternative being that a niche shift has occurred. In this approach, niche shifts will only be detected statistically when there is sufficient data on species occurrences. Non-rejection of the null-hypothesis of NC could be the result of insufficient statistical power and resulting Type II error [78]. When the null hypothesis of NC is not rejected, concluding for NC will require development of a priori analyses of statistical power (Figure 3). The computer simulation of statistical power and application of resampling techniques likely have unfulfilled roles in testing null hypotheses of niche conservatism. Specification of how much niche shift is biologically relevant (the 'effect size' in an hypothesis-testing framework) will be an important part of power analyses and subsequent statistical tests of NC.

Detection error

Why and when might niche shifts go unrecognized? The preceding discussion indicates that insufficient statistical power or ill-formed hypotheses are two potential reasons for failure to identify a particular niche shift. Nonetheless, additional reasons are plentiful. For example, in niche-based species distribution models, a model of a species' climatic niche in the native range may be used to predict the climatic niche and range of the species where it has invaded, or vice versa (e.g., [33,50]). If the model minimizes the rate of omission errors without controlling the rate of commission errors, the application of the model could lead to the conclusion that no niche shift has occurred.

Another situation that may obscure niche shifts happens when the climate niche of a species is modeled over the entire species. Calibration of a model using data from one part of the range may result in a model that does not predict well in another part. While one strategy to fix this is to calibrate models based on pooled data across all parts of the range, this assumes that the cause of model deficiency is insufficient capture of the entire climatic niche (Fig. 2b) [79]. However, a species may show local adaptation to climate in one part of the range (e.g. ecotypes), for example through divergence in life history strategies [80]. This corresponds to initiation of a niche shift, and perhaps ultimately to ecological speciation [81,82]. Hence, various levels of success in predicting a species' invasion could be attained depending on which native population, or selection of populations, are used to fit the models. In the context of invasive species, such information can only be obtained through a well documented invasion history or *a posteriori*, through genetic identification of the source population(s) in the native range. Finally, it is unclear whether the restricted distribution of a clade to a particular range of conditions (e.g., tropical) indicates an active process of climatic niche conservatism (as tested for in [6]) or some other effect, such as dispersal or other evolutionary limitations.

Indeed, ecological or evolutionary tradeoffs might be the drivers of observed niche conservatism. More research is needed on niche shifts/NC in this context.

The ubiquity and influence of the niche conservatism/shift dichotomy in integrative biology has led to diverse scales on which to approach the issue of niche conservatism and change[4], For instance when using SDM's [2], valid tests for niche conservatism during climate-induced range changes will require careful consideration of spatial resolution and size of study area (Figure 2a). For example, in analysis of a potential niche shift during species invasion, artifacts may arise from partial niche fitting in one or both ranges [79,83]. To avoid this, sampling must capture the full distribution extent of the species in each range and GIS environmental maps must be available for the whole study area to allow capturing the full extent of environmental (or trophic) conditions (Figure 2b and [33]). In the case of niche shift following biological invasion, it is important to ensure that the novel conditions experienced in the newly colonized range (i.e., from a niche shift) also occur in the native range, but were not colonized there. Otherwise, an apparent shift might result from the colonization of environmental conditions to which the species was pre-adapted but which did not occur in the native range [22,33]. Overall, the effect of developing models based upon the distribution of the genetically-identified population(s) in the native range from which the invaders come is not fully understood and more research is needed in this area.

Implications of niche conservatism versus niche shift

Niche conservatism has influenced the development of ecology [4]. First, prediction and observation of NC form prerequisites for subsequent development of the niche concept [35]. The niche concept's development and elaboration would not have occurred without some level of stasis of ecological requirements of species. Similarly, biogeographic trends would not exist because species ecological preferences and requirements would constantly and randomly evolve. But while (conserved) niches may distinguish many species, niche evolution is associated with evolutionary novelty and speciation [9]. Niche conservatism and niche evolution are both aspects of evolving life. Thus, it is not surprising to observe NC in some cases (e.g. allopatric speciation [2]) and niche evolution in others (e.g. invasive species [33,50,66,67]).

From the perspective of applied ecology, trends toward NC are critical for the use of species distribution models (SDM's) to predict impacts of climate change on biodiversity and the spread of invasive species. If NC were generally weak or absent, the validity of model-based projections would be questionable [33]. Niche conservatism may indeed remain a valid assumption in projections of impacts of rapid climate change on some species' distributions because populations on the trailing edges of ranges may undergo unique evolutionary processes and be prone to extinction [84,85]. However, the situation regarding spread of exotic species may be different. This phenomenon is of recent origin, being associated with human trade and global transportation. We have few data from past invasions with which to evaluate NC. Further, the documented potential for rapid adaptation and niche evolution in invading species suggests that the assumption of NC is often violated [33,50,66,67]. Thus, the full invasion potential of a species cannot be predicted when post-establishment niche shifts are likely [33,50]. Nonetheless, SDM's may produce informative predictions on candidate areas for successful introduction and establishment of exotic species [33].

Conclusions

This review supports six key points:

- Niche shifts occur in numerous situations.

- Support for occasional niche conservatism exists at several scales.

- Consideration of the consistency between phylogenetic and ecological patterns has provided a null model (niche drift) in association with speciation events.

- Workers who propose SDM tests for NC need to consider the resolution and extent with which data are collected.

- When using SDM's, authors should identify grounds for assuming niche conservatism and specify thresholds for niche shift magnitude, beyond which predictions resting on the NC assumption become misleading or require revised interpretation.

- The framework developed here (Box 2; figs. 1, 3) should facilitate definition, identification and testing for NC and niche shifts.

Future research directions

Advance in understanding, predicting, and taking advantage of NC will require further multidisciplinary work in ecology and evolution, such as coupling field experiments, phylogenetics, population genetics, and SDM's. Further empirical work is needed to determine the degree to which niche shifts may draw into question predictions of species distribution changes and/or extinction as a function of climate change. Identification of the species likely to experience niche conservatism will improve the reliability of predictions of the effects of climate change on biodiversity. Understanding of the relationship between fundamental and realized niches may illuminate the degree to which changes in realized niche generally reflect similar changes in the fundamental niche. Additional experiments are necessary to determine how much the realized niche, inferred from field data, is a valid estimate of the fundamental niche of species. Do realized and fundamental niches generally coincide in competitively dominant species but not in subordinate ones [86]? Research needs identify whether there is a 'force' of niche conservatism that operates against developing ecological differences between related species, perhaps through evolutionary constraints or ecological tradeoffs. We should seek to understand factors promoting niche shifts, to quantify niche shifts at multiple taxonomic levels, and to adapt existing methods (e.g. predictive modelling) to cases where niche conservatism is violated. Resolution of these issues will contribute to understanding the reliability of predictions of species responses to climate change [85] and predictions of risk of establishment of exotic species [33].

Acknowledgements

We thank C. Parisod and N. Salamin, and xxxx anonymous reviewers who contributed to improving the manuscript. This research benefited from support by the Swiss National Science Foundation (grant nr 3100A0-110000) and the European commission (MC-HOTSPOTS and FP6-ECOCHANGE projects).

References

1. Prinzing, A.*, et al.* (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.* 268, 2383-2389
2. Peterson, A.T.*, et al.* (1999) Conservatism of ecological niches in evolutionary time. *Science* 285, 1265-1267
3. Huntley, B.*, et al.* (1989) Climatic control of the distribution and abundance of beech (*Fagus* L) in Europe and North America. *J. Biogeogr.* 16, 551-560
4. Wiens, J.J., and Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution and Systematic* 36, 519-539
5. Graham, C.H.*, et al.* (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58, 1781-1793
6. Knouft, J.H.*, et al.* (2006) Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. *Ecology* 87, 29-38
7. Hawkins, B.A.*, et al.* (2006) Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. *Journal of Biogeography* 33, 770-780
8. Hugall, A.*, et al.* (2002) Reconciling paleodistribution models and comparative phylogeography in the wet tropics rainforest land snail *Ganrosophia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences of the United States of America* 99, 6112-6117
9. Levin, D.A. (2005) Niche shifts: The primary driver of novelty within angiosperm genera. *Systematic Botany* 30, 9-15
10. Rice, S.H.*, et al.* (2003) Ecological niche differentiation in the *Aphelocoma* jays: a phylogenetic perspective. *Biological Journal of the Linnean Society* 80, 369-383
11. Levin, D.A. (2006) Flowering phenology in relation to adaptive radiation. *Systematic Botany* 31, 239-246
12. Thomas, C.D.*, et al.* (2004) Extinction risk from climate change. *Nature* 427, 145-147
13. Thuiller, W.*, et al.* (2005) Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8245-8250
14. Martinez-Meyer, E.*, et al.* (2004) Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology & Biogeography* 13, 305-314
15. Peterson, A.T., and al. (2002) Future projections for Mexican faunas under global climatic change scenarios. *Nature* 416, 626-629
16. Ley, R.E.*, et al.* (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837-848
17. Peterson, A.T., and Vieglais, D.A. (2001) Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem. *Bioscience* 51, 363-371
18. Reichard, S.H., and Hamilton, C.W. (1997) Predicting invasions of woody plants introduced into North America. *Conserv. Biol.* 11, 193-203
19. Thuiller, W.*, et al.* (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11, 2234-2250
20. Dirnbock, T.*, et al.* (2003) Predicting future threats to the native vegetation of Robinson Crusoe Island, Juan Fernandez Archipelago, Chile. *Conserv. Biol.* 17, 1650-1659
21. Dietz, H., and Edwards, P.J. (2006) Recognition that causal processes change during plant invasion helps explain conflicts in evidence. *Ecology* 87, 1359-1367
22. Ackerly, D.D. (2003) Community assembly, niche conservatism, and adaptive evolution in changing environments. *Int. J. Plant Sci.* 164, S165-S184
23. Davis, M.B.*, et al.* (2005) Evolutionary responses to changing climate. *Ecology* 86, 1704-1714
24. Losos, J.B.*, et al.* (2003) Niche lability in the evolution of a Caribbean lizard community. *Nature* 424, 542-545
25. Chase, J.M., and Leibold, M.A. (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. The University of Chicago Press
26. Leibold, M.A. (1995) The niche concept revisited: mechanistic models and community context. *Ecology* 76, 1371-1382
27. Austin, M.P. (1992) Modelling the environmental niche of plants: implications for plant community response to elevated CO2 levels. *Aust. J. Bot.* 40, 615-630
28. Austin, M.P.*, et al.* (1990) Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monographs* 60, 161-177

29. Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution* 19, 605-611
30. Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters* 3, 349-361
31. Holt, R.D.*, et al.* (2005) Theoretical models of species' borders: single species approaches. *Oikos* 108, 18-27
32. Thuiller, W.*, et al.* (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America* 102, 8245-8250
33. Broennimann, O.*, et al.* (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters* in press
34. Ricklefs, R.E., and Latham, R.E. (1992) Intercontinental correlation of geographical ranges suggests stasis in ecological traits of relict genera of temperate perennial herbs. *American Naturalist* 139, 1305-1321
35. Hoffmann, M.H. (2005) Evolution of the realized climatic niche in the genus *Arabidopsis* (Brassicaceae). *Evolution* 59, 1425-1436
36. Roura-Pascual, N.*, et al.* (2006) Niche differentiation and fine-scale projections for Argentine ants based on remotely sensed data. *Ecological Applications* 16, 1832-1841
37. Davis, M.B., and Shaw, R.G. (2001) Range shifts and adaptive responses to Quaternary climate change. *Science* 292, 673-679
38. Allen, A.P., and Gillooly, J.F. (2006) Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters* 9, 947-954
39. Allen, A.P.*, et al.* (2006) Kinetic effects of temperature on rates of genetic divergence and speciation. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9130-9135
40. Holt, R.D. (1996) Demographic constraints in evolution: Towards unifying the evolutionary theories of senescence and niche conservatism. *Evolutionary Ecology* 10, 1-11
41. Orr, M.R., and Smith, T.B. (1998) Ecology and speciation. *Trends in Ecology & Evolution* 13, 502-506
42. Goldberg, E.E.*, et al.* (2005) Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *American Naturalist* 165, 623-633
43. Holt, R.D., and Gomulkiewicz, R. (2004) Conservation implications of niche conservatism and evolution in heterogeneous environments. In *Evolutionary Conservation Biology* (Ferrière, R.*, et al.*, eds), pp. 244-264, Cambridge University Press
44. Antonovics, J.*, et al.* (2001) Spatially explicit studies on the ecology and genetics of population margins. In *Integrating Ecology and Evolution in a Spatial Context* (Silvertown, J., and Antonovics, J., eds), pp. 97-116, Blackwell
45. Kirkpatrick, M., and Barton, N.H. (1997) Evolution of a species' range. *American Naturalist* 150, 1-23
46. Bridle, J.R., and Vines, T.H. (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution* 22, 140-147
47. Abrams, P.A. (2000) Character shifts of prey species that share predators. *American Naturalist* 156, S45-S61
48. Brown, J.M.*, et al.* (1995) The role of natural enemy escape in a gallmaker host-plant shift. *Oecologia* 104, 52-60
49. Cavender-Bares, J.*, et al.* (2004) Phylogenetic overdispersion in Floridian oak communities. *American Naturalist* 163, 823-843
50. Fitzpatrick, M.*, et al.* (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range. *Global Ecology & Biogeography* 16, 24-33
51. Pearman, P.*, et al.* (in prep.) Successful prediction of plant species distributions across six millennia.
52. Silvertown, J.*, et al.* (2006) Absence of phylogenetic signal in the niche structure of meadow plant communities. *Proceedings of the Royal Society B* 273, 39-44
53. Saint-Laurent, R.*, et al.* (2003) Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). *Molecular Ecology* 12, 315-330
54. Filchak, K.E.*, et al.* (2000) Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* 407, 739-742
55. Murphy, S.M. (2004) Enemy-free space maintains swallowtail butterfly host shift. *Proceedings of the National Academy of Sciences of the United States of America* 101, 18048-18052

56. Schwarz, D.*, et al.* (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* 436, 546-549
57. RoelkeParker, M.E.*, et al.* (1996) A canine distemper virus epidemic in Serengeti lions (Panthera leo). *Nature* 379, 441-445
58. Ebert, D. (1994) Virulence and Local Adaptation of a Horizontally Transmitted Parasite. *Science* 265, 1084-1086
59. Grant, P.R., and Grant, B.R. (2006) Evolution of character displacement in Darwin's finches. *Science* 313, 224-226
60. Antonovi.J, and Bradshaw, A.D. (1970) Evolution in Closely Adjacent Plant Populations .8. Clinal Patterns at a Mine Boundary. *Heredity* 25, 349-&
61. Gregory, R.P.G., and Bradshaw, A.D. (1965) Heavy metal tolerance in populations of *Agrostis tenuis* Sibth and other grasses. *New Phytologist* 64, 131-&
62. Jain, S.K., and Bradshaw, A.D. (1966) Evolutionary Divergence among Adjacent Plant Populations .1. Evidence and Its Theoretical Analysis. *Heredity* 21, 407-&
63. Sambatti, J.B.M., and Rice, K.J. (2006) Local adaptation, patterns of selection, and gene flow in the Californian serpentine sunflower (*Helianthus exilis*). *Evolution* 60, 696-710
64. Kittelson, P.M., and Maron, J.L. (2001) Fine-scale genetically based differentiation of life-history traits in the perennial shrub *Lupinus arboreus*. *Evolution* 55, 2429-2438
65. Hairston, N.G., and Bohonak, A.J. (1998) Copepod reproductive strategies: Life-history theory, phylogenetic pattern and invasion of inland waters. *Journal of Marine Systems* 15, 23-34
66. Maron, J.L.*, et al.* (2004) Rapid evolution of an invasive plant. *Ecological Monographs* 74, 261-280
67. Sexton, J.P.*, et al.* (2002) Plasticity and genetic diversity may allow saltcedar to invade cold climates in North America. *Ecological Applications* 12, 1652-1660
68. de Swart, E.O.A.M.*, et al.* (1994) Experimental evaluation of realized niche models for predicting responses of plant-species to a change in environmental-conditions. *Journal of Vegetation Science* 5, 541-552
69. Wright, J.W.*, et al.* (2006) Experimental verification of ecological niche modeling in a heterogeneous environment. *Ecology* 87, 2433-2439
70. Guisan, A., and Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147-186
71. Rehfeldt, G.E.*, et al.* (1999) Genetic responses to climate in *Pinus contorta*: Niche breadth, climate change, and reforestation. *Ecological Monographs* 69, 375-407
72. Araújo, M.B., and Rahbek, C. (2006) How does climate change affect biodiversity? *Science* 313, 1396-1397
73. Benito Garzón, M.*, et al.* (2007) Predictive modelling of tree species distributions on the Iberian peninsula during Last Glacial Maximum and mid-Holocene. *Ecography* 30, 120-134
74. Araújo, M.B.*, et al.* (2005) Validation of species-climate impact models under climate change. *Global Change Biology* 11, 1504-1513
75. Walther, G.-R.*, et al.* (2005) An ecological 'footprint' of climate change. *Proceedings of the Royal Society B* 272, 1427-1432
76. Martinez-Meyer, E., and Peterson, A.T. (2006) Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. *Journal of Biogeography* 33, 1779-1789
77. Thuiller, W.*, et al.* (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography* 27, 165-172
78. Zar, J.H. (1996) *Biostatistical Analysis*. Prentice Hall
79. Randin, C.F.*, et al.* (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography* 33, 1689-1703
80. Hautier, Y.*, et al.* (in prep) Trade-off between sexual and clonal reproduction across altitude in *Poa alpina*.
81. Levin, D.A. (2004) The ecological transition in speciation. *New Phytologist* 161, 91-96
82. Levin, D.A. (2003) Ecological speciation: Lessons from invasive species. *Systematic botany* 28, 643-650
83. Mau-Crimmins, T.M.*, et al.* (2006) Can the invaded range of a species be predicted sufficiently using only native-range data? Lehmann lovegrass (*Eragrostis lehmanniana*) in the southwestern United States. *Ecol. Model.* 193, 736-746
84. Hampe, A., and Petit, R.J. (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* 8, 461-467

85. Botkin, D.B., *et al.* (2007) Forecasting the effects of global warming on biodiversity. *Bioscience* 57, 227-236
86. Vetaas, O.R. (2002) Realized and potential climate niches: a comparison of four Rhododendron tree species. *Journal of Biogeography* 29, 545-554
87. Holt, R.D. (2003) On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research* 5, 159-178
88. Lovette, I.J., and Hochachka, W.M. (2006) Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. *Ecology* 87, S14-S28
89. Ackerly, D.D., *et al.* (2006) Niche evolution and adaptive radiation: Testing the order of trait divergence. *Ecology* 87, S50-S61
90. Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press
91. Parmesan, C., and Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421, 37-42

Table 1 Studies of niche conservatism/shift 1996 - 2007.

Topic	System	Conclusions	Refs.
Character divergence	Theoretical	Character displacement is enhanced with some forms of predation	[47]
Character divergence	Insular birds	Food resource bill morphology change resulting from competition in a drought year	[59]
Host shift	Insect pest	Insipient speciation involves genotype x environment interaction	[54]
Host shift	Viruses/African carnivores	Closely related viruses find hosts in different carnivores	[57]
Host shift	Tephritid fruitflies	Hybrid speciation event associated with host shift	[56]
Host shift	Butterflies (*Papilio*)	Host shift supported by change in predation regime	[55]
Niche conservatism	Theoretical	Higher fitness of individuals in source habitats leads to niche conservatism	[40,87]
Niche conservatism	Genus *Fagus*	Reciprocal modelling of beech in N. America and Europe suggests niche conservatism	[3]
Niche conservatism	Plants and mammals	Non-random association of occurrences with projected range supports conservatism	[14,76]
Niche conservatism	Birds, mammals, butterflies	Sister taxa with adjacent distributions display niche conservatism	[2]
Niche conservatism	Theoretical	Gene flow critical to determining local adaptation and range changes	[45]
Niche conservatism	Plant communities	Niche conservatism shown by correlation in range size for disjunct congeners	[34]
Niche conservatism	Argentine ant	Niche modeling reveals similarity between ants in native and introduced ranges	[36]
Niche conservatism	Plankton diversity	Diversity in tropics explainable by higher speciation rates and niche conservatism	[38]

Niche shift	Plant phenology	Radiation to new habitats often accompanied by altered phenology (review)	[11]
Niche shift	Copepods	Niche shifts are adaptive responses to environmental variability	[65]
Niche shift	Plant invasions	Causal processes change during plant invasions (review)	[21]
Niche shift	Genus *Arabidopsis*	Recently evolved species generally show greater changes in climate space	[35]
Niche shift	Invasive species	Niche shift shown by lack of reciprocal predictability between native and invaded ranges	[33,50]
Niche shift	Plant communities	No phylogenetic signal to occupied plant niches suggests niche shifts	[52]
Niche shift	Introduced plant	Introduced populations rapidly adapt to conditions at local latitude	[66]
Niche shifts	*Pinus contorta*	Populations vary in optima of fundamental niche	[71]
Niche shifts	Bivalves	Shifts supported by species origin in warm waters and adaptation to polar regions	[42]
Niche shifts	Oaks (*Quercus*)	Niche characteristics and species co-occurrence suggest phylogenetic overdispersion	[49]
Niche shifts/speciation	Dendrobatid frogs	Divergence in climatic niche during speciation of Ecuadorian frogs	[5]
Niche size	Edaphic conditions	Niche size maintained by gene flow and strong selection	[63,64]
Niche size	Invasive plant/lab	Plasticity and niche evolution both contribute to invasive ability	[67]
Niche size	Edaphic conditions	Ecotypic variation in soil chemistry tolerance affects modeling of niche	[69]
Phylogenetic and ecological relatedness	Genus *Anolis*	No relationship between clade phylogeny and climate/habitat niches	[6,24]

Phylogenetic and ecological relatedness	New world warblers	Niche conservatism in recently diverged species, shifts at deeper branches	[88]
Phylogenetic and ecological relatedness	European plants	Phylogenetic distance and taxonomic levels explain niche positions	[1]
Phylogenetic and ecological relatedness	*Aphelocoma* jays	Pervasive niche evolution unrelated to phylogenetic distance	[10]
Range limits	Theoretical	Niche limits and demographic process interact to influence range limits	[31]
Species richness gradients	New World Birds	Niche conservatism, extinction in basal clades explain species richness gradient	[7]

Box 1.

A case study: Florida oak niches

Niche conservatism and niche shift are likely not a simple dichotomy. A recent study by Cavender-Bares et al. addresses the distribution of species niches and how this varies at different levels of a phylogeny.

Approach. Cavender-Bares et al. use ecological, physiological and phylogenetic information to address the co-occurrence 17 oak (*Quercus*) species that dominate forests in north central Florida [49]. Cavender-Bares et al. develop a phylogeny for these species using sequences from two nuclear genes. They also evaluate physiological and physical characters of the species as seedlings (in laboratory experiments) and as adults in the field. Additionally, they measure ecological habitat characteristics, including soil moisture and fire return interval, and use them to describe niche overlap among species. They test multiple null models by randomizing (1) the phylogenetic distance scores among species, (2) the distribution of species occurrences to sites, and (3) species' basal area contributions among sites. By comparing species co-occurrence frequencies and niche overlap at different levels of phylogenetic resolution, Cavender-Bares et al. isolate clades at multiple levels in which niche overlap and co-occurrence are either more or less extensive than expected.

Results. Cavender-Bares et al. find that species niches are over-dispersed among the major clades in the phyologeny. This indicates that niche shifts occur extensively within clades and that species within clades are less ecologically similar than would be expected by chance [49]. Nonetheless, species niche distribution within subclades at intermediate resolution showed greater levels of over-dispersion than occurs either at the species level or in the basal clades of the phylogeny. Further a negative correlation between phylogenetic distance and soil moisture preferences suggests niche convergence among distantly related species.

Similar results. Other studies find that patterns in North American wood-warblers (Parulidae) and *Ceanothus* (Rhamnaceae) shrubs and trees suggest niche conservatism in recent speciation but niche shifts deeper in the clades' evolutionary pasts [88,89]. The combination of climatic envelope models and molecular phylogenies of dendrobatid frogs to estimate ancestral states of niche dimensions suggests niche divergence and complementarity in association with speciation [5] (Figure 1d). Thus, recent studies document niche shifts, but niches are not widely divergent in some cases and the degree of shift varies with phylogenetic level within a clade.

Box 2.

A hierarchical framework for evaluating niche conservatism, divergence and drift.

There are at least three situations in which testing for niche conservatism/shift can contribute to improved understanding of ecological and evolutionary processes, and inform application of modelling tools for policy decisions, conservation planning, and management. Testing hypotheses concerning niche conservatism currently comprise: (1) comparisons among two or more populations within a species (e.g. populations of invasive species [33,50]), (2) exploration of collections of (multiple) pairs of sister species in a collection of taxa for niche divergence (e.g., in regions of allopatric speciation [2]), and (3) determination of niche relationships among multiple species within a clade (e.g., within ecological communities [24,49] or concerning a speciation scenario [5]). Thus, the hierarchy consists of comparing entities at increasing levels of biological organization, between populations, between sister species, among two (or more) extant species with hypothesized evolutionary relationships, and between two clades defined by interior nodes at a chosen level in a phylogeny. The common threads among hierarchical levels are the need for quantifying niche divergence and estimating the degree of (relative or absolute) divergence a study can detect (Fig.3).

The establishment of a null model is required to define the uninteresting or random result. With just a pair of taxa, the null model will often be that the taxa have not diverged (i.e., niche conservatism, a shift less than a specified effect size, generally requiring a one-tailed statistical test). With multiple related species, the null model will generally be that of random *niche drift*, a phenomenon with parallels at the community level[90]. There are two alternatives to niche drift: (a) niches have diverged less than expected from niche drift. This implies processes resulting in niche conservatism. On the other hand, (b) niches may have diverged more than expected (over-dispersed niches [49]), which suggests ecological and evolutionary processes that accelerate niche diversification. Both these phenomenon may occur at different levels of the same phylogeny [22,49,88]. One null model or hypothesis will not fit all circumstances. Testing for niche shift in an invasive species' new range may focus on separating signal from sampling error. Mapping of ecological characteristics onto phylogenetic patterns will require a neutral model of character divergence. Thus, before fully answering the fundamental question of whether or not a niche shift has occurred, one must address whether a shift is detectable given the data and tools used, whether the shift has been constrained or enhanced by potentially identifiable processes, and whether it has meaningful or influential consequences in a particular context, say for using niche-based models to predict invasions, for estimating the effects of climate change, or for understanding evolutionary patterns/processes.

Ten issues to consider

1. The hierarchical level(s) of the study has (have) been identified.

2. The niche definition is presented, for example as an edaphic, trophic, or climatic niche, and the definition's use is clearly justified.

3. The putative niche shift is described and discussed (fundamental vs. realized niche, change in species interactions, plasticity, genetic response to selection).

4. An appropriate approach - or combination of approaches – is/are identified to test for NC/niche shift, for example phylogenetic, genetic, species distribution models, and experimental approaches in the field and lab.

5. The variables used to quantify niche divergence are related to identified physiological or demographic processes and find support in the literature.

6. Potential problems of scale has been considered (e.g., Fig. 2; extent of distribution data must be sufficient to allow the full niche of the species to be captured).

7. Sources of potential error variability have been identified and, optimally, evaluated (e.g., sampling error, topological uncertainty in phylogenies).

8. A null model and a testable null hypothesis are developed a priori. They specify the role of all relevant random processes in generating the data.

9. A description of the potential outcomes of the study and their respective interpretations has been elaborated a priori.

10. A power analysis has been completed, with specification of the biologically important effect size, especially if the null result would be used to justify further application of a technique (for example, the use of species distribution models to predict the effects of climate change, a situation requiring NC).

Figure legends

Figure 1 Shifts of realized and fundamental niches. (a) illustrations of the fundamental niche (all ecological situations in which populations of the species can increase) and realized niche (i.e., ecological situations in which populations increase in the presence of competitors). Population growth rate is plotted against an environmental/trophic variable and the presence of populations is plotted along two environmental variables, indicating the same niche concept, with emphasis on population dynamic and ecological requirements respectively. Boxes b-c illustrate three different ecological situations in which niche shifts have been demonstrated, namely (b) invasive species ecology [33], (c) biological response to climate change [91] and (d) phylogenetic analysis [24]. The ecological and evolutionary processes involved in each step are named in italics and correspond to the arrows in the graphs.

Figure 2 Detection of niche shifts and niche estimation. (a) Detecting a niche shift can be sensitive to spatial resolution of data. Sampling at the larger grain fails to detect shift in temperature sensitivity because no new squares are occupied. (b) Niche estimation may be biased by failing to sample the entire ecological gradients along which a species occurs. Sampling in area of low incident radiation S2 shows that the probability of species presence is low in areas with low insolation. Example inspired by [84].

Figure 3 A statistical framework to test niche conservatism and niche shift—an example. The null (H_0) and alternate (H_1) hypothesis and the associated statistical test are defined. Threshold probability values for the rejection of the null hypothesis (NC) are specified. In the case when H_0 can not be rejected, further data sampling based on a power analysis allows estimating the risk of not rejecting H_0 if false (risk of Type-II error) and power. One has support for niche shift if H_0 is rejected. By contrast, niche conservatism (less divergence than a specified effect size) can only be argued if H_0 can not be rejected assuming *a priori* determined statistical power for the test. Two-tailed tests of the null hypothesis of niche drift are constructed similarly. Orthogonal environmental gradients are shown and the framework is adaptable to cases of non-orthogonal axes or other (multivariate) measures of niche position (i.e., marginality).

Figures

Figure 1

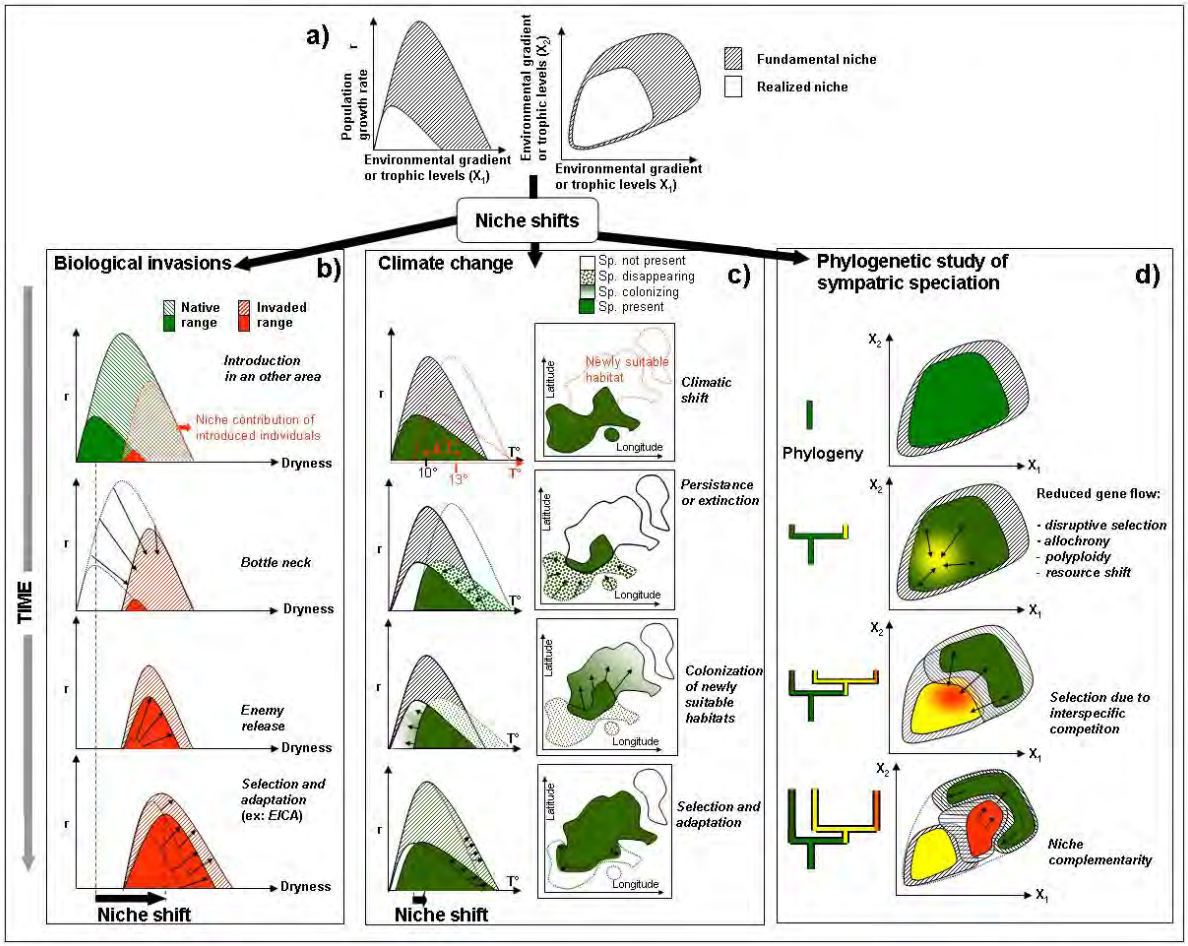


Figure 2

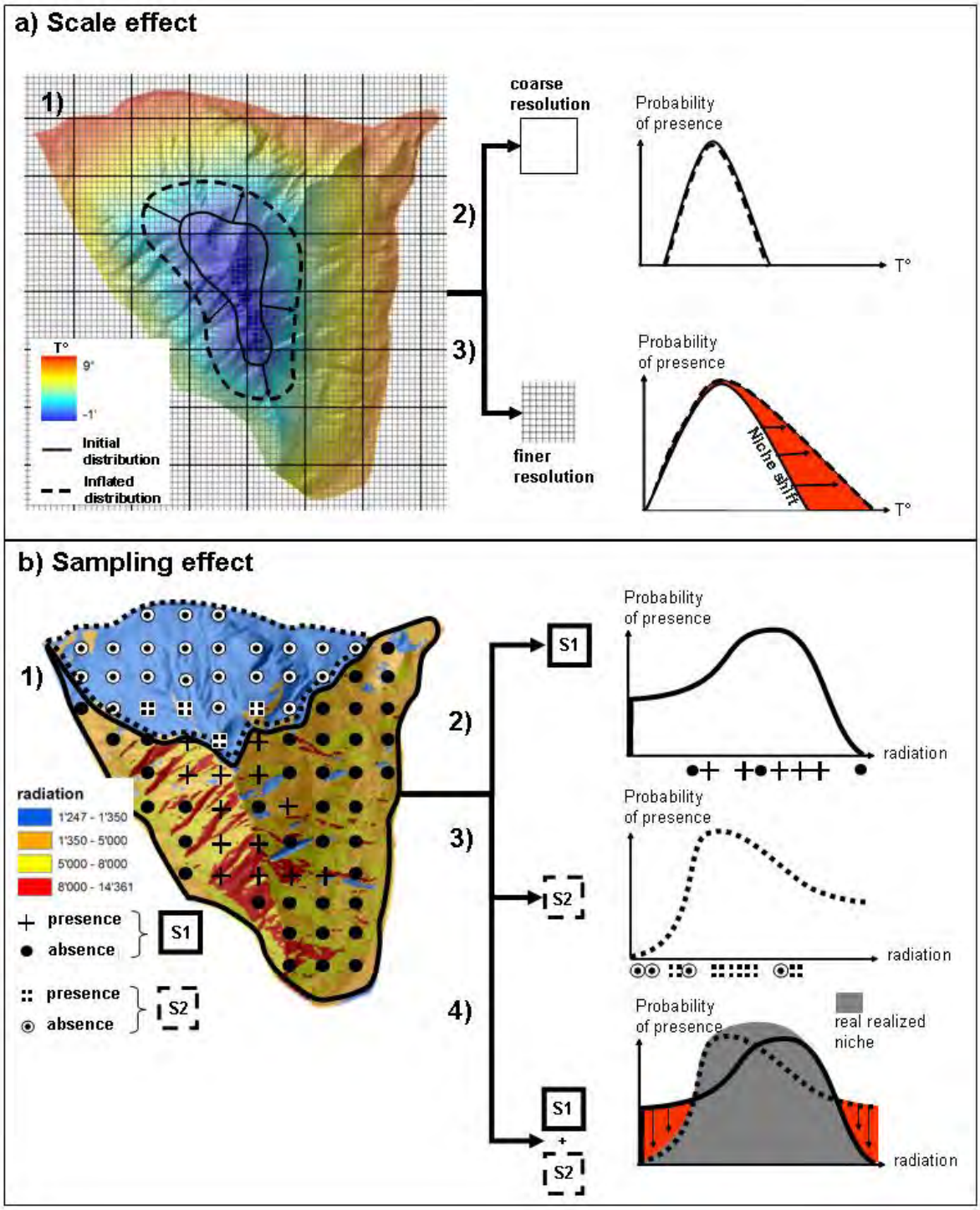
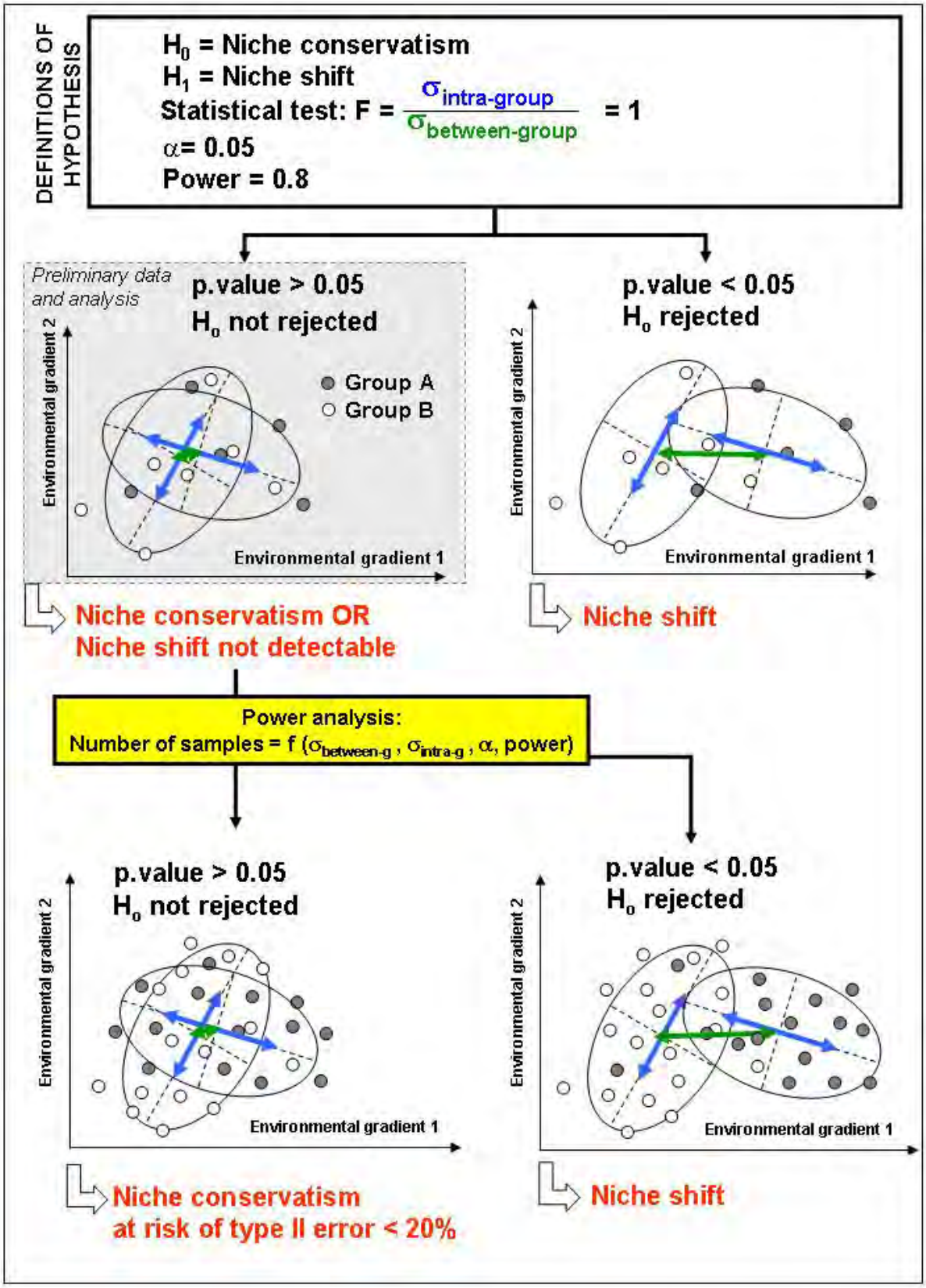

a) Scale effect
1)
T°
9°
-1°
Initial distribution
Inflated distribution
2)
coarse resolution
Probability of presence
T°
3)
finer resolution
Probability of presence
Niche shift
T°
b) Sampling effect
1)
radiation
1'247 - 1'350
1'350 - 5'000
5'000 - 8'000
8'000 - 14'361
presence
absence
S1
presence
absence
S2
2)
S1
Probability of presence
radiation
3)
S2
Probability of presence
radiation
4)
S1
+
S2
Probability of presence
real realized niche
radiation

Figure 3


DEFINITIONS OF HYPOTHESIS
H0 = Niche conservatism
H1 = Niche shift
Statistical test: F = σintra-group / σbetween-group = 1
α= 0.05
Power = 0.8
Preliminary data and analysis
p.value > 0.05
H0 not rejected
Group A
Group B
Environmental gradient 2
Environmental gradient 1
Niche conservatism OR
Niche shift not detectable
p.value < 0.05
H0 rejected
Niche shift
Power analysis:
Number of samples = f (σbetween-g , σintra-g , α, power)
p.value > 0.05
H0 not rejected
Niche conservatism
at risk of type II error < 20%
p.value < 0.05
H0 rejected
Niche shift

Remerciements

Je tiens d'abord à remercier toutes les personnes qui ont participé à cette aventure scientifique à travers les Préalpes.

Les campagnes de terrains des étés 2002 et 2003 ont permis de produire huit des dix chapitres de ma thèse. Ces campagnes ont été financées par le Centre de Conservation de la Nature du Canton de Vaud (SFFN-VD). Je tiens à remercier le SFFN et particulièrement Philippe Gmür et Cornelis Neet pour leur soutien inconditionnel dans ce projet dans les Préalpes vaudoises.

Le Fonds National Suisse de la Recherche Scientifique (FNRS) pour le financement d'une partie du projet MODIPLANT (fonds No 3100A0-110000).

Je remercie aussi le MRI (Mountain Research Initiative), représenté à Lausanne par son directeur Greg Grenwood, qui a financé le premier workshop du consortium Euromont et a ainsi contribué à la publication du chapitre 6.

Toutes les personnes qui ont participé aux campagnes de terrain entre 2002 et 2006 :

Pascal Vittoz, Stéfanie Maire, Dario Martinoni et Valentine Hof ont joué un rôle décisif grâce à leur grande expérience de terrain lors de la mise en place du projet en 2002.

Séverine Wydler, Chantal Peverelli, Thaïs Cornaz, François Bonnet, Julien Chautems, Guillaume Emaresi, Andrey Buryak et Elisa Terrière pour leur aide précieuse.

Pierre Bermane Favrod-Coune et Daniel Cherix qui ont mis à disposition leur chalet « camps de base avancés » pour nos campagnes de terrain en 2002 et 2003.

Anny et Willy Berra, pour leur accueil toujours chaleureux au Refuge Giacomini. J'ai passé des moments inoubliables sur le plateau d'Anzeindaz.

Les étudiants qui ont participé au projet MODIPLANT. J'ai bénéficié de leur enthousiasme et de leur regard neuf sur mes travaux de recherches : Dario Martinoni, Lorenzo De Stefani, Roxane Milleret, Virginie Favre, Grégoire Vuissoz, Stéphanie Spahr, Florian Dessimoz, Emilie Rathey, Jean-Yves Humbert, Hélène Jaccard, Yann Hautier, Jérôme Guélat, David Dubois et Marietta Wilhelm.

Niklaus Zimmermann, Stéphane Joost, Thierry Lassueur, Thomas Dirnböck, Stefan Dullinger, Massimiliano Zappa, Jean-Pierre Dedieu, Wilfried Thuiller, Signe Poulsen et Martin Beniston pour leur amitié et les passionnantes collaborations scientifiques.

Davnah Urbach, Gilles Yoccoz, François Bonnet, Martine Rebetez, Andrey Buryak et Christine Grossen : pour les passionnantes discussions (scientifiques), mais aussi pour les courses en montagne / à pieds, nécessaires à mon équilibre durant ma thèse.

Mes collègues de travail du Département d'Ecologie et d'Evolution de l'UNIL:

Pascal Vittoz encore une fois, pour son aide et ses conseils lors de la rédaction de mes articles et surtout pour toutes les connaissances qu'il m'a apportées.

Peter Pearman : sa vision scientifique, sa grande culture et son esprit cosmopolite m'ont permis de progresser lors des deux dernières années de ma thèse.

Gwenaëlle Le Lay et Thomas Czáka avec qui j'ai partagé le même bureau. Ils m'ont apporté un soutien moral et un savoir technique tout au long de ma thèse.

Thierry Renaud, Robin Engler et Olivier Broennimann, collègues et amis depuis plus de cinq ans.

Ramona Maggini Lehmann et Anthohy Lehmann, pour leur amitié et leur aide lorsque j'ai débuté ma thèse.

Véronique Helfer, Jérôme Pellet, Aline Dépraz, Dorothea Pio et Karen Osmont.

Michel Jaboyedoff, Hans-Rudolf Pfeifer, Alexandre Roulin avec qui j'ai eu le plaisir de collaborer durant ma thèse et qui m'ont permis d'élargir mon horizon scientifique.

Feli Jaquiéry, Corine Bolle, France Pham, Raymond Gruaz, Michel Schuepbach, Patrick Moratal, Joel Cavat, Blaise Pavillard et Franck Chalard qui m'ont apporté un soutien sans faille et à toute heure pour résoudre mes problèmes administratifs et techniques.

Jean-Paul Theurillat Pierre-André Loizeau du Conservatoire et Jardin Botanique de Genève, avec qui j'ai eu le plaisir de partager un peu de leur vision phytogéographique et phylogénétique de la botanique.

Pierrick Dentan et toute l'équipe de Yosemite pour leurs nombreux conseils techniques et logistiques nécessaires lors de mon travail de terrain.

Jean Lechim pour son enseignement qui a été essentiel durant les deux dernières années de ma thèse.

Le Professeur Pierre Hainard, Alexandre Hirzel, Bertrand Gentizon ainsi que Corine et Michel Marie : je n'aurais pas trouvé ma voie aussi facilement sans les avoir rencontrés.

Les Professeurs Nicole Galland, Jacques Hausser, Peter Vogel, Nicolas Perrin, Laurent Keller, Daniel Cherix et Philippe Heeb ainsi que Philippe Christ, Lucas Fumagalli et Michel Chapuisat : grâce à leur enseignement, je n'ai pas hésité à choisir l'écologie et la macro-biologie.

Le Professeur Jérôme Goudet pour sont soutient, ses suggestions scientifiques et sa passion pour les Préalpes qu'il aime partager.

Mes deux experts externes de Jury de thèse, le Professeur Christian Körner et Wilfried Thuiller : le temps imparti pour les questions lors de ma soutenance privée était trop court. J'espère avoir l'occasion de poursuivre ces passionnantes discussions.

Ma famille : ma mère Michèle, mon père Francis, mon frère François et sa compagne Gloria, pour leur soutien et leur enthousiasme autour de ce projet scientifique.

Ruth et Foudil Sidi-Ali : j'ai toujours trouvé un accueil réconfortant à l'Auberge de Sugnens après de longues journées de travail.

Ma compagne Karin, sans qui je n'aurais pas été au bout de ce long parcours académique.

Mon directeur de thèse, le Professeur Antoine Guisan : il s'est enthousiasmé pour chacune de mes idées et m'a aidé à toutes les mener à terme. Il a su créer une stimulation scientifique permanente au sein de son groupe de recherche qui m'a permis de ne jamais douter dans mon travail.